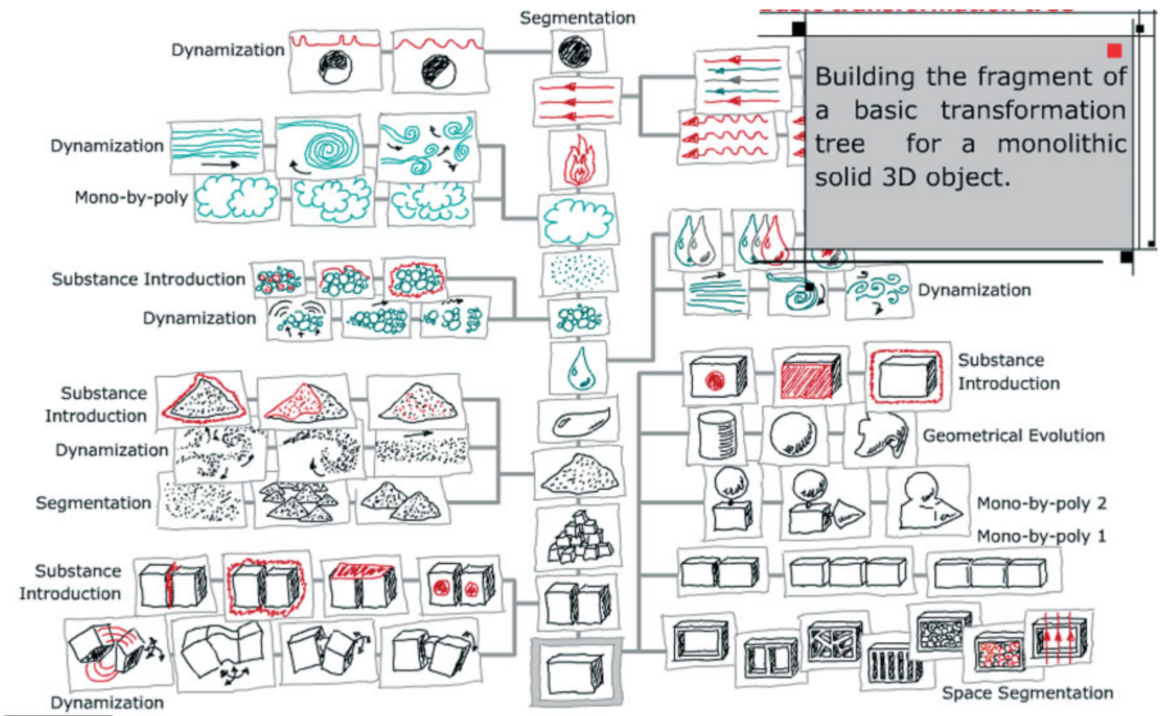
Nikolay Shpakovsky

Tree of Technology Evolution

Ways to New Business Opportunities

Experience of the SAIT TRIZ team
(SAMSUNG Advanced Institute of Technology)



By text of the Book

«ТРИЗ. Анализ технической и патентной информации и генерация новых идей»

Design and Layout by Arkadiy Pankevich

Recommended by Department of Theoretical Foundations of innovation of St. Petersburg State Polytechnical University

Reviewer **Alex Revenkov**, PhD, Assistant Professor of «Aerospace Engineering Manufacturing», Moscow State Aviation Institute (MAI)

Printed by CreateSpace, An Amazon.com Company

Contents

Preface 5

Chapter 1. **Technical Information Structuring Necessity** 8

1.1. Development of the invention methodology 8
1.2. The main propositions of TRIZ 11
1.2.1. Emergence of TRIZ 11
1.2.2. Technical system evolution laws 13
1.2.3. Increasing technical system ideality 14
1.2.4. Revealing and resolving a contradiction 23
1.2.5. Actions on models 28
1.2.6. Considering the psychological peculiarities of thinking 32
1.2.7. Using resources and an information base 33
1.2.8. Inventive problem solving process 38
1.3. Necessity for structuring an information field *39*

Chapter 2. **Objective evolution patterns** 43

2.1. Information organization in an evolution pattern 43
2.2. An example of real system analysis 53
2.3. The main components of the system 58
2.3.1. The functioning system model 58
2.3.2. The functioning system in the context of patent legislation 62
2.4. The functioning system evolution 64
2.4.1. Actions performed at a step-by-step evolution of a system 64
2.4.2. The three-stage algorithm of system transformation 69

Chapter 3. **Description of the Most Relevant Evolution Patterns** 75

3.1. The main evolution patterns of technical objects 75
3.1.1. The «Mono-bi-poly» transition 77
3.1.2. Trimming the system’s composition 78
3.1.3. Expanding — Trimming a System 81
3.1.4. Segmenting objects and substances 83
3.1.5. Evolution of surface properties 86

3.1.6. Internal structure evolution 88
3.1.7. Geometrical evolution 90
3.1.8. Dynamization 94
3.1.9. Increasing controllability 96
3.1.10. Increasing of action coordination 98
3.2. Specifics of constructing and using the evolution patterns 101
3.2.1. Evolution pattern construction recommendations 101
3.2.2. Specifics of use of some evolution pattern steps 103

Chapter 4. **Building Evolution Trees** 107

4.1. Elementary function: how to describe it? 107
4.2. Basic and Specific Evolution Trees 114
4.3. Evolution tree construction recommendations 122

Chapter 5. **Evolution Tree of Display** 124

5.1. Display definition 124
5.2. Producing an image on a surface 126
5.3. Structure of the Display Evolution Tree 128
5.4. From a static to a moving image. Cinema evolution patterns 130
5.5. Advent and evolution of the TV set 136
5.6. Flat display 144
5.7. Segmentation of a display as an object 161

Chapter 6. **Application of Evolution Tree** 182

6.1. Basic principles of work with the Evolution Tree 182
6.2. Search by marking the information field 188
6.3. Using the Evolution Tree for problem solving 190
6.3.1. Expanding the spectrum of preliminary concepts 190
6.3.2. Structural analogy method 195
6.4. New possibilities for circumventing competing patents 210
6.4.1. Main approaches to patent competition 210
6.4.2. Using of Evolution Tree for circumventing a real patent 217
6.5. Effective forecast of systems 224
6.5.1. Using Evolution Trees in forecast 224
6.5.2. Determining the main versions of system's components 228
6.5.3. Examples of prognostic solutions from the display Evolution Tree 230
6.5.4. Computer approach to the analysis of complicated systems 236

Author's afterword 238

Appendix EvolutionTree of display 239

Literature 247

Preface

Winning a battle requires meticulous planning. A commander having no information is blind. Who prepares a battle? Primarily, a reconnaissance unit and headquarters. Reconnaissance — the eyes and the ears of the army — gains information and the headquarters — its top brains — process this information. The need for information will never be satisfied, and the headquarters permanently requires more and more intelligence data.

How can this colossal array of information be analyzed? The military have an effective information organization model — a topographical map of a locality where a battle is to take place. The map is the basis; it displays all the gained information. It is just the map that allows seeing the entire site of a battle at a glance, evaluating the situation and forecasting the result of the forthcoming battle.

To find mineral resources, it is also necessary to collect and analyze much information. A classification structure is needed here as well. It is necessary to compile a map of a given region, plot all relief details and Earth crust structures on this map, determine and plot magnetic fields, mark the climate details, find out where lightning strikes most often during a thunderstorm and see when and what plants grow in the area under investigation. Then it is necessary to analyze all the known signs of occurrence of a mineral in one or another specific area. Thus, the maximally full map of a search area is the principal document while forecasting.

To win a competition, any company works at improving its product. The hard work done in the quiet of labs or at testers' benches results in new ideas, new knowledge which then forms the basis of a new, better product. A huge amount of patent and technical information is collected and processed in order to study the potential modifications of a given product available on the market then develop and manufacture a competitive product. This information also needs some classification structure — a map that would allow seeing all the existing versions of a product at a glance. And it would also be excellent if that wonderful map could tell as much as possible about future, unrealized versions of the product.

The role of this map could be performed by what we call an Evolution Tree. This book describes how to build and use an Evolution Tree. The methodology is based on the tools used in TRIZ — Genrich S. Altshuller's Theory of Inventive Problem Solving that proved efficient in solving complicated technical problems. We think that the essentials of this theory are effective for describing and analyzing different versions of technical systems. The fact is that in developing new machines, designers often deal with disembodied information. The Evolution

Tree-based information-processing method presented in this book allows you to effectively analyze large volumes of data and propose new interesting innovative ideas.

Patent and technical information structured into the Evolution Tree may serve to obtain a correct forecast of the company's product evolution, which will be used to make smart decisions concerning business development. Using the Evolution Tree offers good opportunities for circumventing patents of competing companies and protecting your own technical solutions, including building patent umbrellas. The Evolution Tree is effective everywhere where it is required to see an integrated evolution picture of a technical system, all its principal versions including the most unique and promising ones. This is extremely important in strategic planning of production, innovations and business.

The book is meant for a wide audience interested in or dealing with innovations, primarily for designers and engineers, as well as for students of technical higher education establishments involved in inventive activities. The material presented in the book will also be interesting to chiefs of enterprises, businessmen, and people promoting new goods and interested in their competitiveness.

The book can be especially interesting to patent attorneys. When working with patent and technical information, there arises a problem of structuring an effective analysis. The information obtained through patent search are descriptions of different modifications of the technical system. Analyzing such information is difficult so organizing it into Evolution Trees may be very helpful.

Using the Evolution Tree offers significant advantages at all information processing stages: search, analysis and production of new information. For example, during a search, we can identify the «information field», which immediately increases the retrieved information relevance. The Evolution Tree is simple to analyze due to the visibility and objectivity of the data it presents. The blank boxes of the Evolution Tree can be filled at the analysis stage by acting according to the special methods described in the book and explained with examples. The Evolution Tree also helps obtain new ideas and technical solutions, for example, by using the structural analogy method, as well as other innovation methods.

There are two main practical applications of the Evolution Tree — one for circumventing competing patents and the other for forecasting technical system evolution. Each are illustrated with many examples based on practical experience.

This book saw the light of day thanks to the benevolent and critical attitude of colleagues and friends.

During the initial period of TRIZ establishment in Korea, I was lucky to work in the TRIZ team of Samsung Advanced Institute of Technology together with Vasily Lenyashin and Kim Hyo June. It was a «dream team» which was first to prove the high efficiency of TRIZ at Samsung factories. It was there that the idea of the «Evolution Tree» approach appeared and it is still used at Samsung, largely through the efforts of MiJeong Song.

I would like to express gratitude to Peter Chuksin and Elena Novitskaya who not only helped develop the methods described in the book but also took direct participation in the

writing of sections 3.1.7 and 6.3.2. Many thanks to our Japanese colleagues — MRI (Mitsubishi Research Institute) employees who participated in the work on the first version and particularly to Yoshihisa Konishi.

The author also appreciates the time and efforts of all those who took part in the discussion and refining of the text of the Russian version of the book, first of all his colleagues from the «Invention Machine» project and Samsung corporation, as well as TRIZ specialists of the Belorussian TRIZ school: Georgiy Severinets, Leonid Bachilo and Alexander Skuratovich.

The author would like to particularly mention Viktor Timokhov who participated as a voluntary opponent whose well-intentioned, correct and exacting attitude towards the text helped make considerable improvements.

Special thanks to Viktor Baturin, a creative person capable of generating non-standard, surprising and effective ideas. He gathered a team of creative people, where the author also happened to work, and organized work on the projects that are most interesting and significant for Russia. The atmosphere of cooperation and open creative discussion characteristic of this team turned out to be very effective for work on the book.

Should some questions arise while reading this book, send them to triztrainer@mail.ru.

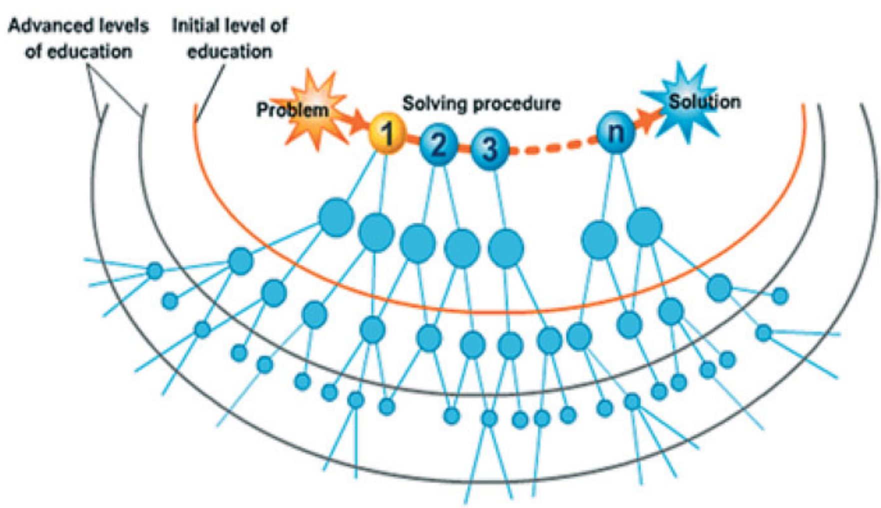
I wish you successful projects and strong inventions!

Nikolay Shpakovsky

Chapter 1

Technical Information Structuring Necessity

Effectively solving complicated inventive problems requires good knowledge of the methods and the ability to organize technical and patent information into a system, which in turn requires finding a principle of building a full and objective information structure.



1.1. Development of the invention methodology

The evolutionary history of technology which forms the foundation of human power is a continuous sequence of inventions, tangible results of humanity's obstinate efforts to seize the gist of the matter. A car, paper, mill, tea pot or a pen — all of them were invented and made by somebody for the first time.

The erection of Egyptian pyramids, construction of vessels, fortresses and military machines — all this required extraordinary wit and persistence. Think, for example, of the process of delivery and erecting of the monument to Peter the First in Petersburg. The monument foundation is a huge stone weighing about 1600 tons which was found in Finland. Gigantic collars, chutes with copper balls for reducing friction, an original method of loading on a vessel (scuttling the vessel, moving the stone onto it and then pumping water away) are but a small part of inventions made in the process. In addition, it was necessary to solve organizational problems: artificers worked the stone during transportation; business problems: crowds of people assembled to observe the giant stone transportation, for a fee, of course. Having no truck tractors, power vessels or modern lifting devices, Russian engineers managed to deliver the block to a distance of 22 kilometers, including 13 kilometers by water [1.1].

Creating a new machine is accompanied by solving numerous simple and complicated problems. Researchers often come to a standstill and a problem seems impossible to solve even by attracting all the then available force and means. To solve such a problem, one needs to step out of the limits of generally accepted ideas.

Imagine a Russian province early in the 19th century. It is necessary to transport a heavy steam-engine across a river. The bridge is not strong enough to allow the passage of heavy loads. A smart person was able to see that the steam-engine included boilers and an engine, that is, several large vessels. Closing the holes will make the machine able to float. No bridge is needed any more and the floating steam-engine can be towed to the opposite bank.

At the same time but in a different place, reshaping of a town square was started. A huge stone hampered the work, so it was necessary to remove it. They gathered horses, harnessed them and tried to dislodge the stone. But no such luck! The task exceeded the ability of the horses. Then quick-witted people understood that there was one more possibility and the stone had disappeared from the square by the morning. The stone was now lying several meters under the ground. They had pushed the stone down into the pit they dug at night.

Such «break-through» solutions are inventions. In the modern world, technology development is so rapid that the invention process has already become a flow production. Any engineer might face a very complicated technical problem which cannot be solved by traditional methods and he should be prepared for solving it.

Inventive activities are not reduced to solving technical problems. A company which is anxious to succeed must run ahead of their competitors. This requires the ability to forecast technology evolution and be first to introduce new solutions. That is, engineers have to solve not only urgent problems but also the problems of future technology.

Patent struggle is another aspect of competition. To feel confident in the market, every company must optimize the patent situation around its products. The company's patent service must be able to circumvent a competing patent by patenting those alternative versions of the product which are not covered by patents of other companies. It is also very important to protect the company's own technical solutions by organizing a kind of patent barrier around it, a so-called *patent umbrella*.

Thus, we have identified three types of problems faced by modern engineers:

- solving urgent technical problems,
- forecasting technical system evolution,
- searching for alternative solutions for patent circumvention and protection.

If there are problems, there should certainly be methods for solving these problems.

Here we run into a great problem: how to hand down not only knowledge but also experience and the art of solving complicated engineering problems to future generations? Traditional mastery techniques were propagated from a master to an apprentice in strict confidence, part of information was lost such as the secret of Damascus steel. Inventors themselves were unable to explain the process of creating new ideas. Certain information about this mystery was accumulated but, all the same, the appearance of new ideas was attributed to intuition, insight and other difficult-to-explain phenomena.

Meanwhile, the need for inventive problem solving methods was becoming increasingly urgent. Considering the technology history, one can notice an abrupt leap in its evolution which happened in the 18th century. That epoch of steam and electricity, rapid development of chemistry and all natural sciences which were themselves the result of many inventions required solving numerous inventive problems. Since the «Great Leap», no stagnation periods have been observed in technology evolution; it advances with a constant acceleration. Useful changes concern not only machines and technologies but also their creation methodology.

Let us acquaint ourselves with some methods of searching for new technical solutions.

The oldest method of finding new solutions to complex problems is searching for versions of possible solutions. «The search for versions method» applies to many different inventive processes. An experienced expert having a good knowledge of a problem and understanding its core must search fewer versions than a newcomer who faces the problem for the first time. This method is effective when the cost of machines and mechanisms is low and the production time is short.

Another well-known method is brainstorming. Here an attempt is made to involve many inventors in solving one problem. During brainstorming, any suggestions are allowed, or, according to Stanislaw Lem, one can «talk any nonsense», because suggestions are not judged in the course of a brainstorming session. Of course, the liberation of consciousness, absence of criticism, especially if the leader of the «brainstormers» is experienced and competent enough, give some positive result, but the effectiveness of this method can and should be improved.

W. Gordon developed the method called Synectics by introducing constructive criticism into the brainstorming method. Synectics employs techniques based on different kinds of analogies; it also requires serious training and psychological preparation of team members.

Other problem-solving methods are also known such as the method of focal objects, morphological analysis. Similar methods are also being developed nowadays. All of them have some positive core and could be improved if their authors and adherents could come to an agreement. This is, however, very doubtful because each author thinks that his is the only true method and considers the rest as a kind of potential useful supplement to his own method.

A reader familiar with the technical system evolution understands that the variety of methods should be inevitably replaced with a single method of a higher level, with an applied «science of inventing». This is true and is proved by the entire evolution course of human intellect: a set of weakly interrelated pieces of knowledge about some object area is gradually enriched, structured and, finally, some theoretical kernel occurs, alchemy turns into chemistry, healing into medicine and astrology into astronomy.

1.2. The main propositions of TRIZ

1.2.1. Emergence of TRIZ

The Theory of Inventive Problem Solving (TRIZ) has become just this sort of science integrating all inventive methods. At present, TRIZ shows the highest efficiency in revealing and resolving contradictions occurring in the course of technical system improvement. TRIZ-based methods imply active use of inventive algorithms and psychological operators aimed at the development of a special thinking style [1.2].

TRIZ generalizes the experience accumulated while creating millions of inventions and the multi-century technology evolution history generously offers amazing examples of solving inventive problems, i.e. problems underlying an aggravated contradiction.

As distinct from his predecessors, G. S. Altshuller was one of the first who set and successfully solved the following problem: *how can a method be found for rapidly reaching the area of effective solutions avoiding numerous trials and errors?*

The search for a solution to this problem was carried out along three directions:

1. Analyzing a great number of patents for effective inventive solutions and revealing the main contradiction-resolving principles used by inventors.

2. Studying the philosophy of resolving contradictions in nature and society, primarily dialectical materialism.

3. Because no methods work outside human consciousness, the third direction was studying the psychological foundations of creativity. The first article devoted to the control of thinking appeared in the «Voprosy Psyhologii» (Problems of Psychology) magazine [1.3].

All this work resulted in constructing a strong and convenient bridge across the abyss separating «high» sciences — philosophy that answers all questions and psychology that understands how to organize human consciousness for making its work more productive. This fulfilled the hopes of engineers, designers or researchers who seek to obtain and realize new ideas.

TRIZ fundamentals were created in 1950-1980 by numerous enthusiast inventors headed by G. S. Altshuller [1.4, 1.5]. Genrikh Saulovich did a titanic work: he laid the foundations of the theory itself, developed various theory application methods and created schools which educated thousands of inventors despite of direct opposition of Stalin's camps or indirect opposition of «invention» authorities in the Soviet bureaucratic system.

The main benefit of TRIZ is organizing the problem-solving process. Instead of waiting for an «insight» or a chaotic search of variants, it offers a preset sequence of actions.

Basically, all problems, regardless of their complication degree, are solved in a similar manner in the human mind. A problem formulated on an object level is immediately transformed by our mind into a kind of abstract model which then undergoes transformation on an abstract level and shows itself as a specific technical solution (Fig. 1.1). An everyday reasoning chain may look as follows: «feet are cold — feet and frost — something warm around the feet — putting on winter boots». When solving most everyday problems, this reasoning is impossible to notice as these thoughts dash swiftly through our mind until there appears a problem containing a strong contradiction: «The feet are cold; the boots are small or are absent at all — what is to be done?» If it happens when you are in town with numerous shoe shops the problem is easily solved. But what's to be done if cold caught you in a snow-covered forest? The contradiction is abruptly aggravated, it is necessary to find some nontrivial way out.

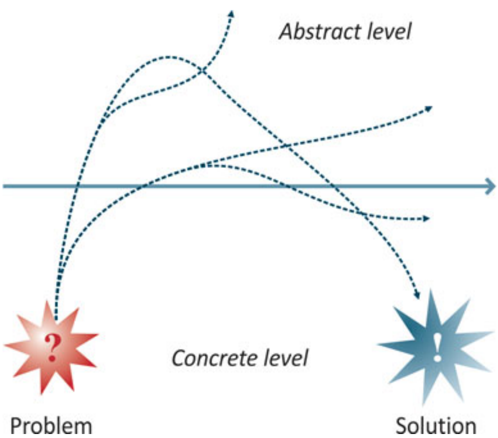


Fig 1.1. Model of thinking operations

Or, for example, to place several units of a machine in a limited space where they can't fit you must solve a technical problem. With the traditional thinking, we will start to construct a chain: «To place all — units and space — no space...» which will never come to an end. This is where reasoning starts and some abstract images are generated in mind; time is wasted waiting for an insight or «an apple falling on the head». The more complicated the problem and the more urgent the necessity of solving the problem, the harder the process.

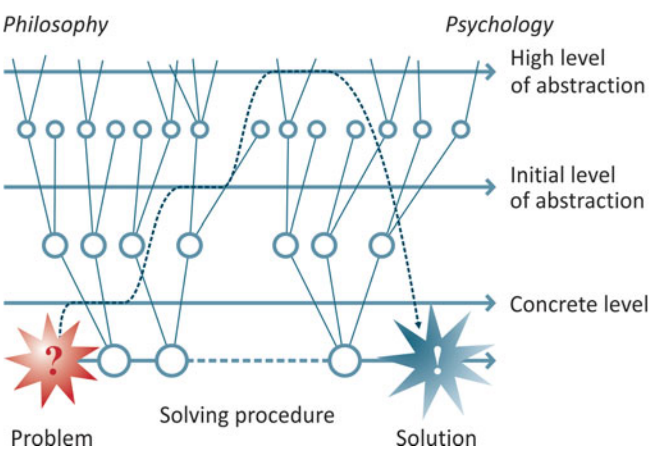


Fig. 1.2. Problem solving by using TRIZ

In principle, a man familiar with TRIZ and understanding how to use it reasons in the same manner while solving a problem. However, his decided advantage consists in that his consciousness does not «float adrift» in the abstract field of thinking, but has definite landmarks and prompts at his disposal, which allows him to find a solution (Fig. 1.2). These prompts are based on the knowledge of the problem solving philosophy, use of concentrated experience of many generations of inventors and understanding of our brain functioning peculiarities. Depending on the problem complexity, moving to a deeper level of understanding both the problem-solving process itself and the sense of actions performed at each specific step of this process is possible.

The TRIZ training system is organized to teach a student:

- the problem-solving technology which allows generating relevant associative prompts helpful in finding a solution,
- the ability to catch and enhance these prompts which often flash across one's mind and may remain unnoticed without special training.

Young as it is and still being elaborated, TRIZ has already acquired very effective tools for analyzing an inventive situation and finding solutions to the better part of complicated problems. TRIZ is based on a series of theoretical provisions:

1. All technical systems evolve not in a random manner but in accordance with objective laws; these laws may be used to elaborate methods.
2. The basic law of technical system evolution is the law of increasing ideality of technical systems.
3. Any system evolves through accumulating and resolving internal contradictions.
4. All actions aimed at technical system transformation are performed on mental models of technical systems.
5. In solving inventive problems, it is necessary to take into account thinking process peculiarities.
6. Technical systems are transformed at the expense of resources.

Let us consider each of these provisions.

1.2.2. Technical system evolution laws

TRIZ employs 9 laws of technical system evolution suggested by G. S. Altshuller [1.6]. They are divided into three groups which were given the conventional names of «Statics», «Kinematics» and «Dynamics». A special place in TRIZ belongs to other two laws: «The law of S-shaped evolution» and «The law of the motive force of contradiction» which are not so much laws as the description of specific mechanisms.

TRIZ development has not changed the set of the constituent laws but their content has deepened. The set of laws forms a certain system. The first part of this system, «Statics» provides three conditions of technical system existence, the second part points to the trend and character of system evolution and the third part supplements the first two.

Table 1

1. The law of system completeness 2. The law of «energy conductivity» in a system 3. The law of harmonization	Statics
4. The law of increasing ideality of technical systems 5. The law of irregularity of system's parts evolution 6. The law of transition to the supersystem	Kinematics
7. The law of transition from macro- to microlevel 8. The law of increasing the degree of substance-field interactions 9. The law of increasing dynamicity, controllability and replacing a man	Dynamics

The book gives a more detailed consideration to the evolution laws and their manifestation in the transformation of a given technical system and now we can say the following. For a system to be reliable, productive, economical, and, finally, competitive, the ideality of each of its subsequent versions must be higher than that of the preceding one. Thus, the basic evolution law is the law of increasing ideality.

There exists a good deal of methods for producing new concepts, new technical solutions [1.7]. However, finding a *new* solution is not enough. It is necessary to find a *new solution which would be effective in a given situation* and which would be *better, more ideal* than a predecessor.

1.2.3. Increasing technical system ideality

The notion of *Ideality* is widely employed in TRIZ for evaluating obtained solutions and is understood as the ratio of a complex parameter characterizing the performance of the given system's function to the system's operation cost [1.8]:

I=F/C, where:

I is the technical system's ideality;
F is the complex parameter characterizing the performance of a useful function;
C is the cost of performing this function.
Based on this formula, the system's ideality coefficient may be increased by two methods:

- Increasing the technical system's functionality by increasing the number and improving the quality of the product produced by this system and/or by providing performance of additional useful functions and production of additional useful products. In so doing, the system's functionality growth should outpace the growth of the costs for performing these functions.
- Reducing the cost of performing the system's useful functions. This is achieved by trimming its structure, removing auxiliary systems, reducing the cost of production, operation, maintenance and utilization. In this case, the number of the system's useful functions

should not be reduced and the quality of performing these functions should not be lowered. In addition, costs may be reduced by patent provision of system existence. For example, circumventing a competing patent may exclude the company's expenditure connected with buying a license.

Over time, Ideality tends to infinity. There arises a situation which is referred to as the «ideal technical system», when the system operation expenses may be neglected compared to the useful action it performs.

For example, a fan used to cool an automobile engine should only be switched on at the moment the engine overheats, its operation at any other time being useless or even harmful. There are devices controlling the fan operation but all of them imply introduction of new components into the system.

How can the ideality of this mechanism be improved?

In this situation, the ideal final result is formulated like that: «The fan blades themselves, without any additional mechanisms, start feeding air when the engine gets overheated.»

The construction corresponding to this formulation looks like that: the fan has blades made of shape-memory material. The blades constantly rotate, but at low temperature they are arranged in the same plane so that no air is fed. On heating above certain temperature, the blades turn and are set at an angle. The fan starts to feed air.

Increasing the system's ideality coefficient according to the above formula may be accompanied by the occurrence of the following three types of problem situations:

- the number and the performance quality of the system's functions suits us, but the expenditures connected with the performance of these functions are unsuitable;
- the number of functions performed by the system does not suit us, it is necessary to provide performance of additional functions;
- the number of functions performed by the system suits us, but the function performance quality is unsuitable.

The first type of problem situation occurs when the costs (denominator) of the ideality formula has an unacceptably high value. The second type is due to a too low functionality (numerator). The problems of the third type arise when both the components of the ideality formula are not satisfactory, i.e. the technical system's productivity is insufficient and the expenditures connected with its functioning are too high.

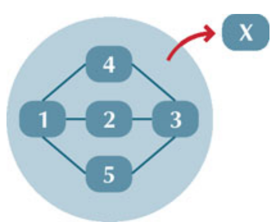
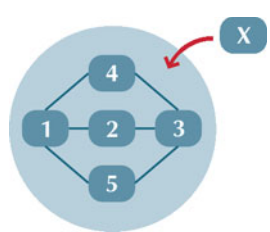


Fig. 1.3. a) Removing an object from a system Fig



1.3. b) Introducing an object into a system

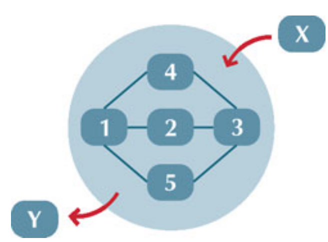


Fig. 1.3. c) Replacing an object

Depending on the situation type, three system improvement methods may be used:

- *Trimming a system*. Trimming implies removal of one or several components from a system (Fig. 1.3, a). The following requirements should be satisfied in this case: «The component should be absent in the system but its function should be performed by means of transferring it to other components of the system».
- *Expanding a system.* When expanding a system, new components — carriers of necessary functions — are introduced into its structure (Fig. 1.3, b). This may be accompanied by the occurrence of contradictory requirements of the system's components.
- *Optimizing a system.* Optimizing a system implies a complex action: both introduction of correcting systems which improve the performance quality of the main useful function and trimming these systems, passing their functions to the components already available in the system (fig. 1.3. c).

In each of the three cases, it is necessary to resolve contradictions resulting from the system improvement.

Trimming a system

Trimming is always directed at simplifying a system and reducing its cost while preserving the number of functions and their performance quality. This is achieved by removing some components which are determined as optional in the course of *value-engineering analysis* [1.9, 1.10].In that case, the functions of removed components are passed on to the remaining components.

The trimming method suggests the following order of actions.

First, the composition, structure and function of the system's components to be improved are determined. For that end, structural and functional analysis of the system is carried out and all the functions are grouped into main, auxiliary and unnecessary ones. Further, a structural-component scheme of the technological process realized by this system is built. After that, a score is given to each component. It depends on the importance of the function performed by a given component as well as on the manufacturing and operating cost of this component. Components having a high score are first candidates for removal from the system.

After finishing the system analysis, a creative stage starts — the technological process of trimming itself. The trimming is aimed at removing all the components — carriers of unnecessary and auxiliary functions and also, if possible, trimming the components — carriers of the main functions. The auxiliary and unnecessary functions should be eliminated while the main functions should be passed on to the remaining components of the system. The following standard formulation is used:

«It is possible not to perform a function if it may be realized a) by previous operations; b) by following operations».

After removing a number of components, a functionally ideal model of the process is built. The next step is building a real construction of the improved product based on this model.

When trimming, it is important to think inventively, to avoid psychological inertia. The more decisive the actions taken by the team working on the problem, the more radical are the changes experienced by the system. Problems occurring in the process usually contain aggravated contradictions which may be solved by using TRIZ tools and other innovation methods. Acting with discretion and only using partial trimming will cause less complicated problems, but will not lead to essential transformations of the system.

Some middle, compromise level of trimming may be selected, but it is better to perform trimming as if on two levels: rational, aimed at system transformation optimal under given conditions, and ideal aimed at obtaining a system possessing high-degree of ideality.

Example of Trimming could be simplifying of design of Ice-maker

Every refrigerator has a device for freezing water into ice cubes(Fig. 1.4). Its simplest version is a tray with partitions separating it into cells. Water is poured into the tray and then the tray is placed in a freezing compartment. After a while, ice cubes are ready and they may be taken out and put into an accumulating vessel. The tray is filled with a new portion of water.

An ice-making machine of a large expensive refrigerator operates practically in the same way, but the process itself is automated. Such a refrigerator generally has two adjacent compartments. One of them is a high and narrow freezing chamber while the other one is the refrigerator itself. An ice-making machine is mounted in the freezing chamber.

There is a cellular tray located in the upper portion of the ice maker. The tray is filled with water, the timer leaves some time for the water to freeze. After that, a special motor with a worm reducer overturns the tray. There is a pusher attached to one side of the tray. When the tray is almost turned over, its second side abuts against a special protrusion. The tray is skewed so that the ice cubes detach from it and fall down into a collector. The process recurs until the collector is filled.

To get a portion of ice, put a glass in a special niche in the refrigerator door and to push a lever. This makes the rod conveyor located in the collector rotate so that ice cubes start to fall into the glass. In some refrigerator models, ice cubes may be crushed into ice pellets.

It looks all right but not all ice cubes are immediately used. Because ready ice cubes need to be stored for some time, an undesirable phenomenon occurs: opening the door allows warm air to get into the freezing chamber and to cause slight melting of the ice cubes. After closing, some cubes may freeze together forming a monolith which is not easy to divide. That is why the rod conveyor which supplies ice cubes performs one more important function. Roughly once an hour the screw rotated by a high-power electric motor turns in the reverse direction, thereby destroying the freezing cubes. This operation is accompanied by a loud rumble.

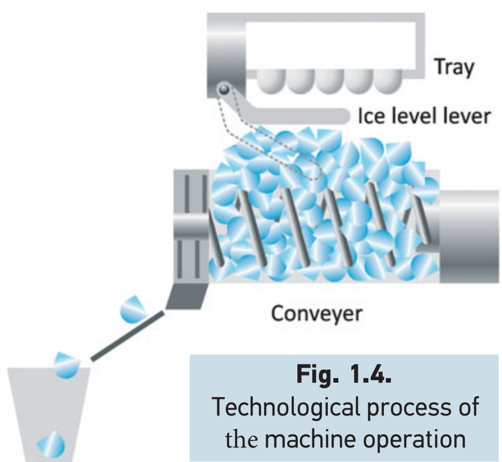


Fig. 1.4.
Technological process of the machine operation

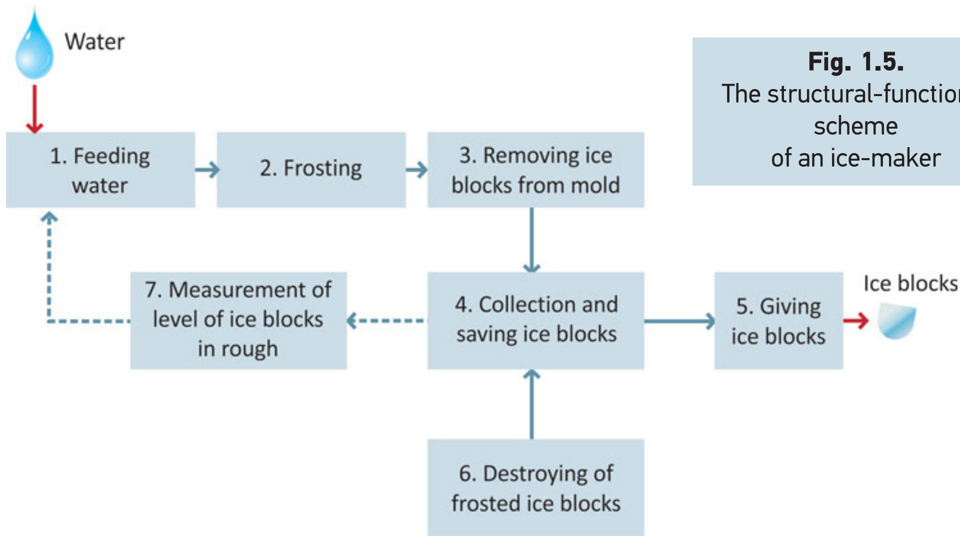


Fig. 1.5. The structural-functional scheme of an ice-maker

Such an ice-making «mini-factory» is very complicated and expensive both in production and in operation. In addition it occupies 1/3rd of the freezing chamber volume consuming the precious internal space.

Trimming was aimed at cheapening the ice maker, reducing its size and energy consumption. Functionally ideal modeling (one of the VEA sections) was selected as a trimming method [1.11].

After building the structural-component scheme of the technological process (Fig. 1.5), analyzing the functions and revealing the undesirable effects related to each component of the scheme, it became clear that the most expensive and «troublesome» component of the system is the ice cube collector.

After analyzing the possibility of trimming technological operations at the expense of previous or subsequent operations, we arrive at a paradoxical conclusion that removing the operation «Regular breaking of frozen cubes» and its carriers — the collector, rod conveyor and electric motor with a reducer — does not require taking the ice cubes out of the tray. To prevent their freezing together, each cube must lie in its cell as a cartridge in a clip until it is used. Then the functionally ideal model of the process will be very simple (Fig. 1.6)

Mutually exclusive requirements are made of the tray: it must occupy little space but the number of ice blocks it contains must be great enough to satisfy the user's needs; ideally, the number of ice cubes must be indefinitely great. How can this contradiction be resolved?

They say that if you want to grant the finite with the infinite property, pass it round, i.e. we need a kind of rotating drum or machine-gun belt rather than a clip. Connecting the ends of such a belt will give a kind of cellular conveyor with ice cubes (Fig. 1.7). Such a conveyor may be placed directly near the

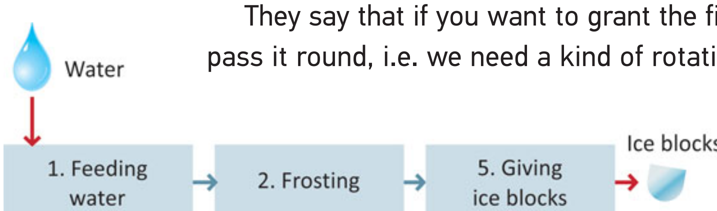


Fig. 1.6. A functionally-ideal model of the ice-making process

freezing chamber ceiling or, better, in the form of a thickened shelf in the upper third of this chamber, where an ice cube feeding device is usually located in a refrigerator.

Now we must solve how to extract ice from the cells. We immediately imagine some rotating extractor which forces out ice cubes with its protrusions (Fig. 1.8). However, this solution reduces our conveyor's ideality. It would be more expedient to find a simpler ice-extracting method, without using an extractor.

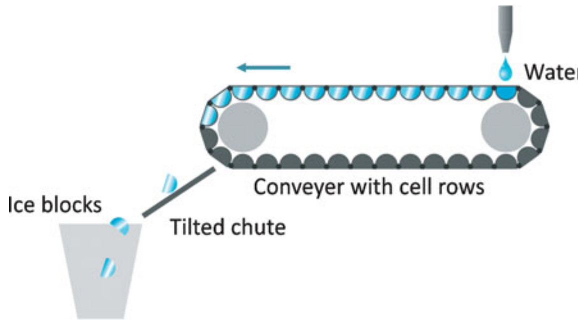


Fig. 1.7. A conveyor-type ice-maker scheme

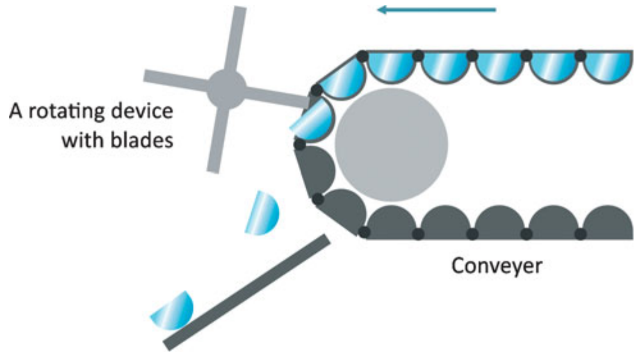


Fig. 1.8. An ice-extracting device

It is appropriate to remember that ice cubes are extracted from the tray of the initial ice maker by skewing. This analogy suggests that it is necessary to create such conditions under which some cells are also skewed when removing ice cubes. Such conditions are easy to provide: cells should be arranged in rows at an angle to the motion direction. When such a cell row runs into a roller, each cell will deform and the cubes will detach from the cell walls. Now all that remains is to prevent the cubes from falling out too early, for example, by mounting a support plate; this will provide a successive supply of cubes during the conveyor movement (Fig. 1.9).

We think of an ideal ice-maker as of a kind of tube having an ice bar inside (Fig. 1.10). When necessary, the bar is moved forward and a crusher separates bits of a required size from it. As the bar is moved forward, the vacant space formed behind is filled with water.

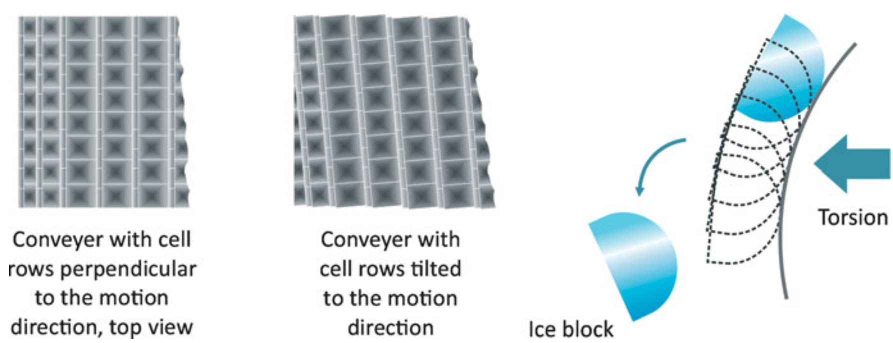


Fig. 1.9. Removing ice bits from a conveyor with a tilted arrangement of cells

Fig. 1.10. An ideal ice maker

Resolving the contradictions which will inevitably occur during the realization of this idea will produce an ice maker having a high degree of ideality.

Expanding a technical system

Expanding a technical system suggests introduction of new components into its structure. It is obvious that such complication of the system is only justified by increasing its ideality, i.e. the functional capabilities of a system grow faster than their realization costs [1.12].

When making a new expanded technical system, it may be supplemented with one or several systems performing *additional functions useful for consumers.* The combined system has several working tools each of which process objects differing in properties.

Such a structure may be illustrated by a personal computer. During its evolution, a computer was equipped with additional peripheral devices. The system was supplemented with detachable memory, speakers, a printer, a scanner, a modem, a webcam and other devices performing additional functions.

It often happens that combined systems unite contradictory properties: the system as a whole is universal because it is designed for treating several different objects, but each working tool may only treat one object. In such systems, a contradiction is resolved either in time, in case of a sequential performance of different operations (replaceable working tools), or in space, in case of their parallel performance (multi-functional devices).

When expanding, a system may be supplemented with new components that add functions thus improving the performance quality of the system's main function. Systems become more complicated and effective, their reliability, service life, and safety grow.

Examples of expanding can be:

A) Automotive brake.

The simplest brake used in first automobiles comprised a fixed braking element (a shoe or a band) provided with a mechanism that pressed it to a rotating shaft. Today, the braking system is expanded and includes brake-cooling and brake pad wear control devices, anti-lock brake system and many others. Of course, all these improvements complicate the initial system and make it more expensive. They, however, improve the performance quality of one of the automobile's most important auxiliary functions — «Stopping the automobile». From the market viewpoint, this repays with interest the braking system's increased cost.

B) Computer mouse.

The cordless mouse manufactured by ClickNJoy will not slip out of your hand even in the hottest weather. There are slits in the upper portion of its casing so that the fan mounted inside can supply air to cool your palm. Several cooling modes are provided.

Introduced systems may also be *doubling* such as the emergency engine-starting system of some aircraft or an emergency parachute. In addition to the main system, *correcting* additional systems are also introduced for removing some harmful effects arising in the main system operation.

C) Engine cooling.

The example of such a system may be the cooling system of an internal combustion engine. First internal combustion engines did not have any cooling capabilities, used to overheat and quickly went out of order. A correcting cooling system was introduced into the engine system. One type of such a system was air cooling: a rotating impeller supplying cool air onto the cylinders. The engine performance reliability increased at once. Because it was an additional system that increased the engine cost, the work aimed at trimming this system started. An injection cooling system that uses the exhaust gas energy is an example of a partially trimmed system. Gases escaping from an exhaust pipe pass through a special jacket to pull cool air in around the engine cylinders. An engine having a heat-resistant ceramic cylinder block is an example of a completely trimmed cooling system.

Combining competing or alternative systems, i.e. systems performing the same function by different methods, may be considered as a special case of expanding system's functions while optimizing its composition [4.11]. All transformation versions obtained during a patent search are alternative systems, that is why this methodology is valuable for dealing with information organized into an Evolution Tree.

The main idea of the methods of combining alternative systems consists not in their mechanical coupling but in a *transfer of the properties* of a more advanced system to an item being improved. To do this, it is necessary to take one of the systems as a basis, provide it with necessary properties of one or several alternative systems and resolve the arising contradictions. The least expensive system modification is usually selected as an example.

D) Washing machine.

Here is an example of combined alternative systems: an automatic washing machine.

A washing machine is a tank filled with a washing solution (water and detergent) into which clothes are placed. During washing, the solution in the tank is activated either by rotating the tank itself or by using special blade agitators. Then the solution is drained and the clothes are wrung out. To do this, the tank starts to rotate quickly and the centrifugal force removes the remaining water.

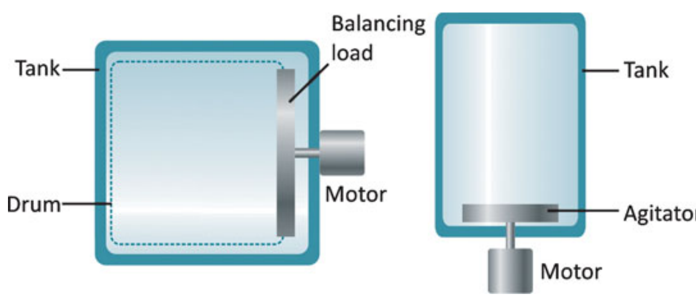


Fig. 1.11. A washing machine structure: a) with a horizontally arranged drum; 6) with a vertically arranged drum

The variety of washing machines may be reduced to two main types:

machines where the axis of revolution of the tank with clothes is horizontal (Fig. 1.11, a) and machines where the tank rotates about a vertical axis (Fig. 1.11, b). Let us present the advantages and disadvantages of both types in a table:

	Washing and rinsing	Wringling water
Washing Machine with horizontal drum	+	—
Washing Machine with vertical drum	—	+

Properly speaking, we have two alternative systems: what is good in one type of machine is bad in the other and vice versa. As we already know, it is necessary to combine them in such a way that the new system only obtains the positive properties of both alternative systems.

- The main contradiction inherent in a washing machine may be formulated as follows:
- the axis of revolution must be horizontal to provide high washing and rinsing quality;
- the axis of revolution must be vertical to provide good balancing while wringing water.

This contradiction can be resolved by allowing additional freedom of motion. Let us imagine an ordinary vertical washing machine hinged to an additional frame (Fig. 1.12). To wash, turn the machine on one side making the tank's axis of revolution horizontal. To extract the water, the washing machine should be reset. The additional frame and the tilting device complicate the machine construction to some extent, but then a blade agitator and a balancing device may be removed.

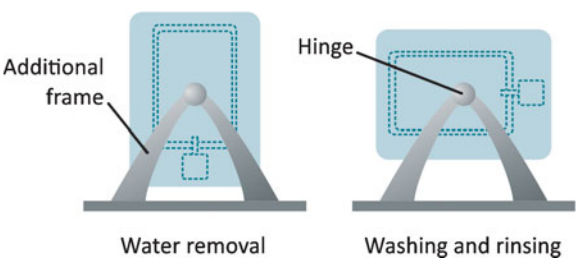


Fig. 1.12. A scheme of turning a washing machine

Trimming the obtained construction in accordance with requirement of Ideality, i.e. removing some unnecessary components from it and optimizing the remaining ones will bring us to a paradoxical conclusion: a washing machine must be spherical in shape (Fig. 1.13).

A spherically shaped tank with water and clothes is mounted on a base containing electric motors and water-supply pipes. There are three or four rollers protruding from the base, each being actuated by its own electric motor. The rollers can spin the tank in different directions and turn it from a vertical to a horizontal position and vice versa. Fig. 1.14 shows the machine positions in the main operation modes.

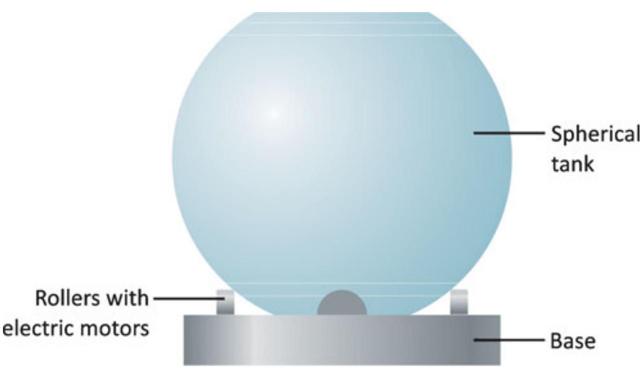


Fig. 1.13. A scheme of a spherical washing machine

The rollers are arranged in parallel and slowly rotate in the same direction so that the tank turns from a vertical position to a horizontal position (Fig. 1.14, a). For the tank's axis of revolution to pass through the washing machine loading door during washing and rinsing in a horizontal position, it is necessary to turn the rollers through 90 degrees (Fig. 1.14, b).

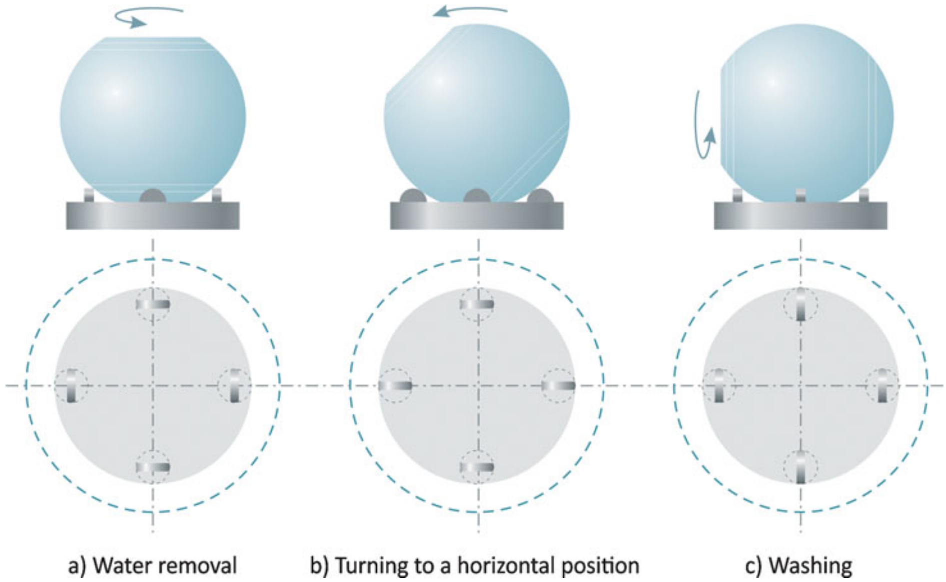


Fig. 1.14. Live rolls position variants

Setting the rollers in such a manner that their axes of revolution cross in the center of the circle (Fig. 1.14, c) allows spinning up the spherical tank for wringing the clothes. Because during rotation water will be pushed to the sphere equator, it is necessary to provide an additional net-like frame for clothes and rotate the drum with stops so as to allow the accumulated water to drain away.

1.2.4. Revealing and resolving a contradiction

Increasing ideality is the general direction of the technical system improvement. However, it is not always easy to achieve and a situation called *contradiction* occurs as a result.

TRIZ defines a contradiction as a conflict between human wishes and a real situation and describes several types of contradictions, the main being: *administrative, technical and physical* contradictions [1.14].

The *administrative contradiction* appears at the top level of a problem because improving a situation requires doing something but the way of doing it is unknown.

The *technical contradiction* is such a situation when improving one operational parameter causes an unacceptable worsening of another parameter. The technical contradiction can be formulated as follows: «Improving the parameter **A** causes unacceptable worsening of the parameter **B**».

The *physical contradiction* means imposing incompatible requirements on an individual part of a system. The physical contradiction is formulated as follows: «To meet the problem requirements, a given zone must have property «X» (for example, be movable) and to perform some function it must have the property «non-X» (for example, to be immovable)».

For example, a supersonic jetliner Concorde which crosses the ocean within several hours did not return the profit expected by its creators. With this formulation, it is absolutely incomprehensible what to do to increase profit. It is an *administrative contradiction.*

Analyzing the situation makes it clear there are many reasons for this problem such as expensive tickets, terrible noise of a sonic boom preventing it from flying at top speed over land, it is only able to land in some airports having a long runway. Having crossed the ocean within several hours, a passenger has to spend a lot of time to reach the final destination point. All these reasons can be reformulated into a more obvious *technical contradiction* such as: using a special wing increased the cruising speed of the supersonic liner but also caused an unacceptable increase in takeoff and landing runway length.

The physical contradiction will sound extremely concrete: a wing must be small to provide a high cruising speed and it must be large to reduce the required runway length.

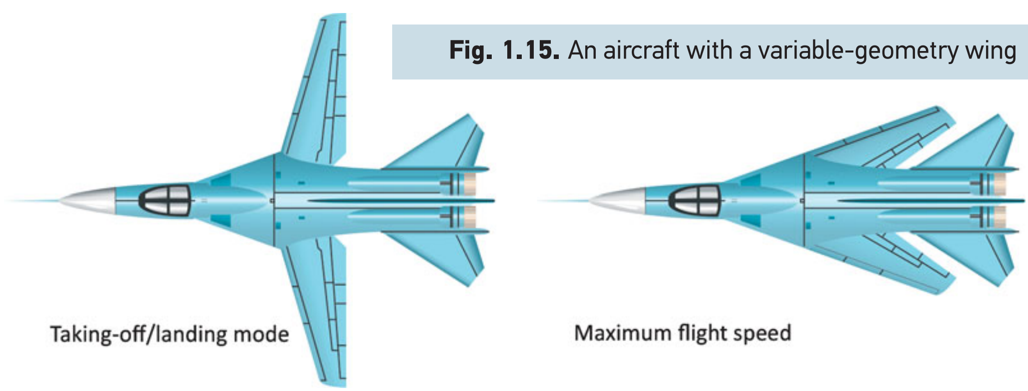
Traditional design methods recommend compromising between the requirements imposed on a system being improved, i.e. they are aimed at smoothing the occurring contradictions. While improving one parameter of a system, other ones are typically worsened. In this situation the best of the options is selected.

TRIZ, on the contrary, recommends an acute sharpening of a contradiction.

Designers usually tend to compromise and develop wings providing the optimal speed value and a shorter runway. Thus we are flying at a lower speed but have a shorter runway. The strongest solutions are found in a different area, it is necessary to resolve a sharpened contradiction. For example, a variable-geometry wing can turn small when flying high and large while taking off or landing. Such an airplane has a high speed when at a height and needs no special long runway for landing (Fig. 1.15).

Grouping the contradictions available in a system into several levels offers new possibilities for situation analysis and the system being improved. In addition, TRIZ has effective tools for removing these contradictions.

Fig. 1.15. An aircraft with a variable-geometry wing



After investigating numerous examples of strong solutions available in the patent collection, G. S. Altshuller revealed a series of special *principles of resolving technical contradictions* [for the list of principles, see 4.17]. The principles only point at the general trend of transformations and guide an inventor to the field of strong ideas. Specific solutions may be found by analogy with the principle or example which illustrates this principle. The same principle may be used for solving problems from absolutely different technical fields.

Here are two examples of solving problems from the field of hydraulic engineering and engine-building.

Warm-up of engine.

Warming-up an engine is an important manufacturing operation. An engine is started light and all its rubbing parts start running. The process is lengthy and consumes a lot of fuel. How can the engine be warmed-up quicker?

Solving this problem is not easy without knowing special principles. Using the contradiction-resolving principle «Blessing in disguise» gives a mighty hint of how this problem can be solved. The principle recommends:

- using harmful factors (in particular, the harmful action of environment) for obtaining a positive effect;
- removing a harmful factor by combining it with another harmful factor;
- enhancing a harmful factor to such an extent that it stops being harmful. A solution that matches the recommendation of point A) is: supplying dust-laden air to the engine instead of purified air. This can make the warm-up several times faster.

Reducing the flow energy.

A water flow running from a mountain possesses a huge destructive force. It can damage waterworks. How can the flow energy be reduced?

We can use the same principle here — «Blessing in disguise». Using the recommendation of point B, we have obtained the following solution: the flow course is separated into several branches which are directed toward each other. The flows collide and mutually cancel the energy.

G. S. Altshuller developed a convenient table for revealing and resolving technical contradictions[1.15]. It has the following structure (Fig. 1.16).

Standard parameters which are to be improved according to the statement are arranged vertically. Parameters which are unacceptably worsened as a result are arranged horizontally. At the intersection of the rows and columns, there are numbers of matching the principles most probably capable of removing the technical contradiction that occurred between the parameter being improved and the parameter being worsened. For constructing such a table, G. S. Altshuller used 40 most effective principles of solving technical contradictions.

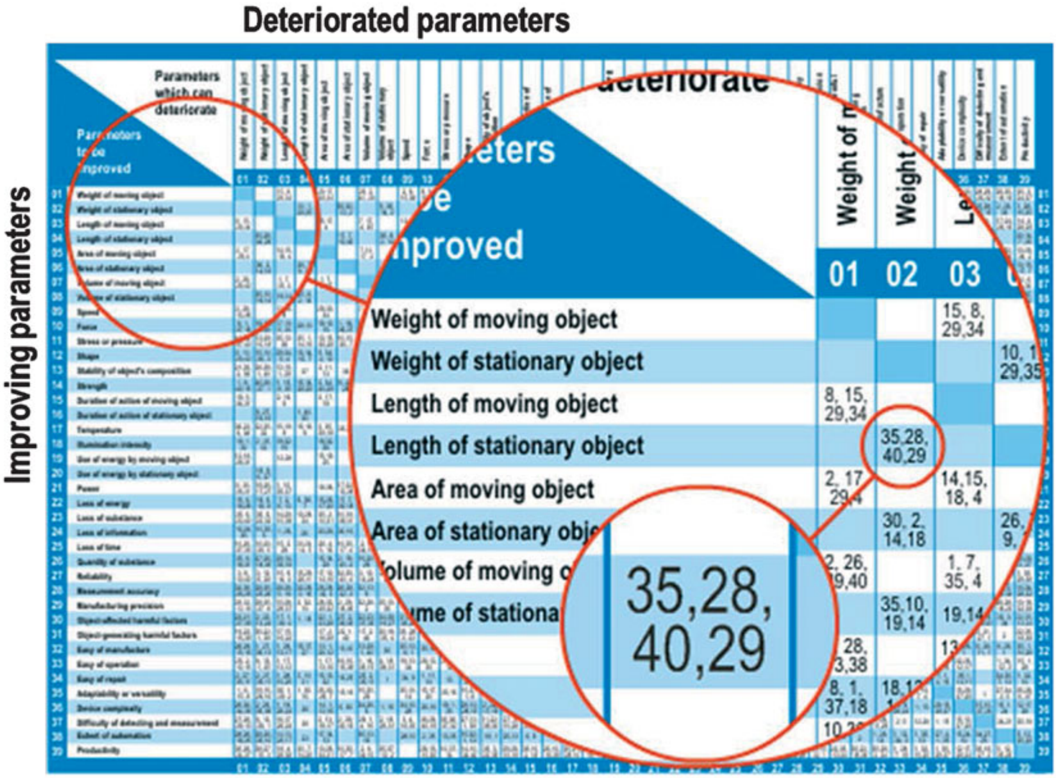


Fig. 1.16. G. S. Altshuller's Table of resolving technical contradictions

Preliminary concepts of solving by using principles may also be obtained without using the Table of Contradictions. To this end, it is necessary to sequentially analyze the possibility of using each of the 40 principles. Every inventor gradually forms his own list of the most frequently used principles. The practical use of the principles of solving technical contradictions has the following peculiarity: the recommendations described in each principle should not be understood literally. The greatest effect is achieved by treating them as a prompt, initial material for reasoning.

Let us take, for example, principle 25: changing color. Understanding this recommendation literally reduces abruptly the field of action. At the same time, treating the principle as a change of surface property in general increases abruptly the opportunities of producing new ideas. In the given case, we may speak about a change of the optical properties of a surface, its roughness, temperature, as well as applying some additional substance, etc.

The *physical contradiction* is a situation when physically mutually exclusive requirements are made of some component or part of a technical system. As distinct from a technical contradiction, a physical contradiction arises not between the technical system parameters, but describes contradictory requirements made of one of its components or even part of a component. The physical contradiction is formulated as follows: «To satisfy the problem requirements, a given zone must have the property «X» (for example, it must be movable) to perform some function and have the property «non-X» (for example, it must be immovable).»

The example of the physical contradiction: a car windscreen must be hard, rigid to resist the oncoming air stream and it must be flexible, elastic not to injure the driver in case of breaking. This contradiction is resolved by using triplex glass having an internal soft layer between two external glasses.

The main principles of resolving physical contradictions [1.16]:

1. If it is required that an element exhibit opposite properties at the same time, such a contradiction is resolved by separating these properties in space.
2. If it is required that an element exhibit opposite properties at the same place, such a contradiction is resolved by separating these properties in time.
3. If it is required that an element exhibit opposite properties at the same time and at the same place, such a contradiction is resolved in a super-system.

Some examples:

Crossing.

How is traffic organized, for example, at road junctions? Not observing any rules will lead to all automobiles trying to cross a road junction at the same time. It also concerns those automobiles which have the right of way (such as ambulances, for example). In this situation, collisions are inevitable because of the physical contradiction: two or more cars are trying to be at the same point of space at the same time.

How is this contradiction resolved in space?

One road runs over the other. Cars cross the junction at different levels and do not interfere with each other.

How is this contradiction resolved in time?

Lights are used. Cars cross the junction according to the lights signals.

How is this contradiction resolved in a super-system?

Special automobiles with switched on signals, for example, an ambulance, have the right of way at the junction. This order is established in a super-system, is determined by special traffic rules and holds good on all roads.

Display.

The screen of any display is made up of tiny squares — pixels. Each pixel can become lighter or darker and generate light of any color. To produce a moving picture, image frames on the screen change 24 times a second, the pixel brightness and color must change with the same frequency.

Thus, the following contradiction arises for a color display: the pixel color should constantly change whereas engineering constraints only allow a pixel of one color to be produced.

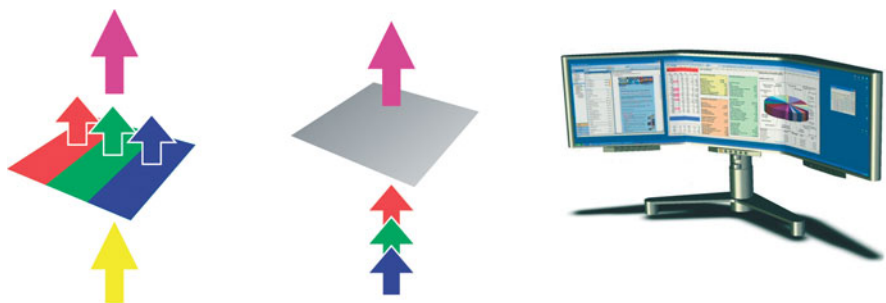
How is this contradiction resolved in space?

A pixel is divided into a number (minimum three) of sub-pixels, each producing only one color — red, green or blue. These are the primary colors of the spectrum and mixing them in certain proportions is perceived by an eye as a needed color (Fig. 1.17, a). The «one displayed frame — one light pulse» rule is observed here.

How is this contradiction resolved in time?

Samsung specialists have developed a special liquid-crystal (LC) screen operation technology called UFS which may be deciphered as a «display of a very high image quality».

With this technology, there is no need to divide a pixel into three sub-pixels. The required brightness and color of a pixel are provided by mounting three backlighting lamps — red, green and blue — behind a liquid crystal filter; the lamps blink alternatively many times during the display of one frame (Fig. 1.17, b). Then a liquid crystal filter capable of opening a window over a pixel controls the formation of a required color.



a) Several color subpixels　b) Pulsating three-color backlighting lamps　c) Using large high-resolution screens

Fig. 1.17. Resolving a contradiction while obtaining a color TV image

If it is necessary to display a red dot, the filter opens the pixel only when the red lamp blinks and keeps it closed when the blue or red lamp blinks. To produce the white color, the pixel remains closed during the entire time of one frame display. Controlling the number of pulsations of different colors allows production of any color of the pixel.

The «one frame — many light pulses» rule is now observed (Fig. 1.17, b).

How is this contradiction resolved in a super-system?

Because the pixel size is limited, increasing the image sharpness requires increasing the number of pixels on a display screen and moving the screen itself away from a viewer. Then the visible size of a pixel will be smaller.

One of the possible solutions is using the principles underlying the Seamless Technology which combines several screens of ordinary size and resolutions into a single high-sharpness super-screen (See. section 5.6). Because the pixel size remains unchanged and the screen size is increased, the image sharpness improves for a viewer.

1.2.5. Actions on models

In problem solving, all actions are performed not on real objects but on their mental models. This is not something unusual: engineering design also deals with models — graphic models.

Models used in each sphere of human activities are continuously improved. It is worth remembering drawings by ancient architects that were used for creating many great pieces of architecture. To explain what exactly had to be done, a great number of drawings were usually created for showing an object from different sides. Developing an effective type of graphical model — Monge diagram (Fig. 1.18) — was a real breakthrough in engineering. Orthogonal projection proposed by Monge, where projections on three mutually perpendicular planes were built, became the main method in diagramming [1.17].

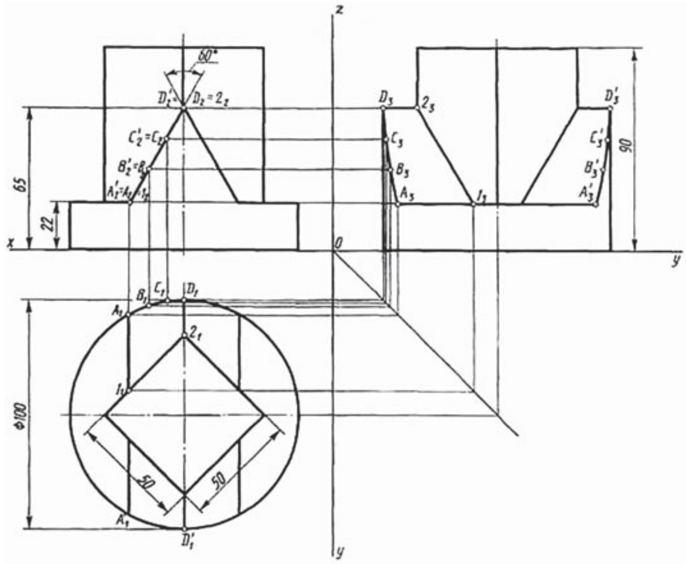


Fig. 1.18. Example of Monge diagram

Graphic design documentation became much simpler and the description quality of the machine shape and structure as well as the drawing visualization improved immediately. The expression, accuracy and convenience of object image measurement improved considerably.

Graphic models are continuously being developed such as: three-dimensional and multiple view diagrams, axonometric projections with cuts and sections which allow better visualization of the shape and structure of a component or unit of any degree of complexity.

Similarly, TRIZ has various types of effective models.

Su-field model

It describes the interaction of a tool and a work object [1.5]. The Su-field analysis got its name from the words «substance» and «field» because substances interact through fields in a Su-field model. TRIZ interprets these conventional terms in a particular way.

In TRIZ, «substance» is something different from substance as it is usually understood in engineering: any material objects, both artificial and natural, i.e. *ice, a shaft, a copper sheet, molten aluminum, a lathe, an icebreaker, vapor, a glass prism, etc.*

“Field» is usually understood as any action of one material object on another object. It is usually a field itself and its derivatives (forces, flows) as well as any interactions between some material objects. For example, odor is substance, not field by nature, but it is a method of interaction between a source of odor and olfactory receptors; therefore, Su-field analysis refers it to the class of fields. Examples of fields according to TRIZ: *field of gravity, friction, pressure, thermal field, ultrasound, magnetic field, odor, flexibility, inertia, etc.*

To build a Su-field model, it is necessary to specify what *substances* interact and determine which substance is a *tool*, which is a *workobject* and what field is responsible for the interaction between them (Fig. 1.19).

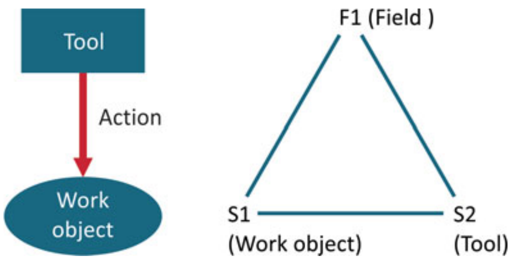


Fig. 1.19. A Su-Field model

The Su-field model is a triangle having the symbol names of substances and fields at its corners:

- S1. System's working tool acting on the work object;
- S2. Work object. Substance exposed to a necessary action;
- F1. Field. Provides the action of the tool on the work object.

Everything that happens in a Su-field model is formulated as «B1 acts on B2 by means of F1». Depending on the problem statement, a Su-field model may include different types of interaction.

An example of building a Su-field model (Fig. 1.20). A magnet holds an iron bar. The Su-field model: S1 (magnet) acts on S2 (iron bar) by means of F1 (magnetic field).

TRIZ provides 76 standard solutions, universal transformation rules for transforming Su-field models. The system of standard solutions is a set of technical system synthesis and transformation rules that follow from the technical system evolution laws. The standard solutions provide a joint action of logical transformation of Su-field models and use of physical and geometrical effects.

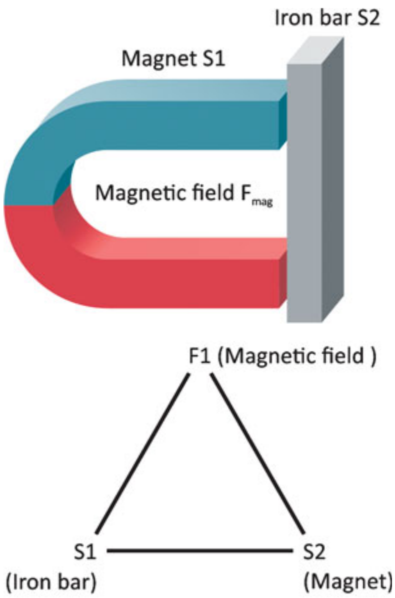


Fig. 1.20. An example of a Su-Field model for the interaction of a magnet and an iron bar

Modeling with little creatures

One more model used in TRIZ is a model composed of «little creatures» [1.26].

While solving numerous problems, the famous physicist Maxwell imagined the studied process in the form of little creatures that could perform all necessary actions. They are known in the literature as «Maxwell's demons». G. Altshuller proposed a similar method of modeling by means of one or several groups of «small creatures» who can perform any action in our mind's eye.

This model allows schematic presentation of contradictory requirements. A large number of little creatures (a group, several groups, a crowd) operate on a conventional drawing (or several sequential drawings). Only changeable parts of the problem model may be presented in the form of small creatures.

Modeling with little creatures is convenient because it allows not only imagining the interaction of a tool and a work object as one-piece components of a technical system, but also maximally taking into account their internal structure.

Example: Torsion bar reliability problem

The mirror of a laser projector is small in size but plays a very important role in the projector work. It oscillates at a very high frequency and provides laser beam scanning on a screen. The mirror is attached by means of two torsion bars and actuated by an alternating electric field (Fig. 1.21, a).

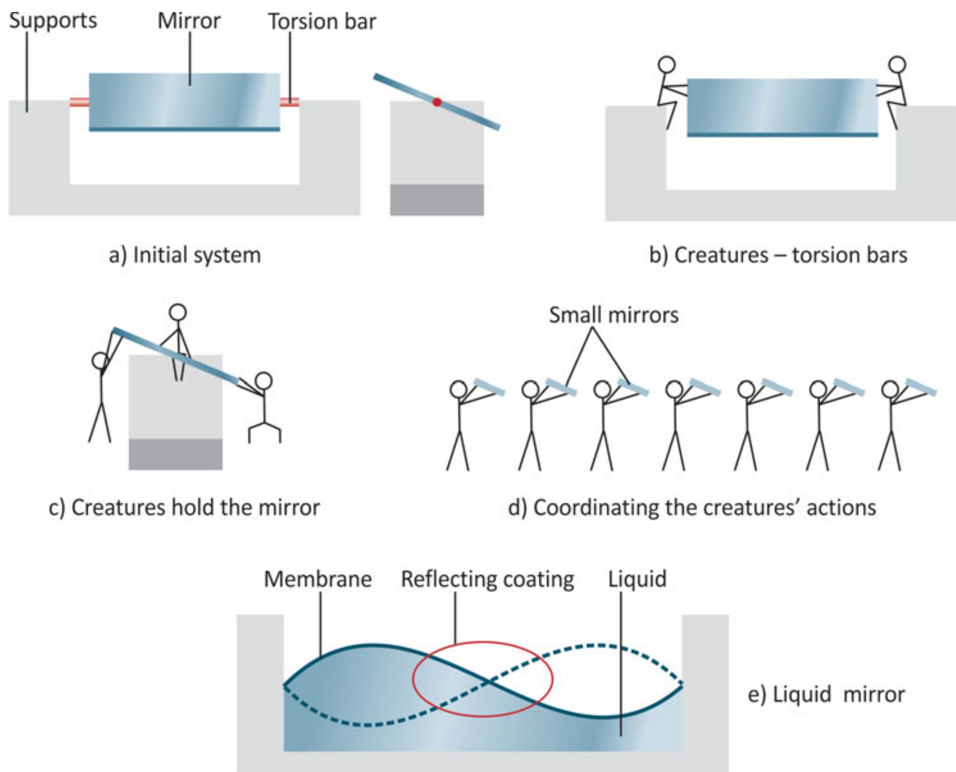


Fig. 1.21. Using a «little creatures» model in solving a problem of a laser projector oscillating mirror

A goal was set to increase the reliability of the torsion bars and prevent their failure.

The mirror model formed by little creatures is shown in Fig. 1.21, b. Two creatures (torsion bars) hold the oscillating mirror with difficulty.

How can we help them? Probably, it is necessary to support the oscillating mirror (see the model in Fig. 1.21, b). We can also facilitate the life of the little creatures by segmenting the mirror and giving each creature one small piece only. Such a piece will be easier to swing; the only thing necessary is to agree on the coordination of actions (Fig. 1.21, d).

The analysis of the obtained model or it is better to say, the game of the little creatures, generated a number of serious ideas, one of which is a liquid mirror. Liquid having special properties is poured into a vessel. There is a flexible reflecting film on the liquid. A standing wave is generated in the liquid; its length is equal to the double length of the vessel (Fig.1.21, e).

An oscillating mirror is formed in the middle portion of the vessel. Probably, the realization of this idea will face technical difficulties, but this mirror is obviously impossible to break.

1.2.6. Considering the psychological peculiarities of thinking

When using any methods, it is necessary to take into account the human thinking peculiarities.

A man is constantly working out algorithms of habitual actions in similar situations, which allows him to perform most manipulations automatically, unconsciously. This is a great benefit because a more rational use of our brain abilities is ensured. Let us imagine a driver who is driving a car for the first time in his life. His brain is working so intensely! He must remember the traffic regulations and instantaneously decide how to change the gears, what to press to start moving, what to pull to stop... In addition, it is necessary to control traffic conditions.

An experienced driver does not think where the change-gear lever is and how to correctly change its position. He performs most control actions automatically so that he can think about his rout, road and optimal speed.

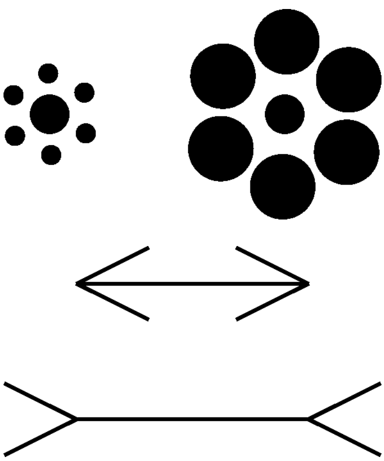


Fig. 1.22. Examples of optical illusions

However, following the developed algorithms becomes an obstacle when it is necessary to solve a new, non-standard problem, which is just what inventors deal with. Starting to solve a new problem, we unintentionally try to apply the already known solutions, methods and notions to it. Thinking inertia makes a man use obvious solutions, whereas the essence of invention requires searching for and resolving paradoxes.

Also, the information processing procedure in the human brain should be taken into account.

Let us take the well-known example — optical illusions (Fig. 1.22). The circles in the center of the top drawings as well as the line segments in the center of the bottom drawings are of the same size. Our brain, however, regards them as different, because it corrects our perceptions in accordance with the surrounding.

If our subconsciousness behaves in this manner in simple cases, then how can it behave when solving complicated problems?

Try to perform a simple experiment. Show your spread palms to somebody and ask, «How many fingers do you see?» And as soon as the answer «Ten' follows, ask quickly, «And how many fingers are there on ten hands?» Very often, the answer is «A hundred». But ten hands have fifty fingers.

To answer another simple question — «How can a deaf-mute ask for a hammer?» — people will show with gestures how a hammer works. Then ask, «How a blind man can ask for scissors?» Very often, the answer will be the clipping gestures with fingers. But a blind man can ask for scissor using words.

These are the jokes of our cognitive apparatus. There is nothing bad about it but we should understand how thinking occurs and learn to guide its work. TRIZ has created a rich arsenal of methods for the development of creative imagination, the main aim of which consists in working out a specific thinking style for controlling thinking operations.

1.2.7. Using resources and an information base

When solving invention problems, it is very important to have a very good information support, because finding some abstract problem-solving idea is not enough. It is necessary to have resources* that may be used for system transformation. Strong solutions are generally obtained by using the seemingly most unexpected resources. Resources are everything that can be used to solve a problem: substances, fields, time, space, information, etc. When analyzing a situation, not only the scheme and its components are considered but also the surroundings. The most successful are usually those solutions where resources were found within the system itself or, even better, in the operational zone.

When dealing with a problem, first it is necessary to try to employ apparent, ready-to-use resources. If such resources do not provide a solution, their transformed versions should be analyzed. The main resource transformation method suggests use of physical and other effects. A search for resources obtained by using effects is especially successful when it is already clear what exactly the sought-for resource must perform in the system. Such a situation may occur at any problem solving stage, but it is most probable after constructing a physical contradiction, when mutually exclusive requirements are made of the same object of a system. Any effect is based on some natural phenomenon. The effect scheme is given in Fig. 1.23.

Fig. 1.23. Effect structure

The model of a physical effect (Fig. 1.24):

Mechanical deformation (Action) of piezo-crystal (Object of action) causes electric potential (Result).

The model of a geometrical effect (Fig. 1.25):

Changing the shape from flat to convex (Action) — Membrane (Object) — Increased strength (Result).

Such presentation of effects is very convenient because a required effect may be easily found based on the resources available in a system.

TRIZ assigns high priority to continually searching for effects which can be used for solving a problem. Some times ago «The Index of Physical Effects and

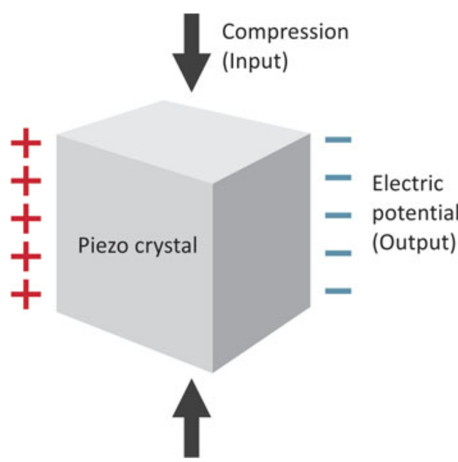


Fig. 1.24. Piezoelectric effect model

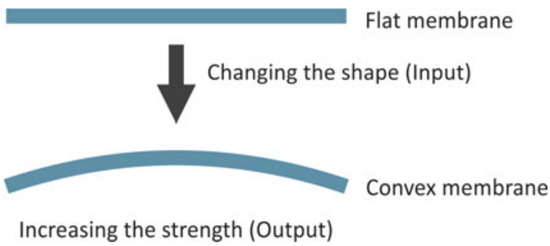


Fig. 1.25. The model of the «Changing the shape» effect

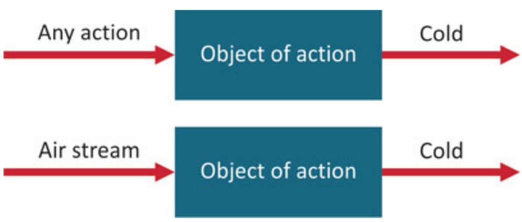


Fig. 1.26. The scheme of the «Product cooling» effect

Phenomena» was compiled based on the analysis of thousands of inventions and study of vast scientific literature [1.18]. This index is a reference for thousands of inventors. It is important to note that the data base of effects is only a background for further independent work aimed at the accumulation of effects for a respective range of problems solved by a given inventor.

Use of effects. For example, you need to provide cooling of a product, transported in a cold storage truck, with minimum transformation of the system and without expending a large amount of fuel.

Let us make a scheme of the required effect (Fig. 1.26, a). First, we consider the resources available in the moving truck. Primarily, it is an oncoming air stream. Our scheme is transformed into a new one (Fig. 1.26, b).

Using the data base of effects, the Index or any other source of information, we find the most suitable effects for making an air stream cold. There are two: evaporation cooling and the Rank effect.

Evaporation cooling is the decrease in surface temperature during the evaporation of a layer of material, for example, a layer of liquid from this surface. Consequently, to operate for a long period of time it is necessary to continuously supply liquid to its surface, which may cause additional problems when using the device on a truck.

The Rank effect is more suitable in this situation. It uses a tangential air stream or any other gas introduced into one end of a pipe so that is moves in a spiral path in the pipe to the other end. The air in the center will be cooler than the air towards the outer area. [1.19]. Such a cooler may easily be used in a moving cold storage truck (Fig. 1.27).

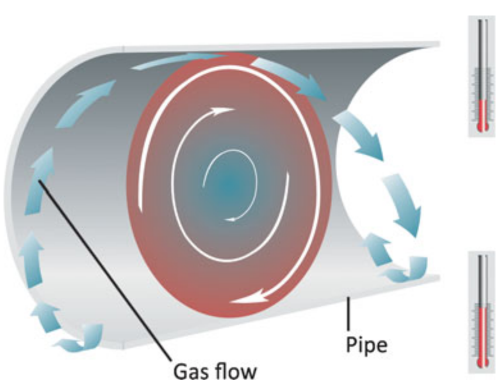


Fig. 1.27. A cooler based on the Rank effect

Clear description and understanding of the effect allows correctly using it and finding its new applications. The effect may be used in various technical systems.

The Coanda Effect was discovered by the Romanian aviation engineer Henri Coanda. A stream of fluid blown out through a flat slit tangentially across a convex surface will adhere to the surface at a relatively large distance from the slit. Using this effect, Coanda developed and patented the world's first «flying saucer» shown in Fig. 1.28 [1.20].

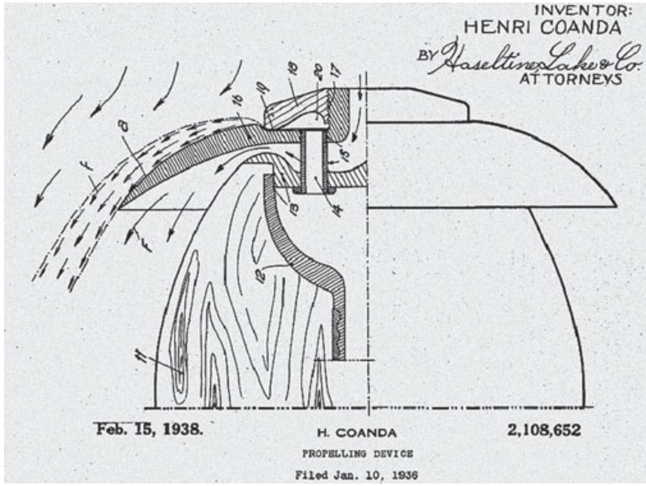


Fig. 1.28. The «Flying saucer». The drawing from the H. Coanda's patent

The Coanda effect should be considered when designing hydro and aerodynamic systems: airplanes, air-conditioners, pumps, submersible vehicles, etc.

One of the examples of using the Coanda effect is the aircraft EKIP developed in Russia (Fig. 1.33). The characteristics of this aircraft are impressive: in the economy mode near ground level, EKIP can fly at a speed of 160 km/h whereas at a height of 8 to 13 km it accelerates to 700 km/h. The aircraft is highly economical and safe, its payload is twice as large and the fuselage payload volume is four times as large as that of a conventional aircraft [1.21].

EKIP has no wings in the usual sense, it only has small stabilizing surfaces. The lifting force is produced by the air flowing about a thick low-aspect wing which is simultaneously a fuselage. Because the curvature of the upper surface of such a wing is much smaller than that of the lower surface, the flow speed at the top of the wing is much higher. Accordingly, the lifting force directly related to the difference of speeds below the wing and at the top of the wing is also higher.

Fig. 1.29. EKIP aircraft

There had been unsuccessful attempts to use a thick profile wing before this aircraft was created: a high-speed flow on the rear part of the wing usually splits into numerous vortices and is detached from its surface, which considerably decelerated the flight (Fig. 1.30, a). EKIP creators analyzed the optimal conditions of a thick wing airflow with the Coanda effect and developed a system of the boundary air layer control which allows the eddy flows occurring at the stern to be sucked into the fuselage. Due to this, the entire upper surface of the wing was flown about by a laminar flow without separation (Fig. 1.30, b), which resulted in a large lifting force at high flying speeds.

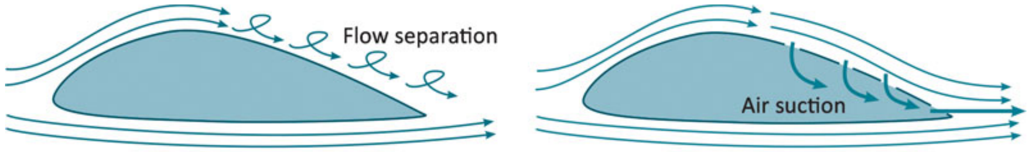


Fig. 1.30. Wing airflow

Physical effects are very suitable for problem solving, but geometrical effects often prove effective, too.

Using of geometrical effects.

Using a spherical shape for the hull of a diving device called a bathysphere (geometrical effect) allowed Auguste Piccard to reach the bottom of the Mariana Trench — the deepest point of the World Ocean. A shell of the same thickness but of any other shape would be crushed by the huge pressure.

Such a geometrical figure as the Moebius band — a one-sided surface — is widely used in engineering. For example, making a band saw in the form of the Meobius band doubles its durability (Fig. 1.31).

Great benefits have been achieved using microlevel effects. Using these effects resulted in the appearance of all «smart materials» which avoid complicated construction. Solving a problem on the microlevel is always better than using some mechanisms. Comparing LEP displays (LEP-technology — See p. 5.7) and an ordinary CRT (See p. 5.5) shows it is undoubtedly simpler to construct. *Body armour* containing liquid filler capable of instantaneously hardening upon a strong impact is one of the new inventions obtained by using the microlevel effects. Under normal conditions, this material is flexible, so it may be used for making trousers, sleeves and gloves [1.22].

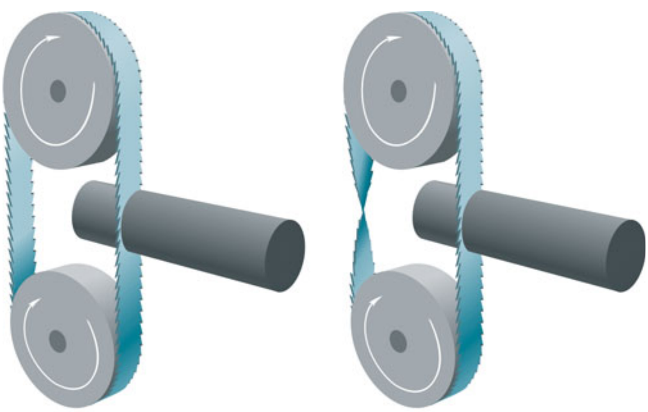


Fig. 1.31. A band saw

Example: Regenerating polymer

Causing damage to the shell of an object often threatens with serious accidents. To increase shell reliability, its wall should be made thicker, which increases the product weight and cost. How can this contradiction be resolved?

One of the possible ways is using self-regenerating materials. Shells are manufactured from a special kind of plastic containing evenly distributed micro-capsules with epoxy resin and hardener micro-granules. When a shell is damaged, the capsules in the damage zone burst, the resin escapes from them filling the cracks or scratches. This causes a chemical reaction with the hardener. As a result, the damage is repaired. If the Russian space station «Mir» had been built of such a material, it would have been safe from cracks and it would have been unnecessary to bury it in the ocean [1.23].

Microlevel effects produce a high level of improvement to processes. Lately, the development of nano-technologies has produced impressive results in a wide range of applications.

Nanocatalysts can prevent losses and increase efficiency of many technological processes, for example, production of oil products. Almost 20% of crude oil remains unprocessed because of the imperfect oil refining technology. Refining efficiency increases dramatically by using special ceramic catalyst filters with nano-pores. These filters can retain even one molecular chain. Passing crude oil through such a catalyst will increase the processing efficiency to almost 100%. [1.24].

Metal rubber. It is possible to create nano-materials having very special properties, such as, current-conducting rubber. The American company NanoSonic has developed a polymer that is flexible and elastic as rubber while being as electrically conductive as metal. A plate of metal rubber is grown literally by molecules. A substrate (for example, glass) is alternatively immersed in tanks of chemical solutions. The first solution contains positively charged metal ions and the second one negatively charged polymer molecules. They combine with one another from layer to layer, thereby creating a structure where metal and polymer are mixed at the molecular level. The company hopes that metal rubber will find applications in different fields of technology — from the aerospace industry to electronics [1.25].

Another resource is biological effects [1.26]. Sometimes, a biological object — an animal, a plant, a colony of microorganisms or any other living organism — may be helpful in solving a technical problem. Animals' behavior obeys some laws conditioned by their instincts or training. These behavioral peculiarities may be studied and used for certain purposes.

Using of biological effects.

For example, trained dogs can perform functions that are impossible by the most sophisticated instruments. Dogs can sniff explosives and drugs, find people under ruins of collapsed buildings. Guide-dogs help blind people to walk on busy streets. A cat can help pass a wire through a pipe of the most complicated shape, hamsters and other animals feel earthquakes well in advance, etc.

The evolution of technology based on simpler biological objects, such as plants, is also predictable and may be used to solve inventive problems.

A gas indicator. A natural gas transmission pipeline has a huge length. It is difficult to predict where a microcrack can develop through which gas will leak out and it is not easy to monitor the condition of a multi-kilometer pipeline. In Czechia, they found a simple and beautiful solution to this problem. They plant alfalfa over a pipeline. The smallest amount of gas causes alfalfa to change height and color. These changes are easily seen from a helicopter.

1.2.8. Inventive problem solving process

Several algorithms of different degrees of complexity are available for solving inventive problems.All of them are based on the main principles of inventive problem solving developed by G. S. Altshuller [1.27]. According to these principles, solving an inventive problem starts from analyzing a complicated problem situation and formulating the problem condition itself. A problem situation can be improved by different methods. A lot of problems can be formulated, but it is necessary to select the key problem from this variety. This problem allows cardinal improvement of the situation and the transformation of the technical system, participating in the problem situation, in accordance with the *technical system evolution laws.*

Then we continue with problem solving itself. The main action is understanding which solution is the most desirable for us. We call this the IFR *(ideal final result).* After formulating the IFR, one should understand whether it can be achieved without transforming the system and the acceptable degree of compromise.

Generally, it is impossible to achieve the IFR without changing something. There are always some limitations in a system which prevents doing this. There arises a situation called a *contradiction*: we want to achieve the IFR but there are circumstances that hinder doing this. Contradictions must be resolved. Our problem can be transformed into several *models* developed within TRIZ. They are:

- The Su-field model with a set of standard solutions
- The model of technical and physical contradiction
- The model of small creatures, etc.

Each kind of problem model has some rules for transforming it into a solution model. After obtaining a solution model and analyzing available resources that may be used for changing the system as desired, we obtain a number of intermediate concepts. Sometimes one of the concepts is immediately accepted as a final solution to the problem but usually a set of solutions is used as «building material» for constructing a final solution to the problem according to certain rules.

Then the final (principal) solution is transformed into a technical solution which takes into account all of the real conditions. After that, the parameters of the new system are calculated to obtain an estimated solution.

At the end of the solving procedure, it is necessary to check whether the problem has been adequately improved and the problem solving process can be stopped or whether it is necessary to repeat the entire cycle.

* * *

It is impossible to give an exhaustive idea of TRIZ in such a short chapter and this was not the author's objective. Specialized literature in TRIZ is available in abundance today. In addition to the proposed references, we can recommend the books by Altshuller for independent study, [4.34–4.36], as well as the books by other authors [4.37.–4.41]. We also recommend the following web sites:

Archive site of G. S. Altshullerwww.altshuller.ru
American TRIZ Journal www.triz-journal.com
Generator www.gnrtr.com
TRIZland www.trizland.ru
TRIZ Encyclopedia www.triz.port5.com

1.3. Necessity for structuring an information field

TRIZ-based methods are very effective for generating problem solving concepts. Using these methods allows you to comparatively quickly find an «area of strong solutions» and obtain several solutions to your problem. However, this is not always enough in dealing with real projects. One should be sure that problem-solving opportunities have been amply investigated and the full scope of useful changes of the technical system is explored.

During our work we deal with three types of tasks: problem solving, forecasting, and circumventing of patents. In each case, we have to deal with a huge array of information pertaining to different aspects of the same technical system (device, unit, part, etc.). Note that each problem-solving stage requires work on sets of information.

For example, at the analytical stage, it is necessary to collect information about analogous devices to investigate their advantages and disadvantages. We must be sure that all known basic versions of a technical system of interest have been revealed. The completeness of the data collected will prevent us from «re-inventing the wheel» so that we can focus on truly new, patentable solutions to our problem. Further, in the process of solving our problem, we deal with an array of unique information already at the concept synthesis stage. Complicated problems may have tens and even hundreds of preliminary solutions, which hampers selection of the best ones. But there is still no guarantee that we have not missed the most promising ideas.

A forecast project also starts by collecting information on all existing embodiments of a system under investigation. This information is largely derived from patents, scientific, and technical publications. The amount of information (descriptions of technical system embodiments) can be so large that without some visual organization structure, convenient and fast navigation is just impossible.

In circumventing a patent and creating a patent umbrella, we are again faced with a necessity to know everything or almost everything about the system we are going to protect with patents. This means that a maximally full search is needed, derived information should be structured and the obtained structure be thoroughly analyzed. Only after this is done, is it possible to reliably find unpatented versions of the system, check them for practicability and patent.

In each of these cases, we get into a vast information field. Moving in such a field in search of new ideas and concepts without being able to view it as a whole is difficult and inefficient. It is much easier and more convenient if known system versions are arranged in an order suitable for analysis and we can see all of them at one time. In this case, an effective information structure itself should show what versions it lacks and suggest new versions to be found in the first place. Then the information structure becomes a kind of «map» and helps us navigate the information field.

How can we build an information structure that best suits our purpose?

First of all, it is necessary to find an objective classification criterion. An alphabetical list does not add any new information to that already collected through the search. Classifying by patenting companies or by issue years will give us new information such as patenting dynamics, and identification of leading companies in the production of a given product. The value of the information increases.

There are a great many of such classification criteria: device complexity, mass and material consumption size, cost, principle of operation, etc. Almost all of them make sense and are suitable for achieving certain goals. However, subjectively selected criteria are of no value for obtaining new ideas while solving inventive problems, finding new system versions while making a forecast or circumventing competing patents.

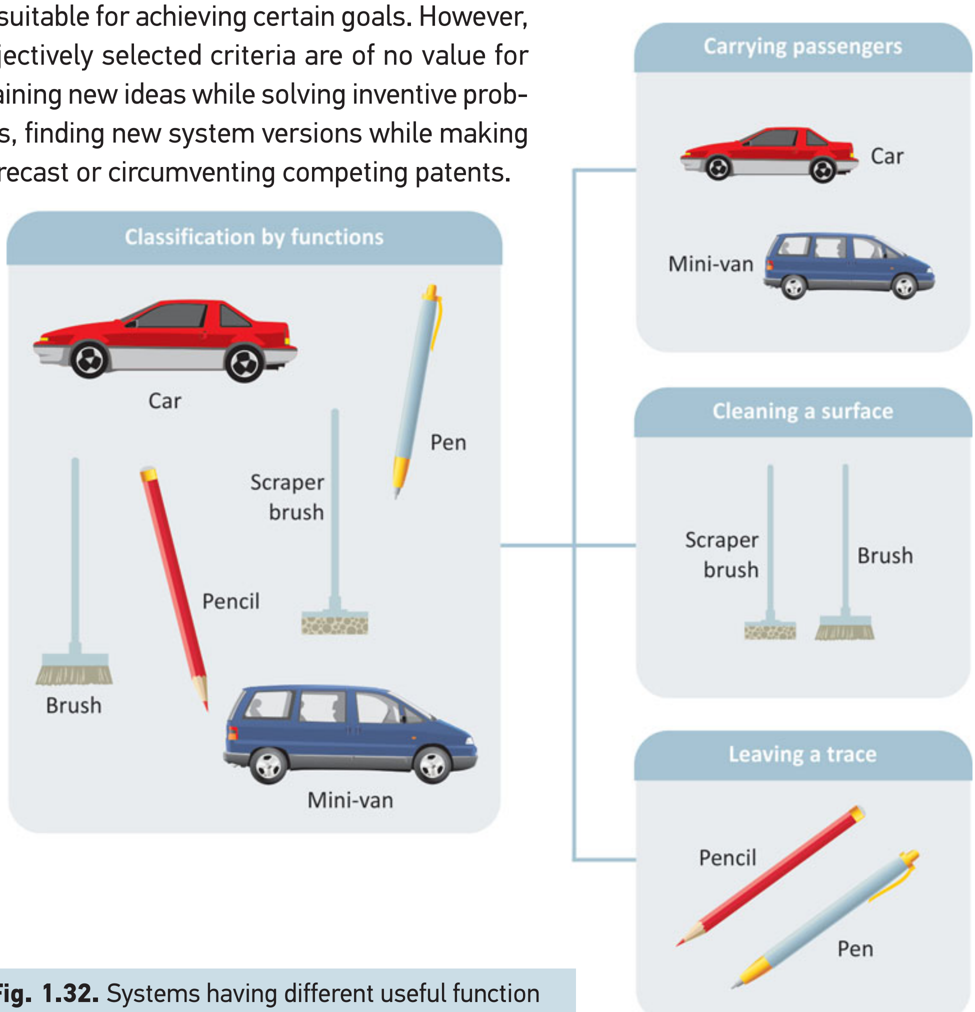


Fig. 1.32. Systems having different useful function

Could we use the same classification principle as the one applied in patenting where much attention is focused on this research direction [1.3]?

Patent classification is quite logical at the level of sections representing *different technical systems*. The main classification criterion for technical objects is their designation, main useful function. We can easily classify different technical objects, for example, cars, pens or brushes. Cars are designed for «carrying loads and passengers», pens for «leaving a trace on a surface», while the intent of brushes is «cleaning a surface (fig. 1.32).

How can technical systems performing the same function be classified? For example, different brushes (Fig. 1.33)? By material? By manufacturer? By size? By the head shape?

Unfortunately, within sections which include *different versions of the same technical system*, patent structuring largely has a random, chaotic character. It may be based on design features, principle of operation, characteristic property and even on the uses of a technical system described in a patent. All these classification methods are rather subjective. As a result, information organization and search for information in a patent base are hampered.

For example, section B 60 C «Tires» of the International Patent Classification (WIPO) classifies tires by material and general design (cordless, with inflatable inserts, with sealing devices between beads and a rim, sectional tires, etc.), shape and size of cross-section (closed toroidal, asymmetrical, folding (for storage), etc.). Devices for inflating, removing and mounting a tire are also considered as tires though differ from the tire itself both functionally and structurally.

Thus, classification by function used in the patent branch does not help us navigate our information field which includes descriptions of versions of the same technical system. Classification by the principle of operation or by a characteristic property is also far from being objective.

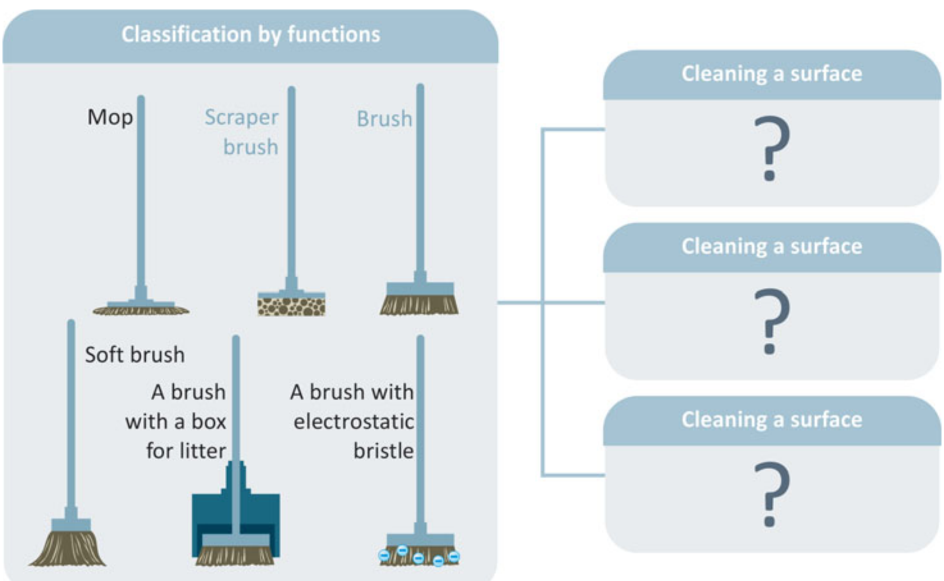


Fig. 1.33. Systems having the same useful function

However, building a really effective information structure that would serve our purpose is not restricted to a search for an objective criterion. There are also other requirements described below.

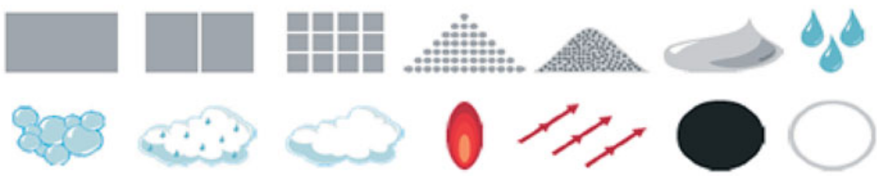
- Classification should be based on objective criteria (*objectiveness* requirement).
- All significantly differing versions of a transformed system should be provided (*fullness* requirement).
- The classification method should be universal for all technical systems and, at the same time, be suitable for describing transformations of a specific system (*generality* and *specificity* requirement).
- The information presentation method should be maximally visual and show the presence of gaps produced by the patent search (*visualization* requirement).
- The structure should contain data on missed information sufficient for obtaining a concept of their practical fulfillment (*informativity* requirement).

In the next chapters, we will try to explain how to build an information structure capable of meeting all these requirements.

Chapter 2

Objective Evolution Patterns

The patterns produced by analyzing the evolution of a large number of real systems are an effective tool for describing the sequence of the system versions.



2.1. Information organization in an evolution pattern

Objectiveness is the first item in our list of requirements of a patent and technical information classification system. To satisfy this requirement, it is necessary to find a method for arranging the technical system versions in certain sequences built according to objective criteria. Generally, a great many of such sequences can be built. The question is whether each of them can be considered objective.

Is every such sequence objective?

Watching cars in the street, one can quickly build a large number of sequences, even if just by the order driving on a street. A truck, a car, a trailer, a bus, a jeep — isn't it a sequence? A dump truck, a tank truck, a snow removal machine, a bus... You may order motor vehicles by color, for example, by the spectrum of colors: red, orange, yellow, green. We can search for some sense in the arrangement of color variants by some feature: by better visibility on a road or by visibility of dirt. However, such sequences, built by a random feature, are generally of no use for technical system improvement.

A similar problem is considered in TRIZ based on the study of the objective evolution laws of technical systems. G. S. Altshuller's idea was to build objective «evolution patterns of technical structures» [1.5]. Evolution patterns are used in TRIZ as a tool for viewing the dynamics of an obtained technical solution. In addition, organizing information into evolution patterns looks very promising for describing the evolution of a technical system or its components. Let us consider evolution patterns in more detail.

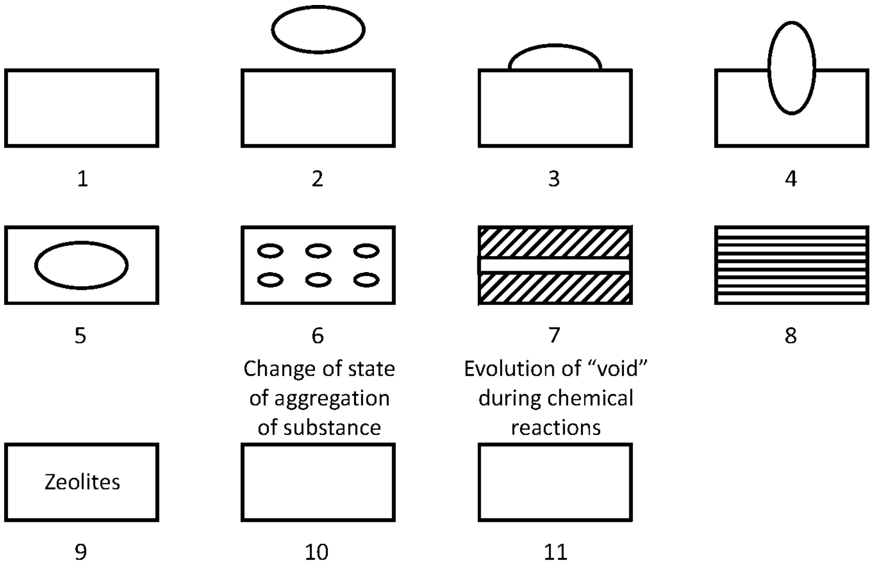


Fig. 2.1. The pattern of increasing void according to G. S. Altshuller:

1 — solid object;
2 — «void» beyond direct contact with an object*;
3 — «void» contacts an object;
4 — «void» partially penetrates an object;
5 — «void» within an object;
6 — segmented «void»;
7 — through «void» (a «pipe» within a solid object);
8 — capillary structure;
9 — zeolite structure (tubes formed by molecules);
10 — «void» appears in an object as the result of a physical effect (for example, the occurrence of bubbles in a boiling liquid);
11 — «void» is released during chemical decomposition of substance (for example, evolution of gas during decomposition reaction)

Evolution patterns used in TRIZ were constructed by using the following approach. First, a large number of different technical systems were analyzed and evolution patterns of their components were built by a certain feature. Then the patterns were compared to find frequently repeated steps which were then included in a generalized pattern applicable to the entire class of systems and their components.

For example, the study of a large number of existing systems proved that in the course of evolution they became increasingly dynamic, mobile. Several hinges turned a rigid ladder

*) It should be noted that the term «void» is defined in TRIZ as some area filled with less dense than the primary medium substance.

into a folding one, a computer keyboard and mobile telephone sets also became foldable, rigidly attached car seats got some hinges and turned into folding seats, a bicycle also «learned» to fold. In other words, the evolution of each of these systems resulted in increased mobility. Only after accumulating sufficient statistical data, researchers drew a conclusion that the step of introducing a hinge into a rigid construction may be included into the generalized pattern of mobility increase. It may be expected that the evolution of most systems will follow this trend.

At present, TRIZ offers multiple evolution trends obtained in the course of studying real technical systems. «Increasing Void» is one of such trends suggested by G. S. Altshuller and I. M. Vertkin [2.2] (Fig. 2.1). It shows transformations of a solid object caused by adding the void* to it. The void may be located near an object, on its surface, or a closed cavity may be formed within an object. Then follow transformations of the internal cavity: segmentation, making a through cavity, use of capillaries and microcapillary structures. An example of transformations described by the increasing void pattern may be transformation of a gear wheel,

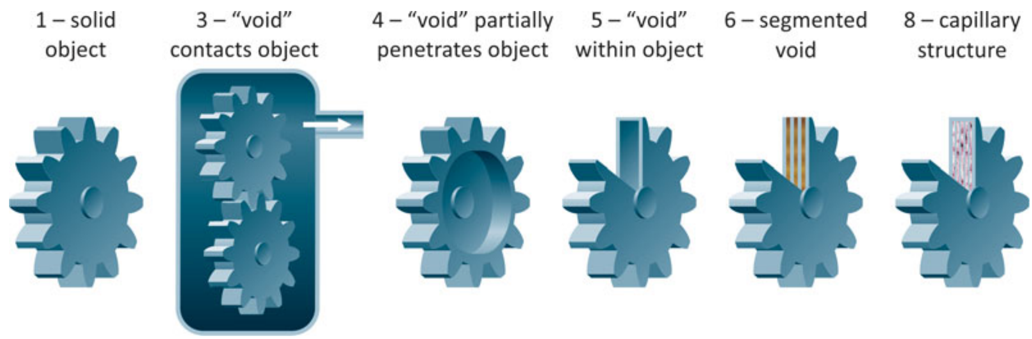


Fig. 2.2. Some transformation versions of a gear wheel according to the increasing void pattern

pinion, which is a disk with toothed flank surface and a hole in the center for putting on a shaft. Here are several versions of a gear wheel transformed in accordance with this pattern (Fig. 2.2.).

The starting version — «solid object» — is a solid pinion having flat end faces without any additional holes or indents. Introducing the void that contacts the pinion (version 3 by fig. 2.1) may be provided by enclosing the interacting wheels in a hollow case with a vacuum in the case. This method is well suited to high-speed reduction gears where resistance of air entrained by the pinions is of a significant value. For the pinions to be lighter and for the other components to be more conveniently placed in the reduction gear, the «void» may partially penetrate the flank surface of the pinion (version 4). In version 5, the void is within the pinions: hollow pinions have elastic, adaptive working surfaces which cushions impacts or varying loads.

A segmented «void» within a pinion can be made by a structure of radial channels, through which lubricant is fed from a wheel hub to teeth (version 6). Capillary structures (version 8) are also frequently used in gear wheels. They may be wick-type lubricant-feeding systems or noise-absorbing fillers of hollow pinions.

Here is one more pattern, «Segmentation» [2.1]. G. S. Altshuller suggests the following transitions (Fig. 2.3): an internal partition appears in the initial object, then the number of partitions grows and the formed compartments are partially separated from each other. A hinge or flexible coupling occurs between the compartments which becomes increasingly mobile and dynamic during evolution until it transfers to the field* level and then to coupling by program or command methods. Then the link disappears completely. V. M. Petrov [2.3] proposes his interpretation of the «Segmentation» pattern (Fig. 2.4). His version first of all means transformation of the object itself, but not the links between the object parts. A solid object becomes flexible, is segmented into smaller parts, changes to gel, liquid and gas, and then transformations down the field level follow.

Such interpretation of the «Segmentation» pattern seems more logical than the previous one. However, in our opinion, the «Flexible Object» step does not fit the pattern. The main

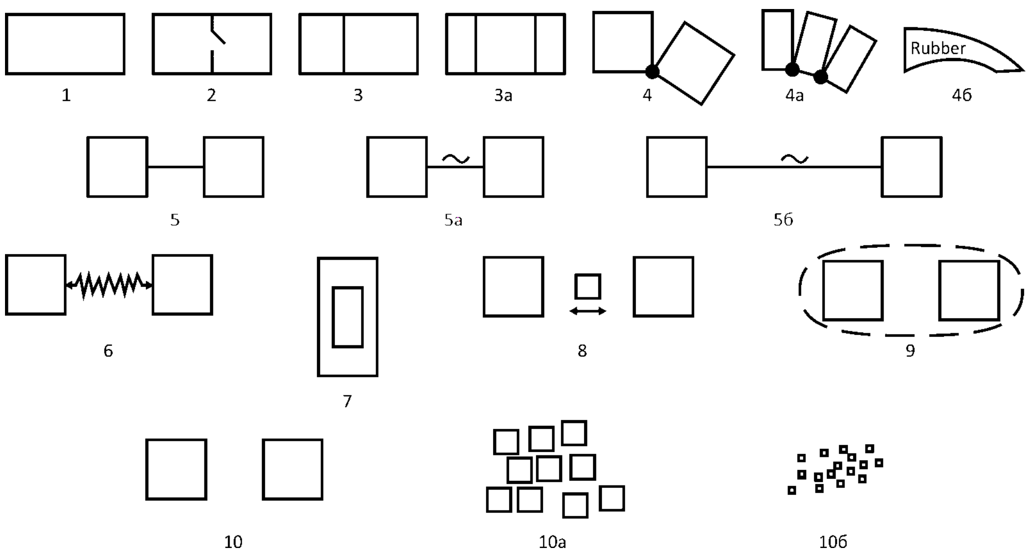


Fig. 2.3. Segmentation pattern according to G. S. Altshuller [2.1]:

1 — initial object;
2 — a partial internal partition appears;
3 — the partition becomes full;
3a — the number of partitions grows;
4 — partial separation of the formed tightly coupled or hinged compartments;
4a — the number of hinged compartments grows;
4b — ultimate increase in the number of hinge links — rubber;
5 — a weight-type construction: parts are rigidly connected by means of a rod the length of which is of the same order of magnitude as the size of the parts.
5a — coupling becomes more dynamic; flexible coupling;
5b — the maximum increase in length and flexibility of the couplings: for example, cable systems;
6 — field coupling;
7 — structural coupling (one part moves freely inside the other);
8 — «shuttle-type» coupling (may be illustrated by a motor boat plying between two ships) — substance or field coupling;
9 — program coupling (no coupling, but each part moves according to a pre-coordinated program);
10 — zero link;
10a — zero link in a poly-system;
10b — zero link in a limiting poly-system

property of a flexible object is mobility, ability of taking any shape. In so doing, the object itself remains solid, is not segmented into parts whereas the building logic of the segmentation pattern implies sequential segmentation of an object into increasingly small parts.

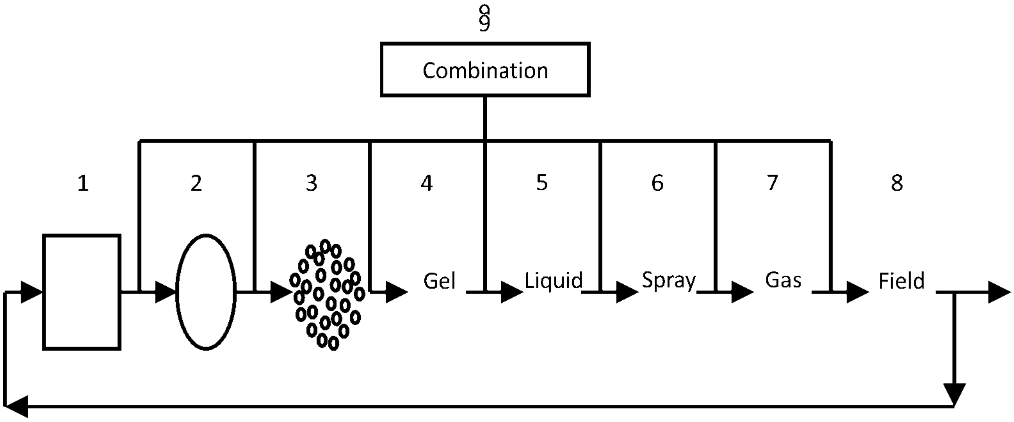


Fig. 2.4. The segmentation pattern according to V. M. Petrov [2.3]:

1. Solid monolith system;
2. Fully flexible (elastic) object;
3. Powder-like object;
4. Gel;
5. Liquid;
6. Aerosol;
7. Gel;
8. Field**

Comparing the two «Segmentation» evolution patterns — G. S. Altshuller's and V. M. Petrov's — proves that even the most similar sequences revealed by different researchers have significant differences, which is the evidence of a subjective approach to the selection of generalized transformation versions based on the study of evolution of real systems.

The «Segmentation» pattern according to V. M. Petrov and, partially, according to G. S. Altshuller may be illustrated by a transmission that transmits rotation from an engine to an executing mechanism. Thus, the «Solid Monolithic System» version implies power transfer by means of tight coupling. Flexible, elastic coupling is provided by using a belt drive or a flexible drive. According to the logic of this pattern, a powder-like object may be illustrated by a chain-drive, where a flexible drive is segmented into elementary links — «granules». Liquid is a working medium in various designs of hydraulic couplings and fluid converters.

Gas as a working medium is occasionally used to transmit rotation. One of the examples may be the original «air transmission» used in windmills for transmitting rotation from a rotor to a pump or generator. Because a windmill is generally mounted at a great height and a driven aggregate at the ground level, transmission is too extended and complicated. To simplify rotation transmission, the rotor blades are made hollow and each blade has a hole at the end. The mast is made tubular and the slewing tower of the windmill is made hollow (Fig. 2.5).

As the blades rotate, the inertia force moves the air within the blades away from the center of rotation and ejects it outside. Reduced pressure is formed within the blades, air starts

**) The term «field» in TRIZ refers to a force or energy acting on the object.

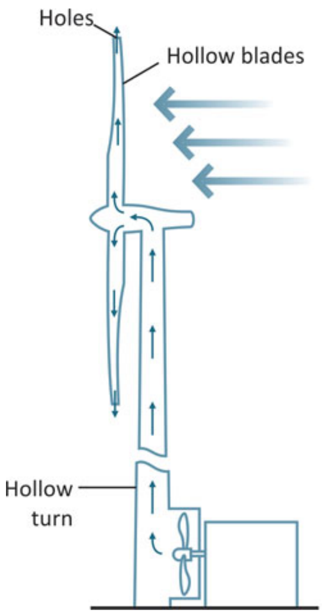


Fig. 2.5. A windmill with an «air transmission»

to be sucked in from below, through the hollow mast. Due to this, a strong vertical air flow occurs within the mast. An impeller installed at the mast entry will start rotating. Now connect the impeller to the driven aggregate to set the latter in rotation.

A similar principle is used to ventilate homes and industrial buildings. Ball-shaped impellers are attached to the ends of ventilation ducts. Each impeller serves as a very simple and effective exhaust system (Fig. 2.6). The impeller is attached to a ventilation duct by mean of light bearings and blades are installed so that even a slight wind sets them in rotation. The blades moving in a circle spin the air within the sphere eject it outside, thereby creating rarefaction within the duct so that the air actively moves from the room into the ventilation duct. This device is widely used in Korea. It increases several fold the ventilation system effectiveness.

Fig. 2.6. A vent-pipe activator

As for the «field transmission» of a windmill, it may be illustrated by a wind-wheel where a generator is combined with a rotor. Electric energy is taken from the generator and may be transmitted by wires at a considerable distance.

The study of technical systems evolution also helped improve the methods of building their evolution patterns. More informative patterns more adequately reflecting the essence of technical systems transformation were found. One of the important developments in TRIZ that had a significant effect on the pattern building methods was a system of technical systems evolution laws. Checking the patterns for compliance with these laws made them more logical and complete.

Let us examine the relationships between «patterns», «trends» and «technical systems evolution laws».

- The TRIZ *laws of technical systems evolution only* give a general description of the links between phenomena [2.5]. The evolution laws are not easy to use as a problem solving tool because they are too general; therefore, the action of these laws is described through trends and systems evolution patterns.

- *Trends* show general evolution directions of system components according to the objective laws of technical systems evolution. The trend may be graphically presented in the form of a vector.
- The *evolution pattern* is the specification of some trend, the evolution regularity of a given object or process. It is not just an evolution direction, but a detailed «route map» with the indication of characteristic transformation versions of a technical system or its components.

For examples, the law of increasing ideality is instantiated by two trends:

- technical systems generally become simpler and less expensive in the course of evolution;
- as a system evolves, the number of useful functions it performs generally grows.

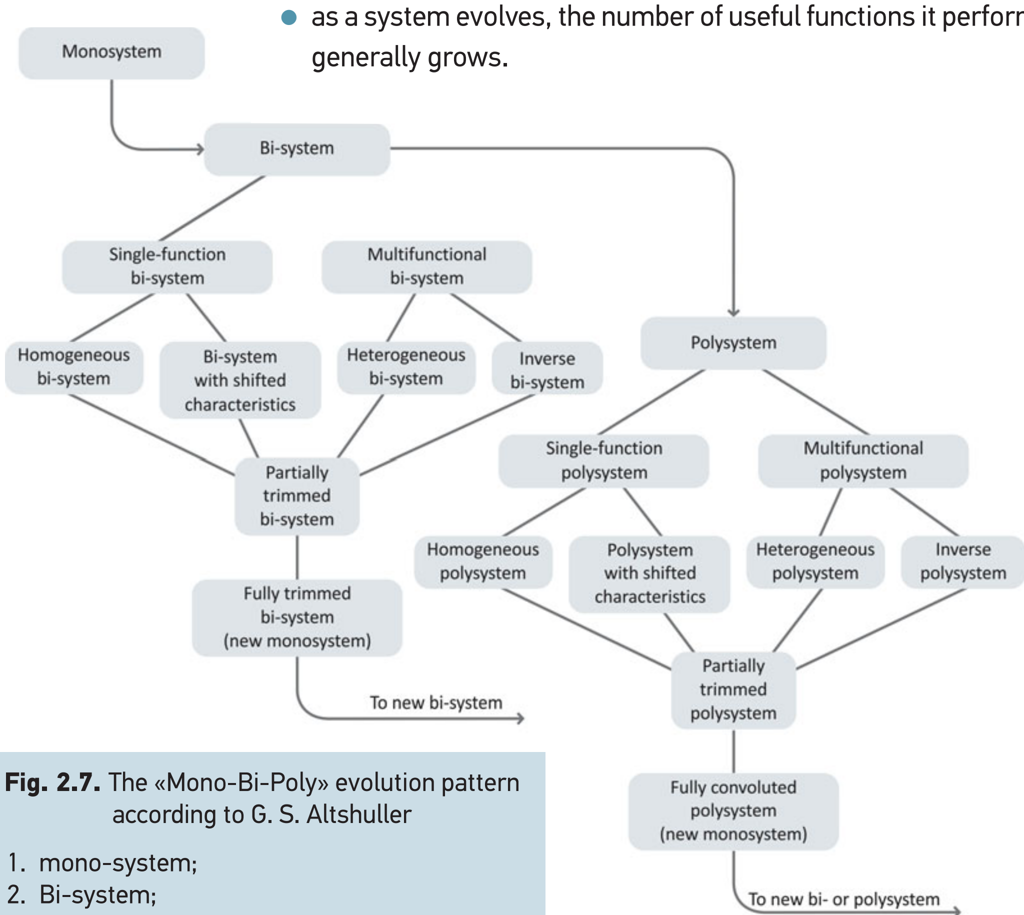


Fig. 2.7. The «Mono-Bi-Poly» evolution pattern according to G. S. Altshuller

1. mono-system;
2. Bi-system;
3. Single-function bi-system;
4. Multifunctional bi-system;
5. Homogeneous bi-system;
6. Bi-system with shifted characteristics;
7. Heterogeneous bi-system;
8. Inverse bi-system;
9. Partially trimmed bi-system;
10. Fully trimmed bi-system (new mono-system);
11. poly-system;
12. Single-function poly-system;
13. Multifunctional poly-system;
14. Homogeneous poly-system;
15. poly-system with shifted characteristics;
16. Heterogeneous poly-system;
17. Inverse poly-system;
18. Partially trimmed poly-system;
19. Fully convoluted poly-system (new mono-system)

Accordingly, two evolution patterns based on these trends are known: «Trimming» and «Mono-Bi-Poly».

The pattern, in fact — is the trend with put on them step by step in concrete ways transformation of technical system (for basic or theoretical pattern), or variants of these transformations (for a pattern of a specific technical system).For example, G. S. Altshuller obtained the «Mono-Bi-Poly» pattern [2.1] by studying and interpreting the law of the technical system transition to a super-system (Fig. 2.7). This evolution pattern is interpreted in the following manner.

As a system evolves when it exhausts its resources it tends to combine with another system to form a more complex bi-system structure. Also, several initial systems may combine and form a poly-system. The main condition for the system transition to a bi-system or poly-system is the need to improve the operational characteristics of the initial system and to introduce new functions which can be performed by a combined system.

Bi- and poly-systems may be single-function or multifunctional. Single-function bi- and poly-systems consist of similar or different systems which can perform similar functions. Multi-functional structures may comprise heterogeneous systems performing different functions and, as a particular case, inversion systems having opposite functions.

After the system has combined with other systems, all the components of the produced poly-system start combining into a mono-system of a higher level.

Here is an example of a «Mono-Bi-Poly» pattern of the «*Spanner*» system (Fig. 2.8).

A single-head spanner is only suitable for loosening or tightening one-size nuts. We also cannot loosen a «bolt+nut» connection with one spanner, because it is necessary to hold the bolt head and to loosen the nut on the other side of the bolt simultaneously. This is impossible to do with a single spanner, a transition to a bisystem — two spanners — is needed.

Placing two different sized heads on one handle gives a partially trimmed bi-system. Such a spanner can be used to loosen or tighten nuts of two different sizes.

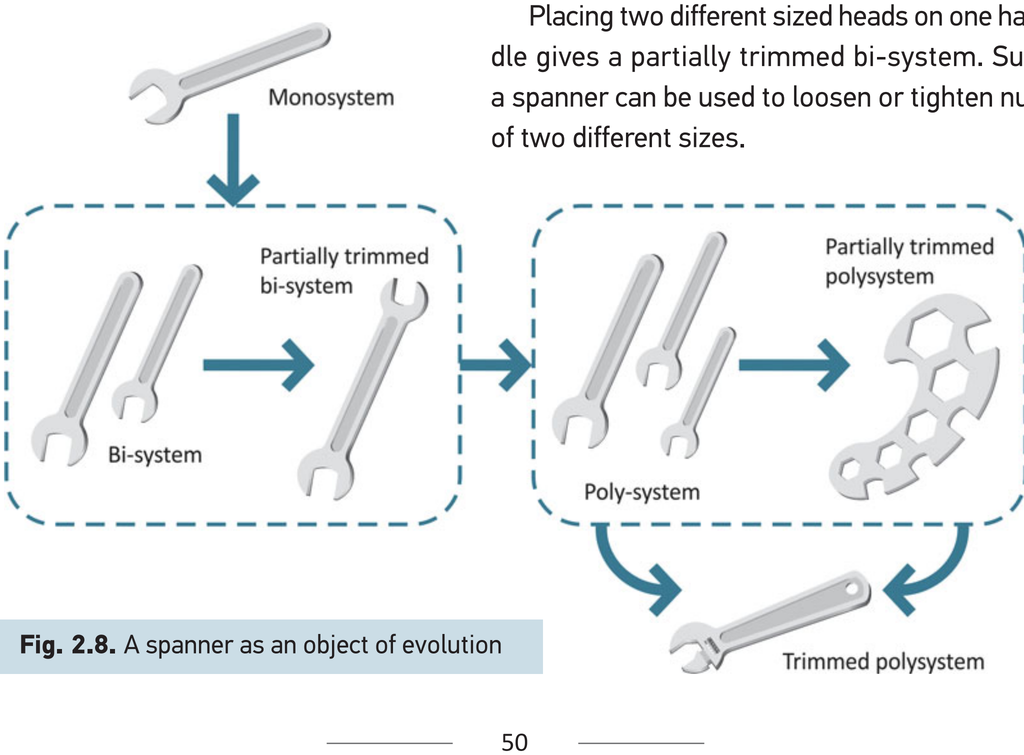


Fig. 2.8. A spanner as an object of evolution

The next spanner type is a number of spanners forming a set. A set of heads and a tap driver for rotating the heads may serve as a trimmed version of such a set. One more example is a traditional bicycle spanner — a shaped plate having many heads inside and even shaped grippers for special nuts.

Increasing the number of heads beyond some limit is irrational, because the spanner will become too big and inconvenient to use. Consequently, as the «Mono-Bi-Poly» pattern suggests, such a polyspanner should be somehow transformed into a monospanner of a higher level.

The almost fully trimmed spanner is an adjustable-wrench with a single head that the size can be changed.

The adjustable-wrench is a higher-level system compared to the initial single-grip spanner. Using the screw-wrench, one can loosen or tighten a nut of any size within some range. It should be noted, however, that two spanners are still needed to hold both a nut and a bolt while working with a nut screwed on a bolt, i.e., the next evolution stage will be a transition to a bi-system composed of two spanners.

Such well-defined prognostically strong patterns as «Mono-Bi-Poly» are effective for analyzing the technical system evolution. TRIZ offers a large number of evolution patterns that describe changes in various parameters of a system. This may be illustrated by the patterns developed by B. Zlotin [1.2] (see diagram on fig. 2.9).

Also worth mentioning are numerous interesting evolution patterns suggested by Yu. Salamatov [2.4], V. M. Petrov and other researchers, but one thing should be taken into account. In literature, evolution patterns are not generally accompanied by any comments but listed as sequential technological objects or versions. This implies that practical use of evolution patterns may be made more efficient by developing effective mechanisms to describe the progression.

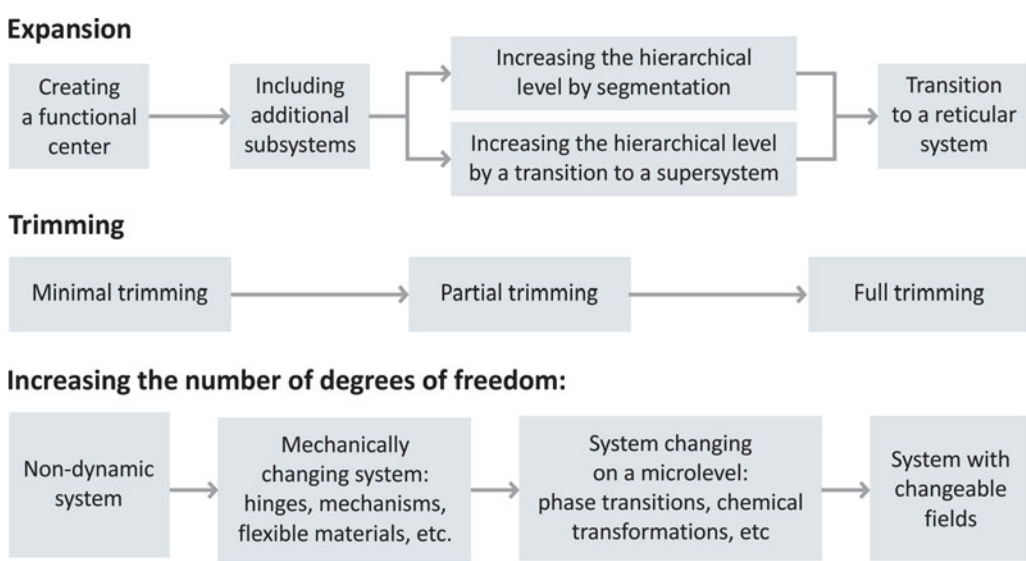
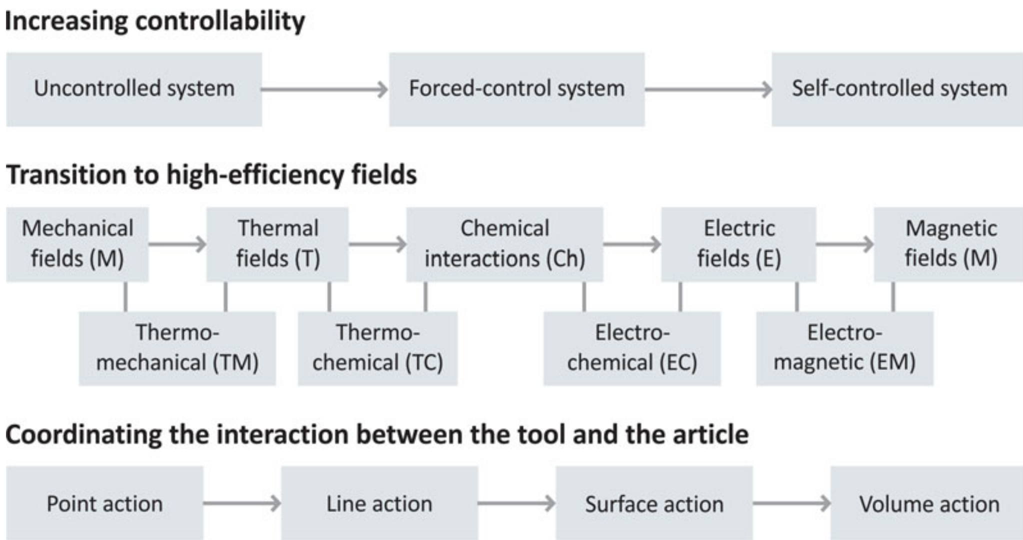


Fig. 2.9. Some evolution patterns



Continiuos of Fig. 2.9.

In addition, a correct description of an information field is impossible using sporadic patterns, even if these are based on objective laws. The problem is that the evolution of some object or system is not described by a single pattern. One can trace several patterns simultaneously in the evolution of a real system. There arises a question: how can one pattern be separated from another?

Let us consider a simple object — a ruler (Fig. 2.10). The ruler should be long to measure large sizes, but, at the same time, it should be short to be convenient to carry. How can this contradiction be resolved?

Segmenting these parts with hinges turned a bulky meter long ruler into a carpenter's folding ruler. At the next stage, the ruler mobility increased: a new flexible measuring-tape that could be reeled up for storage. It measures quite precisely a distance in a straight line (in plan view). Further evolution in this direction led to the appearance of a curvimeter —

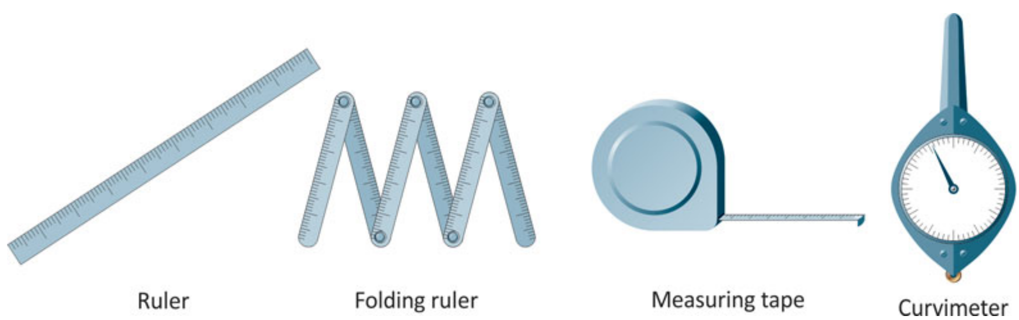


Fig. 2.10. The pattern of increasing the measuring tool compactness: «ruler — folding ruler — measuring reel — curvimeter»

a rotating wheel connected to a counter and running on a pattern being measured. The measuring device became even more dynamic, mobile; the curvimeter may be used to measure a distance in a curved line regardless of the surface shape. Then our ruler makes a transition to the «field» level such as a laser ranger). But let us expand on the first four versions of the measuring tool.

Trying to build a logical evolution pattern would immediately bring us to a dead-lock. First, we had a solid ruler, then we divided it into numerous parts. It uses *segmentation*, separation into parts. Next, a segmented ruler becomes folding. Hence, *dynamization*, increasing of mobility is next.

At the same time, there are many parts. Hence, the «*Mono-Bi-Poly*» pattern is traced. A folding ruler better fits the carpenter's pocket size than a meter-long ruler, not to speak of a measuring reel. Consequently, the *increasing of coordination* occurs here.

We may also say that dynamization means segmenting a ruler into parts and connecting them with hinges. Then we mean not a simple transformation, but some more complicated action. In addition, a joint coupling of object parts does not always provide mobility. Let us segment a ruler into three parts and hinge the parts (Fig. 2.11). The produced construction will not be dynamic and will remain rigid. Researchers often combine different actions performed on a system, such as segmentation and dynamization of its parts into a single mixed pattern [2.2].

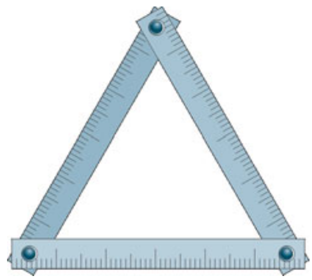


Fig. 2.11. Rigid hinged construction

The readers can build their own evolution patterns of any real technical system. The process is very interesting and cognitive. To begin with, let us try to build several patterns for an aircraft wing and find general layout principles.

2.2. An example of real system analysis

The simplest aircraft — an unguided glider — was created by combining a wing, balancing planes and a kind of fuselage into a single construction [2.6, 2.7].

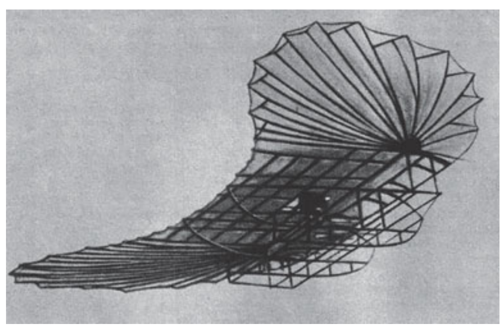


Fig. 2.12. The Etrich glider — a flying wing shaped into a zanonia seed

Precoordinating the shape, size, mass and layout of lifting surfaces, a fuselage and a pilot allows a glider launched from a mountain to fly stably, gradually lowering and not falling down. For example, the Etrich glider (Fig. 2.12) formed according to the «flying wing» design was shaped like a winged seed of having aerodynamic

Fig. 2.13. The Lilienthal's glider

Fig. 2.14. The «flying wing» B-2 airplane

characteristics that provided good flight stability. However, such a glider did zanonia***, not have any controls so a pilot could not change the wing position and shape to control the flight direction

The next step is the Lilienthal's balancing glider (Fig. 2.13). It was a guided glider, because the pilot could change the wing position. In the Lilienthal's glider, the pilot was not attached to the wing rigidly, as in the Etrich's glider, but dynamically, by flexible coupling, and could displace his body weight relative to the wing. If the pilot bent his arms and moved his body forward, the wing-setting angle decreased, the glider lowered down its nose and gathered speed; if the pilot moved his body backward, the wing became steeper with respect to the oncoming air and the glider slowed down its flight.

Such a control system proved very effective for light-weight aircraft; for example, this control strategy is successfully used today in hang-gliding.

Heavier aircraft of the «flying wing» type, such as the Northrop's invisible bomber B-2 (Fig. 2.14) have the same principle of providing the initial stability as the zanonia seed, whereas the operational control of the aircraft is realized by means of aerodynamic rudders. Flight control by changing the wing shape is performed in the following way.

For the aircraft to fly in a straight line, it is necessary that its wing be comparatively flat (Fig. 2.15, a).

For the aircraft to ascend, it is necessary to deflect the rear edge of the wing upward. The oncoming air stream pressure will deflect the rear part of the «flying wing» down, the nose up and the aircraft will start ascending (Fig. 2.15, b).

To direct the aircraft downward, it is necessary to deflect the rear edge of the wing down. The oncoming air stream will start forcing the rear part of the aircraft up so that it will start descending (Fig. 2.15,c).

***) Resilience seed Zannoni in flight by the fact that its ends are referred back wing and bent slightly upward. Because of this when you lean forward the seed wing tip is located at a large angle to the direction of flight and tip the back of the seed down, aligning the flight. When you angle the seed back the main surface of the wing loses lift and related back ends of the wing, by contrast, are set at the optimum angle of attack, creating lift, tilt forward a seed. Constant game of aerodynamic forces ensures stable flying seed Zannoni.

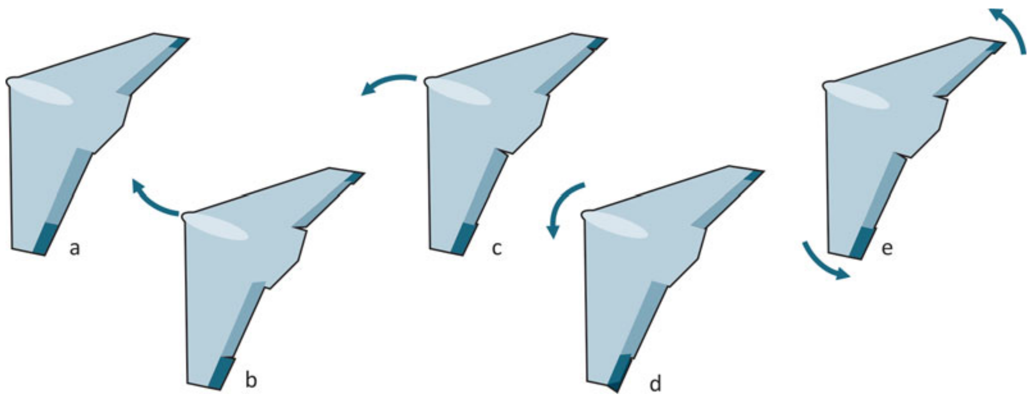


Fig. 2.15. The «Flying wing» control scheme

For the aircraft to roll, for example, to the left, it is necessary to deflect the rear edge of the right half of the wing down and that of the left half of the wing up. The left half of the wing will start lowering, while the right one will go up and the aircraft will roll to the left (Fig. 2.15, d).

What is necessary to make the «flying wing» turn? The following solution was found for this type of aircraft: attaching an airbrake of split air rudders to the end of each half of the wing. For example, to turn to the left, it is necessary to slow the left half of the wing, i.e., to increase air resistance to its motion, and the aircraft will start turning to the left (Fig. 2.15, e).

To change the wing shape, its construction was supplemented with air rudders hinged to the rear edge. Using controls, the pilot could deflect the rudders up and down providing a required coordination of the wing shape and the flight conditions. In addition to rudder surfaces, B-2 has airbrakes — additional surfaces flaring up and down and installed at the wing tips. On large aircraft, air rudders are deflected not manually but hydraulically, which allows automatic control of the aircraft on stable segments of flight. As is seen from the B-2 example, the coordination of the wing shape and the flight modes is achieved by introducing additional movable components into the wing design. To provide a low take-off and landing speed, the wing cross-section should have a great curvature, while at flying speed the cross-section should be flatter, i.e. the front part of the wing (the leading-edge flap) and its rear part (the trailing-edge flap) must be able to go down and take the initial position due to the action of the control mechanism. To this effect, they must be dynamized, movable.

The next step is a still higher coordination of the wing shape and flight conditions. The problem is that air flows poorly around a composite wing which causes high resistance to the aircraft motion. It is necessary to pass from discretely moving components to changing the entire wing shape. To achieve a controlled change of the entire wing shape, it is necessary to increase its dynamization degree and, ideally, to make each part of the wing movable. Preserving a good wind shape of a wing plays an important role in this situation.

The example of an elastic wing is the NASA-developed wing of a so-called transformer supersonic aircraft F/A-18 [2.8] provided with elastic wings (Fig 2.16). The design of this

Fig. 2.16. Transformer airplane F/A-18

wing is based on rigid components — longerons — installed along the wing length. Several parts of the wing are flexibly attached to the longerons and are moved by means of hydraulic mechanisms. The composite wing formed of several parts is coated with a flexible shell, which allows the leading-edge and the trailing-edge assemblies to be deflected without actually disturbing the contour smoothness.

The decision to give up traditional ailerons, leading-edge flaps and trailing-edge flaps considerably increases the aerodynamic properties of the wing, as well as the aircraft maneuverability and control reliability. The seamless smooth wing has no auxiliary mechanical surfaces so air flows around it well. The flexible wing control drive is simpler and lighter than that of the traditional wing. In addition, the flexible wing is easier to make adaptive, capable of automatically changing its shape depending on the flight conditions. Coincidentally, the transformer aircraft trials were carried out on the eve of the 100th anniversary of the first flight of the Wright brothers' aeroplane. It is interesting that the world's first aeroplane was also controlled by bending the lifting surfaces, when the pilot had to pull cables attached to the wing to deflect its trailing edge down. This technology was used under necessity at that time. The aspiration for maximally simple design in combination with the limited possibilities of very flexible and light weight wings with a fabric covering dictated their conditions.

With aviation advances, the speed and load-carrying capacity of aircraft grew considerably. That required the transition to rigid wooden and then metal construction of wings that were impossible to bend. So the advanced solution — control by changing the surface curvature — had to be replaced with hinge mount of rudders and ailerons. With the appearance of a transformer wing, the development of the aircraft control tools completed a full coil of a spiral. Long after, designers went back to the old, almost forgotten technical solution already at the level of modern materials, technologies and means.

After collecting a sufficient amount of information on the aircraft wing design (most of this information is beyond the scope of this book), one can construct a number of evolution patterns. Take some component — a wing itself, its geometrical shape, surface or internal space — and align the versions of this component in the order of change of a selected parameter.

By coordination (component: wing surface; action: coordinating the shape and flight conditions):

- a fixed wing of unchangeable shape and position
- a solid wing capable of changing its position
- a wing with a deflectable rear part

- a wing with deflectable rear and front parts
- a flexible wing of a changeable shape

By controllability (component: control system; action: simplifying of control):
a wing without operational control

- changing the wing setting angle by displacing the pilot's weight
- manual control of air rudders by means of flexible cables
- manual control of air rudders by means of rods
- semi-automatic control by means of pneumatic mechanisms
- semi-automatic control by means of hydraulic mechanisms
- autopilot

By dynamization degree (component: wing; action: increasing of mobility):

- a rigid wing
- a wing with hinge-deflectable air rudder
- a dynamized wing: movable flaps, leading edge flaps and airbrakes
- a changeable sweep wing
- a flexible adaptive wing with flexible coupling of the front and rear parts
- a wing changing to the rotation mode during taking-off and landing

By the number of surfaces (component: wing; action: adding of new aerodynamic surfaces):

- a wing
- a wing and an air rudder
- a wing and several air rudders
- a wing, several air rudders and leading-edge flaps
- a wing, rudders and leading-edge flaps, airbrakes, trimmers
- all these components trimmed into an adaptive flexible wing

In addition, a number of other wing evolution patterns may be traced, for example, changing the wing shape: rectangular, elliptic, triangular, swept, sweptforward, etc. or complicating the internal structure: placing of fuel, arms, ammunition load, control system components within a wing, etc.

The wing surface evolution may be illustrated by the following sequence: a smooth wing, corrugated wing, swept wing with knife-shaped protrusions, wing with microhollows forming a whirling «air lubrication», etc.

The wing coloring: an ordinary paint, camouflage color, coating with a reflecting mirror layer, coating with a radar-absorbing black paint, chameleon paint, etc.

By the number of wings: a monoplane, biplane, triplane, five-winged aircraft… polyplane, etc.

Perhaps, it is better to stop at this point and to admit that the number of evolution patterns for an aircraft and glider wing starts growing in an uncontrollable manner without bringing us nearer to the answer to the following questions:

- How many evolution patterns can be built for a single component of a system?
- Are all the patterns equally significant?
- Which patterns can be considered the «most essential»?

Such transformation sequences can be built by analyzing the evolution of components of practically any technical system. In each case, we start understanding quickly that the most important thing is absent — the hierarchy of the evolution patterns themselves: it is not clear which patterns are dominating and which of them are subordinate. This resulted in that the number of evolution patterns used in TRIZ started to grow rapidly. According to some researchers, they have developed hundreds and even thousands of evolution patterns. Hence construction of patterns is in many ways a subjective process, which leads researchers to the unstable ground of the trial-and-error method.

Using the evolution patterns for prompting only, searching for real problem solutions by analogy with already known technical solutions described in those patterns is easy enough. But if our objective is obtaining a clear information structure required for high-quality system evolution prediction, we need a strict definition of each evolution pattern, separation of one pattern from another and selecting those of them which describe in the most adequate manner the transformation of the technical system components. This may be achieved by:

- identifying a technical system model corresponding to a really operating system and making a list of the main components of this model;
- understanding how the system described by this model evolves and determining the character and sequence of actions which change its components while passing from one version to another.

2.3. The main components of the system

2.3.1. The functioning system model

When improving technical objects, the most effective approach is considering the object as an organized system with relationships between the components.

The system is a set of interrelated components which interact with the environment as an organic whole [2.4]. The criterion of the system is structuredness, interrelation of component parts, subordination of the entire system to a certain purpose. The purpose of the system is to perform action on a work object to obtain some result. Such an action is called the system's function. A system can perform multiple useful, harmful and just unnecessary functions during operation. The system is characterized by the main *useful* function, i.e. the function for which it was created and designed. The system (for example, an automobile) is often composed of thousands of parts.

To analyze the system, it is necessary to build the correct model. We need to thoroughly determine the system's components (subsystems and components) and to identify links, interactions (or relations) between those components. The principal difficulty in building the system model consists in that dividing the system into parts has a relative, conventional character. Such segmentation depends on the modeling purpose. This concerns not only the boundaries between the system parts, but also the boundaries of the system itself. Also the degree of detail to divide components, i.e., to determine what part of the system can be considered elementary, indivisible for the purpose of analysis, is relative.

What model of a system designed for performing a required function will best match the real technical system?

When improving machines, one has to deal with a special type of systems which are often called technical systems. The «technical system» illustrated by the model suggested by G. Altshuller [2.5] is widely used for many different applications. According to this model, any technical system comprises an engine, a transmission, a working tool, and a control system (fig. 2.17).

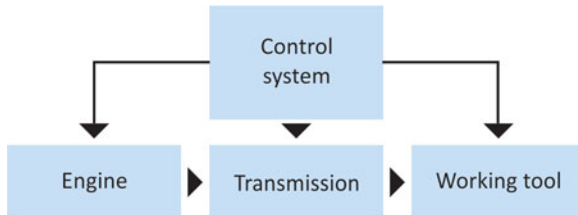


Fig. 2.17. G. S.Altshuller's model of the technical system [2.5]

Here the following circumstance should be taken into account: for correct comparison of different technical systems, these should be considered under similar conditions. It is desirable to choose conditions providing the fullest and most objective description of the system.

During its lifecycle, a system passes through all stages: design, manufacture of the system's components, completing the system using these components. Then follows the system operation, repair and maintenance and, finally, utilization of the system's components after their lifetime has expired. One can obtain certain information on a system by analyzing each of its stages. We are interested in those system's features which manifest themselves during its operation, when a system performs its useful function, produces the product for the sake of which it was created. Tests are the most important moment for everyone who took part in the system creation and preparation for work.

Only by observing a working system and measuring its results can we positively determine the components and structure of a system and the interaction between its components. This is the information needed to make future improvements of the system. So, we must determine the composition and structure of a functioning system model.

Thus, the term «technical system» is better to use for defining an organized set of technical objects. As for a system considered during its operation, let us use the term «functioning technical system».

A functioning technical system is a system that combines all the components required to perform a necessary function that is considered and analyzed in the course of the system operation.

For example, a functioning technical system that performs the function «to carry a cargo» can comprise the following components: a truck, fuel, a trained driver aware of the destination point, air, road, force of gravity and many other things. It is clear that the work object (cargo) is also included in the system, because defining the system itself is impossible without defining the cargo.

The composition and structure of a functioning technical system is shown in Fig. 2.18.

The functioning system described by this model exists at two levels: *information and objective*, interacting through an operator or a control system. At the information level, the operator or control system should be provided with information about the sequence of performing the technological process that allows the function of this system to be realized. In addition, the operator needs knowledge, skills of control of the system's object part* and, in

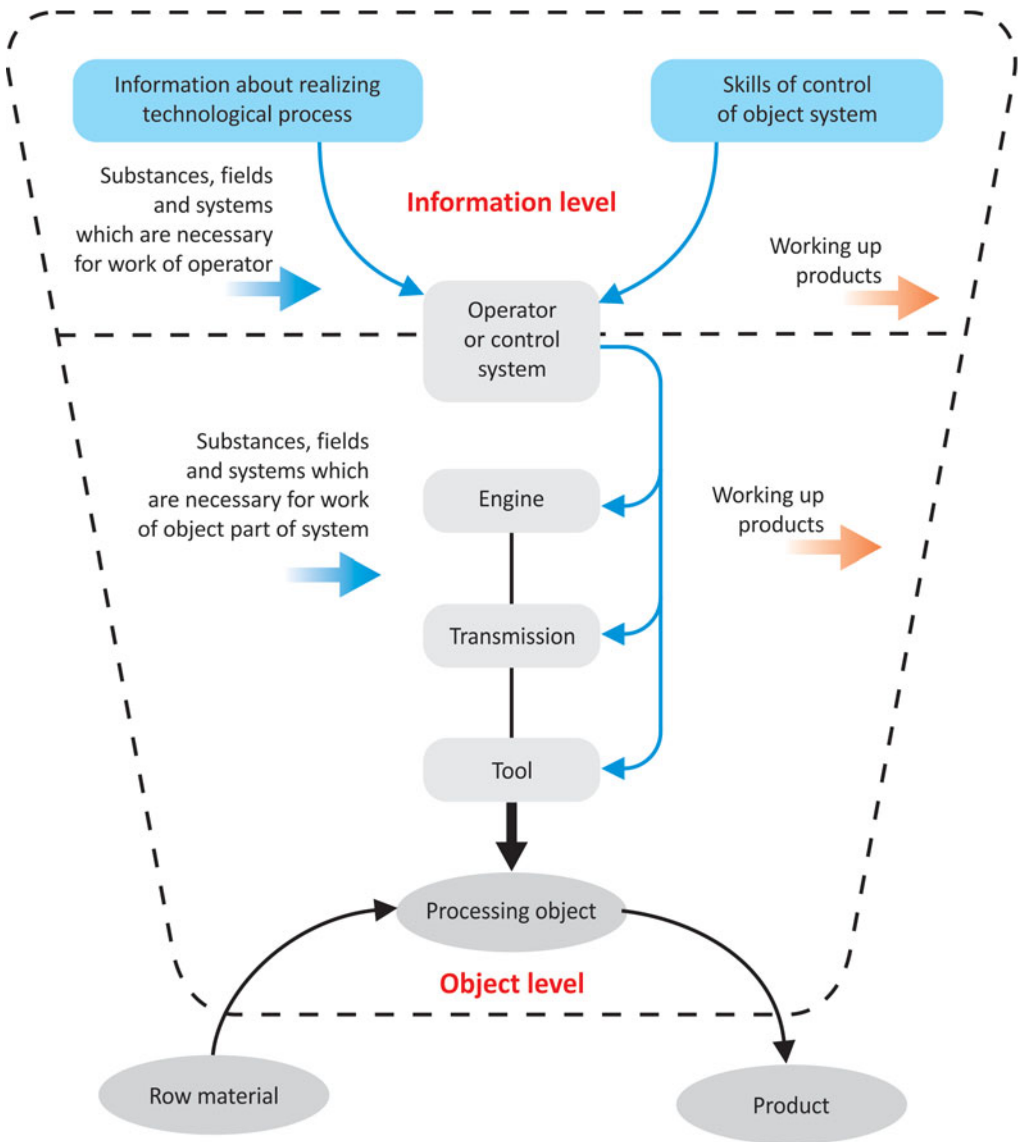


Fig. 2.18. The functioning technical system model

case a special system is used for control, it is necessary to provide a certain action algorithm for this system. The operator performs intellectual operations and, based on the decisions taken, performs certain controlling motions. The control system receiving signals for executing some controlling motions or acting in accordance with some algorithm stored in its memory operates in the same manner.

The object part of a functioning system is a set of all material objects used for providing the system operation. At the object level, the same parts as those of the technical system may be included in the model of a functioning technical system (see fig. 2.18): an engine, a transmission and a working tool acting on a work object.

The character of interaction between the working tool and control system depends on the composition of a specific system. The object part of a technical system is often a machine — an assembled set of technical objects, such as, for example, an automobile. Here we have all the components of the classic four (By G. Altshuller): an engine, a transmission, a working tool and control system components. The operator exists in this system as an externally introduced component.

In a great number of functioning systems, the operator's capabilities are used to a fuller extent, as, for example, in a «system for drawing» or a «system for tightening a screw». The employed material objects are a pencil or a screw-driver and the functions of a transmission, engine and other functions, i.e. the control system functions are performed by an operator himself. In this situation, it is impossible in principle to determine the classic composition of a «machine» at the expense of technical objects only.

A functioning system may employ any artificially created technical object, for example, a hammer in the nailing system. Natural objects, for example, using a stone for nailing, may also be employed for organizing functioning systems. In specific cases, a functioning system may only comprise an operator. An electrician replacing a burn-out light bulb may serve as an example. A man performs not only the function of an engine and a transmission, but his arms also serve as a working tool.

It would be expedient to supplement a functioning system with a work object and consider this object in dynamics: from a blank part through operations at different processing stages to a finished article.

Another significant component of a functioning system is substances, fields and additional systems necessary for the functioning of an operator and machine. For an automobile, it may be fuel, lubricants, air, heat and many other things, even a road; for a ball-point pen, it may be ink, heat, gravitation. For a copying machine — electricity, powder, document to be copied. An operator needs air, heat, food and water, clothing, footwear, goggles, etc.

In addition, when considering the composition of a functioning system, it is necessary to take into account the processing products of supplied substances, fields and systems and to think how these can be removed or used, for example, exhaust gases of an automobile should be removed from a driver, and the heat produced during engine operation may be used to heat the cabin.

As mentioned above, it is necessary to differentiate between a functioning system and a system of technical objects — a machine. A functioning system *performs* a certain function at a given moment of time. As for a machine, it is initially *designed* to perform its main function i.e. a machine may be part of a system that performs the same function as the main function of this machine, but not necessarily. For example, an automobile is designed and manufactured for carrying loads and passengers. If it is used in a functioning system having the same designation, the working tool of the automobile will be its body or passengers' cabin and the work object will be cargo or passengers.

If somebody wants to use the automobile for compacting a garden path, the functioning system will be quite different. The system's function will be «to compact soil», the wheels will be its working tool and garden soil will be the work object. An automobile may also have other applications: in action movies, they break a gate (work object) with a car bumper (working tool).

The roar of a poorly tuned engine «processes» the ears of neighborhood residents at night.

A driver resting in the shade of his truck will be a work object while the truck body itself will be a tool providing the shade.

And, finally, an automobile standing in a garage does not perform any function, but, on the contrary, is itself an object processed by the garage.

All the functioning systems will be different in this case, though they use the same system of technical objects (machine), the automobile.

Any system of technical objects has a quite different status depending on what functioning system it belongs to and where it is applied.

The simple conclusion follows from here:

- A machine, technical or natural object may serve to organize numerous functioning systems.
- Different versions of a functioning system may include different machines, technical or natural objects.

2.3.2. The functioning system in the context of patent legislation

Because we are dealing here with the handling of technical and patent information, it is important that the functioning system concept be matched with the models used in patent legislation. In addition, checking for matching will allow us to make a conclusion about the correctness of a suggested model. The thing is that the patent legislation is based on one of the best developed information organization methods and can serve as a good landmark for checking any proposed system's models, including the above-mentioned one, for adequacy.

Comparative analysis proves that the above-described model of the functioning system matches well the models used in the patent legislation. For example, in addition to the description of a patented device itself, the obligatory part of a patent application is a description of this device «in operation». This allows better understanding of how the patented device will operate being a part of some functioning system.

In addition, the functioning system's composition matches well the list of objects of patenting provided in the patent legislation (Fig. 2.19).

According to p. 2, Art. 4 of the Patent Law of Russian Federation [2.15], objects of invention are *a device, a method, a substance, a strain of a microorganism, cell cultures of plants and animals, as well as use of a known device, method, substance or strain for a new designation.*

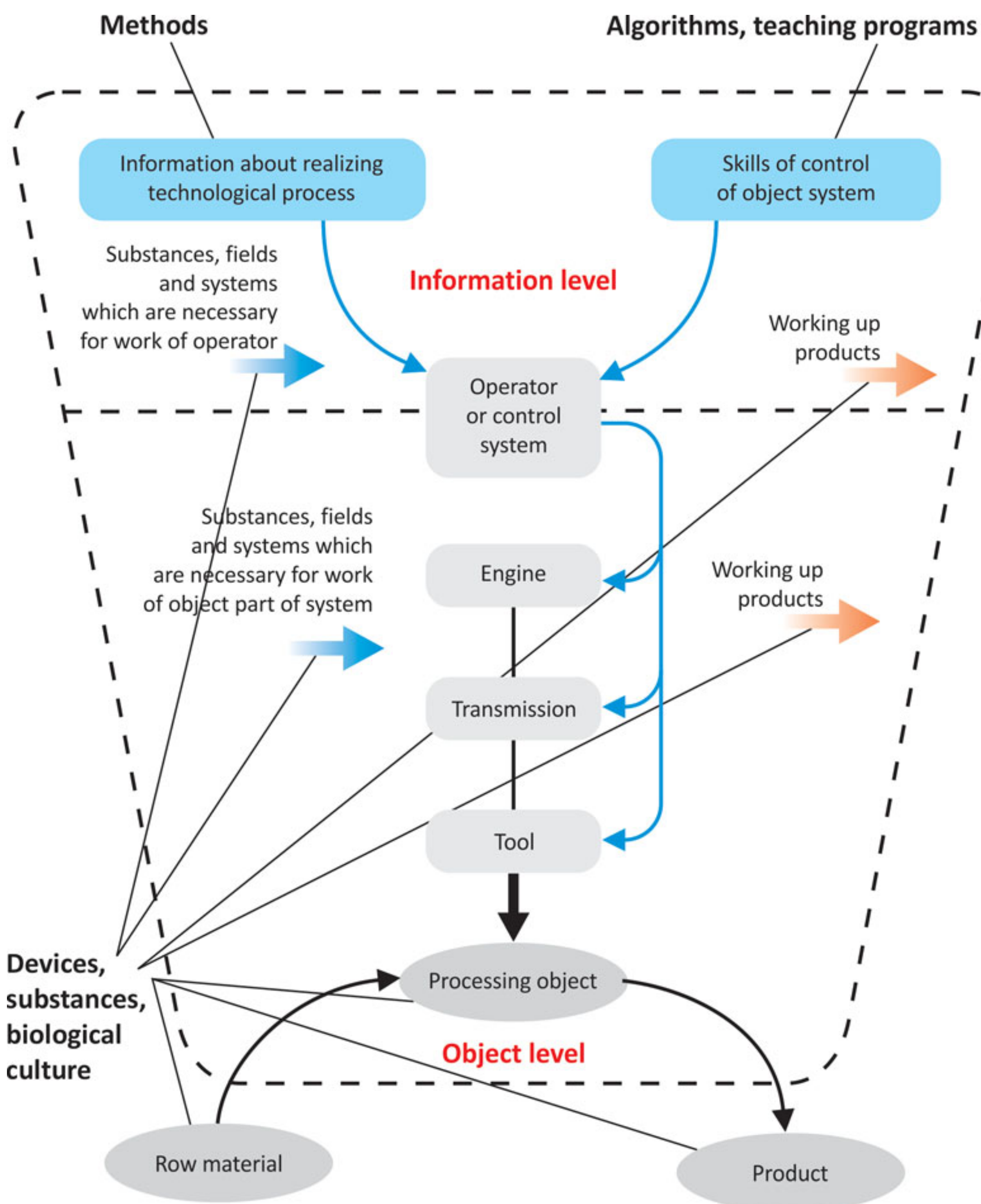


Fig. 2.19. The functioning system as a set of objects of invention

The Patent Law interprets «a *device*» as a system of components arranged in a space and interacting with each other in a special manner. Of importance are the availability of structural components, links between the structural components, shape and relative position of structural components, material and other parameters.

In the proposed model of the functioning system, the notion of a «device» belongs to the object part of the system and may be primarily applied to a set of technical objects controlled by an operator or a control system. Devices may also be used to provide a machine and operator with necessary resources.

The notion of «*method*» defined in the Patent Law as one of the objects of invention only refers to the information part of the functioning system and is a description of the process of interaction of the system's material objects that should be kept in the memory of an operator or a control system. Use of a known device, method, substance or strain for a new purpose, or in a new manner, is another special type of patent described in a separate law.

A «substance» as a set of interrelated components, strains of microorganisms, cell cultures of plants and animals (sets of cells characterized by similar features) stipulated in the Patent Law can also be parts of a functioning system. These are primarily substance resources necessary for the operation of the system's object part and operator, as well as processing products.

As to the model component «*object part control skills*», the object of patent protection is training courses, instructions, knowledge obtained by training, as well as algorithms and computer programs for auxiliary and principal automatic control systems.

Because the set of legally stipulated objects of patenting and the set of components constituting the functioning system model are practically the same, we can draw a conclusion that our model of a functioning system is correct.

Hence the most adequate and effective will be those evolution patterns which describe transformations of the main components of the functioning system model. They are:

- a set of material objects,
- a description of the process of their interaction,
- substances, fields and systems necessary for the work of the system's object part and operator,
- substances, field and systems as processing products,
- algorithms and programs for automatic control systems,
- training programs and courses.

2.4. The functioning system evolution

2.4.1. Actions performed at a step-by-step evolution of a system

In section 2.3, we identified the main components of the system which, when transformed, may allow transfer from an existing modification of a system to new, improved ones.

As was shown in section 2.2., performing an action several times on some components of the system gives us a sequence of embodiments of these components which can be combined into an evolution pattern. It is necessary to understand which actions should be performed in in what sequence on the main components of the system while passing from one system's modification to another.

Evolution of biological systems and evolution of technical systems have much in common, but there is one basic difference. In biological systems, transformation mechanisms are laid down in an organism by nature and to start evolution, it is only required to provide the initial impulse and to maintain favorable conditions. All the rest will be done by a plant or an animal itself. As for technical systems, each transition from one version to another is only performed due to the external interference of a man, a subject. In the absence of man, a system generally evolves only towards failure of constructions, gradual loss of nuts and bolts and other fine parts.

One of the main postulates of the modern technical philosophy states that all systems evolve according to objective laws [2.16, 2.17]. These laws reflect the existing, stable and recurring interactions between system components, between systems themselves and the environment. TRIZ uses 9 laws of technical system evolution [2.5] proposed by G. S. Altshuller (see table 1.2).

Some researchers [2.1, 2.3 and 2.4] think quite reasonably that the three first regularities united in the «Statics» group (table 2.1) are *criteria or conditions* necessary for the occurrence and existence of any technical system. It should be noted that this statement suits in the best way possible the functioning system we are considering now. In this connection it seems well-founded to divide the evolution laws into two big groups [2.2], the first of which — «*Criteria or conditions of occurrence and existence of a functioning system*» — comprises the laws of system organization (laws 1 to 3) while the second one — «*The evolution laws of a functioning system*» describes the system evolution laws (laws 4 to 9).

In the context of such division, the system existence cycle may be illustrated by the following diagram (Fig. 2.20). According to this diagram, every functioning system can be in two states:

- at the creation or transformation stage («in a workshop») — at this stage, laws 1 to 3 (see. the Table) manifest themselves;
- at the stage of performing its function and evaluation of operation parameters («on a testing ground»). Here, we can trace the operation of laws 4 to 9.

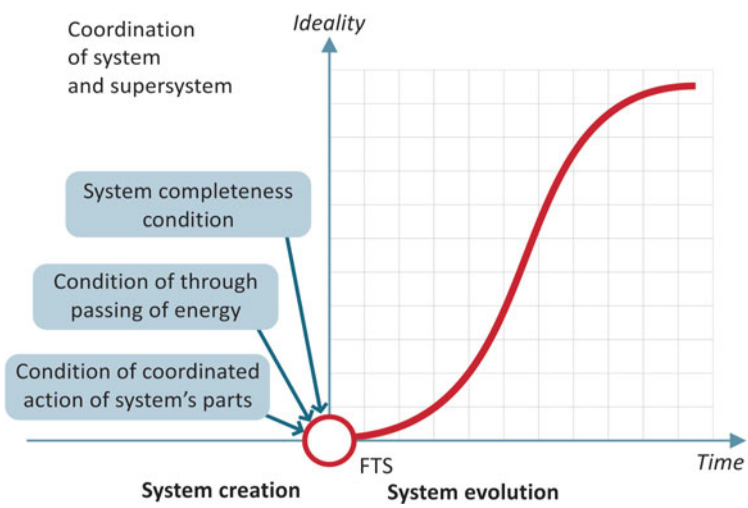


Fig. 2.20. Creation and evolution of a system

These two states differ fundamentally and have considerable qualitative distinctions. When a system is disassembled and is being updated «in a workshop», actions necessary for preparing it for testing are performed on it. This may be accompanied by changes both in the set of components of the system and the arrangement and relations between the components. Thus, at the creation and transformation stage, we can only speak about some set of components of a functioning system, but the system itself will only appear after all the steps to assemble the system are completed. The result of system transformation is coordination of the parameters of all its components, primarily of the working tool and work object. Only after such coordination has been achieved, we can say that the system is ready for transition to a new qualitative state, from a mere «set of components» to a «functioning technical system», which allows testing its new embodiment and comparing its new operational parameters with those previously demonstrated.

The evolution of a functioning system can be evaluated by its *ideality* by comparing its different embodiments. We will consider ideality in detail in section 4.3.1, now we will use the short definition: the more high-quality products the system produces and the fewer expenses this needs, the more ideal the system is.

Increasing the ideality is the main condition of the system viability during its *evolution*. The system ideality is achieved by improving the coordination of parameters and modes of operation with the operation conditions. During operation, the system's disadvantages are being revealed: either the main function performance quality is poor or the system functioning cost is too high; therefore, the work on system transformation and creation of its new modifications is permanent.

During the system evolution, the change of its operational parameters and, hence, the change of its ideality are illustrated by an S-curve (Fig. 2.20). At the initial period of system evolution, its operational parameters improve slowly, then follows their explosive growth and subsequent deceleration as some object limit is being approached. After reaching this limit, the system is replaced with a new one and the cycle recurs.

Deciding whether the new version of a functioning system is improved compared to the previous one is only possible after testing it. The designers' efforts are focused on the process of creation of new versions of the technical system. The system evolution illustrated by an S-curve is the visible result of the process. Thus, the continuous system evolution process is realized in a step-by-step manner, through a sequence of discrete transformations of the system structure (Fig. 2.21). As the system evolves, the available improvement resources are becoming fewer, evolution decelerates and the system inevitably transfers to a new level which provides new resources for further evolution.

It is clear that transformation of the system itself needs some actions performed on its components (Fig. 2.22). What actions exactly?

In literature, there are available lists of actions designed for transformation of technical system components, for example, principles for technical contradictions elimination and a set of standard solutions to inventive problems by G. S. Altshuller [2.19]. There have been revealed 50 principles, part of which are collected in a special table which allows constructing a contradiction available in the problem and offers hints for resolving this contradiction. Also known are 76 standard solutions to problems where actions necessary for the transformation of technical system components are described through so-called Su-field models* [2.5]. Each principle or standard is an indication of how technical system components should be transformed. Here are specific examples.

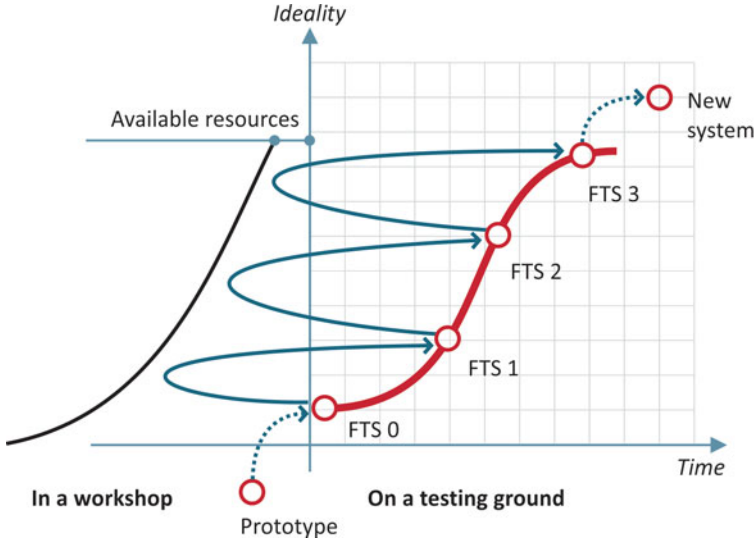


Fig. 2.21. Step-by-step system evolution

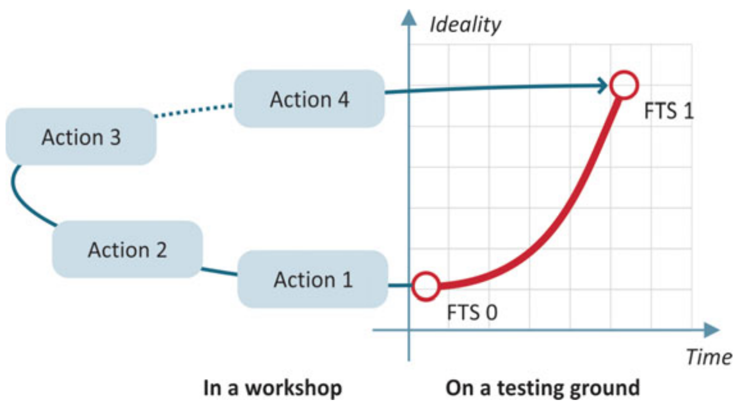


Fig. 2.22. System transformation as a sequence of some actions

From the list of Altshuller's principles:

Principle 3. «Local quality principle»:

a) transition from a homogeneous structure of an object (or environment, external action) to a heterogeneous structure;
b) different parts of an object must have (perform) different functions,
c) each part of an object must be in the most favorable working conditions.

Principle 37. «Use of thermal expansion»

a) use of thermal expansion (compression) of materials,
b) use of several materials with different coefficients of thermal expansion.

From the list of standard solutions to inventive problems:

Standard 2.1. Transition to more controllable fields.

If a Su-field system is given, its effectiveness may be increased by replacing an uncontrollable (or poorly controllable) working field with a controllable (well-controlled) field, for example, by replacing a gravitational field with a mechanical field, electrical field, etc.

Standard 2.2.5. Structuring a field

If a Su-field system is given, its effectiveness can be increased by transition from homogeneous fields or fields having a disordered structure to heterogeneous fields or fields having a certain spatial structure (constant or variable).

Another source of information about actions aimed at system component transformation may be the Interindustry Collection of Heuristic Object Transformation Principles compiled by A.I. Polovinkin [2.20] and comprising over 250 principles divided into 15 groups:

- *quantitative changes;*
- *shape transformations;*
- *spatial transformations;*
- *temporal transformations;*
- *motion transformations;*
- *material transformations;*
- *transformations through elimination;*
- *transformation through addition;*
- *transformation through replacement;*
- *differentiating;*
- *integration;*
- *use of preventive measures;*
- *use of reserves;*
- *transformation by analogy;*
- *combining and complex synthesis;*

The Collection includes a variety of transformations pertaining to the parameters of technical system components, examples of their technology, solving methods of technical, and inventive problems, etc.

Examples of the heuristic principles from the «Transformation through replacement» section:

9.2. Replacing of sliding friction with rolling friction.

9.3. Replacing of machining with working without metal removal.

. . .

9.11. Converting one physical magnitude into other ones.

. . .

9.17. Replacing an object (component) with a much simpler one.

Analyzing the three above-mentioned lists of transformations — principles of resolving technical contradictions, standard transformations of Su-field models and heuristic principles — shows that a series of actions (with inessential variations) practically double one another. Nevertheless, it is possible to select about 150 significantly differing transformations which may be used, with a certain degree of conventionality, to transform technical system components. They are not few, however. To use such a database for practical purposes, it is necessary either to reduce it or to structure it identifying several main actions and using the rest of them as auxiliary ones.

Let us try to understand which actions aimed at transformation of the functioning system components are main, essential and which of them may be considered as auxiliary. To answer this question, let us consider in detail the process of system transformation from one version to another.

2.4.2. The three-stage algorithm of system transformation

According to V. M. Petrov, while transforming a system into its new, improved version, it is necessary to satisfy three existence conditions of a complete, fully operable functioning system.

The appearance of a new modification of a system is due to a stepwise fulfillment of the following conditions:

- *providing the «conformity of the system's composition with the system's function» [2.21];*
- *establishing links between the system's components;*
- *coordinating the parameters and modes of action of the system's subsystems.*

Each of these stages is accompanied by the accumulation and organization of resources required for the fulfillment of the next stage. As the system existence conditions are being fulfilled, the *coordination* between the parameters and operating modes of the system's interacting parts is also growing. Eventually, it allows a better coordination between the working tool and the work object, as well as a better coordination between the system and the environment.

Fulfilling these three conditions of a functioning system may be presented in the form of a three-stage algorithm (Fig. 2.23) that shows what actions are to be performed while transforming the system's components and organizing its new version.

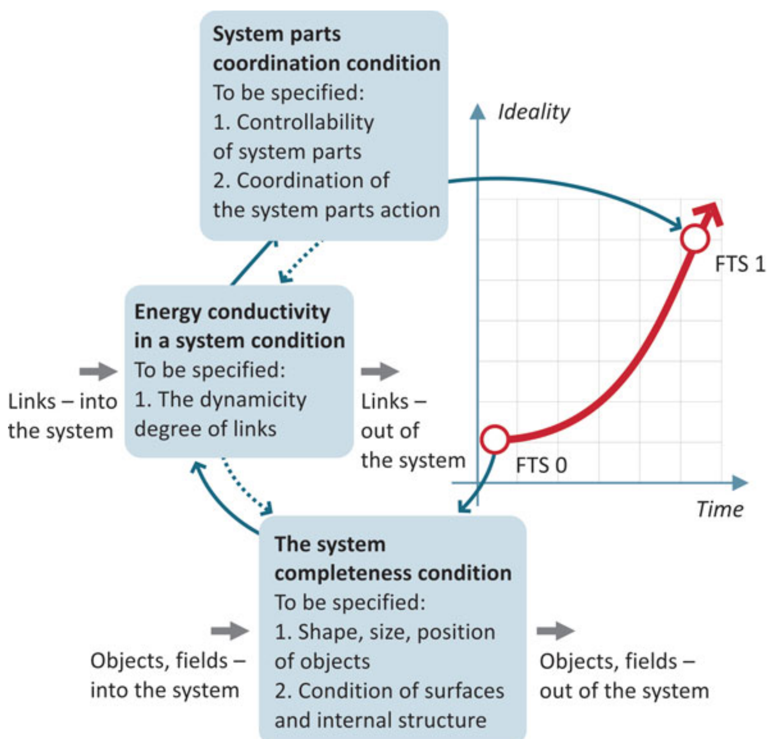


Fig. 2.23. The three-stage system transformation algorithm

Stage 1

At the first stage, the main system existence condition is satisfied — the availability of all system's parts required for function performance. When designing a new version of a system, we can add, remove, or replace components such as: objects, fields and processes. A special method of producing new objects is segmentation, dividing initial objects into a larger or smaller number of parts. At this work stage, the system is supplied with the major portion of substance and field resources.

Having coordinated the number and type of the added components, it is necessary to coordinate their shape, size and mutual position and then to specify the properties of their material, properties of their internal structure and the system components' surface parameters. For fields, it is necessary to coordinate the position, size and shape of the field spread area, as well as the field parameters.

Hence we may speak about the following actions to be performed within the framework of this stage.

- introducing components into a system;
- removing components from a system;
- segmenting the system's components;
- changing the shape and size of the components;
- complicating the internal structure of the components;
- changing the surface conditions of the components.

Stage 2

The next condition to be provided is the *passage of energy between the functioning system's components*. This condition is fulfilled by organizing internal links within the system. Actions to be performed while organizing these links are similar to those performed while providing the system completeness: we can introduce or remove one or several links or replace one link with another.

After introducing the links where necessary, the character of these links should be specified. The links may be both rigid and flexible, with different degrees of freedom, for example, different types of hinge joints. Material objects may be linked by fields. It is necessary to provide the main direction of energy transfer in the system: from an engine through a transmission toward a working tool. In addition, it is important to establish an energy link between the system's parts and the control system.

Establishing the rigid and flexible links between the system's parts results in the formation of some logically connected frame, the system's base for attaching movable, dynamized components of the system by means of flexible links. The greater the number of flexible components in the system and the higher their degree of motion freedom, the more dynamic and more adaptable to working and environmental conditions the system is. No doubt, this concerns both material objects and fields. Light, electromagnetic and acoustic fields are dynamic by nature and their parameters are not difficult to change. There are, however, «inertial» fields and their parameters are not easy to change promptly, for example, heat field or radiation field.

Within the framework of this stage, we can speak about the following transformations:

- introducing components-links into the system composition,
- removing components-links from the system composition,
- providing the mobility of links
- providing field parameter changes

Stage 3

This stage checks the system functioning for coordination with the fulfillment conditions of the main operation — the action of the working tool on the work object. Because the composition, structure and character of interaction of the system's components have been substantially determined, such a check allows optimizing the system operation and specifying the parameters of its components.

Here it is necessary to make sure that the system components' parameters are sufficiently coordinated *during the fulfillment of the first two conditions*. If necessary, the final coordination of the system components' interaction should be performed in accordance with the technological process of its operation.

Also, it should be determined whether the system needs operating control during functioning. Most of the system's components were coordinated at the previous stages and have rigidly set parameters providing the system operation. The parameters of some components need coordination by adjustment, that is, by regularly eliminating their deviations from values which are optimal for a given system. Parts of the system's components require *operating control* — constant change of parameters in accordance with the changing operation conditions.

A system having a high degree of initial coordination is comparatively easy to control, almost no operating control is needed. For example, roller-coaster cars are well coordinated with the motion conditions and only need to be accelerated at the beginning and stopped at the end of the path (Fig. 2.25). When coordination is mainly due to operating control, the operator's job is not easy. A device for Buggy-Rollin may serve as an example of minimal coordination. This device is composed of a large number of rollers attached directly to a human body: to knees, feet, hands, elbows, etc. The rollers are not attached to each other and are only combined into a system by the roller-skater's efforts. It is not easy to Buggy-Roll because one should not only follow the path but also to keep each roller in the necessary position.

Checking the system's components for the coordination**** degree and final coordination of all the parameters and actions of the components are the finishing design stage which includes the following transformations:

****) Matching, coordination (from the Latin: *Co* — together and *ordinatio* — ordering) — alignment of a number of different processes, such as coordination of the fuel to the movement of the piston in the cylinder internal combustion engine.

- providing operating control
- introducing new operations into a technological process
- removing operations from a technological process
- dividing operations into smaller ones
- combining several operations

The algorithm of reverse sequence of system-transforming operations:

A) Specifying the desired result

1.1. Coordinating the mutual action of the working tool and the work object

This requires that the working tool parameters be changed depending on the requirements imposed by the work object.

1.2. Coordinating the actions of the remaining components of the system.

To provide coordinated action of the working tool and the work object, it is necessary that all the system's components act according to a preset algorithm.

B) Operations on the system's components

2.1*. Introducing objects, processes and links and coordinating their parameters*

To provide dynamization, it is necessary to have several objects or processes and links between these. If there aren't enough objects, they may be obtained by adding objects or by segmenting the existing ones and transforming the available links.

The geometrical shape, location and size of newly introduced objects, as well as the surface properties and internal space structure should be coordinated. For new links, it is necessary to coordinate their character with the requirements made of the interacting objects.

2.2. Providing dynamization

To provide sufficient controllability, the technical system's components and processes in them must be dynamical enough.

2.3. Providing controllability

For coordinated action, the technical system's components must be controllable, i.e. must preserve or change their parameters in accordance with the operator's or control system's commands.

2.4. Check coordination.

Ensure that the required parameters of the working tool match the conditions of its functioning, and can easily and quickly change with these conditions.

Thus, the coordination degree of the parameters and operating modes of the system's components with the technological process performed by the system increases while following the three-stage algorithm. Three levels of such coordination may be identified.

- *Initial coordination* of the system's parameters with some averaged operation conditions; it is performed at the design and manufacturing stage. This coordination may be static — a car is stable both when it stands and when it runs. The example of dynamic coordination is a bicycle that is stable only when moving.

- *Regular coordination* by compensating adjustments occasionally performed during the system operation.
- *Final coordination* of the system's parameters with variable operating conditions, which requires a possibility of operating control, i.e., quick and comparatively simple change of system's parameters with a change of its functioning conditions.

In practice, the three-stage algorithm may be started from any stage. As is evident from the scheme shown in Fig. 2.28 each transition between the three main stages of the algorithm imply a possibility of going back to the previous stage. This is necessary if performing a required transformation becomes impossible at some algorithm stage due to lack of resources. Then the analysis should be repeated within the framework of the previous stage and necessary additional resources should be obtained, i.e., the sequence of operations prescribed by the algorithm may be both forward and reverse.

At the beginning of the system transformation process, the reverse sequence of operations is preferred. It will be more efficient to start work from a detailed specification and correction of the system's components coordination, special attention being paid to the working tool and work object. Doing this work will better allow understanding which components should be changed and the manner they should be changed such as: which components should be additionally dynamized or need dynamicity reduced; which additional components should be introduced into the system; or how their position, shape and size can be coordinated.

In other words, to design a system, it is necessary:

- first, to understand how exactly we want to improve the coordination between the working tool parameters and the work object.
- then to find out how to provide coordination between all the remaining parts of the system, i.e., to specify the technological process of system operation,
- further, to make sure that the system has necessary components and to introduce missing ones, if necessary,
- next, to check whether the introduced components are sufficiently dynamized to provide an effective control,
- to dynamize these component, if necessary.

After completing the entire analysis cycle, it is good to repeat it until the designer obtains a satisfying result and resolves the occurring contradictions.

Bringing together all the actions performed at each stage for transforming the components of the object part of a system will give us a list of ten basic actions to be performed in order to transform the system.

To create a new version of a system, the following ten actions must be performed in a certain order:

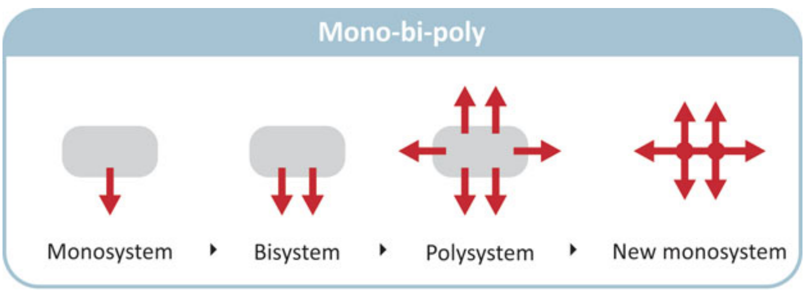
- *Introducing components and links into the system composition.*

- *Removing components and links from the system composition.*
- *Replacing components and links of objects with others*
- *Dividing the system's components into parts*
- *Changing the shape and size of the system's components*
- *Changing the internal structure of the system's components*
- *Changing the surface condition of the system's components*
- *Providing the mobility of links between the system's components and possibility of changing other parameters of the system*
- *Providing and simplifying of operating control*
- *Checking and improving the coordination of the system's components operation.*

Chapter 3

Description of the Most Relevant Evolution Patterns

The evolution patterns produced by analyzing the evolution of a great number of real systems are an effective tool for describing the sequence of the considered system versions.



3.1. The main evolution patterns of technical objects

This is how the list of evolution patterns was formed. First, a database of evolution patterns of technical structure components known in TRIZ were collected [1.2., 2.1 — 2.4, 2.22 — 2.26]. Then a single table was created (table 3.1) containing in its left column the main actions to be performed to transform a system according to the presented list.

After that, suitable evolution patterns were selected for each action. For example, performing the «*Segmentation*» operation several times in a row will give a set of versions of this component to form the «*Segmentation of objects and substances*» pattern. The resulting list of the ten basic evolution patterns is given in the right column of table 3.1. A pattern is a set of system versions and description of the transitions between these versions. Thus, constructing a correct model of a system and analyzing the sequence of actions aimed at its transformation produced a list of evolution patterns most adequately describing the transformations of a technical system and its components. Recall the functioning system model (see Fig. 2.18) and you will notice that the evolution patterns mostly describe the evolution of material components of a system — substances, fields and forces.

For a technological process, topical will be the following evolution patterns which describe transformations of technological operations:

- *introducing new operations into a technological process;*
- *removing operations from a technological process;*
- *replacing operations with others;*
- *segmenting operations into smaller ones;*
- *combining several small operations.*

Let us confine ourselves to considering the evolution patterns of system's material objects as our interest centers primarily in the evolution of technical objects. More detailed consideration of the evolution patterns for processes is beyond the scope of this book.

The main ten evolution patterns presented in Table 3.1 can be used to build an information structure. Below we are giving a detailed consideration to them.

Table 3.1

	System transformation actions	Corresponding evolution pattern
1	Introducing of elements and links into the system composition	Transition from a mono system to bi- and polysystem (Mono-bi-poly)
2	Removing of elements and links from the system composition	Trimming a system
3	Replacing one elements and links with others	Expanding-trimming a system
4	Segmenting the system's elements	Segmenting of objects and substances
5	Changing the shape and size of the system's elements	Geometrical evolution of objects
6	Changing the internal structure of the system's elements	Object structure evolution
7	Changing the surface state of the system's elements	Object surface mtcrorelief evolution
8	Providing mobility of links between the system's elements and possibility to change its other parameters	Dynamization
9	Providing and simplifying of operating controll	Increasing the system elements controllability
10	Checking and improving the coordination between the system's elements operation	Increasing the action coordination of elements

In the literature, one can come across a number of technical system evolution patterns developed at different times. These patterns may be used in problem solving, but they are generally formed as particular cases or more developed versions of the above-mentioned ten patterns.

3.1.1. The «Mono-bi-poly» transition

This pattern describes system versions obtained by adding components which are either analogous to those already available in the system or new and which can perform additional useful functions.

If some system does not cope with its function, it would be logical to supplement it with one more similar system. Such a combined system is called bi-system. In addition to increased productivity, a double system can have absolutely new properties different from those of a single system, which often allows previously impossible operations to be performed.

«With one marker, you can outline everything in the world except this marker itself. With two markers, one can outline everything in the world». This funny example shows how combining two systems into a new bi-system leads to the appearance of new capabilities. Combining is understood not only as mechanical combination of two systems. It rather means organizing their joint use for performing a required function.

The number of components that may be combined into a system is unlimited. The example may be a squadron composed of several war vessels under a single command. Such a system formation, *poly-system,* has much more capabilities than an equal number of ships not united into a system and acting independently.

Build-up of similar components of a system generally proceeds to some limit. Having reached the limit, there often occurs a transition from a poly-system having several similar components to a partially or fully trimmed *mono-system of a higher level*. Such a transition may be illustrated by a sailing vessel. The number of sails increased with the perfection of the vessel design. A five-mast vessel had more than fifty large and small sails, which were very difficult to control and required a well-trained crew. That continued till the advent of a steamship, when numerous masts and sails were replaced with a higher-level mono-system — a steam engine equipped with paddle wheels.

Arranging the system's versions obtained by introducing new components in the order of expanding the system's composition produces the «Mono-bi-poly» evolution pattern. The trend traced in the steps of this pattern shows a sequential build-up of identical (or very similar) components in the system. Introducing not only objects similar to those already available in the system, but any other necessary objects — carriers of new functions — may be considered as the development of the «Mon-bi-poly» pattern.

Adding new components is aimed at increasing the technical system productivity, reliability and technological process performance quality. The sequence of system's versions in the «Mono-bi-poly» pattern is described in detail in section 2.1 (see Fig. 2.7). Also, the simplified version of the pattern may be used for practical purposes.

The sequence of generalized steps of this pattern may look like this (Fig. 3.1):

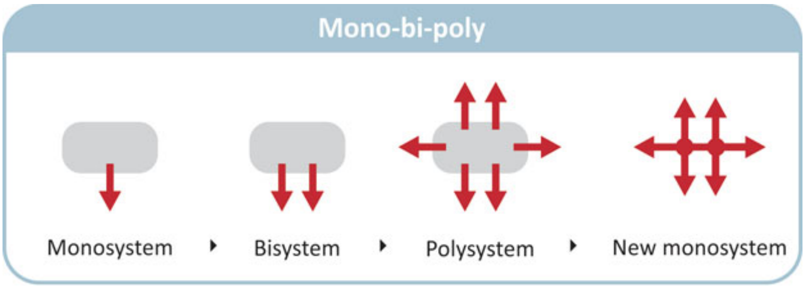


Fig. 3.1. The «Mono-bi-poly» pattern steps

The initial version of the «Mono-bi-poly» pattern is some *single object or system*. Next steps may be:

- *introducing a single additional object,*
- *introducing several additional objects,*
- *transition to a higher-level mono-system.*

It is generally implied that an object analogous to the one already available is introduced into a system. This, however, is not required. New objects, fields and forces performing additional functions may also be introduced into a system. The principal rule of system transformation according to this pattern is providing functional expansion of the system, getting a chance to better perform the system's main function or some additional useful functions.

Example: street lamp (Fig. 3.2). A single lamp can illuminate a limited territory, for example, part of the roadway or sidewalk. A twin lamp illuminates both the roadway and the sidewalk. Large areas in town squares are illuminated with lamp comprising many lamps arranged in a circle, which are enough to create a large light spot. However, in any case the light spot size will be limited by a support span.

Sometimes it is necessary to illuminate a large territory but its size is not known in advance. Such situations may occur during natural calamities or technological catastrophes, industrial or man-made disasters. Continuing to follow this pattern by increasing the number of lamps requires a great many lamps which must be delivered to a needed region, installed and powered... This would take very much time and considerable expense.

Considering the «Mono-bi-poly» pattern, we can see that it ends with a transition to a super-system, to a higher-level «lantern». In our case, it is a transition to a single high-power light source which is more effective for illuminating large areas. The illustration is a space solar mirror made of superfine light-reflective film. The mirror is launched to space and, if necessary, is oriented on some earth region which needs illumination [3.1].

3.1.2. Trimming the system's composition

The «Trimming» pattern describes system modifications obtained by removing components from the system.

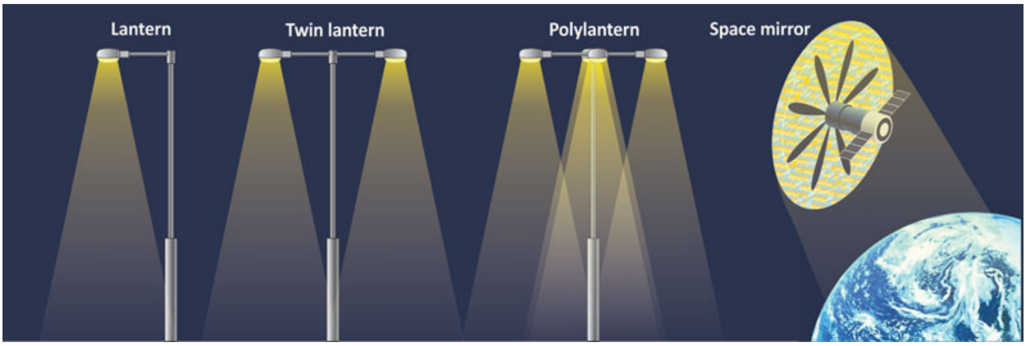


Fig. 3.2. Evolution of territory-illuminating systems

When perfecting a system, it is often necessary to make it less expensive. Cost may be reduced by using less expensive materials, optimizing the shape of system's parts, or lowering the machining quality. However, resources of this type are few. TRIZ treats trimming — one of the main ways of reducing expenses necessary for performing a required function — as a reduction of the system's composition and optimization of its structure. This means that a trimmed system provides a good functioning quality at a smaller number of components than that of the initial system.

For example, drawing a color picture needs a complicated system — a set of colored pencils. Each pencil has a lead of one color enclosed in a wood holder. Because there are many pencils, a special box is needed to keep them. To simplify the system, dispensable parts should be removed from its composition. One holder may be used with several leads, i.e. an automatic multicolored pencil may be created. Such a pencil will be more complex that an ordinary one, but definitely simpler than a box of pencils.

The maximally trimmed modification of the system will be a pencil having a multicolored lead. It does not differ from an ordinary pencil by appearance, but its lead consists of four to six segments of different colors. Turning the pencil while drawing allows choosing the required color.

In trimming, one can always trace a trend toward technical system simplification while the quality of the useful function performance is preserved. This process consists of removing a component from a system and transferring its function to the remaining components. It sometimes happens that all the components can be removed from the system and conditions are provided that allow its function to be performed, for example, by some other system operating nearby. In this situation, we can say that the trimmed system turned into an «ideal system»*.

One more trimming method implies replacement of complex, expensive components with simpler and cheaper ones.

Trimming is aimed at reducing the manufacturing and operational costs of a technical system.

*) The ideal system — a system that does not exist, but its function is executed. Like a model of a black body used in physics.

The initial version of the «Trimming» pattern is some *initial structure of a technical system* that needs simplification by removing a number of components from its composition.

The pattern may include the following steps (Fig. 3.3):

- *removing a single object from a system,*
- *removing several objects from a system,*
- *transition to a maximally trimmed system,*
- *using an ideal system.*

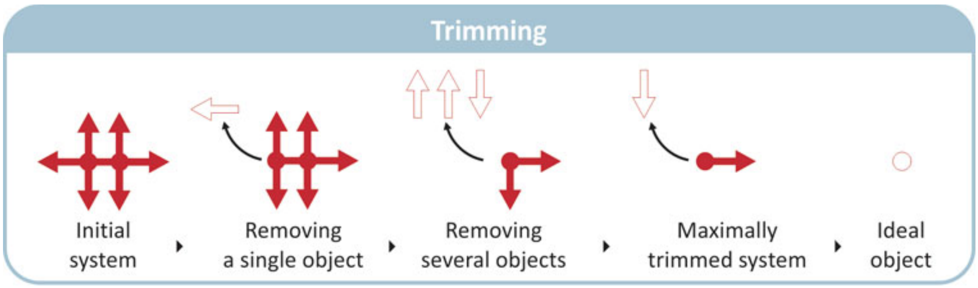


Fig. 3.3. The steps of the «Trimming the system's composition» pattern

Example of trimming: Jablochkoff candle (Fig. 3.4)

Prior to the appearance of a light bulb having a filament, an electric arc was used. The arc was formed by passing current through carbon electrodes. The ends of the electrodes were placed a distance from each other. As the electrodes burned down, the gap between them increased and the arc died out.

To maintain steady-state illumination, a special man had to bring together the ends of the electrodes by means of a screw mechanism. Such a gap-adjustment system performed its function, but was complicated. Then a clock-controlled device was created. It gradually

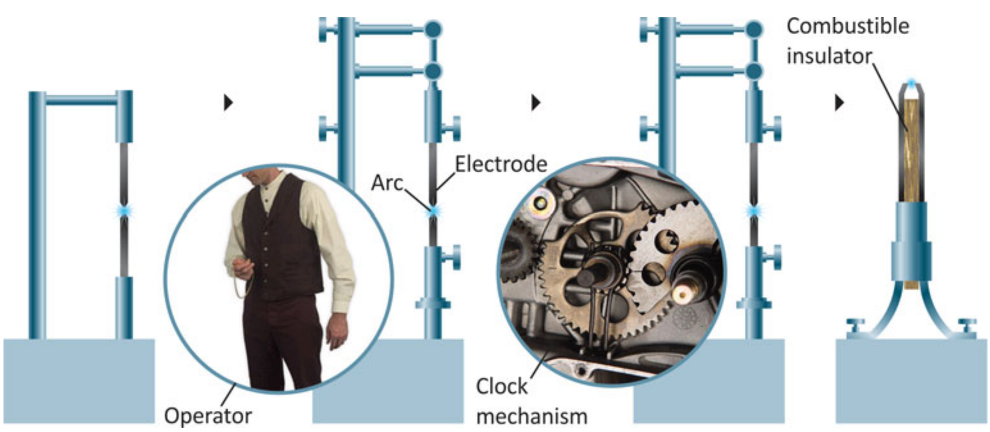


Fig. 3.4. Trimming the electrode gap adjustment system

moved electrodes toward each other. As a result, a man-operator was removed from the adjustment system.

The adjustment system was completely trimmed by the Russian physicist Jablochkoff. He removed the adjustment system, i.e. replaced it with an «ideal system». Yablochkoff mounted electrodes not opposite each other but in parallel and separated them with a layer of combustible insulation material. As the arc burned, the insulation material also burned down, so the gap between the electrodes was always the same.

3.1.3. Expanding — Trimming a System

This pattern is the result of sequential performance of the two actions described in the previous patterns (see sections 3.1 and 3.2): at the initial stage, new components are introduced into a system and the system is expanded; at the final stage, unnecessary components are removed from the system. The «Expanding — Trimming» pattern is often found when system analysis covers a considerable timespan and describes the entire life cycle of a technical system starting from its appearance till the moment it is replaced with a new technical system performing a similar function.

Any technical system first appears in its simplest form — as some *functional nucleus* only including components necessary for the performance of the main function of the system. At this evolution stage, the system has a low degree of ideality because of its low operability. After the minimum functioning of the simplest technical system has been ensured, the system starts to expand. New components are introduced which can perform additional functions facilitating a high-quality performance of its main function. Accordingly, the system's cost starts growing because additional components have to be introduced into the system. This part of the «Expanding — Trimming» pattern is similar to the «Mono-Bi-Poly» pattern, the only difference being that not only objects similar to those already present in the system are introduced, but also any other necessary objects — carriers of new functions.

System expansion prevails until the function performance quality satisfies the user. The system becomes maximally expanded, complete and has the maximum number of components. Designers try to remove unnecessary components from the system and pass their functions to the components remaining in the system. This part of the «Expansion — Trimming» pattern is completely analogous to the «Trimming» pattern.

The «Expanding — Trimming» pattern starts from a single object or system and can include the following steps (Fig. 3.5).

Expanding:

- *formation of the functional nucleus of a system and providing the system's minimum operability*
- *introducing a single additional object into a system,*

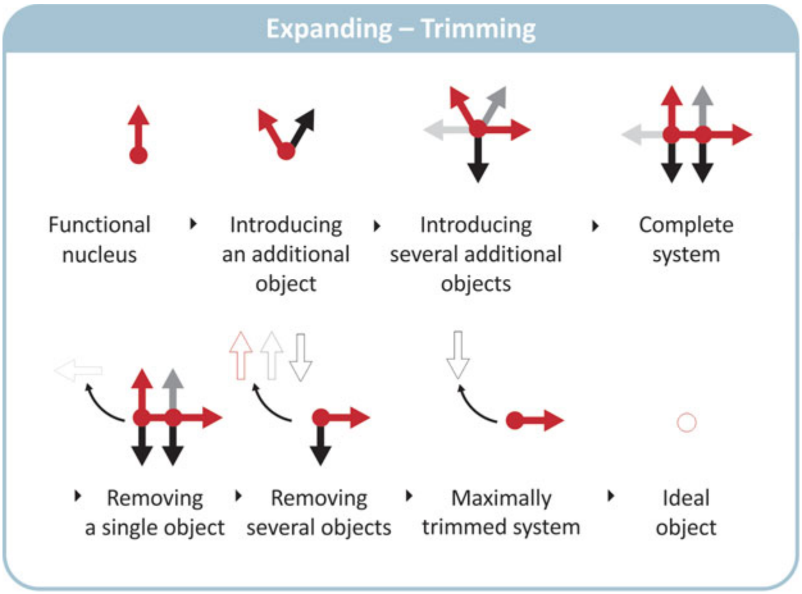


Fig. 3.5. Steps of the pattern «Expanding — Trimming a System»

- *introducing several additional objects into a system,*
- *forming a complete system and providing its adequate operability.*

Trimming:

- *removing a single object from a system,*
- *removing several objects from a system,*
- *transition to a maximally trimmed system,*
- *transition to an ideal object.*

Example of «Expanding-trimming» pattern: car body

The first car was an ordinary cart equipped with an engine and a control system. It was the functional nucleus of the technical system «Automobile» and provided its minimum operability.

The functional nucleus of a car may be illustrated by a sporting cart (Fig. 3.6). The design of this quick tiny car only comprises the most necessary units and components made at the up-to-date level.

First cars actually had no cabin, its function was performed by separate components of the car: a frame, seats surrounded by a low side. The design of first cars was expanded by introducing new components — carriers of additional functions. There appeared headlights so that the car could move at night. Then a roof came into existence as well as a windshield and a passenger cabin so passengers became protected against rain. By the 1930's, the car design had acquired practically all components required for satisfactory performance of its main useful function.

Fig. 3.6. The car body evolution

Since then the designers' efforts have been largely aimed at the trimming of the car design components (certainly, the process of adding new functions did not stop but this trend ceased to be prevailing). First, the frame, roof and cabin were combined to form a single unit — a carrier body. The functions of the removed components were passed to that body. Then headlights and signaling lights also moved into the body, which allowed their own housings and mountings to be removed. The trimming process is still going on, for example, the windshield wiper acquired the function of a radio aerial, etc.

As to an ideal car, this notion makes sense first of all relating to the useful function a functioning system which includes a car performs at a given moment. If one only needs to go somewhere to communicate a message, it is an ordinary telephone that may be considered an «ideal car».

3.1.4. Segmenting objects and substances

This pattern describes system versions obtained by segmenting system components, which results in the acquisition of absolutely new properties by these components.

Segmentation, alongside with the introducing of new components into a system, serves as the main source of resources for subsequent dynamization of this system. Segmentation is the system evolution trend which consists in sequential dividing of one-piece, monolithic objects into parts. Such segmentation can be endless until the object turns into vacuum and further into an «ideal system». As a matter of fact, the «Segmentation» pattern illustrates the system transition from the macro- to the microlevel.

At the object level, system's components are divided into parts, then into smaller parts down to the powder and fine dust level. Segmentation transfers to the molecular level, the object may be liquid or gas or combinations thereof. Then the level of atom's parts (ions and component particles) follows. At the microlevel, objects completely disappear, only interac-

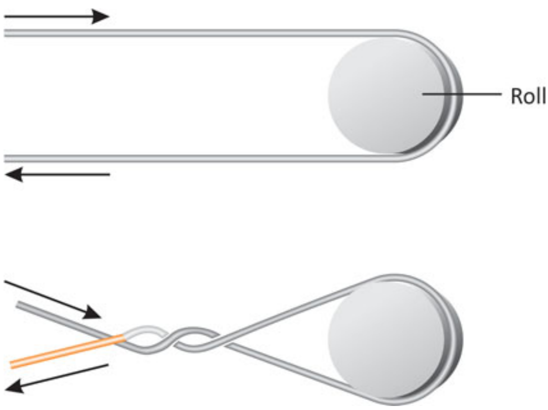


Fig. 3.7. Wire self-cleaning in a coiling machine

tions, fields remain. This pattern ends with a transition to vacuum and complete disappearance of objects from the system under investigation.

This can be illustrated by a sequence of modifications of a tool for cleaning different surfaces. A *one-piece metal scraper* will be the starting point of this transformation sequence. Then the scraper is divided into parts thus turning into a *brush*, next it is «crushed» down to powder and replaced with a *sandblasting machine*. According to the «Segmentation» pattern, a sand blast in its turn is replaced with a water jet supplied under high pressure. Our «scraper» becomes increasingly small turning into a plasma jet, then a laser beam or *magnetic field* may come. The final step may be use of an «*ideal scraper*», when parts are cleaned by rubbing on each other. Such a method is used in removing scale from wire while reeling it up. On one of the tensioning idler pulleys, wire is twisted several times (Fig. 3.7). As the wire passes through the pulley, its coils rub on one another so that the surface is cleaned with minimum energy consumption and without any additional fixtures [3.2].

It should be noted that a new cleaning method does not deny the use of the previous one, it just offers additional opportunities for increasing the surface cleaning quality. The choice of a suitable working method or a combination of methods from the above-described group depends on the kind of surface to be cleaned and the kind of dirt to be removed.

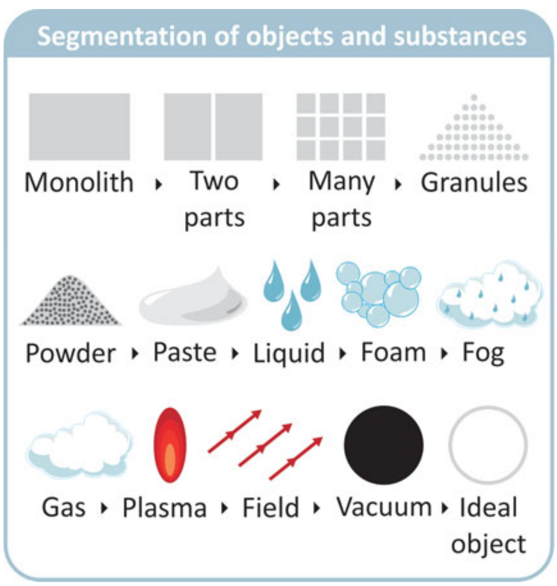


Fig. 3.8. Steps of the pattern «Segmentation of objects and substances»

The pattern «Segmentation of objects and substances» starts with a single one-piece object and may include the following steps (Fig. 3.8):

Segmenting an object into parts:

- *transition to an object segmented into two parts,*
- *transition to an object segmented into many parts,*
- *transition to granules,*
- *transition to powder.*

Segmenting an object down to the molecular level:

- *transition to a paste-like substance,*
- *transition to liquid,*
- *transition to foam,*
- *transition to fog,*
- *transition to gas.*

Segmenting an object down to the level of atom's parts:

- *transition to plasma.*

Segmenting an object to the field level:

- *transition to field interaction.*

Segmenting an object to vacuum:

- *transition to vacuum.*

Using an ideal system.

At each segmentation level of this pattern, a large number of characteristic states of substance or field may be identified. For example, at the molecular level, the above-presented transformation versions may be supplemented with gel, vapor, aerosol, suspensions in gas medium, etc. In addition, different types of transformed substance may be mixed resulting in a huge number of combinations, for example, pastes having different ratios of molecular and object phases.

As an example we present an aircraft propulsion unit

The engine of a flying aircraft activates a propulsion unit — a propeller or a reaction jet which propels the aircraft, that is, the propulsion unit is what exactly sets an aircraft in motion either by thrusting off the surrounding medium or by reaction forces (Fig. 3.9).

A single-blade propeller may be considered as the initial version of an aircraft unit of propulsion. This is how Leonardo da Vinci imagined the propeller of his helicopter (Fig. 3.10). Today, a single-blade propeller looks rather exotic and is only used for light-duty target airplanes and model airplanes. The next modification is a widely-spread two-bladed propeller. Then follows a multi-blade and, finally, a double-row propeller where the front and the rear row blades rotate in opposite directions. The following regularity has been revealed concerning equal thrust propulsion units: the more blades a propeller has, the smaller is its diameter and the size of a blade itself.

Next comes a jet-prop engine which is a mixed version: the propulsion unit is represented here by jointly operating propeller blades and a reaction gas jet. A complete transition to the molecular level has been realized in a jet engine which sets an aircraft in motion by means of a gas jet.

As aircraft evolve, the propulsion unit becomes more and more integrated with the engine itself. Further segmentation brings the system to the level of atom parts which serve as a working agent for ion and plasma engine — propulsion units. The propulsion unit based

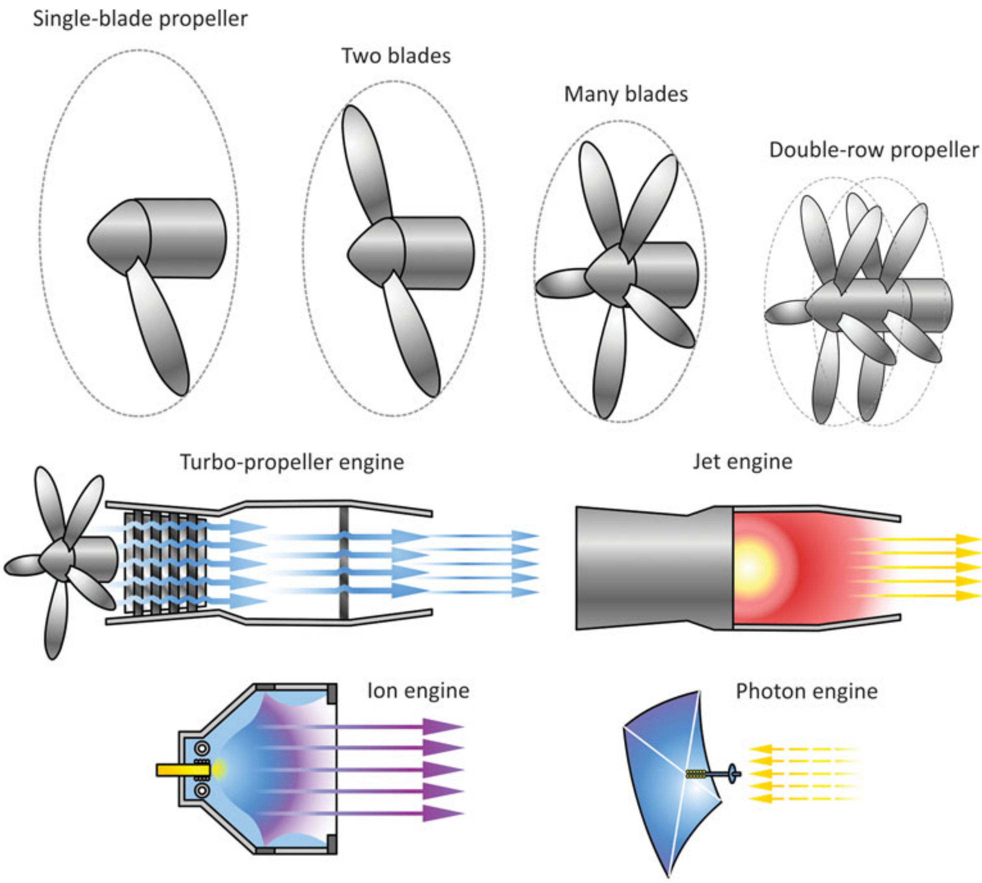


Fig. 3.9. Evolution of the aircraft unit of propulsion

on the field principle of operation may be illustrated by the use of a laser beam for aircraft acceleration from the Earth or solar wind in space. The hypothetic photon engine may also be assigned to this level.

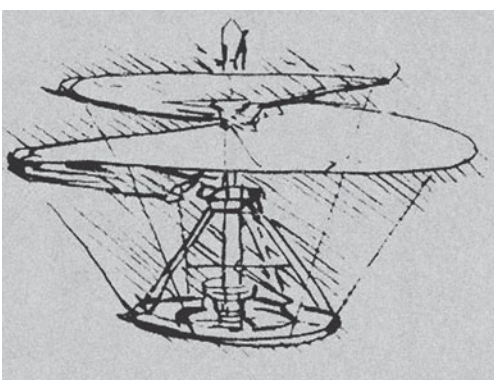


Fig. 3.10. A single-blade propeller proposed by Leonardo da Vinci

The «propulsive unit» of a glider may be considered as the ideal version because it does not exist from the traditional point of view while its function is performed by incoming air interacting with the wing [3.3].

3.1.5. Evolution of surface properties

This pattern presents component transformation versions obtained by changing the shape of its surface micro-relief.

Objects are usually in surface contact with each other. Therefore, the surface of each system's component is an important resource for system perfection. Changing the surface micro-relief and properties allows controlling the friction, cohesive force and adhesion between objects as well as other effects objects exercise on one another.

It often happens that the evolution of some object starts from a modification having a smooth surface. This means that the micro-relief of the object surface will start to become more complicated. This idea may be illustrated by a car or bicycle tire. The first pneumatic tires invented by Dunlop were simply hose pieces glued together to form a ring. As long as they were used for bicycles, they provided a sufficient grip on the road. With a car, however, the smooth surface was unsuitable. As a result, the tire was provided with a casing having *protrusions* on its surface. Subsequent evolution of the wheel tire tread consisted in complicating its micro-relief.

The tread of a modern tire has a *complicated raised pattern* — a set of longitudinal and transverse grooves. Wide longitudinal grooves prevent skid on a wet road while transverse grooves and small rifles impart good controllability to the wheel and provide effective braking. An interesting method of providing good grip on the road is used in race cars. Their wheels have no tread at all but are coated with a special sticky rubber. A wheel grips on the road due to adhesive power.

In the «Evolution of Surface Properties» pattern, one can trace formation of protrusions and hollows followed by a reduction in their size and complication of their shape. At the end of this pattern, a transition to the microlevel is observed, that is, to a surface having special properties provided by adding fields and forces. The trend illustrated by this pattern is aimed at better coordination of the interaction between the system's components.

The «Evolution of Object Surface Properties» pattern starts with the object modification having a smooth surface and may include the following steps (Fig. 3.11):

- *formation of protrusions and hollows,*
- *formation of a finely profiled surface,*
- *use of a surface having special properties.*

The practical use of this pattern offers an ample scope for specifying each of its modifications. A large number of different types of protrusions and hollows can be provided on the object surface: longitudinal, transverse, in the form of pits, etc.

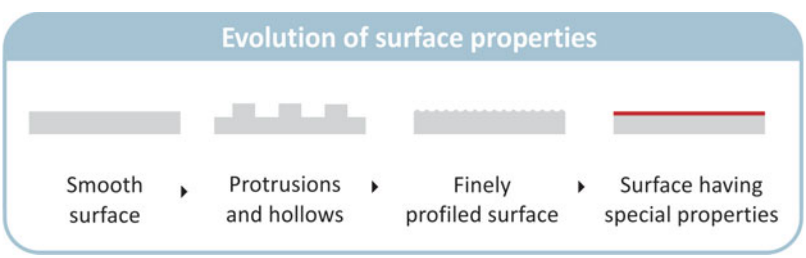


Fig. 3.11. The steps of the «Evolution of surface properties» pattern

A surface having special properties may be obtained in different ways. One of them is using different fields, described in the mini-algorithm MATChEM (see below, section 2.6.2), and combinations thereof. Such material properties may be used as elasticity, control of adhesion to other surfaces, different light-reflecting properties, or other coatings having special properties.

Example — steering wheel (Fig. 3.12)

Initially, steering wheels of first automobiles had a smooth surface. A smooth steering wheel was not easy to hold, therefore, the growing engine power was accompanied by the appearance of protrusions on the steering wheel surface. The steering wheel became easier to hold but there occurred another problem — comparatively high protrusions did not allow the driver's hands to slide freely on the steering wheel during fast maneuvering.

Fig. 3.12. Steering wheel

To remove this disadvantage, the steering wheel surface began to be made rough and flexible. Now the driver can reliably hold the steering wheel without clasping too strongly its elastic rough surface. At the same time, the driver can unclasp his hands a little to allow the steering wheel without large protrusions to slide obediently in his hands.

The surface of a modern steering wheel often has special properties, for example, it may be provided with a device for heating in winter or cooling in hot weather.

3.1.6. Internal structure evolution

This pattern presents component transformation versions obtained by changing of its internal structure. The inside of any component of a system is an important resource for improving the system. First of all, the system's compactness may be increased by placing some components in the cavities of other components as is the case with the Russian nested doll («Matryoshka»). New properties may also be obtained by transforming the inner space, making it in the form of porous-capillary structures, introducing fields and forces.

The «Internal structure evolution» pattern has much in common with the «Introduction of voidness» pattern considered in section 2.1. It shows the sequence of modifications obtained through the formation of internal cavities, separation of these cavities into parts followed by reduction in size. Further a transition to mini-cavities, for example, pores and capillaries may be observed, which is followed by a transition to the microlevel — space having special properties obtained by addition of fields and forces.

This may be illustrated by an ordinary fiber used in fabric manufacture. An important problem is increasing the heat-insulating properties of fabric for making clothes. A fabric made of solid fibers of dense material will be durable but not warm enough. Loose fibers will make the fabric warmer but not durable. How can we obtain a durable fabric that would be light and warm at the same time? It turned out that each hair of wool is actually a miniature tube. This is what allows a bear not to freeze even in severe frost. Clothes made of *hollow fiber* fabric also turned out very warm, but the fibers used to crumple at crease lines and lose their internal space.

To remove that disadvantage, a porous structure was introduced into the fiber preventing it from crumpling. The warmest will be clothes made of fibers having a porous internal structure capable of conducting *electric current,* with the surrounding tube made of insulating material. Now connect to a battery to make the fibers release heat.

The pattern starts from the object modification having a solid internal structure and may include the following steps (Fig. 3.13):

- *introduction of voidness,*
- *formation of several volumes,*
- *segmentation of space into multiple volumes,*
- *introduction of fields and forces.*

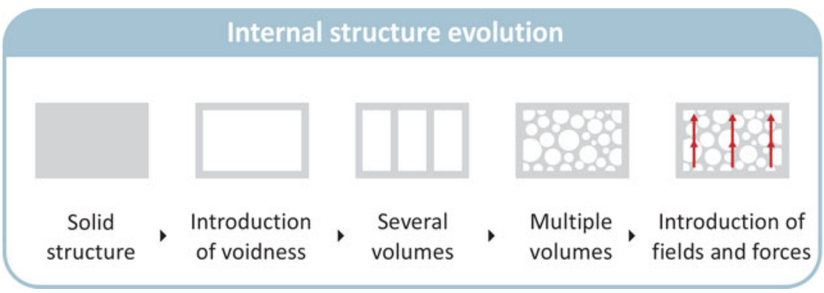


Fig. 3.13. Steps of the «Internal structure evolution» pattern

This pattern may also be essentially supplemented by specifying the transformation versions it contains. Introduction of various typical structures is possible, for example a string, a column-type structure, a structure made of separate components as well as various porous and capillary structures.

In addition, different fields and forces as well as substances — carriers of these fields and forces may be introduced into the object's internal structure. Combined structures can also be formed within an object. Such combinations often form rather complicated systems.

This may be illustrated by a heat pipe that includes internal void, a capillary-porous structure, liquid and vapor.

Example — car bumper (Fig. 3.14)

Bumpers of first cars were rigid, because they were formed of a solid thick metal strip. The next step was a bumper having a closed or unclosed cavity inside. Such a bumper was lighter in weight and provided better collision energy absorption because it could deform on impact.

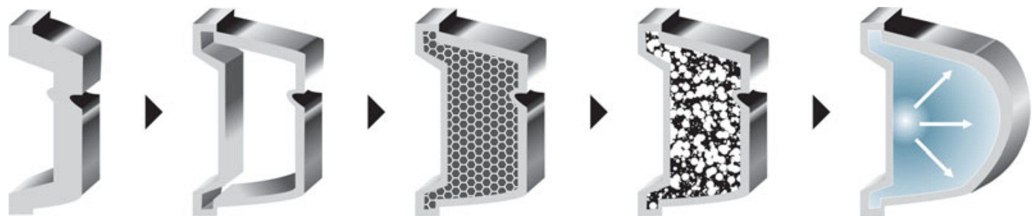


Fig. 3.14. Evolution of the internal space of a car bumper

To increase safety of traffic, the bumper walls began to be made thinner and the inside began to be filled with honeycomb filler. Such a bumper could collapse on impact, thereby absorbing part of the energy. A modern bumper is a 3D plastic shell the interior space of which is filled with a shock-absorbing porous material. The shell strength is designed to yield on a comparatively weak impact, thereby cushioning it; on a strong impact, the shell fails, actively absorbing the impact energy.

Further evolution of the bumper passes through the stage of imparting active properties to the bumper's interior space. This may be illustrated by an active bumper operating like an airbag. A system of transducers senses the car speed and distance to an obstacle. If the speed is high and the distance is small, an on-board computer gives a command to activate protection. The flexible shell of such a bumper is instantaneously filled with compressed air and absorbs the impact energy [3.4].

3.1.7. Geometrical evolution

This pattern is a sequence of transformations of the object, resulting from changes in its geometry.

Complication of components' shape is typical for the evolution of any system. One of the technology evolution trends is making a system compact and to place more parts within a limited space. An important resource is the geometrical shape of technical objects.

For example, a horn is a several meter long bell-mouthed pipe. At the same time, it is fairly compact because it is not straight but has a form of a complicated coil (Fig. 3.15).

Fig. 3.15. Compact horn

Another example is a car. First cars were angular and had a simple shape. But air resistance increased with the growth of speed and the car bodies began to be made streamlined. That required transitioning from flat panels to more complicated shapes as well as reduction of the car size, which was hampered by two circumstances: firstly, technological capabilities which restricted designers' imagination; secondly, the necessity to place an engine, transmission, passenger cabin and other necessary components in the car body.

To meet all the requirements, the shape of the external and internal surfaces of the car body and other car parts had to be made more and more complicated. For example, the fuel tank of a modern car that needs to be inserted into the free internal cavity of the car body has a very complicated shape; the shape of the body itself is also very complicated. It is the result of permanently compromising between stream-lining and providing simple access to the most important parts of the car, passenger comfort, entry/exit convenience.

Complicating the shape of geometrical structures is possible for linear structures, surfaces and 3D Figures. A linear structure becoming increasingly curved & complicated. Surface transformation proceeds from flat to cylindrical surfaces that can be unrolled to form a flat surface. Then transition to spheroidal surfaces is observed. Such surfaces cannot be made flat without deforming the material. Next comes a complex-shape surface composed of several different types of geometrical surfaces.

The shape of 3D structures also becomes more complicated during evolution: from structures formed by flat bounding surfaces to sets of complex-shape surfaces.

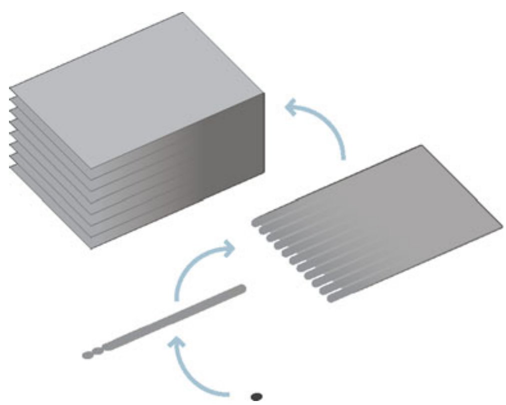


Fig. 3.16. Basic geometrical shapes as a set of points

In addition to complicating the shape of linear structures or surfaces, there occurs a transition from one type of geometrical objects to other types, for example, «point — line — surface — volume». This transition is actually an increase in the number of single geometrical components (Fig. 3.16). The point is a single geometrical object. A sequence of a great number of points forms a line. Surface is formed by a great number of lines and a 3D structure is a set of surfaces arranged in layers. Each transition of this kind provides new resources for perfecting the system operation.

The «*Geometrical Evolution*» pattern starts from a *single geometrical component (point)* and may include the following steps (Fig. 3.17):

Transition from point to line:

- *straight line,*
- *line curved in one direction,*
- *line curved in two directions,*
- *compound line.*

Transition from line to surface:

- *flat surface,*
- *cylindrical surface,*
- *spheroidal surface,*
  combined surface.

Transition from surface to volume:

- *prism,*
- *cylinder,*
- *sphere,*
- *complex 3D structure.*

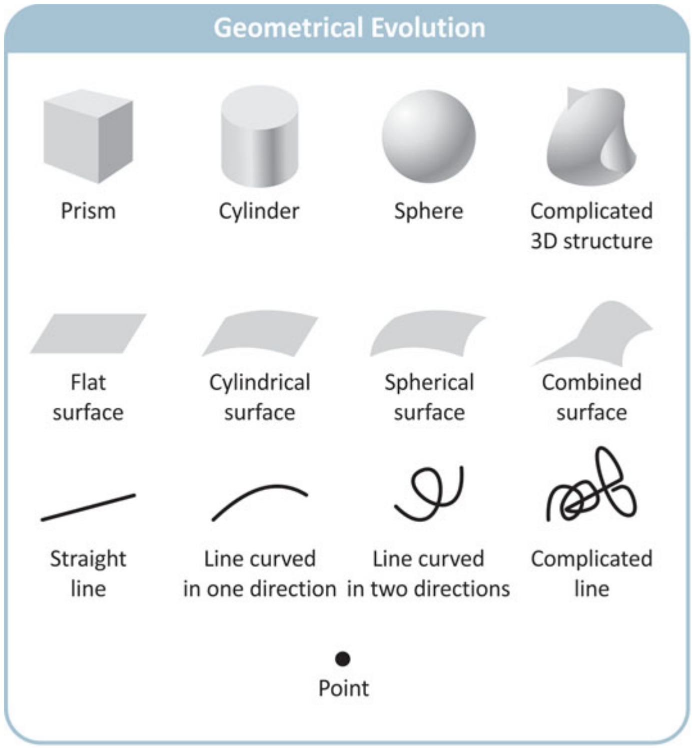


Fig. 3.17. Steps of the «Geometrical Evolution» pattern

For practical application of this pattern, it is expedient to collect a database of various types of linear, surface and 3D constructions. For linear structures, different types of units widely used in nautics and mountain climbing can prove useful. The example of an interesting surface is the Moebius band and that of a 3D structure is the Klein bottle. Each of these forms paradoxically has only one surface.

Example of «Point — line — surface — volume» transition — bearing contact spot (Fig. 3.18)

The «point — line — surface — volume» pattern may be illustrated by the shape evolution of the contact point between friction-reducing components and a race in different types of bearings. For example, in a ball-bearing, a point contact is provided between the ball and the race surface. In a roller bearing, a cylindrical roller and a race have a line contact. This ensures a higher loading capacity of the bearing. In a slider bearing, a movable shaft and an immovable bearing contact through the finest film of oil so this case may be regarded as surface contact.

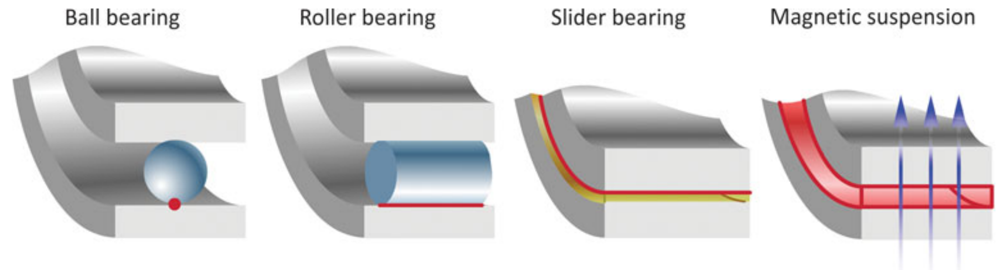


Fig. 3.18. Contact spot of different types of bearings

Volumetric interaction takes place in those types of bearings where the shaft surface and the seat work surface are separated in space and make contact, for example, through a field as in a magnetic bearing, or through an air jet as in an air bearing.

Example of complicating the geometrical shape of a pattern — roller bearing contact line (Fig. 3.19)

In different modifications of a roller bearing, one can trace complication of the geometrical shape of the bearing/race contact line.

In bearings having cylindrical and conical rollers, the contact between the roller and the race occurs along a straight line. Barrel-shaped roller bearings and the race have a curve contact. Such bearings withstand higher axial loads and, in addition, have the self-aligning feature. In rotary groups, where a seat needs to be elastic in the radial direction, bearings having flexible cylindrical rollers with a spirally formed surface are used. Here, interaction occurs along a compound curve.

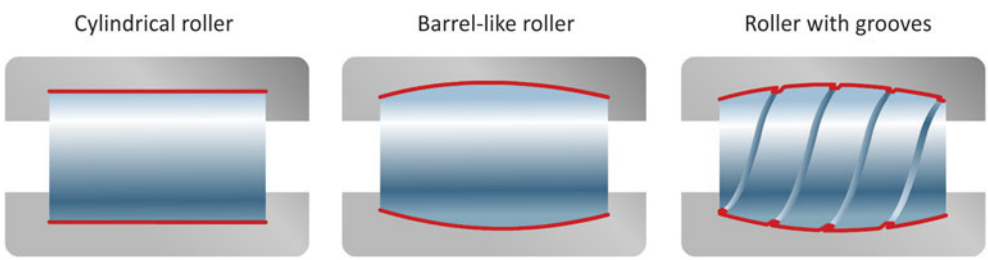


Fig. 3.19. Contact line of different roller bearings

Example of complicating the geometrical shape of surface — computer mouse (Fig. 3.20)

The first computer mouse invented in 1964 by Douglas Engelbart was shaped into a parallelepiped so all its surfaces where flat. As such the shape was far from the ergonomic ideal, its front surface began to be made as part of a cylinder. That shape was not perfect, either. As a result of evolution, the mouse body shape began to be formed of several surfaces of double curvature. This shape is well-adapted to the user’s hand.

Fig. 3.20. Evolution of the computer mouse shape

It was only the front surface of the mouse body contacting the user's hand that changed in the course of evolution. The bottom surface remained flat. This evolution trend is caused by the necessity of coordinating the top surface with the hand shape and the bottom surface with the table surface.

3.1.8. Dynamization

This pattern describes system transformations achieved by increasing the mobility, dynamics of its components through changing the character of links between these components and gaining an opportunity to change the parameters of the components themselves.

In the general case, dynamicity implies a change in some parameters of a system — temperature, pressure, speed, motion freedom, etc. In the simplest case, we may speak of dynamizing a system by increasing the mobility of its components. Dynamization makes a system controllable or adaptable to changes in operation conditions. It becomes possible to adjust system's components to the optimal operation mode, to match more accurately its parameters with the changing requirements imposed by the environment.

Just at this transformation stage, it is necessary to check whether the parameters of the system's components can be changed. The dynamicity degree of the parameters is selected depending on specific operation conditions of a component. In case of need, tight couplings are replaced with movable, flexible; fields are replaced with more dynamic ones. For example, a permanent magnetic field may be replaced with a variable field produced by an electromagnet.

To ensure the mobility of a system's parts, it is necessary to have resources, i.e., if a system consists of one component, this one can only be dynamized by changing some parameter characterizing the operation of the entire component.

Additional dynamization possibilities appear when a system has several objects and it is possible to provide their mobility relative to each other. Introduction of such resources into the system must be provided by performing actions illustrated by patterns 3.1.1–3.1.7. Thus, system dynamization, along with the provision of operational control, is the most important action in the hierarchy of transformations; it is directly responsible for the preparation of full coordination of all system's parts with each other and the system itself with the environment.

Let us consider an ordinary door to illustrate the hierarchy of mental actions aimed at system transformation. This is what we need to do to obtain a door: to separate a fragment, equal in size to the future door, from a wall («Segmentation»), to make it thinner and lighter («Coordination of shape, size and arrangement of components»). Rigidly attaching the door to the formed opening will not allow it to open. Hence we must perform *dynamization*, i.e. to provide a hinge mount of the door.

Then it is necessary to find a method for opening and closing it (*Controllability*) and to determine when the door should be open and when it should be closed (*Coordination of system's components operation*). The first step of the «Dynamization» pattern corresponds to the system's modification wherein the system's parts are tightly coupled with each other (Fig. 3.21). The pattern may comprise the following steps:

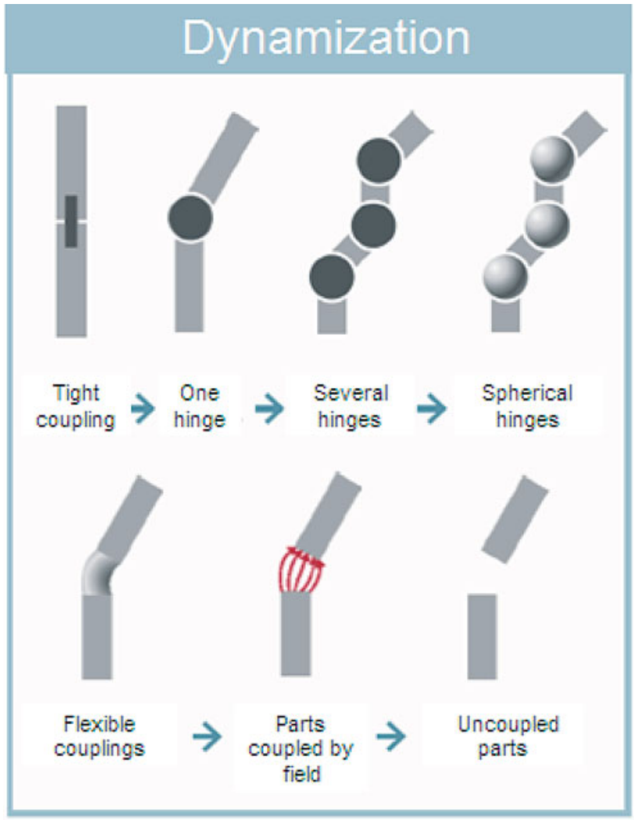


Fig. 3.21. Steps of the «Dynamization» evolution pattern

- *transition to a system that is movable in one direction,*
- *increasing the degrees of freedom of the system's components,*
- *transition to flexible couplings,*
- *transition to a system having field-coupled parts,*
- *transition to a system having separated parts.*

To specify the performance of each transition, a designer must constantly collect information about different types of couplings. For example, for the «flexible coupling» transformation type, they may be couplings having both different degrees of flexibility and different degrees of freedom.

Example of dynamization — toothbrush mount (Fig. 3.22)

A monolithic, rigid toothbrush head/handle mount is replaced with a hinge mount which allows bending within a small range.

The next variant is using two hinges, which considerably increases mobility. Then follows an «accordion» — a corrugated portion of plastic that increases the coupling flexibility.

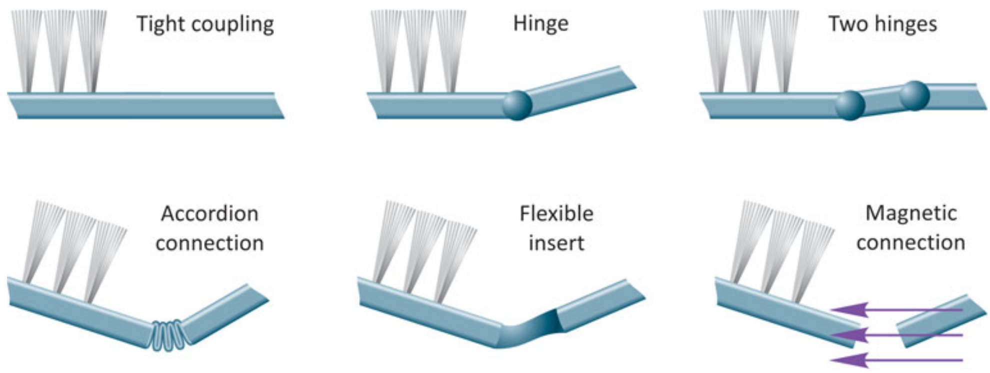


Fig. 3.22. Dynamization of the toothbrush head/handle mount

To prevent dirt accumulation between the accordion pleats or in hinge cavities, the coupling portion is made monolithic, smooth, but the material used is flexible and elastic.

The final step of this pattern may be a toothbrush head physically separated from its handle but held and controlled by a magnetic field. Such a head will be the most adaptable and movable version. At first sight, use of magnetic field for mounting a toothbrush head looks unrealistic. However, such brushes are known [2.31], they provide a soft and gentle effect on teeth. A similar magnetic coupling of a cleaning component and a control component is used in a device for cleaning windows on top floors of buildings. A cleaner moves a magnetic handle on the internal side of the glass while a cleaning sponge which also has a magnet moves synchronously on the external surface of the glass.

3.1.9. Increasing controllability

This pattern presents system versions obtained by simplifying the interaction between an operator or control unit and the object part of the system.

Most parameters of a system have been matched during manufacture and do not require any change during operation. However, there are a number of parameters which need to be changed during operation. Doing this is the function of an operator or a control subsystem.

System controllability is a possibility of a simple and effective change of the system's parameters when an operator or a control device changes the system's functioning conditions. The essence of the «Increasing controllability» pattern consists in sequentially simplifying the process of the operator's or control device interaction with the object part of a system.

In a conventional vacuum cleaner, the presence of manual control is obvious, because a man holds and moves its nozzle on a surface. The new generation of vacuum cleaners represented by Electrolux «Trilobite» robotic vacuum cleaner does not require manual control. In principle, the operation of such a robot may be semi-automatic, when an operator sends commands by means of a radio transmitter.

The robotic vacuum cleaner can also operate in a fully *automatic mode*. It orientates itself in space with the help of ultrasound and easily detects where walls, chair legs, door apertures and steps are located. When cleaning a room, a kind of «ground map» is formed in the control unit, which optimizes the vacuum cleaner path. In addition, the robot may be connected to an electric power supply for recharging the batteries.

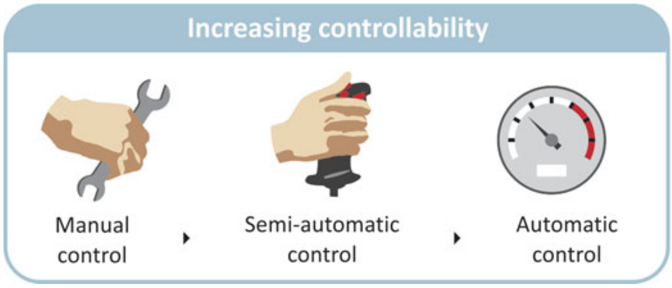


Fig. 3.23. Steps of the «Increasing controllability» evolution pattern

Transition from manual to semiautomatic control considerably simplifies the operator's actions: small physical efforts are needed to control heavy mechanisms. Transition to automatic control reduces the operator's function to programming the control unit and supervising its work. Simplification of control increases the number of objects having operatively changeable parameters, which creates conditions for a fuller coordination of a system.

The *«Increasing controllability»* pattern starts with the version where the system's components *only have preliminary coordination*, *but are not operatively controlled by an operator or control program.*

The pattern may include the following steps (Fig. 3.23):

- *manual control,*
- *transition to semi-automatic control,*
- *transition to automatic control.*

When specifying the modifications of this pattern, it is necessary to take into account the variety of types of control mechanisms. They may be mechanical, hydraulic, pneumatic, electric, etc. In the general case, any action may be used to control system's components. In addition, there is a large number of ways to provide semi-automatic control of systems, the information on which should be collected and used.

Example of Controlability pattern — trailer train wheels (Fig. 3.24)

A heavy-duty truck or trailer train carrying a heavy load has up to eight wheel axles. So many wheels are needed to provide a good support of a loaded vehicle. An unloaded truck

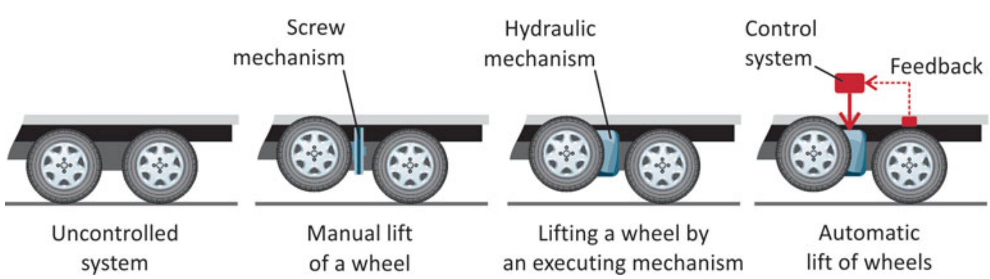


Fig. 3.24. Evolution of the truck train wheel-lifting mechanism

does not need that many wheels. Moreover, each running wheel experiences resistance, which eventually leads to increased fuel consumption.

To avoid this situation, some of the truck wheels began to be lifted over the road by means of a screw mechanism. That immediately reduced fuel consumption by an unloaded vehicle, but the wheel lifting and lowering operation itself took much time and needed much driver's effort.

The next step is using a hydraulic or pneumatic drive to lift and lower wheels. Now a driver only needs to press a button to make wheels lift over the road. The process has become much simpler.

At present, automatic control of the wheel lifting and lowering operation employs the feedback principle. A sensor determines the loading factor of a vehicle and gives a command to an executing mechanism. No participation of the driver is required.

An interesting system for regulating the wheel/road contact is traditionally used in Czech «Tatra» trucks (Fig. 3.25). The rear dual wheels of a truck are installed at an angle. As a result, only the outside wheel runs on the road when the truck is empty, thereby providing minimum rolling resistance. When the truck is fully loaded, shock absorbers are compressed, the wheels assume a perpendicular position relative to the road and the load is distributed between all the wheels. Such a system has neither sensors, nor drive. In addition, the size of the wheel/road contact spot changes not in steps but smoothly, depending on the truck load.

Fig. 3.25. Tatra truck: wheels are installed at an angle

3.1.10. Increasing of action coordination

The pattern shows system versions obtained by increasing the coordination of all the parameters, characteristics and actions of the system components with the performance specifics of the main function and external conditions.

Coordinating all the parameters, characteristics and actions of system's components with the main function performance specifics — interaction with a work object — is what all the designer's work is aimed at. Coordination is checked at each transformation stage of a system. The system matching capabilities are best seen when the system is already designed and ready to use. In this situation, it is necessary to make sure that the system is well coordinated with its functioning conditions, to check how the system's components will interact with each other, to see whether further coordination of their action is possible.

- It would be expedient to check the system for matching according to different criteria [1.2, p. 62]:
- coordinating the *functions* performed by the system's components with its main useful function;
- coordinating the *system composition and structure* in order to leave only those system's components, which are really indispensable for performing the system's function, and organizing them into an optimal structure;
- coordinating various *parameters* of the system's parts: shape, size, surface condition and internal structure — both between themselves and with the environment;
- coordinating the *rhythm*, system parts' operation sequence, which allows tuning the system operation in the same rhythm, intensifies the system parts' action due to the operation in a resonant mode;
- coordinating the *materials* used in the manufacture of system parts with the complexity of these parts, which allows selecting the optimal manufacturing and functioning technology of the system.

Proper coordination increases the ideality of any system, but designers often ignore this possibility. For example, the main component of an automobile water cooling system — a heat sink — is blown with a special fan. Fans of the first automobiles were mounted directly on an engine shaft and worked *continuously* even though an engine only needs cooling when overheated. If the engine is cool the fan operation is senseless and even harmful as it causes rapid wear to the engine.

The next step was an *on/off fan* which only operated when an engine needed cooling. Even though the technological development existed from the days of the first automobiles, a system equipped with an on/off fan only appeared in the 1950s. Think about how much fuel was wasted during that time. Engineers could have made that improvement but didn't.

Then a way to *disconnect the heat sink itself* was added. This directed a cooling liquid flow through an annular pipe back into the engine. A special valve — thermostatic switch — only opens the door to the heat sink when the liquid is heated to a certain temperature, and the fan switches on only if necessary.

One of the effective coordination principles is *use of pauses during operation* for performing auxiliary operations. With regard to an engine, it may look as preheating when starting in frosty weather. A crankcase is equipped with an electric heater for heating oil in the crankcase.

The initial version of the «Increasing coordination» pattern corresponds to a system having *uncoordinated or poorly coordinated components*. The parameters of the system's components are coordinated mainly by preliminary or regular coordination, but final coordination is absent.

The pattern may include the following steps (Fig. 3.26):

- *transition to a step-wise change of parameters,*
- *transition to a gradual change of parameters,*
- *recuperation (repeated use) of energy.*

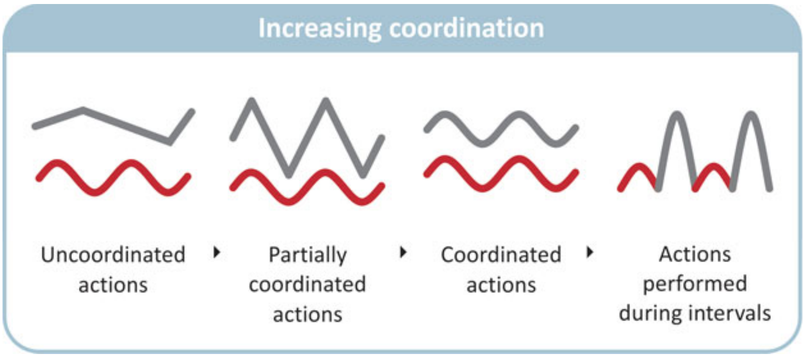


Fig. 3.26. Steps of the «Increasing coordination» evolution pattern

The trend underlying this evolution pattern predetermines a fuller matching of technical system parameters to its functioning conditions. All the efforts of designers, manufacturers and operators must be aimed at providing this trend.

Example of coordinating of parameters: a gear-box (Fig. 3.27).

The components of first automobiles were coordinated with each other only to such an extent that an automobile could move. They did not have gear-boxes and a driver could only change speed by increasing or decreasing fuel supply into the engine. The automobile speed changed within a small range, which was very inconvenient.

The next step was invention of a gear-box. It contained several gear wheels of different diameters. A driver could move the gears on the shafts and bring them into engagement changing the gearing ratio. That immediately increased coordination between the wheel rotation speed and the car motion conditions.

First gear-boxes had two or three speeds, and then their number grew to five. Special vehicles such as agricultural tractors have 10-17 and more speeds. Quite natural was the appearance of a continuously varying transmission widely used in modern automobiles. Such a gear-box changes the transmission ratio gradually, in accordance with the operating conditions.

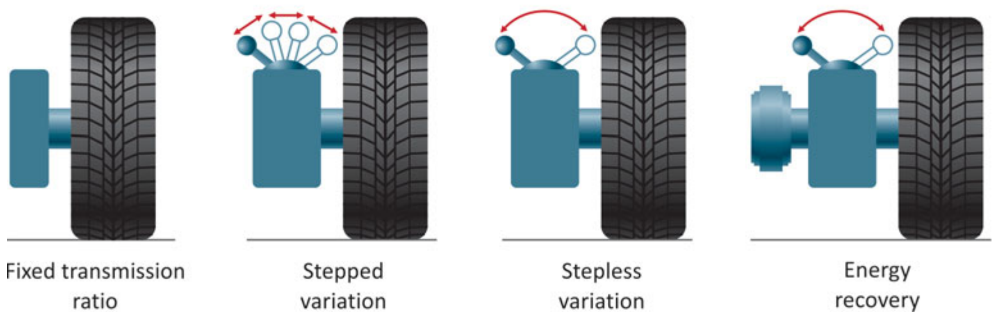


Fig. 3.27. Evolution of automobile transmission

Further increase of coordination may be illustrated by a hybrid drive. Such a drive consists of a traditional internal combustion engine and an electric generator which feeds electric motors built in the wheels. The motor rotation frequency may be varied for each wheel separately within a wide range, thereby providing good coordination with the motion conditions. In addition, when a motor vehicle brakes or moves downhill, the wheel motors start operating in a generator's mode. They produce electric energy and recharge batteries.

3.2. Specifics of constructing and using the evolution patterns

3.2.1. Evolution pattern construction recommendations

The ten basic technical system evolution patterns considered in section 3.1 do not at all preclude the possibility of independently constructing additional patterns. As was mentioned above, the ten described evolution patterns may be considered as basic ones, i.e. these patterns are enough for describing all changes occurring with the technical system's components as the system changes from one version to another. All other evolution patterns describing the system changes will be particular cases of these basic patterns.

In the practical use of the evolution patterns, it is undoubtedly very important to develop both new transformation versions within the patterns described in this chapter, and derivative patterns for various specific cases. In doing so, the following rules should be observed:

1. *unity of the object's property being transformed and the transformation type,*
2. *completeness of action and transformation hierarchy,*
3. *check for coordination,*
4. *optimal generalization of information.*

With respect to the *first rule*, it is necessary to understand that a correctly constructed pattern describes the results of consistent application of only one action, the action being aimed at a change of only one property of the object under transformation.

Concerning surface transformation, one pattern should describe a consistent change of one of the surface properties, for example, its shape. A second pattern may describe the evolution of the surface properties and yet another pattern may be dedicated to the dynamic properties of the surface, etc.

According to the *second rule*, all actions performed on the system's components during the system formation should be performed in a certain order. Each subsequent transforming action may only be performed after necessary conditions and resources have been provided by performing the previous action.

Hence the hierarchy of actions necessary for transforming a system and evolution patterns concerning these actions will be as follows:

- Introducing of new and segmenting of available objects, processes, fields and forces.
- Coordinating the shape, size, surface, internal structure properties of the system's components, process parameters, fields and forces.
- Providing dynamization of sets of objects, processes, fields and forces.
- Providing controllability of system's components.

The result will be coordination of the actions of the system's components.

The set of the presented actions must give a new functioning modification of the system. The results of these actions can be illustrated by the development of lines arranged in the same order. (Fig. 3.28).

The third rule results from the previous two.

After each transformation cycle, the system is checked for coordination. The sequence of checking is reverse to that of the sequence of actions performed on a system in accordance with rule two:

To *coordinate* the actions of the system's components, these components should be controllable. For the system's components to be controllable, it must be possible to change the parameters of these components during operation, i.e. the components should be *dynamized*. Dynamizing a rigid monolithic object or a process having invariable parameters is very difficult. Accordingly, it is necessary either to supplement the system with some objects (or processes) having different parameters or to segment the available objects (or processes) into parts and *coordinate the parameters of these parts* with each other.

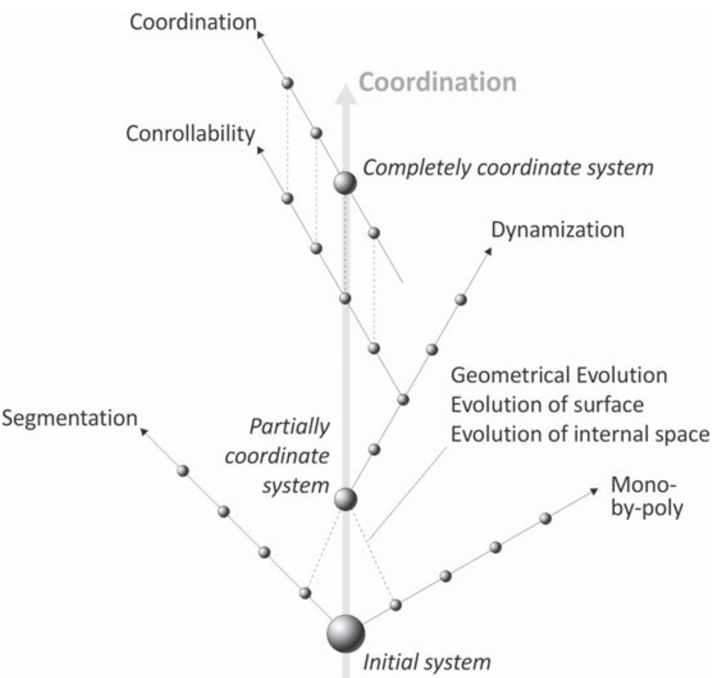


Fig. 3.28. Hierarchy of Evolution Patterns obtained as result the actions

For example, to provide controllable motion of an aircraft, it is necessary to change the wing shape. The wing shape must be controlled from the pilot's cabin. To provide controllability, it is necessary to make the wing dynamized, i.e. the entire wing or its parts should be movable. Dynamization resources can be obtained by separating the rear portion from the main part of the wing. It is also possible to add additional rudders to the rear edge of the wing, coordinate their shape and introduce links, for example, hinge mechanisms.

The *fourth rule* determines the *degree of generalization* of information. Depending on the analysis purpose in each specific case, it is necessary to find the optimum degree of difference of transformations, because excessive specification increases the number of insignificantly differing versions and hampers analysis while excessive generalization is not instrumental. It is essential that transformation differ in *quality*.

For example, for the «*Segmentation of objects and substances*» pattern such transformations may be (Fig. 3.29):

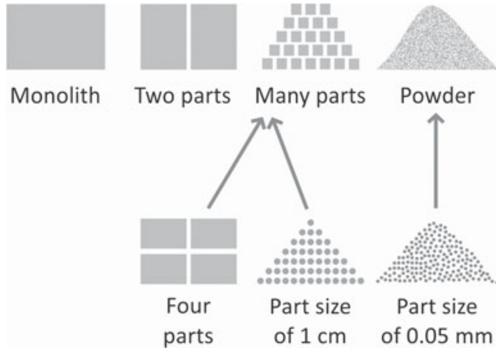


Fig. 3.29. Optimum degree of difference of transformations

- *monolithic object;*
- *object segmented into two parts;*
- *object segmented into many (conventionally countable) parts;*
- *objects segmented into a plurality of finest (conventionally uncountable) parts.*

Intermediate versions are: three parts, four parts, object fragment size of 1 cm or 0.05 mm — do not have essential value; they may be referred to one or another transformation.

3.2.2. Specifics of using some evolution pattern steps

A special case of transformations of system's components are so-called amplifiers, mini-algorithms, which define the specifics of using certain evolution patterns or their steps. Let us consider several examples.

Operator of introducing additional components:

This operator determines the rules of introducing objects, forces, fields, etc. when expanding a system. The rules are built by analogy with the «Introduction of voidness» evolution pattern (see section 2.1) and may look as follows (Fig. 3.30):

- *introduced object is adjacent to the main object,*
- *introduced object contacts the main object,*
- *introduced object is arranged about the main object,*
- *introduced object is partially arranged within the main object,*
- *introduced object is placed within the main object,*
- *object is introduced into the main object by parts.*

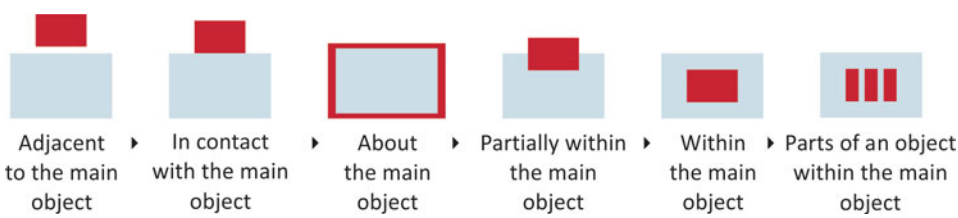


Fig. 3.30. The «Introduction of additional components into a system» operator

This algorithm of introducing additional objects allows an inventor to not miss possible arrangement versions of introduced objects while solving a problem. The set of rules may be expanded independently by supplementing it with special cases of locating an object being introduced that arises in the course of practical problem solving.

Field introduction operator:

The MATChEM operator shows the sequence of fields which may be used for system transformation [1.2]. The MATChEM abbreviation is made up of the initial letters of the names of the principal field types (Fig. 3.31):

- *Mechanical — friction, impact, pressure, vibration...*
- *Acoustic — acoustic vibrations, ultrasound, infrasound...*
- *Thermal — heating, cooling...*
- *Chemical — interaction on atomic and molecular level...*
- *Electric — static electricity, Ampere forces, Lorenz forces...*
- *Magnetic — magnetization, magnetic interaction...*

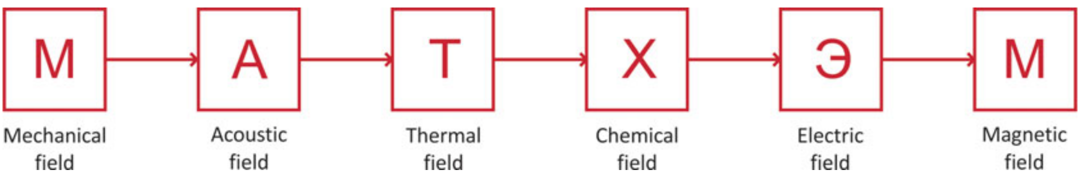


Fig. 3.31. The MATChEM operator

The list of fields may be widened by combining them, for example, by using thermo-mechanical, electrochemical or electromagnetic fields. In addition, the list may also be supplemented with other field actions, such as nuclear radiation or gravitational field [3.5]

The MATChEM operator is used in the following manner. First, a list of fields useful for solving a problem is made. Then one of these fields is selected and the model of a system being transformed is rebuilt either by fully changing its principle of operation or by introducing the selected field as an additional field. This operation is applied to all the fields on the list.

Selecting the improved system modification:

In work with the evolution patterns, the following question often arises: Which direction should we move to improve the system under analysis if it is located in the central portion of the pattern (Fig. 3.32)?

At first glance, it seems obvious that the versions lying farther from the prototype in the evolution pattern will be more perfect. Let us see whether it is really so.

The evolution pattern is a sequence of system modifications arranged in the ascending order of one of the parameters. Each evolution pattern exists at two levels: generalized (basic pattern, see section 2.1) and specific (evolution patterns of a specific system).

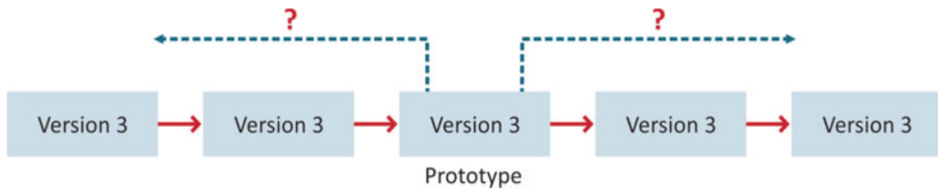


Fig. 3.32. Selecting the improved system modification

Generalizing many specific evolution patterns forms a basic pattern of system evolution. This trend is based on an objective law of technical evolution. Systems corresponding to the descriptions located closer to the end of the evolution pattern are as a rule more advanced and more perfect than those based on the principles described in the initial transformation versions.

A laser cutter is more perfect than a circular saw, a jet aircraft is more perfect than a propeller airplane, a hydraulic transmission has some advantages over a mechanical one whereas an electric drive is more dynamic and better controlled than a hydraulic one, etc.

A different situation arises when we analyze the evolution pattern of a specific system and must decide which of its modifications will be preferred in a given case. To determine the motion direction, it is necessary to take into account which parameter of the prototype system we want to improve.

If we want to improve productivity while working with the «Mono-bi-poly» pattern, it is necessary to move ahead, towards the increase in number of the system's components. As to the «Trimming» pattern, to improve the productivity parameter, we must pay attention to the modifications located before the prototype, because preserving and increasing the productivity requires that no functioning components be removed from the system.

If we want to make our prototype less expensive and simple, the situation reverses. The «Trimming» pattern shows us a direct way towards the system cost reduction; as for the «Mono-bi-poly» pattern, less expensive versions are located at its beginning.

The second important rule of choosing the motion direction within the evolution pattern is the necessity to analyze specific operation conditions of a system being perfected and resources which may be used for its improvement. *Airplane armament.*

At the beginning of the First World War, pilots of airplanes equipped their machines with different weapons. The most widely spread variant of a battleplane was a two-seater with a machine-gun mounted in the gunner's cabin on a swivel which allowed the gun to turn toward the enemy. However, the operational effectiveness of such arms was low, it was practically impossible to hit an enemy aircraft with a swivel gun.

The airplane only turned into a dangerous weapon after rigidly attaching a machine-gun to the fuselage. It became possible for a pilot to aim the airplane at an enemy, being sure that the gun had the same direction as the airplane nose. The French pilot Garrot who was the first to fire machine guns through the propeller blade (the blades were provided with a metal plate for bullets to rebound) amazed German pilots by his Fighting effectiveness.

At the same time, a transition from a hinge mount to a rigid joint is a step back from the «Dynamization» pattern standpoint.

Example: De-mating of spaceship stages

Designers of the first Soviet multistage space vehicle «Vostok» could not find any technical solution for the de-mating unit of the electric circuit between the spaceship and separating stages. The electromagnetic lock was unreliable and worked badly at low temperatures.

The solution that was finally found was an obvious step back in the «Segmentation» pattern — from field to a segmented object. The magnetic lock was replaced with a mechanical device — the simplest plug and socket joint. A separating stage took its pins out of sockets just like we do to take a plug out of a socket.

Example: Aircraft shape

Aircraft were becoming more and more streamlined with evolution.

This statement has already become commonplace but... let us look at one of the state-of-art aircraft — the American Fighter plane F-117 (Fig. 3.33). The angular, poorly streamlined construction does not correspond at all to the «Geometrical evolution» trend — the shape simplification is obvious.

The fact is that this Fighter plane was designed and built to be invisible, but making an airplane invisible is not an easy job, it requires use of numerous fundamentally new technical solutions: from special materials, mechanisms and devices to application tactics. One of such solutions is using an angular fuselage the surface of which reflects radar beams randomly. This hampers airplane detection.

It is worthy of note that any system improvement versions should be thoroughly calculated based on the ideality requirement. Each new system modification may only be considered improved when it effectively performs its function at lower expenses under specific conditions.

The next chapter deals with the organization of the above-described patterns into a higher rank structure — the Evolution Tree.

Fig. 3.33. The angular shape of a stealth aircraft F-177 dissipates radar beamsExample: Aircraft shape

Charter 4

Building Evolution Trees

Two main principles allow describing the performance of any function: morphological analysis and organization of information into a tree-like structure made up according to the rules from technical objects evolution patterns.

4.1. Elementary function: how to describe it?

The problem of optimizing the information about the arrangement of the surrounding world remains urgent for researchers. Systematizing the accumulated knowledge has always included creation of information structures, models. The fuller and more objective the model, the better is the understanding of the evolution specifics of a phenomenon of interest.

As for meeting the *completeness* requirement of a structure designed to organize technical and patent information (see the list of requirements in 1.3), it is necessary to determined technical systems' components to be described in the first place. The model of a system for performing a needed function looks as follows: there is a tool (a component that performs an action), a work object (a system's component exposed to the action) and interaction between these two components. This is the minimum set of acting components, but their number may be increased (Fig. 4.1).

Generally, any function performed by two interacting components (a tool and a work object), can be resolved into elementary *functions*. These functions are an action resulting in a single simple transformation of the object's parameters, which generally can't be further divided into qualitatively different actions.

When performing the system's main function, not only the tool and the work object work. Each part of the system performs its auxiliary function to help the tool produce the required product. If an auxiliary function needs to be considered instead of the main function, the model should be rearranged to focus on that function. Some examples of functions are: *«a cutter cuts an object», «an air stream cools an object»*, etc. Another example, a solder bit performs two actions during soldering: heats an object and applies a solder layer to its surface. To be analyzed, this action should be resolved into two elementary functions: «*a solder bit heats an object*» and «*a solder bit applies a solder layer to a surface*».

Fig. 4.1. Performing a function

Thus, we must describe an elementary system as some structure composed of two components: a working tool and a work object, which immediately suggests the use of morphological analysis [3.1].

First attempts to carry out morphological analysis were made in the Middle Ages.

In the 12th century, Raymundus Lullius, philosopher and mystic, theologist and missionary, used a special method and a device called «a machine of the true» for generating thinking aspects. The device was composed of several concentric circles with nine general notions such as «Sky», «God», «Man», «Virtue», «Truth» and the like written on each circle. How did that «machine» function? Randomly turning the circles relative to each other produced different combinations of these notions which were regarded as new thinking aspects. The article «The Brilliant Master Raymundus Lullius» [3.2] gives more information about the morphological approach.

Nowadays, the Swiss astronomer Fritz Zwicky has developed Lullius's idea and proposed a contemporary-styled morphological box [4.1]. That enabled information about complex material objects, including technical systems, to be analyzed. In the thirties, he applied the morphological approach to solving of astrophysical problems and predicted existence of neutron stars. During the Second World war, Americans invited Zwicky to participate in the development of rockets and he carried out the morphological analysis of a rocket engine [4.2].

His morphological box allowed obtaining a great number of modifications of a rocket engine, including the design principles of the secret German flying bomb V-1 and V-2 rocket. Much time has passed since then, but even today the developers of new types of rocket engines can find their prototypes in the half-century old morphological box!

The idea of the morphological box is as follows.

If we have some volume of information, to deal with, this may be done in the following manner. First, it is necessary to identify basic notions that are contained in the information under investigation and answer the following question: «*What exactly*?» Then we must analyze the available information, trying to answer the question: «*How is it organized?*» i.e. how different modifications of the basic, key objects are made, and determine their principal features. After that, it only remains to plot the selected basic notions on one axis and the specifics of their execution on the other axis. The obtained structure — the morphological box or, the morphological table — is simple and illustrative, which allows its qualitative analysis (table 4.1).

Table 4.1

	Object 1	Object 2	Object 3	...	Object M
Modification 1	Object 1/1	Object 2/1	Object 3/1	...	Object M/1
Modification 2	Object 1/2	Object 2/2	Object 3/2	...	Object M/2
Modification 3	Object 1/3	Object 2/3	Object 3/3	...	Object M/3
...	...	...	...	...	...
Modification N	Object 1/N	Object 2/N	Object 3/N	...	Object M/N

The maximally full enumeration of all possible, even theoretical, versions of the system's parts may result in that such a morphological table will very likely include the construction or process modification we are seeking.

Table 4.2

Body colour	Righting unit	Body material	Ink supply method
Red	Ball	Steell	Capillary
Blue	Porous stem	Plastic	By graviti flow
Black	Pen-point	Aluminium	By a pump
Yellow	Tubule	Wood	Manual
Silvery	Goose-quill	Rubber	Dry ink

For example, a very simplified morphological table for a pen may look as follows (Table 4.2).

Combining the four factors chosen as basic ones for a pen, we obtain hundreds of modifications, which will undoubtedly include new ones. There are too many combinations to carry out an in-depth analysis of each one. But at the analysis stage, our aim is obtaining new data for filling the information structure; therefore, the large number of versions is rather an advantage that helps satisfy the completeness requirement, than an obstacle.

This example illustrates one more essential advantage of the morphological approach which consists in that a system may be described by parts, changing at least one of which may lead to the appearance of a new system modification.

However, the same example reveals disadvantages: it is very difficult to find all basic versions of the technical system's main parts, because they are often selected subjectively and through intuition. In addition, the morphological box is often comprehensible only to those, who built it; therefore, it is not easy to use for teamwork. For example, it is not clear what the expression «ink supply method: manually» implies. One can only guess whether it means the use of a manual pump, dipping a pen into an inkpot or just pouring ink into a pen from some bottle. But the most important thing is that it is difficult to determine which of the obtained versions will be the best and which one the worst. This is due to the fact that the choice of the main features (with respect to a pen, they are color, writing unit, material, ink supply method) is often random and subjective, only the analysis purpose being more or less clear.

The morphological box is simple and visual, but its problems, just like those of all currently popular simple methods of inventive problem solving, are rooted in that simple systems are always more difficult to control than complex ones.

This is also the case with technical systems. Mowing a lawn seems to be a very simple job. All you need is taking a grass-mower, switching it on and wheeling it back and forth on the lawn, discharging the grass container when needed. The process becomes much more complicated as soon as you try to do the same job with a scythe which is much simpler in design. One can learn to use a grass-mower by reading a guide while mowing with a scythe demands good control of the body to provide the required motion of the tool. Comparing the process of drawing a circle with a pencil and doing the same with drawing-compasses or a computer will bring us to the same conclusion. Complex systems have more developed control organs which take over most control functions only leaving the simplest ones to a man.

The same is with the morphological box. The morphological box is extremely simple so, to successfully use it, an effective control system is needed. Such a control system is the algorithm for selecting features placed on its axes and their realization versions.

There exists a series of practical recommendations on making one's choice more objective. However, the subjectiveness degree in building the morphological box still remains high enough and the selection of versions is still incomplete. The problem of objective and comprehensive choice of possible versions of technical system's parts is very important and is permanently under study.

Further development of the morphological analysis method may be found in the works by Yu. M. Chyapyale. The combinatorics method proposed by the author suggests more precise rules of selecting the axes of the basic notions of the morphological box: a) the notion of the working tool; b) the notion of the working medium; c) taking account of the state of aggregation of an object (object's parts) and medium; d) the feature of the working tool's geometrical shape; e) the features of the working tool's structure (relative position and mobility of parts); f) the features of mutual coupling of the working tool's parts on the macro- and microlevels. Here a certain hierarchy of notions is already observed, a kind of subordination of axes, which orders the constructed morphological table [4.3].

G. S. Altshuller [2.1] showed that there exists a certain structure of evolution patterns, a kind of a general evolution scheme of a technical system. A similar evolution scheme, also made up of evolution patterns, but based on other principles, may be found in the works by Yu. P. Salamatov [2.4]. B. I. Goldovsky [4.4] proposes an analogous idea. He suggests that the evolution patterns be used for full description of technical system's transformation versions in the morphological box. Similar ideas were also expressed by other researchers, for example, A. I. Skuratovich and V. G. Lysenko. The idea looks sensible, but no specific realization of such an information structure has been observed so far.

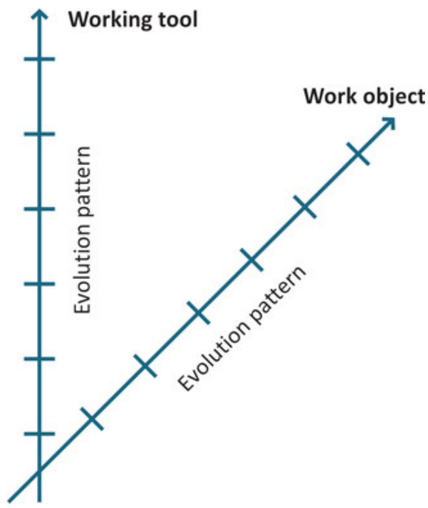


Fig. 4.2. A simple morphological box for a tool and work object

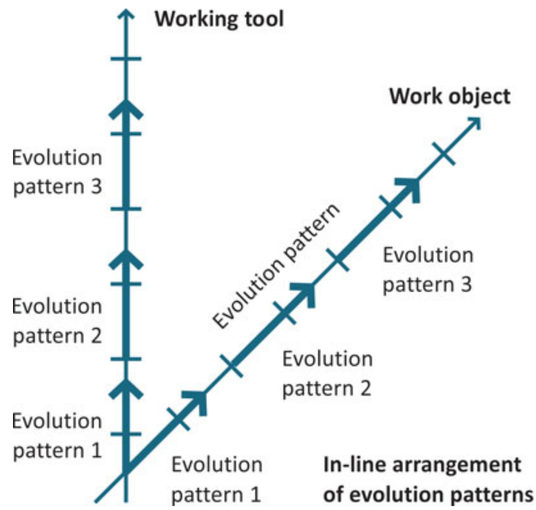


Fig. 4.3. In-line arrangement of evolution patterns

A. I. Polovinkin has proposed a solution for analyzing the structure of a technical system [4.5] — a so-called hierarchical tree («AND-OR-tree»), which represents all possible versions of each of the system's components. The hierarchical tree is actually a morphological box for the system's structure presented not in the form of a standard table, but as a tree-like structure.

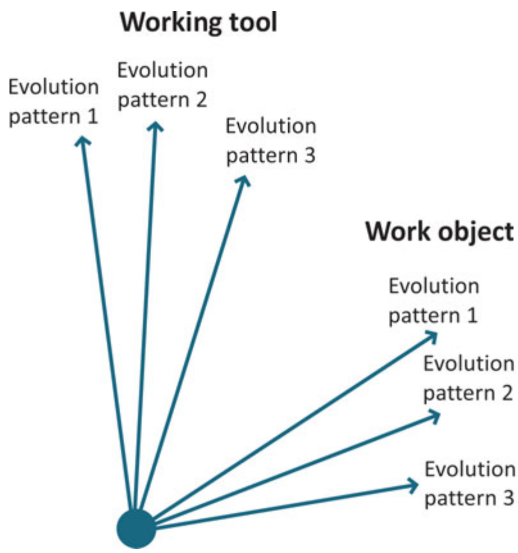


Fig. 4.4. Evolution patterns start from one point

Thus, the performance of an elementary function may be described by using the morphological box having two axes: one for tool modifications and the other one for the work object modifications (Fig. 4.2). To fill the morphological box cells with all possible versions of interacting components, the evolution patterns of these components may be used. There arises a question: what is the best way of arranging the tool evolution patterns on the axes of the morphological box?

The morphological box shown in Fig. 4.2 would satisfy the problem solution if all the modifications of components were only described by one evolution pattern. However, as shown in Chapter 2, this requires several evolution patterns; therefore it is necessary to find a rational way of their arrangement on the axes. The evolution patterns may be arranged in-line, one after another (Fig. 4.3) or may emerge from one point (Fig. 4.4). However, such a simple arrangement will also be effective enough.

As is shown in Chapter 3, the evolution patterns illustrate the results of system transformation actions performed in a certain sequence: each subsequent action is performed based on the results of the previous one. This causes a certain arrangement hierarchy of the patterns (see Fig. 3.28), according to which a new pattern may emerge from any point, any transformation version. Proceeding from this condition, it can be said that arranging the patterns in the form of a tree-like structure reflects most adequately the essence of actions aimed at the transformation of the system and its components.

We called the structure, where evolution of a system is studied through a change of some parameter, the *Tree of Technology Evolution.*

Of course, all technical systems, just like everything in the world, evolve in time. There exist a lot of models, describing the historical aspect of evolution of some machines. But this is only one type of model, the closest one to the real situation. Sometimes, it is more effective to «roughen» the situation and build some other, not a temporal model, and to see how the system or object changes within this virtual model. As distinct from temporal models describing the historical evolution aspect of some machine, we have obtained a virtual model according to which a system evolves toward the change of some feature. Properly speaking, the evolution tree is a set of system's alternatives ordered according to the technological hierarchy.

The structure of a real evolution tree is presented in Fig. 4.5.

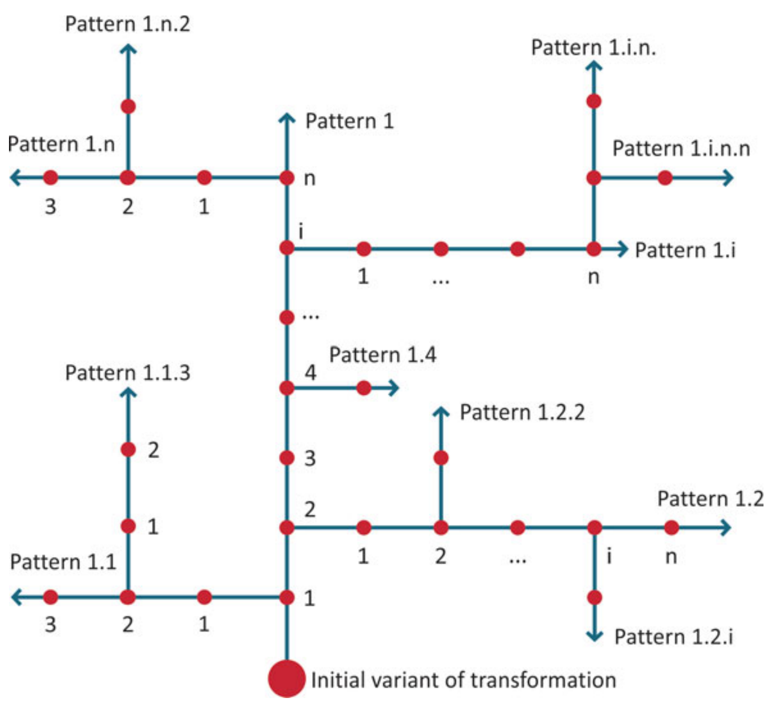


Fig. 4.5. The Evolution Tree structure

Each branch of the tree is an evolution pattern of one component of a system considered in the context of a certain objective evolution trend. The tree always contains the main — axial — pattern, which starts directly from the initial version of a technical object. The object modifications placed at each point of the main pattern may serve as a starting point for side patterns of the second order. Each subsequent level of the tree hierarchy represents patterns built on the previous level's patterns. A set of thus arranged patterns forms the simplest tree structure. When constructing real evolution trees, it is necessary to observe a certain action sequence and take into account some rules discussed below.

Using a tree-like structure makes possible transformation versions of system's components more illustrative and considerably simplifies navigation of this diversity of versions. In this case, each axis of our morphological box for a tool and a work object will be not a set of several, often casual transformation versions of components, but an evolution tree with all basic system modifications logically arranged on its branches (Fig. 4.6).

Considering the «tool and work object» pair by sequentially combining the modifications represented on the tool and object evolution tree you may find all basic methods of performing the function under research. This allows carrying out a qualitative analysis of information for receiving prognostic technical solutions.

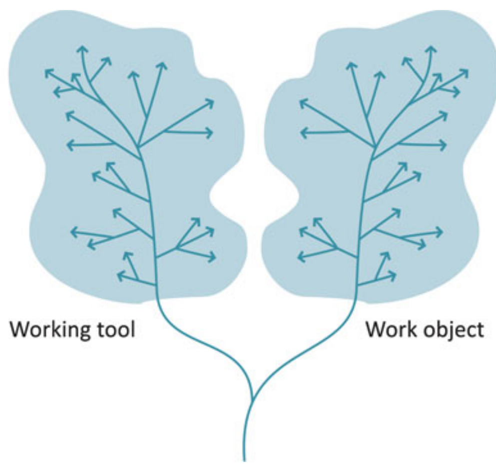


Fig. 4.6. A morphological box having tree-like axes

Using a tree-like structure makes possible transformation versions of system's components more illustrative and considerably simplifies navigation of this diversity of versions. In this case, each axis of our morphological box for a tool and a work object will be not a set of several, often casual transformation versions of components, but an evolution tree with all basic system modifications logically arranged on its branches.

Considering the «tool and work object» pair by sequentially combining the modifications represented on the tool and object evolution tree you may find all basic methods of performing the function under research. This allows carrying out a qualitative analysis of information for receiving prognostic technical solutions.

In principle, when constructing an evolution tree, there is no restrictions neither concerning the number of hierarchy levels, nor concerning the number of used patterns, i.e. the evolution tree of any technical object may be continued endlessly both in breadth, by the number of evolution patterns of one hierarchical level, and depth, by the number of these levels. It may be said that the evolution tree has the properties of infinite fractal structures (Fig. 4.7). Fractal is a self-similar set, i.e. a set having similar structure both on a large scale and on a small scale [4.7]. The theory of fractals offers objective prerequisites for mathematization of information processing by using evolution trees [4.6].

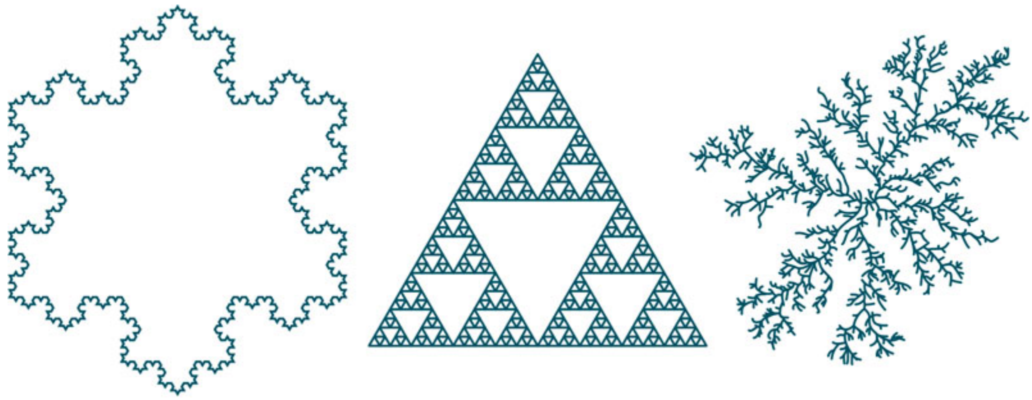


Fig. 4.7. Examples of fractal structures

4.2. Basic and Specific Evolution Trees

Looking at the descriptions of the evolution patterns mentioned in Chapter 3, one can easily notice that each of them looks as if it exists at two levels. One level is abstract and serves for describing sequences of generalized transitions, transformations of objects within a pattern. At the abstract level, these transformations are similar for a large number of evolving technical objects. The second level is specific; it shows technical system's modifications obtained by using generalized transformations. We can build an evolution tree on two levels. The tree having branches in the form of abstract descriptions of evolution patterns will be referred to as *basic* while the tree built for a real technical object will be called *specific*.

It should be noted that in building an evolution tree of an object it is necessary to abstract away both from the function for which it was initially *designed* and from the function for which it was *used* during operation. It is appropriate to remind that any function may be performed by using a large number of various technical objects. Similarly, any technical object may participate in the realization of many functions. This causes the notion of a «main function» to be senseless when building an evolution tree.

When collecting information needed for building the evolution tree of a technical object under analysis, it is essential to understand the role it can play in the realization of the elementary function being studied.

A Basic Evolution Tree is an organized set of evolution patterns of generalized, abstract features of technical objects (Fig. 4.8).

This is, in fact, a huge Evolution Tree which takes into account all possible modifications of a generalized technical object having features characteristic of all other technical objects. The object's features may be the presence or absence of additional components, their geometrical shape, micro-relief and surface state, internal structure, links between the compo-

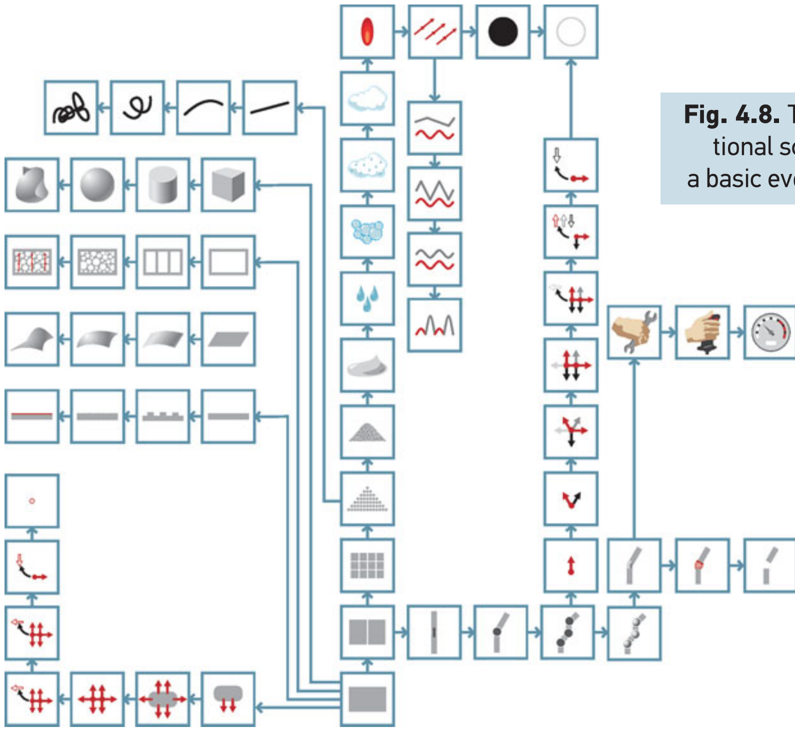


Fig. 4.8. The conventional scheme of a basic evolution tree

nents of complicated objects, parameters characterizing the object state as a whole and the state of its individual components. For technological operations, processes, a set of features may be absolutely different.

In the general case, building a basic Evolution Tree for a material object may be started from the simplest transformation version. It may be a monolithic, rigid object of a simple shape confined by flat surfaces and having no internal structure, a kind of building block or brick. Transforming such an object in accordance with definite evolution patterns will provide basically different modifications of this object.

Any evolution pattern may become the «trunk» — the main axis of the Evolution Tree, but it would be more convenient to use the patterns providing essential transformations of objects, such as «Segmentation of objects and substances» (see Fig. 4.9). According to this pattern (see p.3.4), our generalized object may be presented as segmented first into two then into several parts, and then down to powder. Further the object segmentation passes to the molecular level — to liquid and gas as well as to their combinations. This is followed by plasma, field and vacuum. The system's modifications according to this pattern have considerable qualitative differences between each other, which allows obtaining the most effective resources for transforming objects within other evolution patterns. The system's modifications within the «Segmentation» patterns are the starting points for the building of patterns.

The Tree's «branches» — evolution patterns emerging from the trunk — are selected based on the following considerations. Ideally, several patterns arranged as shown in Fig. 4.9 may emerge from each point of the Tree. To this end, each object modification located on

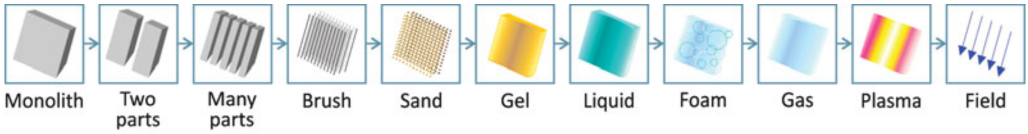


Fig. 4.9. The vertical (main) axis the Evolution Tree of a cleaner

the vertical «trunk» of the Tree is analyzed in the context of resources available for building these patterns. For example, an object segmented into two parts has the following resources: two parts of the object, their surface and interior space. According to the hierarchy of actions performed on a system, the following patterns may be built: «Geometrical Evolution», «Internal Structure Evolution» and «Evolution of Surface Properties».

Then we may check whether it is possible to build the «Dynamization» pattern by consistently describing introductions of limited mobility couplings (hinges) and transfer to more mobile, dynamic couplings. Thus, the purpose of each set of transformations is making a system more dynamic, adaptable to the operation conditions and environment. This will create prerequisites for good controllability of a produced system and its fuller coordination with the external requirements.

The «Increasing Controllability» and «Increasing Coordination» patterns may only be presented on the Basic Evolution Tree in the most general form because tracing the controllability growth of a working tool is only possible on a real system where the links of all its components with the operator or control system have been determined. The same applies to evaluating coordination of system's parameters with the operation conditions. It can only be evaluated by clearly seeing all operation details and immediate environment. The Basic Tree implies direct dependence of these parameters on the system's dynamicity.

Possibility of using the «Dynamization» pattern is checked with respect to all the steps of the Tree's «trunk». For example, for the initial object version, dynamization is possible by changing only the parameters of the entire object. To increase the object's dynamicity even more, it is necessary to introduce resources needed for that purpose. For example, transforming an object according to the «Mono-Bi-Poly» results in a system composed of several objects. The parameters of each, such as position in space, may be changed independently of each other. Such a system can be made more dynamic than a solid object.

This approach suits all the levels of the Tree's main pattern. Suppose we have an object. To build patterns, the object should be checked for the availability of resources needed for further transformations; in the absence of resources, it is necessary to obtain them. First of all, it is necessary to introduce or remove additional objects, processes and links. This may be done using «Mono-Bi-Poly», «Trimming», «Expansion-Trimming», «Segmentation of Objects and Substances» patterns. Next follows coordination of the parameters of the introduced components of a system. Here the «Geometrical Evolution», «Evolution of Internal Structure» and «Evolution Surface Micro-relief» patterns may be traced. As a result, we get dynamization opportunities providing mobility i.e. the «Dynamization» pattern works.

Moving up the «trunk», the transformed object becomes more and more coordinated and dynamic. For example, liquid consists of a great many of molecules weakly connected with each other. It is a very dynamic structure maximally coordinated with the shape of the vessel where it is contained. Its dynamicity, however, can also be increased by changing its molecular composition (replacing viscous liquid with easily-flowable liquid, for example, by using kerosene instead of glycerin) or supplementing it with additional liquids having different properties, for example, chemical reagents or ferromagnetic particles. In that case, it becomes possible to repeat the entire transformation cycle: «*Introducing components — coordinating parameters — dynamizing*» to obtain a multi-component dynamic liquid medium at the end of this cycle.

It may be said that new objects, fields, substances or processes can be introduced at any point of the Tree, even where no transformation possibilities are seen at first glance. This will allow use of the entire set of transformations for object development and further infinite expansion of the Evolution Tree. This is one of the reasons why the Basic Evolution Tree is impossible to present in a finished form. Although the transformations described in the patterns are correct and unique, the Basic Evolution Tree is multivariate. It looks more like a set of evolution patterns subject to certain application rules in each particular case, which is quite enough for practical use.

A Specific Evolution Tree is an organized set of transformation versions of an object under examination.

The Evolution Tree will be unique for every object depending on the specifics of the problem to be solved, availability of information, research problem distinctness, etc. However, a Specific Evolution Tree will be built in accordance with the same principles as the basic one.

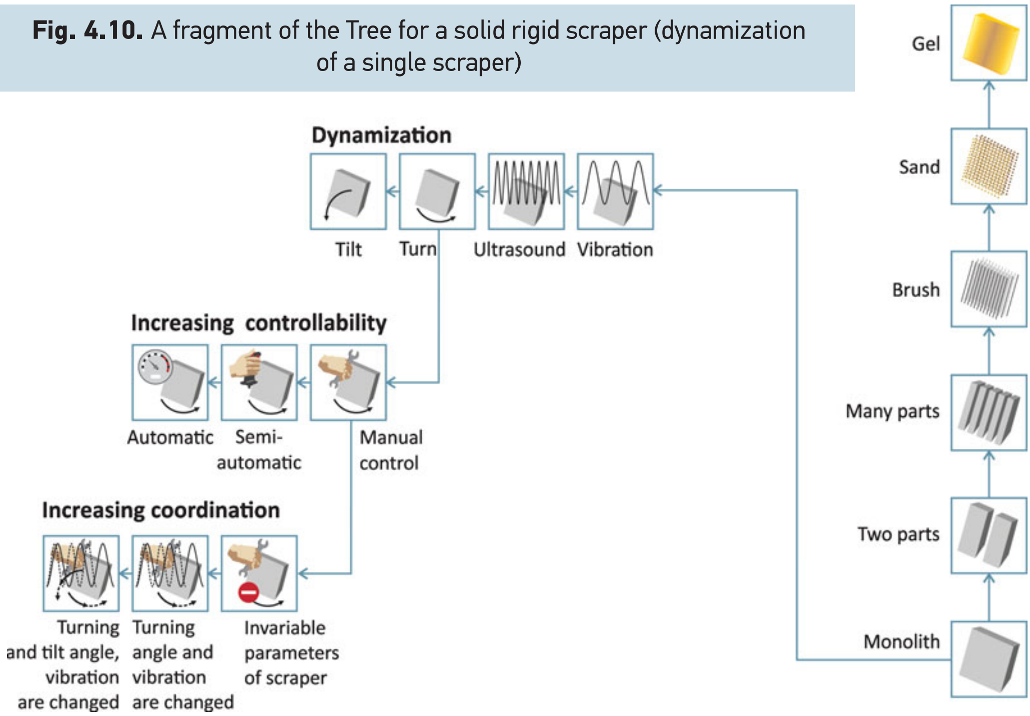
Let us illustrate the details of constructing a specific Evolution Tree with a fragment of the Tree created for a tool performing the «*Cleaning the object surface*» function (Fig. 4.10). High-quality performance of this function is essential, for example, in microelectronics, because manufacture of microcircuits requires multiple operations of applying and removing layers of different materials.

Before collecting information on different versions of a surface-cleaning tool, let us determine the technical object for which we are going to construct an Evolution Tree.

The cleaner's function is «*To remove a layer of material from a surface*». Thus in search for information we will select those systems and objects, which allow this function to be performed in a certain manner. They may be a scraper and a brush, a jet of water and adhesive substance, plasma and electric charge, i.e. everything that can be used to remove a layer of material from a surface. The starting point is the simplest version from the point of view of technical system evolution trends — a single monolithic solid rigid object having flat surfaces, a scraper (see Fig. 4.9).

For the vertical axis (the «trunk» of the future Tree) we select the «*Segmentation of objects and substances*» pattern. According to this pattern, the monolithic scraper is segmented into several narrower scrapers and turns into a brush, a jet of sand, adhesive substance, liquid,

Fig. 4.10. A fragment of the Tree for a solid rigid scraper (dynamization of a single scraper)



foam, gas, plasma. Then follows segmentation down to the level of field and vacuum, for example, a vacuum cleaner.

Side «branches» — second order patterns — emerge from each step of the vertical pattern and describe versions of the transformed scraper. The number and composition of patterns for different steps differ because they depend on the available resources. Important resources for the scraper are its cutting edge, front surface and internal space.

It should be checked whether there are any possibilities to *dynamize* the object modification represented in the first order pattern. Dynamization possibilities of a single scraper are available but they are very scant (Fig. 4.10). They are increasing the speed and force of the scraper pressing as well as changing its temperature, sharpness, angle of setting and other practical parameters. In addition, dynamization may be achieved by introducing an additional action — vibration, ultrasound, etc. Such a scraper may be controlled manually, semi-automatically and automatically ("Increasing Controllability» pattern).

Providing convenient control of a dynamized scraper allows fuller *coordination* of its parameters with operation conditions. For example, for strong surfaces with hard contaminants, the scraper pressing force may be increased, vibration and ultrasonic disintegration of the dirt may be used. With a less strong surface and weak dirt, different scraper operation parameters are set so that its action on the surface is softer.

Adding one more or several scrapers to the initial one offers new possibilities of dynamizing and coordinating the scraper's parameters with the operation conditions (Fig. 4.11). For example, in case of vibration, several scrapers can vibrate either all together as a block or

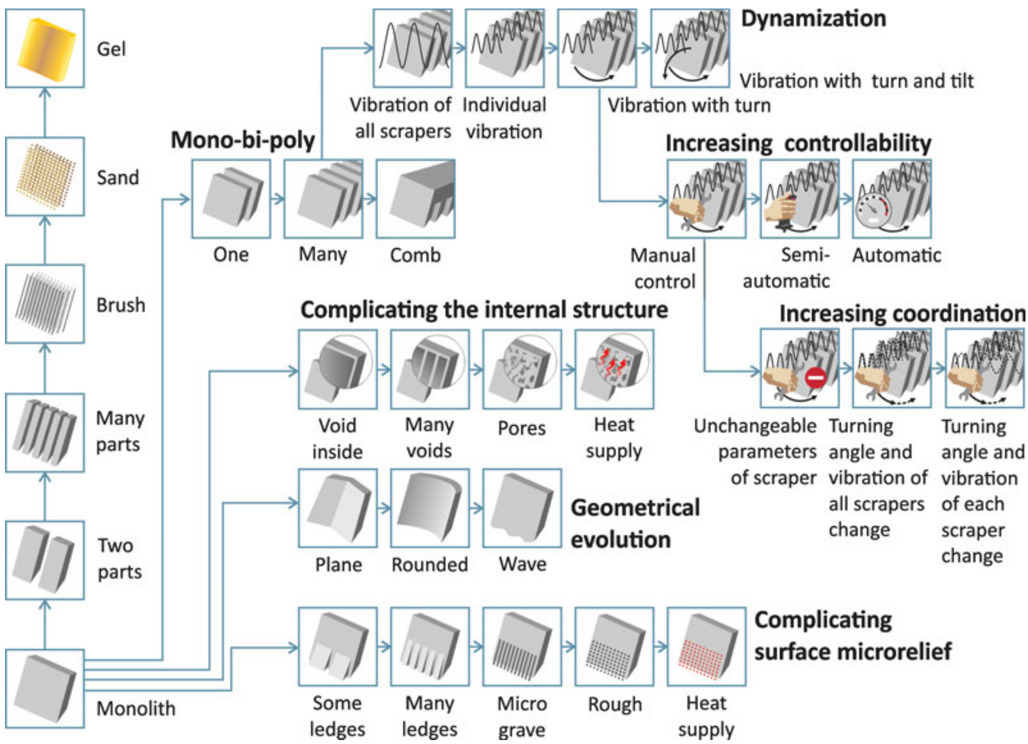


Fig. 4.11. A fragment of the Tree for a solid rigid scraper (introduction of several scrapers)

separately. The frequency and vibration force of each scraper can change, each scraper can be set at a different angle or be lifted and lowered independently of others. With a good control, such a scraper has much more coordination capabilities.

In addition, a system composed of several scrapers may also be dynamized by rotating the scrapers mounted on a common base. Then both the scrapers themselves and the plate uniting them into a single block may be made flexible. Such an elastic band excellently matches surface roughness. Closing the band and forming a circle will produce a kind of belt cleaner having a much higher productivity and operation quality.

One can easily notice that the patterns dealing with the description of the scraper's geometrical shape, surface state and internal structure stand apart. This is due to the fact that transformation versions described by these patterns are similar for all kinds of scrapers and may be referred to any modification: both to a single scraper and to a bi- and poly-system. The set of patterns including «Geometrical Evolution», «Internal Structure Evolution» and «Evolution of Surface Properties» serve as a intensifier of the concepts resulting from the analysis and show a possibility of preliminary coordination of shape, size, surface state and internal structure of the system's components. Using the «intensifier» allows the following transformation versions to be obtained:

- making the scraper concave for collecting dirt or convex for throwing it sideways;
- making the scraper's edge wavy;

- forming guiding protrusions and roughnesses;
- complicating the scraper's internal structure, for example, by introducing a void for supply of liquid or air into the cleaning edge.

When building the Evolution Tree of a real object, it is enough to show these patterns only once for a group of a single-type object modifications having in mind that transformations described in the intensifier patterns may be applied to all these modifications.

Another opportunity to obtain new coordination resources is introducing new cleaners in addition to the scraper (Fig. 4.12). It is a jet of water supplied behind the scraper, a brush that removes remaining dirt and a jet of hot air for drying the surface. The resulting set of resources offers the widest possibilities to dynamize the component of the produced compound cleaner.

For example, vibration of the scraper itself, its ability to turn and tilt may be supplemented with the water jet pulsation. A brush mounted behind the water jet is itself a sufficiently dynamized structure but its activity may be increased by using additional vibration.

With a good control, such a dynamic structure will be very effective and will allow the most different types of dirt to be removed from different surfaces. The thing is that here it is possible to use either all the components of the produced system together and some of them, which suit best a given type of surface.

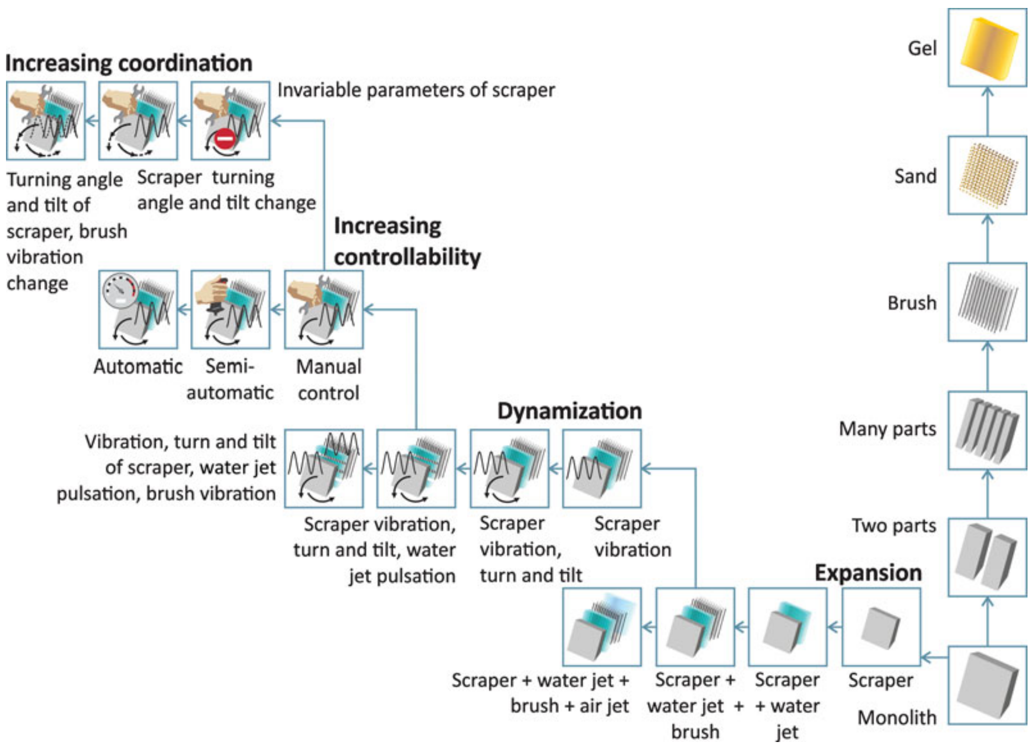


Fig. 4.12. A fragment of the Evolution Tree for a solid rigid scraper (introduction of additional cleaners)

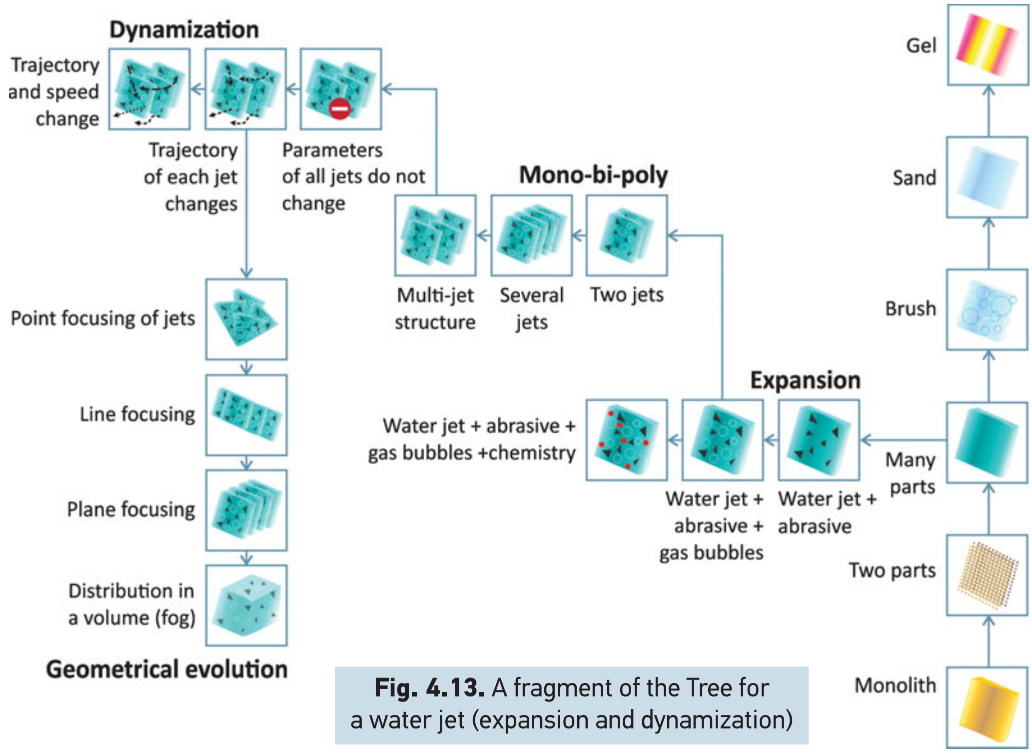


Fig. 4.13. A fragment of the Tree for a water jet (expansion and dynamization)

Tree fragments for other object modifications located on the main axis are built in a similar way. For example, a water jet is a scraper segmented down to the state of «liquid» (Fig. 4.13).

A jet of water is not a monolithic object but a set of segmented components (molecules), which in itself suggests a high degree of dynamization. A mobile water jet can easily change its shape: it may be made flat, concave or convex or in the form of a hollow cylinder. A water jet can be made pulsating with different frequency up to ultrasonic frequency or resonant with a given surface, which abruptly improves dirt removal.

New components introduced into a water jet can also increase the cleaning quality. They may be abrasive particles, bits of ice, gas or vapor bubbles, chemically active substances, etc. To create preconditions for increasing the dynamicity of initial or modified jet of liquid, the «Mono-Bi-Poly» pattern may be employed — two jets, several jets or a multi-jet structure. Then, in addition to the dynamization of each jet, we may speak of their dynamization on a new system level — on the level of their whole. Here a possibility of changing the trajectory and speed of each jet independently is traced.

The jet cleaner evolution may be described through the parameters of not only a water jet itself, but also through the parameters of a contact spot. Several jets may be focused in such a way that they hit the same point or a line. The «Geometrical Evolution» pattern recommends distributing the jet action over a plane or in a volume. The latter variant may correspond to the creation of fog for fine cleaning and absorption of harmful chemical substances.

Thus we have constructed fragments of the Tree describing the evolution of a scraper and its transformed version — a water jet. Continuing this work and taking into account all transformation versions descried in the Basic Tree, we will determine all the basic versions of a cleaning working tool. But it is often more important to focus on a single portion of the Tree and devote more attention to it.

Meeting the requirements of the classification structure:

The classification method of technical system transformations in the form of the Evolution Tree satisfies to a considerable extent the requirements formulated in sections 1.3.

1. To organize information, we have selected a tree-like structure that allows visual presentation of descriptions of *all basic* known versions of an object under examination.
2. The Evolution Tree is an organized set of objective evolution patterns based on the analysis of the evolution of many technical systems. Hence the construction of Evolution Trees suggests use of an *objective classification criterion.*
3. Every evolution pattern includes a set of generalized descriptions of transformation versions and transitions between them and may be illustrated by a transformation example of a specific technical object. Hence the *requirement of generality and specificity is satisfied*.
4. Information presentation in the form of a tree-like structure allows a designer to see all the basic transformation versions simultaneously and to distinctly *trace their structure*.
5. The availability of the Basic Tree allows foreseeing all *significant transformation versions* even if the information available on the versions of a system under consideration is scant or fragmentary.

4.3. Evolution tree construction recommendations

The Evolution Tree may be built both manually by pasting numerous cards on a large sheet of paper, or using a computer. These two methods do not have any vital differences: a large poster is just more convenient than a computer because it allows you to embrace all the modifications of an object under consideration with your eye.

In building the Evolution Tree, the following sequence of actions should be observed:

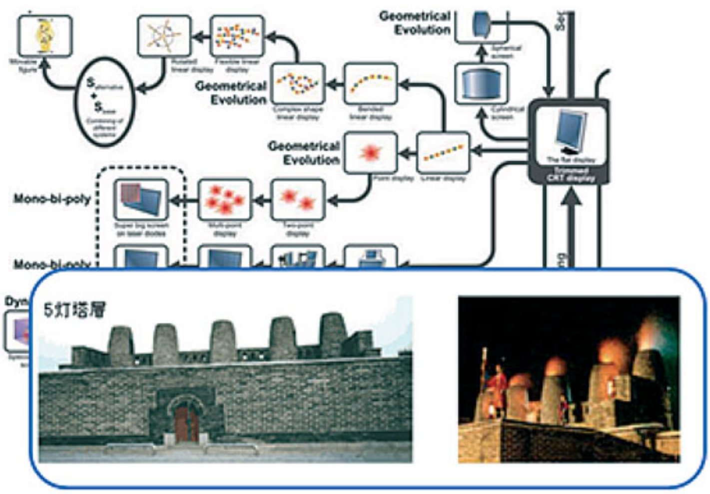
1. Determining the elementary function performed by the object of interest, clarifying and formulating its role in the performance of this function.
2. Collecting information on similar objects which either are known to perform the same role in the realization of the same elementary function, or can be adapted to the performance of this function. Making a short description of each modification of the object, paying special attention to the essence of the transformation which resulted in the appearance of this modification. Finding the initial transformation version of the object, the simplest one in terms of the technological evolution.

3. Selecting the main evolution pattern — the trunk of the future Tree. It may be any of the evolution patterns, but using those patterns where transformations of components are especially significant, such as «Segmentation of Objects and Substances» or «Mono-Bi-Poly» would be more convenient. Building the main evolution pattern, the frame of the future Tree by placing cards with the description of corresponding versions of the object under consideration.
4. Constructing of second-order evolution patterns keeping to the following rule: constructing *dynamization* patterns of object modifications if possible; if it is impossible to obtain dynamization resources, first building the patterns which provide resources — «Mono-bi-poly», «Segmentation» and «Expansion».
5. Checking whether it is possible to build second-order patterns which describe transformations of object's shape, surface and internal structure. These patterns are: «Geometrical Evolution», «Internal Structure Evolution» and «Evolution of Surface Properties». To optimize the Tree structure, it is better to add these patterns only if they reflect object transformations which are important for subsequent analysis.
6. Checking whether it is possible to build third-order patterns — «Dynamization» — after the «Mono-Bi-Poly», «Segmentation», «Expansion» patterns. Constructing these patterns in significant and indicative places of the Tree.
7. Constructing the «Increasing Controllability» patterns placing them after the Dynamization patterns. These patterns should only be built for characteristic and significant cases of controllability. For all other cases, the controllability of objects is clarified by analogy. Building the «Increasing coordination» patterns in characteristic and indicative places of the Tree.
8. Carrying out an additional information search, supplementing and specifying the Tree structure.

Charter 5

Evolution Tree of Display

Two main principles allow describing the performance of any function: morphological analysis and organization of information into a tree-like structure made up according to the preset rules from technical objects evolution patterns.



5.1. Display definition

Technological evolution tree of the real object is constantly growing and developing in a similar way to a living tree, changing its shape and bringing fruit in the form of new predictive ideas.

The Evolution Tree for a specific technical object — display — is based on the above described algorithm. It should be noted that building a finished Evolution Tree right away is impossible. Methodically continuing information search and structuring will cause the Evolution Tree to constantly grow and evolve as a real living tree, changing its conFiguration and giving fruit in the form of new prognostic technical solutions. The poster attached to the book (Appendix 1) is just a potential example of the display Evolution Tree for some construction stage.

Depending on the designer's purpose, a «magnifying glass may be applied» to a respective section of the Tree for a more thorough examination. To increase the tree crown thickness, it is necessary to add new transformation versions for constructing new evolution patterns.

The display Evolution Tree is an example of how *informationmay be organized* for further processing. When building it, we did not have any selected prototype to be improved; our aim was studying the display evolution over a large historical stage.

Since people first learned to create images, two trends in the evolution of fine arts have been traced: *art* and *techno*. The art trend describes the perfection of artistic skills. Techno deals with the evolution of technologies and technical capabilities for creating images. Looking at masterpieces of fine arts, we can't but admire the heights artistic skills can reach. Our analysis will proceed in the techno direction. The proposed research will deal with the evolution trends of technical means which allowed creation of illusory reality — a mobile picture developing in space and time — instead of a conventional flat immovable image.

In accordance with the rules of Evolution Tree construction, prior to collecting information, it is necessary to understand in the performance of what elementary function the display takes part and what is its role in the performance of this function. The function realized with the aid of the display may be formulated in the following manner: to *visualize information.* The realization of this function causes some images in the human consciousness. Thus, the initial definition of the display may be as follows: *the display is an information visualization device.*

It is important to note that *«visualization»* as transformation of the indiscernible to the human eye into what it can discern — is too wide a notion. Accordingly, there exist a great number of devices for visualizing this indiscernible information. Such a wide definition allows both a sheet of paper with printed letters and a telescope, binocular, microscope and other devices enhancing the capabilities of human vision and visualizing what is inaccessible to a naked eye to be considered as a display.

The initial definition of the display may be specified by analyzing the function performed with its participation. Performing any function requires two components: a *Tool* and a *Work Object*. What is a tool and what is a work object with respect to the function «*To visualize information*»?

To visualize images, it is necessary to coordinate the human eye (naked or intensified with a special device such as eye-glasses, a binocular or microscope) with a device presenting information in the form discernible to the human eye. The work object is an information-perceiving sensor, i.e. a human eye and the tool is a device which presents this information (Fig. 5.1).

It should be kept in mind that information may be obtained not only by means of an eye: forming a notion of some object is also possible through tactile sensations, by touch. It is just the principle on which displays for the blind and visually impaired are based.

In terms of the above-stated, the display definition will be as follows: t*he display is a device for providing a man with visual information in the form perceivable by his sense organs.*

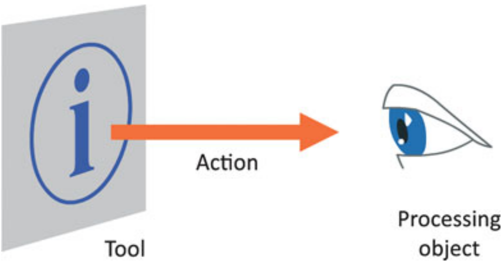


Fig. 5.1. Performing the function «To visualize information»

This definition separates vision enhancement means from information presentation means. However, based on this definition, a sheet of paper containing an image, a clay tablet, knotted strings of the quipu writing may be fairly considered as a display. It may be said that the entire visible world around us is a huge display. Such a definition of the display is undoubtedly interesting in terms of philosophy, but it does not suit design purposes.

It is just why we are going to use the name of display only with respect to *an artificially created objectspecially designed* for performing the role of a tool in the realization of the function «*To visualize information».*

When analyzing the display, it is important to take into account the inequivalence of its components with respect to the function performance. The display is a complex system including at least an image-generating device, a control unit for control of this device and a screen. In addition, the system may also contain other components, such as a casing. We are primarily interested in the system's component which serves as a *working tool*.

What is the display's working tool? It is the system's component which directly performs the main function of the system, fulfills the system's designation and cannot be removed from the system. In our case, the working tool is the screen where the image appears. Carrying out the analysis, we will focus just on the screen evolution without losing sight of other important display components.

What makes a huge army of scientists, designers and engineers work on the development of the display? One of the most important requirements is to get a clear image quality as close to the real one. The second requirement: reducing the cost of the display and the cost of its operation. The dilemma is the desire to get a good picture, but at the same time spend less money. Achieving this challenging goal causes many companies around the world to be permanently active seeking a solution.

Having formulated the display definition, we can start collecting information on devices coming within this definition,.those devices which are used or may be used for providing a man with visual information in the form perceptible by his sense organs.

5.2. Producing an image on a surface

After collecting information on different types of displays, it is necessary to identify the initial version of this system, i.e. the simplest one from the technological evolution viewpoint, and to understand how the simplest image can be produced on a surface.

One of the methods is similar to that widely used in redrawing an image on a «cell-by-cell» basis. In the general case, two operations should be performed:

1. first, dividing the surface into segments;
2. then, imparting properties, perceived differently by an eye, to each of the produced segments.

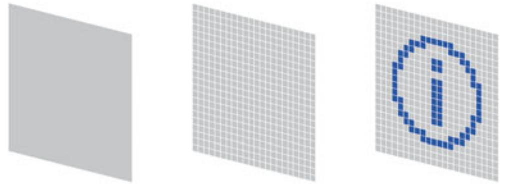


Fig. 5.2. Forming an image by surface segmentation

This will result in the appearance on the surface of a multitude of dots having different optical properties and forming an image (Fig.5.2). The greater the number of dots and the smaller their size, the sharper the image.

How can the optical properties of an individual surface dot be changed?

The surface color in this segment can be changed, for example, by painting or treating the surface with special fluorescent chemical substances. The surface shape may be changed by making it convex or concave. The surface smoothness also influences its optical properties, for example, a rough surface and a mirror surface are differently perceived by an eye. In addition, it is possible, without changing the surface shape, to change its position in an individual segment relative to the main surface by making it concave or, on the contrary, inclined. Changing the shape, position and properties of the surface segments allows a high-quality image to be produced under specific illumination conditions. Such an image can only be observed from definite points (Fig. 5.3).

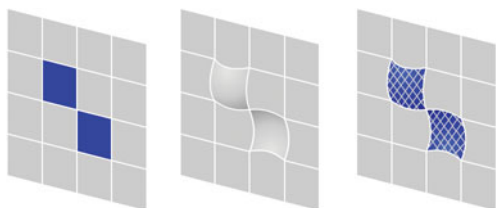


Fig. 5.3. Changing the optical properties of a surface

These methods may be combined and the produced effect may be enhanced, for example, by illuminating a transparent surface, coated with a semi-transparent paint, from inside. Such elementary points may be used to compose an image as it is done in photography or cinematograph. Examining a photo or film strip under a microscope, we can see a multitude of grains of light-sensitive material forming an image. The smaller the grain on a film, the higher the image quality.

The image on a TV screen or a computer monitor is also formed of the finest particles — *pixels*. The TV set design must provide controlled change of color and brightness of each pixel. Reducing the pixel size and increasing the image sharpness is one of the most important trends in the development of electronic displays.

Another method of producing an image arises from the technological evolution of objects according to the «Geometrical Evolution» pattern. From this viewpoint, producing an image may be presented as the expansion of the number of optical components in accordance with the «Point — Line — Surface — Volume» transition (Fig. 5.4).

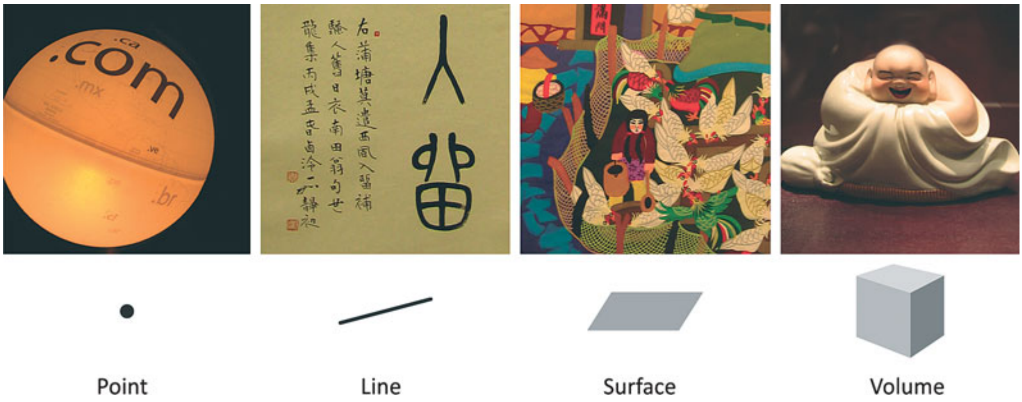


Fig. 5.4. Producing an image by increasing the number of elementary points

Indeed, one elementary geometrical element — *point* — is not enough for producing an image. To make a simple outline drawing, it is necessary to use a *line* — a set of a large number of sequentially arranged points. A picture, a multicolor continuous image, can be obtained by arranging a vast number of lines one near another to generate *a surface*. The next step may be *a 3D image* — a sculpture which may be presented as a set of surfaces arranged one over another.

Thus, both with the first and with the second image-producing method, the conclusion is that the simplest version of the optical display is a single elementary optical element — point visually standing out from the surface and transmitting certain information.

5.3. Structure of the Display Evolution Tree

After selecting the initial version of the display, we can move to the next Tree construction stage. At this point, it is important to understand which evolution pattern will be basic and to determine the key points of this pattern which is the Tree «trunk».

The preliminary search resulted in several display modifications. After analyzing those modifications, the following key ones were selected:

- elementary optical element;
- image on a surface;
- cinematography;
- TV-set with a cathode-ray tube;
- flat display;
- needle display;
- fog display;
- air display;
- display inducing image directly in the consciousness.

Transformations of any technical system performed over a long period of the system existence are usually described by the «Expansion — Trimming» pattern. Thus we may expect that this trend will be realized within the general structure of the Display Evolution Tree (Fig. 5.5).

Indeed, the sequence of the first three modifications will show the *expansion* of the display composition, because an image on a surface is a set of elementary optical elements and cinematography is a set of alternating images. Because a cinematographic image may only be seen in the dark, there appeared a motion-picture theater which comprises a hall with seats for viewers, a projector and a screen.

Cinematography was expanded due to the introduction of new elements, for example, sound. Since the advent of sound color cinematograph, the satisfactory performance quality of the main useful function was achieved and the viewers' requirements were conveniently met. The image-producing system became fully expanded.

Further evolution of the system is mostly due to the *trimming* of the elements — system's parts. The hall with seats is removed from the system. The remaining components — the projector and screen — are combined in a single cabinet forming a TV set with a cathode-ray tube. Further trimming of the system leads to the appearance of a flat display, where a projector and a screen are integrated into an elementary point — pixel. Different constructions of the flat display are well-optimized and its resources for further trimming are depleted to a considerable extent.

To obtain resources for further improvement of the display, the display evolution should pass to the microlevel. Such a transition may be performed in accordance with the «Segmentation of Objects and Substances» pattern. The last four transformations well suit this pattern covering such transformations as «*Segmented monolith*», «*Fog and vapor*», «*Gas*» and «*Ideal Object*».

Thus, let us consider the «Expansion — Trimming» pattern passing then into the «Segmentation of Objects and Substances» pattern, as the Tree trunk.

After drafting the main axis of the Tree, a repeated information search was performed. It resulted in additional transformation versions. For example, in addition to the system modifications presented in Fig. 5.5, display versions corresponding to the transformations of the «Segmentation of Objects and Substances» pattern were found so that the upper portion of the Tree looks like this (Fig. 5.6).

First goes the needle-type screen and electron paper which basic transformations are called «*Segmented Monolith*». The powder display corresponds to the «*Segmentation to Powder*» step, the liquid crystal display to the «*Segmentation to liquid*» step. The bubble and fog displays illustrate the «*Segmentation to Foam*» and «*Segmentation to Fog*» transformations. The display that projects an image directly in the air is an example of the «Segmentation to

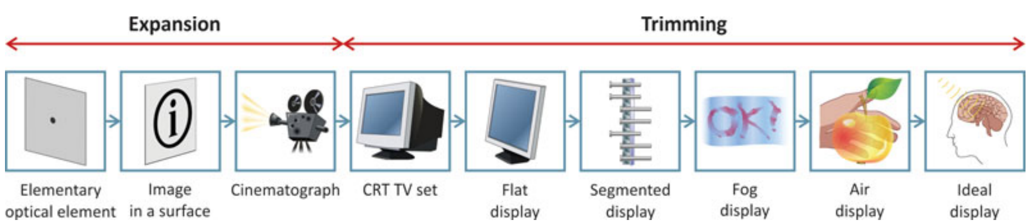


Fig. 5.5. Expansion and trimming of the display during its evolution

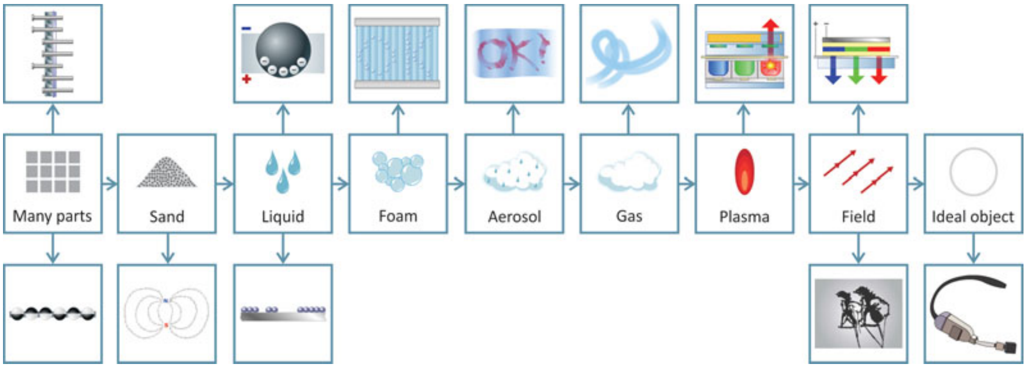


Fig. 5.6. Transition of the display evolution to the microlevel

Gas» transformation. Then follows the plasma display — «*Segmentation to Plasma*» step. It is followed by a series of displays which form an image due to the minimum transformation of fields, for example a display based on the principle of light emission by special polymers, which corresponds to the «*Segmentation to Field*» transformation.

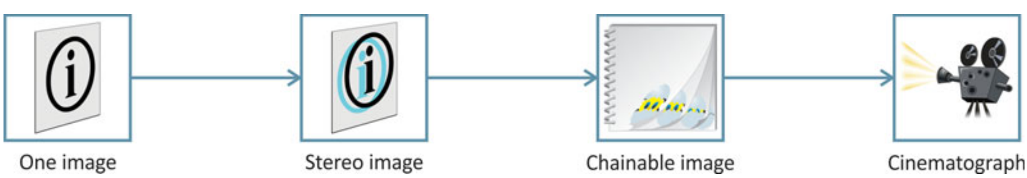
At the top, our Tree has displays which maximally comply with the basic transformation «*Ideal Object*». They are portable displays and displays which project image directly onto the retina. The most ideal version of the display is a hypothetic device that brings about visual information directly in the user's consciousness.

Horizontal side «branches» emerge from each transformation version situated on the Tree's «trunk». They are patterns according to which the «trunk» display transformation versions evolve. The number and composition of these patterns is determined by available resources. For example, an important resource of the display is its working surface (screen), light-emitting system as well as the control system of image-composing pixels.

In accordance with the Evolution Tree construction rules described in section 4, the evolution of components of the display are described by such patterns as «Geometrical Evolution of Screen Surface», «Complicating the Screen Surface Microrelief», «Mono-Bi-Poly of Pixel Color», «Display Dynamization» and others. Let us consider in more detail the evolution patterns of the Evolution Tree and display transformation versions.

5.4. From a static to a moving image. Cinema evolution patterns

Fig. 5.7. Mono-bi-poly of image



A considerable amount of time has elapsed since the appearance of the first cave drawing. During all that time, painters continuously worked on increasing the pictorial capabilities of the created images. New paints were invented, every master had his own special method of applying them to a surface.

The attempt to impart 3D illusion to a flat image by combining *two images* into one may be considered as one of the first major steps on the path of technological evolution. A single image is replaced with a double image for producing a stereoscopic pseudo-volumetric picture (Fig. 5.8). A stereo image may be obtained by viewing two drawings separately — one with a left and the other with a right eye, for instance, through a nontransparent screen. After a short practice, you will succeed in obtaining a single 3D picture from two separate images. It is much easier to obtain a stereoscopic image by using polarized light and special eyeglasses. However you do it, in the context of technological evolution a single image on the surface is transformed into a double image, i.e. *a bi-system*.

Next comes a *poly-system* composed of several components. A sequence of drawings is generally used to illustrate evolution of events in time. The simplest example is the sequence of drawings used in comics.

Thus, a «Mono-bi-poly» transition is traced (see section 3.1) according to which multiple static pictures turn into a single mono-system of a higher level — a moving dynamized image.

An example of such a moving image may be hand-made cartoons that were popular at the beginning of the last century. Simple pictures with sequentially changing position of characters were drawn at the edge of each sheet in an ordinary notebook. The notebook pages were rapidly turned over producing the impression of a moving picture (Fig. 5.9). The notebook with pictures is but a funny toy, while cinematography may be considered as a higher-level mono-system, the final version of the «*Mono-bi-poly*» pattern.

The invention of cinema allows the following contradiction: paintings are motionless, static. The viewer wants to see moving pictures, more adequately conveying reality.

Cinematograph

A great many images placed on a transparent moving strip are projected onto a screen. Single pictures alternate rapidly unseen by one's eye, thereby forming a moving image. Cinematography is one of the greatest achievements of mankind, offering huge opportunities for depicting the surrounding world [5.1].

December 27, 1895 — the day when the Frenchmen Louis and Auguste Lumière organized the first paid public film show in the Boulevard des Capuchines in Paris — is considered the birthday of cinematograph. However, the projector design was the joint product of many people such as the famous American inventor

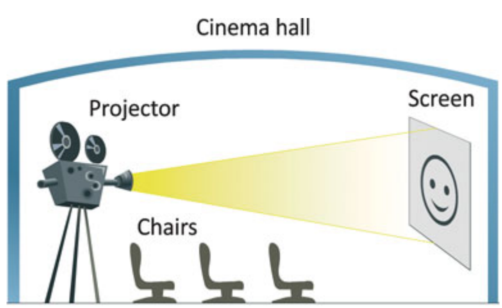


Fig. 5.8. Cinematograph

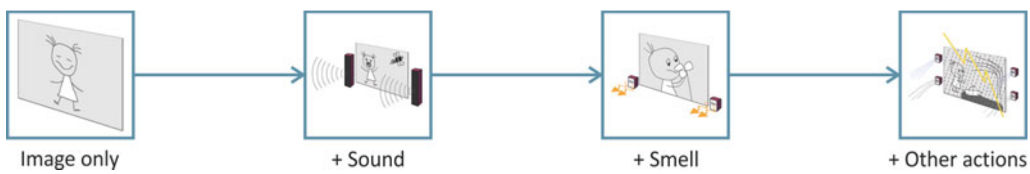
Edison, whose «cinetoscopes» began to be sold on Broadway in 1894. In Germany, it was Max Skladanovsky who experimented with cinematrograph. He made several short films and demonstrated them at Wintergarten Theater in Berlin.

Among those who contributed to the creation of cinematograph, we should mention Hannibal Goodwin, the inventor of celluloid film; Louis Laprence who invented perforation — four holes on both sides of a frame, which is still a standard; William Freeze Green, the inventor of the cinecamera; Birt Acres and Robert Paul, the creators of the projector. One of the most important components of the projector — the mechanism of intermittent motion of film which makes motion-picture apparatuses chirp amusingly — was invented by the Russian inventor I. Timchenko.

A large number of people participated in the creation of cinematograph and it is just the result of these forces that gave rise to cinematography — the land of dreams of the 20^{th} century.

The cinema hall is a huge display (Fig. 5.8). The system for demonstrating «moving pictures» generally includes a projector, a screen, viewers' seats and a «housing» — the walls and roof of the cinema building. Viewers are found inside this display.

Fig. 5.9. Image expansion



Initially, only *image* was demonstrated to viewers in the cinema. Therefore, actors tried to convey all information about a described action by means of facial expression, gestures and other purely pictorial means. If it was impossible to do without verbal comments, intertitles were used. Music served for emphasis. At that time, however, music was an *external* component with respect to image. Musical background was provided by pianists present in the cinema hall.

Sound was the first serious supplement to image. Sound recording invented by Edison and sound reading from the film by means of a photosensitive component invented by Stoletov provided musical background fully synchronized with the screen image. Because we perceive the lion's share of information through eyesight and hearing, cinematograph satisfied the audience for many years creating almost fully naturalistic illusion.

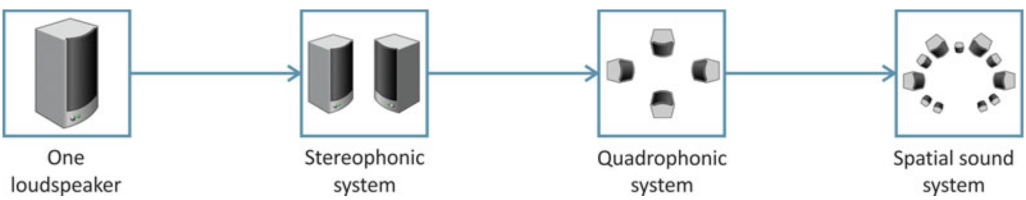
There were also attempts to use other sense organs of viewers: sense of smell, tactile receptors, vestibular apparatus. One of the attempts to perfect the naturalistic illusion is introduction of *odor*. Unfortunately, it is not easy to render all the range of odors which may accompany the action on the screen. First, inventors tried to use background odors such as the odor of the forest or the sea. In the fifties, the Swiss professor Hans Laube, developed the Smell-O-Vision system and presented it in 1960 during the demonstration of the film «Scent of Mistery». Special tubes passing from the control panel to the cinema hall supplied gas with odors synchronized with the action in the film. The idea was not a success because removing the previous scent from a hall required a very effective ventilation system [3.10].

The individual approach — giving a special «Odorama» card to each viewer — turned out to be much more promising. As the play progressed, a viewer had to scratch marked disks in succession and smell them. Another alternative is a small box hanging on the viewer's neck. Scent samples inside the box were activated by means of a radio signal.

At present, the idea of using odors for information conveyance is being regenerated on a new level — on the basis of computer technologies [5.3]. For example, Tsuji Wellness and France Telecom have declared about their joint project called «Kaori (scent) Web». A small additional block is connected to a computer through a USB-port and produces certain scents to a user's command. The scents are produced by mixing the vapors of six different smelling gel-like substances. A set of these substances is stored within the device and scents are spread by built-in fans.

The display modification with the addition of *special effects* to the image also belongs to this evolution pattern. A sea battle on the screen is accompanied by water splashing on the audience. Wind blowing in the hall enhances the effect produced by the image of a horrible storm on the screen. Pillars of fire can arise in front of the screen or laser effects may be used. We may expect further inventions in this direction.

Fig. 5.10. Mono-bi-poly of accompanying sound



Additional components introduced into a system under consideration may also be analyzed by means of technical system evolution patterns.

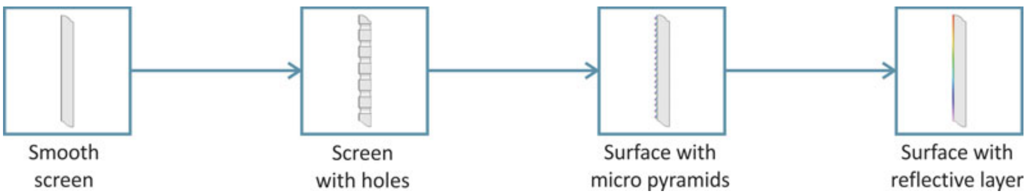
Let us take, for example, the image-accompanying sound. The advent of the sound film after the silent film was a real breakthrough. The first sound system only comprised *one speaker* mounted behind a screen. The mono speaker system allows fully employing the human hearing capabilities and producing a full presence effect; therefore sound generation systems continued to develop.

The stereophonic system which removed that disadvantage was composed of *two speakers* placed at the sides of the audience. Filters split a signal supplied to the speakers into a high-frequency and a low-frequency component. Correctly positioning the speakers relative to the walls and other sound-reflecting components of a building produces a pseudo-3D sound.

The quadraphonic sound system employs *four loudspeakers* positioned in front of, behind and on the sides of the audience. Different methods are used to impart a volumetric effect to the sound: sound delay on one of the channels, varying the sound frequency and loudness on each channel. This results in new interesting effects.

The volumetric sound system may include ten or more speakers. Most of them are positioned near a screen and the rest of them are placed on the sides. This provides not only a volumetric natural sound but also sound with different special effects, dynamic sound capable of moving synchronously with its source on the screen.

Fig. 5.11. Complicating the screen surface micro-relief

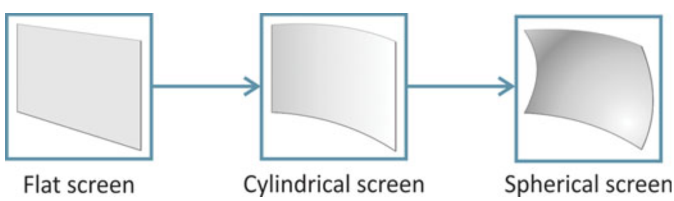


Initially, the screen surface was *smooth*. Yet with the appearance of sound, it became evident that with the speakers mounted behind the screen the sound was loud, the screen started to vibrate, thereby degrading the image quality. To remove this harmful phenomenon, perforations on the smooth surface of the screen were introduced. Their function was to dampen vibrations.

The next step was caused by the necessity to house more people in a cinema hall. With a smooth screen, only viewers sitting close to the central axis of the hall could enjoy a high-quality image whereas for those sitting closer to the sides the image quality was much worse. To remove that disadvantage, the screen surface began to be covered with *microscopic pyramids*. The faces of the pyramids reflected part of incident light, which improved the image for people occupying side seats.

A screen having a surface with an additional *reflecting coating* may be considered as the final transformation version within this evolution pattern. It may be a layer of special plastic, but at the dawn of cinematograph evolution, a better image was produced just by watering the screen.

Fig. 5.12. Geometrical evolution of the screen



The screen is most frequently made *flat*. However, the flat screen does not allow a high-quality wide-screen picture to be produced. This is because the film in the projector is flat and the distance from the projector lens to the screen center and screen edges is different. With a small image format and a large distance from the projector, this circumstance does not affect the image quality.

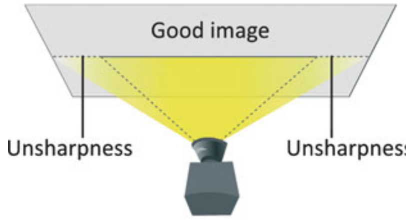


Fig. 5.13. Flat screen

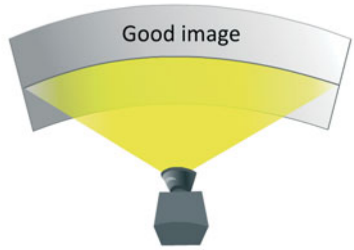


Fig. 5.14. Cylindrical screen

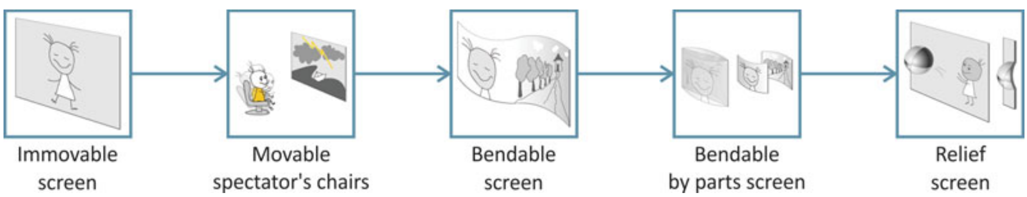
However, the larger the image size and the smaller the distance from the projector to the screen, the greater the image distortion. On a wide flat screen, sharpness is practically impossible to adjust: if the image is sharp in the center, it will be blurred at the edges and vice versa (Fig. 5.13).

This defect of the screen was removed by transitioning to a *concave* cylindrical surface. The radius of the screen surface bend was selected such that the distance from the projector lens to all the points on the screen was the same for a given cinema hall (Fig. 5.14). The image at the screen edges was less exposed to distortion. Viewers liked the cylindrical shape of the screen too, because it was much more convenient to observe the action on a wide bent screen «embracing» the audience.

The next shape modification within the frame of the *Geometrical evolution* pattern is the spherical screen. The spherical screen offers quite new qualitative capabilities for image display. For example, the spherical screen of a planetarium allows the image of the entire starry arch to be observed.

In addition, the spherical screen may be used to demonstrate special kinds of films where action surrounds the viewers creating the presence effect.

Fig. 5.15. Screen dynamization



Screen dynamization offers new possibilities to film creators. An immovable screen is replaced with a movable one. Moving the screen itself is unnecessary. The motion illusion may be produced by changing the position of the viewers' seats. The seats tilt, vibrate, jump up synchronously with the film action, which allows the audience to feel themselves as participants of the events on the screen. Such cinema halls are especially suitable for dynamic films which one can see at «Futoroscop» in the French city of Poitiers.

The Japanese corporation NTT developed this idea and demonstrated the system of control of human motor reflexes which employs the effect of galvanic vestibular stimulation (GVS). A viewer wears earphones which broadcast not sounds but direct current pulses. There is an anode connected to one ear and a cathode to the other. Supplying an electric signal to the mastoid behind the ear will cause a man to lose balance. «Twitching» in the left ear will shake him to the left, «twitching» in the right ear will shake him to the right. GVS-earphones can make the sensations of those training on a race car simulator practically indistinguishable from real-life sensations. The driver will even be tossed to and fro at virtual turns.

Following the logic of this pattern, it may be proposed to change the *shape and position* of the screen according to the film maker's idea during the film demonstration.

The screen curvature may be changed from concave through flat up to convex in accordance with the image character (for example, a convex screen for a closeup picture, concave for a distant view or in some other manner).

The screen shape may be changed *by parts*, for example, half of the screen is flat while the other half is convex or half of the screen is concave and the other half is convex. The screen may be wavy.

Following this pattern brings us to the idea that the image may be emphasized by changing the screen surface relief shape synchronously with the image. As a result, the viewer will see a movable color bas-relief image.

5.5. Advent and evolution of the TV set

From the technological evolution viewpoint, the appearance of the TV set is obviously the *trimming* of cinematograph. Really, cinema includes a building, viewers' seats, a projector and a screen. Trimming completely removed the building from the system and the viewers' seats were moved beyond the display. The projector and the screen, initially separated in space and unconnected, were combined into a single compact device which scans and projects image on a screen (Fig. 5.16).

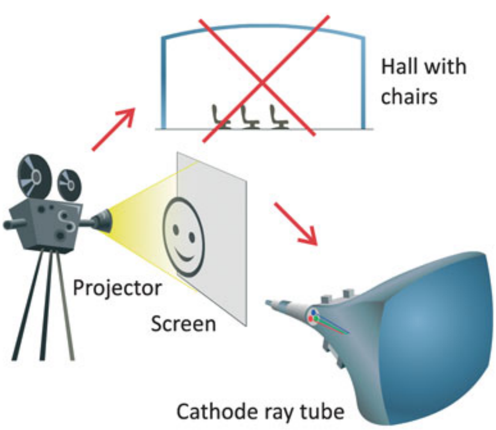


Fig. 5.16. Trimming of cinematograph down to a TV set with a cathode-ray tube

Invention and evolution of television was enabled by the development of a cathode-ray tube, ionoscope, kinescope and many other components responsible for image transmission, reception and demonstration. All this resulted in the creation of the TV set [5.4.].

In first TV sets, scanning was performed mechanically by means of the Nipkow disk. The rotating disk had some spirally arranged holes through which dots of a primitive image were projected onto a screen.

Mechanical scanning provided poor quality of image; therefore developers focused their efforts on the development of electronic scanning methods. The scanning cathode tube was invented by the German researcher Karl Ferdinand Braun in 1897. First cathode-ray tubes found application in oscilloscopes.

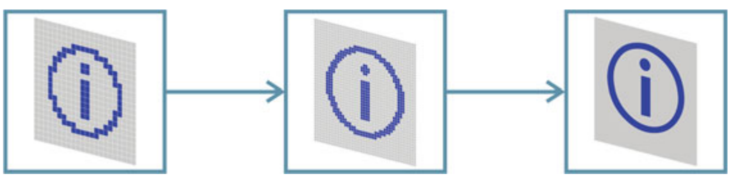
The CRT did not have any mechanical components and image was scanned by magnetic deflection of an electron beam generated by an electron gun. In 1932, the RCA research laboratory in America, headed by the hardware engineer Vladimir Zvorykin, demonstrated the first electronic TV set. For some time, electronic and mechanical television existed in parallel. But in the middle of the thirties, mechanical television producing a grainy and blurry image and having the scan limit of 30 lines was replaced with electronic television with the number of scan lines amounting to 170–400 [5.5].

Fig.5.17. A TV-set having a Nipkow disk (the window on the right is a small screen)

Television did not equal cinematograph both in image size and image quality. But the possibility to get prompt visual information and see films without leaving one's home allowed television to survive in the competition with cinematograph and to occupy its technological and market niche.

CRT television evolution patterns

Fig. 5.18. Display surface segmentation



An image on a photo or a cinefilm is formed of small grains of light-sensitive material. Photographers always seek to produce a sharper image so that one could see small details, which can only be achieved by reducing the grain size. An image produced on a TV display is built in a similar way. The display is segmented into small portions the properties of which can be separately controlled. These portions are referred to as *pixels*. Reducing the size of a single pixel, that is, segmenting the display surface into increasingly small parts is one of the main TV display evolution trends.

The display of first mechanical TV sets based on a Nipkow disk was formed of only 30 lines which produced about 800 pixels. Looking at such a display, one could understand what was shown — a human face or a landscape and to distinguish a man's face from a woman's face. But small details were not seen on such an image at all.

The next step was the advent of electronic television exhibiting a 170-400 line scan which provided an acceptable picture quality, especially if a producer and a camera man tried to avoid demonstration of small details. A standard for a cathode-ray tube — 525 lines — was gradually introduced. It still remains the basic standard for economy class TV sets.

Further increase in image sharpness was achieved after the appearance of High Definition TV (HDTV). Currently, two resolution standards are employed: HD720 (720×1280 pixels) and HD1080 (1080×1920). This format allowed the finest details to be sharply displayed and improved the image quality.

An interesting method of increasing the sharpness of a displayed picture is employed in the HD1080 format display when a received signal has a lower resolution. With the ordinary technology, each received pixel is uniformly distributed among 4 display pixels and the image on the display just becomes coarser. GHDV technology provides a high picture sharpness while viewing ordinary programs. This is due to every received pixel is only brightened on one display pixel while the color and brightness on the remaining three pixels are formed by the display computer depending on the color and saturation of the neighboring screen portions.

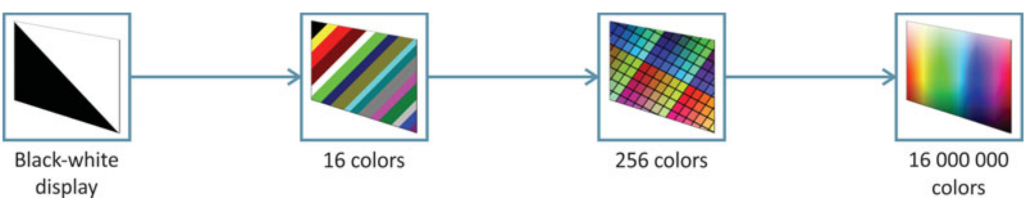
The display sharpness improvement pattern will continue even if it surpasses the human eye perception threshold, because there exist numerous displays where it is necessary to discern the finest image details, which is achieved by multiple zooming.

Increasing the display sharpness is hampered by the fact that producing a color image requires dividing a pixel into three portions, which reduces the image sharpness threefold. Transition to other types of displays, such as, for instance, liquid crystal displays, offers new resources for solving this problem. One of the promising displays called UFS (which can be deciphered as an ultrafine sharpness image display) [3.40] allows a color image to be produced without dividing a pixel into three parts.

The required brightness and color of a pixel are provided by placing at the rear of a liquid-crystal filter three highlighting lamps: red, green and blue, which flash a great number of times during the display of a single frame while the formation of a needed color is controlled by the liquid-crystal filter which can open a window in front of a pixel. This increases the display sharpness at least threefold, the pixel size being the same.

As for the prospects of the display sharpness improvement, a transition to a molecular and further to a field level is possible. For example, use of iridescent effect which is responsible for the intricate color patterns of soap bubbles looks promising.

Fig. 5.19. Mono-by-poly of screen colors



Another contradiction that developers are constantly emerging displays, there was a discrepancy of technical possibilities of creating colors the users need.

The number of colors grew with the display evolution. Initially, it was a *two-color* black-and-white image. Such an image could even be called single-color because in cinematograph a picture is produced by projecting black color of different depth onto a white background. The image is produced on the screen of a black-and-white TV-set by combining a dark background with bright dots.

Then a color screen appeared where the needed color was the result of mixing the basic colors — red, green and blue.

The number of colors presented to a viewer on a computer monitor was limited by the display card capabilities. The screens of the first color monitors could only give *16 colors.*

Fig. 5.20. Mono-bi-poly of pixel colors

Perfecting the computer monitor and software design increased *the number of colors* to *256*. Modern displays provide a very precise color rendition by using *16 000 000* or more colors.

The number of active colors composing a pixel was growing with the display evolution.

In black-and-white displays, only one color — white — was active while the black color was produced as a background color at the expense of the dark color of the screen surface. Then a transition to a color image followed. It is interesting that the first color TV sets employed the principle of mechanical scanning of colors analogous to that used in black-and-white TV sets. Such a color TV set was an ordinary cathode-ray black-and-white tube having a rotating three-color light filter installed in front of it.

Further evolution of the color television was identical to that of the black-and-white television and subsequent modifications employed electronic scanning instead of mechanical. The number of active colors in a color TV set pixel increased compared to the black-and-white TV set. Each pixel began to be segmented into three subpixels of different colors: red, green and blue (RGB). The white color of the pixel was formed by simultaneously switching on all the three subpixels and the black color was formed as the background color of the screen after switching off.

Samsung engineers noticed that to produce the white color of a pixel, it is necessary to maintain all the three subpixels switched on, which causes excess energy consumption. But the white color is often used in computer-generated images. Therefore, the company developed a new manufacturing technology of liquid crystal displays. The company's engineers suggest using a fourth white pixel in addition to the three main subpixels. Introducing the white color into a pixel will reduce energy consumption by at least 50%. This is particularly important for a portable computer on rechargeable batteries when the time of self-contained operation is one of the most important characteristics of mobile devices. At the same time, according to the developers, the brightness of the new monitors will be at the same level as that of the ordinary ones.

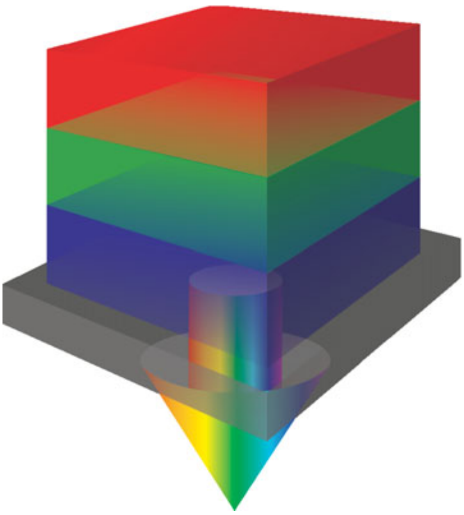


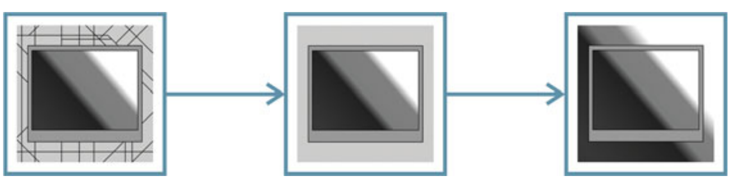
Fig. 5.21. In the SOLED system, light-reflecting layers are arranged one over another

The next version is a pixel containing *six color subpixels*. The Israeli company Genoa Color Technologies has developed the technology according to which the three basic colors — red, green and blue — are supplemented with three additional colors — cyan, purple and yellow. The joint emission of all the pixels allows rendering 95% of the color gamut perceived by a human eye [5.7].

This pattern will be logically concluded with a pixel capable of generating any required color, yet it is not clear how to make such pixel in accordance with presented tendency — introduce more subpixels. There arises a contradiction: cells should be many to provide more precise color rendition but the pixel size is limited and subpixel cells should become increasingly small. It is obvious that increasing the number of colors in this manner has its limit. One of the possible methods of solving this contradiction may be a display where each pixel is capable of generating any color.

Such a transformation may be illustrated by one of the display modifications developed according to OLED (organic light-emitting devices) technology (Fig. 5.21). The new display called SOLED was developed by researchers from Princeton and University of South Carolina, USA. Transparent light-emitting layers in each pixel of this display are arranged one over another, sandwiched. Activating all the three signals with an electric signal in a necessary proportion may produce any color of an active pixel [5.8].

Fig. 5.22. Coordinating the image with the surrounding background



It is important not only to obtain a sharp, color-saturated picture on a display, but also to make it easy-to-perceive by a user without any quality loss, which is not always easy. For example, an important role belongs to the place where a display is installed and to the image coordination with the background. The most non-coordinated version is a place illuminated by the sun where a display is surrounded by objects which are in sharp contrast with the image. It is better to place a display against a grayish background of a wall illuminated with a gentle glow. In this case, viewers' attention will not be distracted from the action on the screen. In addition, some displays are equipped with a device for automatic adjustment of the screen brightness depending on external illumination.

Fig. 5.23. The image on the display is not coordinated with the surrounding

Fig. 5.24. The image of the display is coordinated with the surrounding

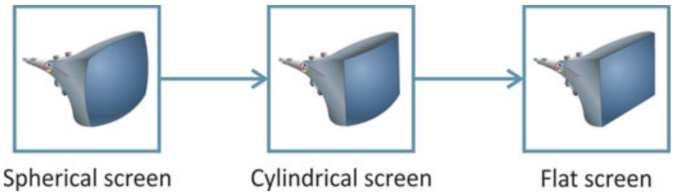
Philips moved even farther towards image coordination with the surrounding and offered a TV set providing background illumination — Ambilight, which immediately won popularity on the market. The idea was to make use of the peculiarities of human vision. Looking at some object, we only see clearly the place at which our sight is directed while the rest of the space is perceived by peripheral vision. The same is here: the wall around the display producing a sharp image was highlighted with special lamps, their color being matched with the averaged color of the image (Fig. 5.24).

First models of Ambilight TV sets illuminated the wall on the screen sides, the light color being the same on the right and on the left. In Ambilight Full Surround TV set, the screen was highlighted on four sides and a processor selected color and brightness for each side individually.

The natural continuation of the side-illumination display evolution toward the image/background coordination was creation of Amibilight Spectra system. The glow intensity of one hundred fifty light-emitting diodes of three primary colors arranged along the screen perimeter produces a tremendous participation effect.

It should be noted that the following contradiction is being gradually aggravated here. Both the background illumination and the display picture are dynamic in character whereas the screen edge frame circumscribing the picture is static. To add visuality, they should fuse forming *an organic whole. One would expect that it is the boundary area between the image and background illumination that should become more dynamic and help to bring the image perception on the screen close to that of the real world.*

Fig. 5.25. Geometrical evolution of the screen



The screens of the first TV sets had a work surface in the form of part of *a sphere* with a comparatively small radius. The image on such a screen was slightly distorted, which was not quite convenient for viewers. That shape was caused by some production details and insufficient strength of the glass used in the kinescope production (Fig. 5.26). The fact that the shape of the tube, made from fragile glass should be as close to a sphere as it keeps outside air pressure well. The image on the spherical screen turned slightly warped and it was not convenient for the spectators.

Perfecting the kinescope production technology was accompanied by the development of new, stronger types of glass and the invention of strengthening metal frames. As a result, screens were becoming increasingly flat, the curvature in the vertical direction being made smaller and smaller, while in the horizontal direction it remained large enough. The surface of the screen having a horizontally curved section and practically smooth vertical section can be considered *cylindrical* in a certain sense (Fig. 5.27).

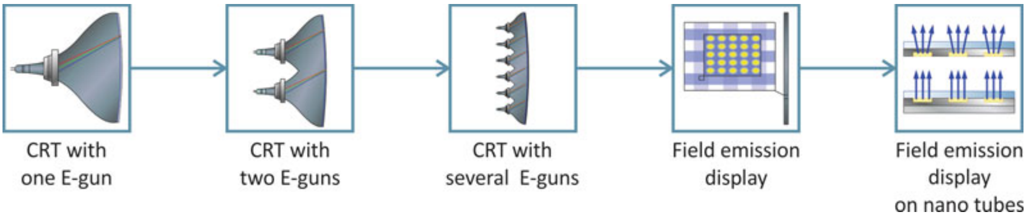
The screen curvature was reducing with the technology development. Finally, absolutely *flat* screens, such as Sony's «Trinitron», appeared on the market. Such screens are the most convenient for image perception and have practically replaced computer displays of other shapes by now (5.28).

Fig. 5.26. Spherical screen

Fig. 5.27. Cylindrical screen

Fig. 5.28. Flat screen

Fig. 5.29. Segmentation of the cathode-ray tube



In perfecting the TV set construction, designers tended to increase the screen size and reduce the housing depth. However, it was not an easy task with respect to the CRT display. The problem is that one of the basic disadvantages of the tube is its large length, because the electron beam drawing an image on a screen has a relatively small angle of deflection.

The designers' efforts aimed at reducing the CRT length were focused on the increase of ultimate deflection angles of the electron beam. However, the beam deflection may be increased to a certain limit: the ratio of the screen diagonal to the cathode-ray tube length has limits. Hence making a large screen increases the overall size of the TV set, including its thickness. Of interest is the attempt to increase the screen size by segmenting the CRT into two tubes of a smaller size. The screen size remained unchanged but the image was produced by two electron guns of a smaller length, each of which «served» its half of the screen.

This solution immediately reduced the display thickness and increased its compactness. It is clear that using *several small electron guns* for one screen may reduce the screen thickness and bring the display housing shape nearer to flat. This trend, however, was not further developed because the problems of joining images on two neighboring screens «served» by different guns was not solved at that time.

Curiously the CRT segmentation trend has recently been continued [3.14]. The technology developed by Candescent Technologies Corporation for flat displays is based on the principle of image formation in traditional kinescopes. Field-emission displays as they are called by this company, have several microscopic electron guns in each pixel. To produce the image, the electron guns act on a layer of phosphor (Fig. 5.30). It turned out that segmenting an electron gun into a number of *small guns* was sound, just the optimal segmentation level of the initial electron gun was not achieved. It should have been segmented into a number of small guns at least equal to the number

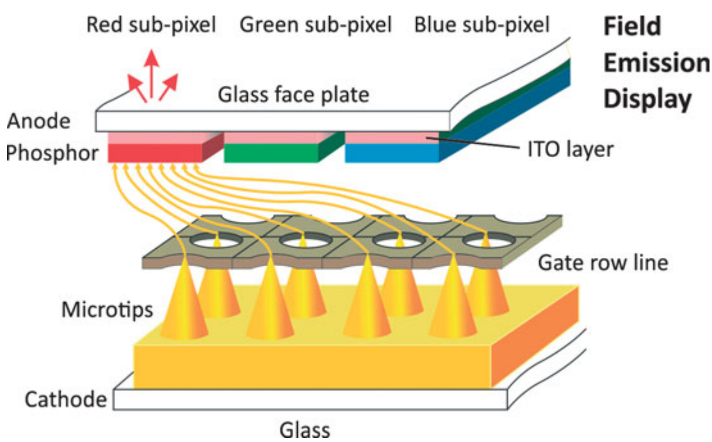


Fig. 5.30. Field emission display

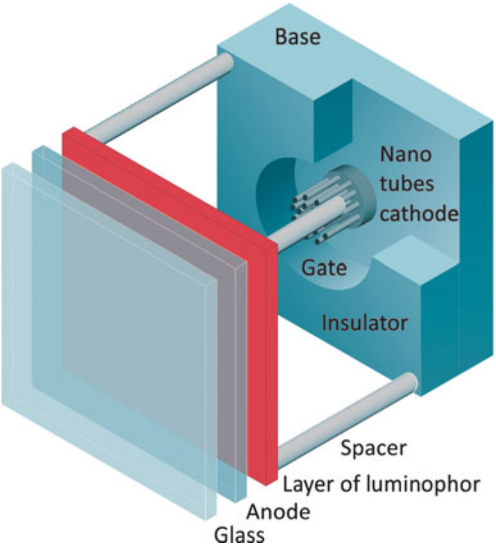


Fig. 5.31. Display on nanotubes

of pixels on a screen as is done in field-emission displays.

The next step is using silicon nanotubes. These original cylindrical macromolecules having approximately half nanometer in diameter and several micrometers in length are placed on a cathode in the direction of the upper anode with phosphor (Fig. 5.31). As voltage is supplied to the cathode, each nanotube actively emits a beam of electrons which excite phosphor emissions. In this case, a *great number of microscopic electron guns* is located *in each pixel of the display.*

Due to the use of carbon nanotubes Applied Nanotech jointly with six Japanese electronics companies created the prototype of a 25 inch TV set having a brighter and sharper image than that of modern models [5.9].

Such displays have low energy consumption but their cost is high because of expensive silicon nanotubes needed for their production. Nanotubes, however, are likely to become less expensive due to the active research carried out in this direction.

For example, the New York company NaturalNano has developed the method of extracting nanotubes from a special kind of clay called halloysite and having a tubular structure. This method may be considered as an alternative to the artificial creation of nanotubes. The natural nanotube diameter is 40 to 200 nanometers and their length is one micrometer [5.10].

5.6. Flat display

According to the ideality concept, an ideal display is the one that is absent while its function is performed. A TV or computer display approaching the ideal is something absolutely flat, occupying almost no significant space; properly speaking, it is just a screen. For example, a picture on a wall occupies minimum space. The picture is a window to another reality called art. The display is also a window — a window to the virtual world (Fig. 5.32), but as distinct from the picture, it shows not one dead fragment, but the dynamics of events.

According to the ideality requirement, the display evolves towards the decrease of its relative thickness. The cathode-ray tube itself was becoming increasingly shorter mainly due to the increase in the electron beam deflection angle. That, however, did not lead to a significant reduction in the display thickness (Fig 5.33). A sharp reduction in thickness only became possible after the liquid crystal display and other types of non-scanning displays had been invented. In such displays, all image generation systems *are trimmed* to a single pixel. Properly

Fig. 5.32. A picture and a flat display

speaking, each pixel simultaneously contains a device that performs the functions of an electron gun, a device responsible for the control of light signal formation and a screen. Trimming the cathode-ray display into a flat display was carried out by removing the electron beam scanning system and maximum miniaturization of the light signal generation device (Fig. 5.34).

The development in 1967 of the first prototype of a *liquid crystal display* by the RCA engineer Heilmeier was the beginning of the flat display era. Sharp Corporation simplified the panel manufacturing technology and introduced the liquid display to the market as a replacement for bulky light-emitting-diode displays used in calculators. In 1973, the EL-705 calculator with the world's first liquid crystal screen appeared on the market. Ten years later Sharp produced Crystaltron, the first color TV set having a three-inch liquid crystal screen. Other types of flat displays continued to develop concurrently with the development of liquid crystal panels. They are plasma panels, displays employing *light-emitting polymers* (LEP) and *organic substances* (OLED), displays based on the *field emission* principle, etc.

Flat displays are rapidly growing in size. Here the leadership belongs to the South Korean Samsung.

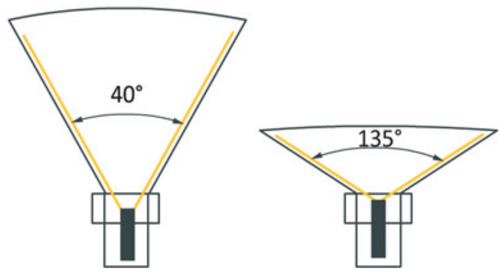


Fig. 5.33. The display thickness decreases with the increasing beam deflection angle

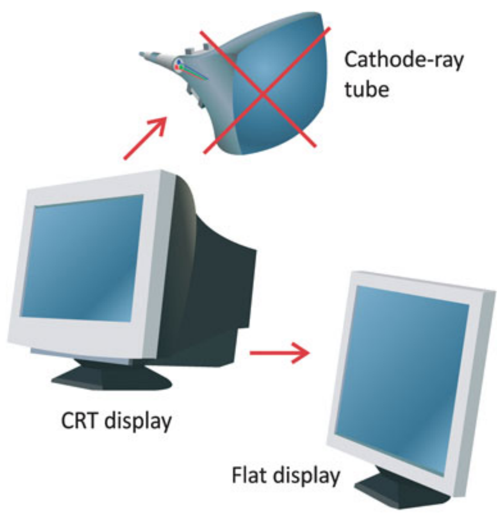


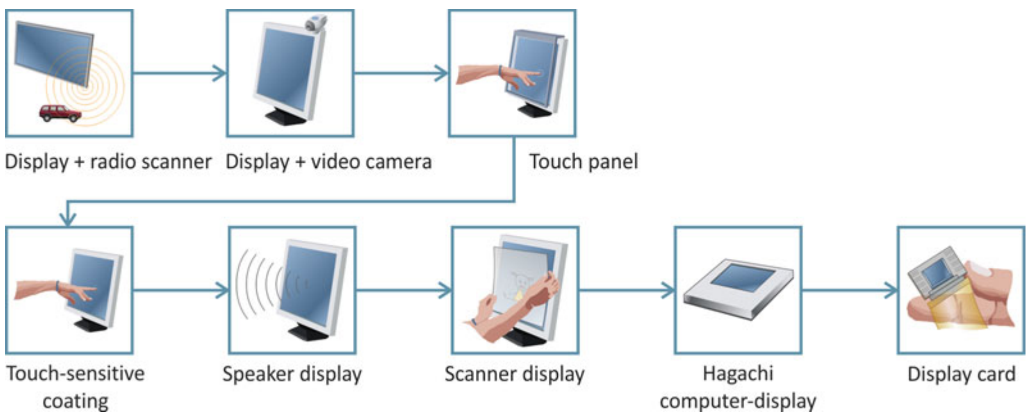
Fig. 5.34. Trimming the cathode-ray tube display

Several years ago, the company announced launching of a liquid crystal display with the record diagonal size equal to 57 inches. At the beginning of 2005, this company beat its own record and started commercial production of a display being 72 inches in diagonal. The plasma panel declared by Samsung in 2004 is 102 inches in diagonal. The company's promising display based on the use of organic light-emitting substances already has the diagonal size of 21 inches.

The appearance of flat displays called into being an entire range of electronic devices. While the replacement of CRT with a flat display in a TV set or table computer offers few advantages to users, such devices as a notebook, palm computer or mobile phone with a small screen would be very difficult to create without using a flat display.

Flat display evolution patterns

Fig. 5.35. Expanding and trimming of flat display systems



Considerable display transformation resources are provided by supplementing a display with new systems having other functions, which results in the display expansion. The most effective construction is obtained by trimming the components of the combined system after introduction of a new system.

The advertising campaign of Alaris Media Network has combined the *display* and the *frequency scanner* which determines the frequency of a radio receiver. Now, huge billboards — displays produced by the company — can determine to which radio stations most drivers of bypassing cars are tuned [5.11]. Each US radio station generally has its own audience. A control unit built in a billboard receives a message and displays a corresponding image. For example, if radio stations predominately preferred by men are detected then fishing implements are advertised on the display. If, on the contrary, most drivers are women, the display switches to advertising cosmetics and ladies' wear.

American researchers from Duke University combined a *display* and a *web-camera.* They created a device called Face-Off which determines whether a user is looking at a monitor.

A display only remains switched on when there is a user sitting in front of it. If there is nobody near the monitor, the display automatically switches off, which saves power [5.12].

At the CeBIT America 2003 trade fair, KEYTEC presented a device called «Magic Touch». This new product looks very much like protective screens mounted in front of a computer monitor. However, as distinct from such screens, the panel serves not so much for reducing the radiation level as for providing a user with additional computer control capabilities. In particular, after installing this device, any liquid-crystal or ordinary CRT monitor can perform the functions of touch displays. To open a file or perform some other actions, it will be just enough to touch a place on the pane [5.13].

Such a system can be trimmed by combining a sensor panel and a screen. To this end, simply transfer the touch panel function directly to the monitor screen. This may cause a problem of reducing the display image brightness because the transparency of most existing touch film screens does not exceed 76%. However, a breakthrough is expected in this direction. Fujitsu has developed a *touch screen* having a transmission factor of 96%, which practically causes no reduction of the display brightness [5.14].

The high transparency has been achieved by giving up the resistive film used in most touch screens. Acoustic waves are generated on the screen surface. The change in their characteristics serves to determine the touch point. In other words, the resistive film and the screen itself were practically completely *trimmed*; the only thing that remained of the film is its useful property — sensing the user's fingers touching the screen surface.

Trimming may be illustrated by those cases when a display takes the functions of other computer or TV set components. For example, for a compact flat display, there exists a problem of speaker location. Sharp and Hosiden Electronics demonstrated an experimental prototype of a liquid crystal panel which itself can be used as *speakers* [5.15].

In this construction, vibrating elements are built directly in the liquid crystal panel. Similar developments of other companies are known. Their technologies, however, consist in mounting an additional vibration film in front of the panel whereas in the «speaker display», sound is emitted by glass itself. It should be noted that the main problem was preventing image distortion during sound reproduction and the developers coped with it.

A hybrid of a *display* and a scanner was suggested by developers from Toshiba Matsushita Display Technology [5.16] who presented a device called Input Display where each pixel of the screen surface is provided with a microscopic optical sensor, i.e. a scanner and a display are trimmed into one system — a screen with sensors. If a sheet of paper with a picture is put on such a surface, optical sensors read information about image brightness at each point of the sheet and transmit it to a computer. Apply a picture to the screen to obtain its scanned image on the screen.

For the time being, such a scanner only allows black-and-white images to be displayed on a screen, but a color scanner is on the way.

Further trimming of computer's auxiliary components may be observed in a portable computer, for example, Hagaki PC (the Japanese work «hagaki» means «card»). The «digital card» is a display located directly on the housing of the world's smallest personal computer

Fig. 5.36. Hagaki computer

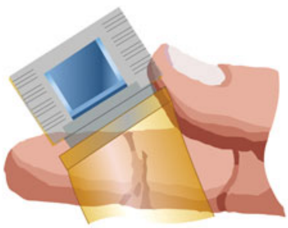


Fig. 5.37. Display card

capable of supporting desktop versions of the Windows operating system (Fig. 5.36). Hagaki computer lays claims to the intermediate niche between the notebook and the palm computer and is quite capable of eliminating the latter devices as a class [5.17].

Sharp Corporation developed a screen with a *microprocessor built directly in a transparent substrate*, which allows producing so-called display cards. This card, the size of a calling card, performs the function of a portable memory. It differs from an ordinary flash-memory by that the «display card» operates like a small computer, which may be used not only for data storage but also for data viewing (Fig. 5.37).

This mini-display may be used for work with different devices: from mobile phones to car navigation system. Such a display with a built-in microchip suits well the «wine bottle of the future» developed by the British students from the Saint Martins College of Arts and Design (England). «The self-cooling animated bottle» has a small display which offers full information about the wine it contains: a short film offers information on where the wine was grown, how it was bottled, what taste it has, etc. [5.18].

Fig. 5.38. Mono-bi-poly of flat displays

One screen is sometimes insufficient for effective work on a computer. The problem may be solved by combining *two displays* on a single base. The TopHead company offers the model TM150, a patented liquid crystal monitor equipped with two independent displays (Fig. 5.39). The main display of the double monitor is 15 inches in diagonal and produces a picture featuring high resolution and good light rendition. The second screen is auxiliary. It has less impressive characteristics: 6.4 inch diagonal, maximum resolution of 640x470 pixels at a smaller color depth [5.19].

A computer for brokers may comprise up to *nine displays*, so that it is possible to observe multiple deal-making processes.

The number of combined displays may be increased even more. This allows solving a very complicated problem — making a super-large high-definition screen. Such a screen is not easy to manufacture as a whole because of some limitations. Therefore, an apparent way is composing it from *several displays* of a smaller size. A control unit distributes an image between all the screens, but the boundaries between the screens reduce the image quality. The screens may be mounted immediately adjacent to each other but this entails the problem of matching the image parts on adjacent screens. According to the «*Mono-bi-poly*» pattern, it is necessary to *combine* all screens into a single monosystem of a higher level.

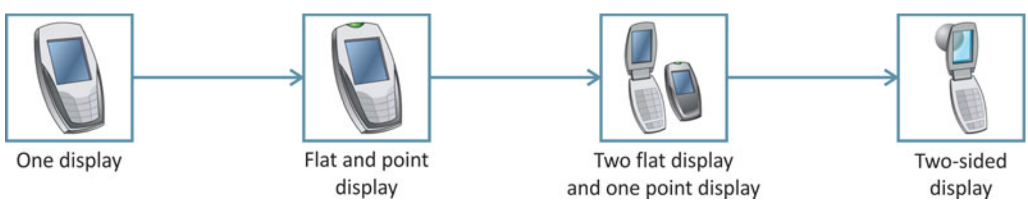
Fig. 5.39. TopHead's double display

The example of such a system is the display developed by Seamless Display. This display is composed of several screens but is capable of producing a single image using the seamless display technology (Fig. 5.40). This technology was developed at the engineering department of Oxford University. A portion of video information undergoes digital compression near the screen boundary and then special lenses transfer it through the boundary to the neighboring screen. Simultaneously, the glow brightness of all displays making the «seamless» display is matched. The image produced on the large screen does not have any visible boundaries [5.20].

According to the previous pattern, a large-size display is formed by combining a large number of light signals — pixels. According to this pattern a super-large display can be formed by combining several normal size displays into one screen.

Fig. 5.40. A «seamless» display

Fig. 5.41. Mono-bi-poly of mobile phone displays



Ease of use of mobile phones after providing them with a display considerably increased. Initially, using only *one display* placed on the phone face was a standard. It showed the number of a called subscriber or some other information. In addition, the display started to glow when ringing up a wanted party, which considerably facilitated communication.

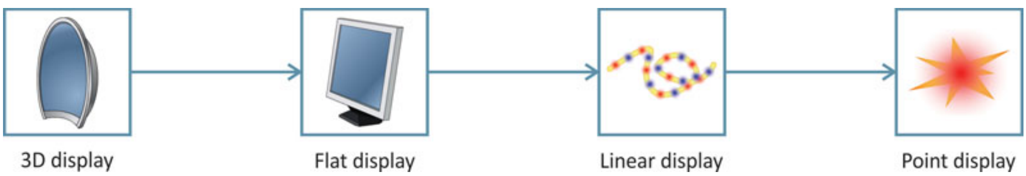
Then the main display began to be supplemented with a point display — an *additional light indicator* placed on the top side of the phone. This turned out to be very helpful when the phone was worn in a special case on the waist and the main screen was not seen well enough.

Further improvement refers to folding phones. In a folded phone, the main display was not seen; therefore, a small additional screen began to be placed on the reverse side of the face panel. The internal display is used to navigate Internet, view graphic images, work with the menu, etc. The external display is mainly used to display service information, such as date, time and phone number of a calling party.

Together with the signal indicator, such a phone has *three displays*.

The trimmed version of a poly-system (several displays of a mobile phone) is a *two-sided display*. There exist several modifications of such displays, for example, those produced by Mitsubishi. This company developed a new type of liquid crystal displays capable of displaying an image on two sides simultaneously. The two-sided display is much less expensive than two presently used one-sided displays because two displays have many common parts, such as a highlighting source. Such displays are expected to be used not only in folding mobile phones, but also in video cameras.

Fig. 5.42. «Point — line — surface — volume» transition for the flat display

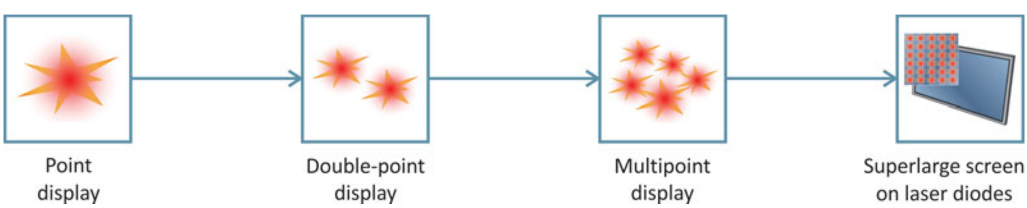


In the display evolution, we can identify a pattern corresponding to the trend of transition from a point to a line and then from a line to a surface and volume. This is a kind of expansion of geometrical elements. Tracing the same pattern for the display in the reverse order will illustrate the *trimming* of the number of elementary optical elements.

The initial step is a *3D* display such as the one shown in Fig. 5.45. The next step is a flat display and then follows a transition to a *linear* shape display. It may be a transparent tube with sequentially arranged single computer-controlled light sources. Such tubes are widely used in the creation of advertising slogans.

Next comes a *point* display — a single source of light. Such a display can convey a certain type of information by alternating the duration of flashes and pauses between them as well as by changing the signal color and brightness. A sea beacon is a characteristic example of such a display.

Fig.5.43. Mono-bi-poly of point displays



The previous pattern (reverse to the «*Point — line — surface — volume*» transition) ended with a point, an elementary optical element. The end step of this pattern gives an opportunity to transfer to another pattern — *expanding a point display* into more complex structures. Really, two-signal street lights for walkers may be considered as *two point displays.*

A display composed of *several* observed points is a display panel providing information about arrivals and departures at airports and railway stations, information in the form of a «crawler» message, etc.

It is interesting that such systems were used as early as in ancient times. For example, in ancient Korea, they used a system of fortresses with signal towers. Conventional information could be sent by lighting signaling fires in a certain order in such towers (Fig. 5.44).

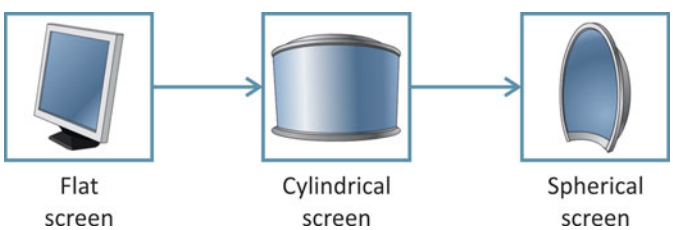
Taking into account that an observed point may be a *pixel*, screen segmentation into a large number of independent points, pixels, will remove technological restrictions that hamper production of super-large screens.

Fig.5.44. Five fires turn

An example is a screen composed of light-diodes mounted on a common panel. Individual light diodes may be used without any serious problems for composing a screen of any size.

For example, the world's largest outside high-resolution display of the Diamond Vision series manufactured by Mitsubishi Electric is mounted in Atlanta (US). The screen included in the Guinness Book of Records is 21.64 m high and 24.09 m wide. It weighs 50 tons and contains over 5 million LED cells capable of reproducing up to 1 billion different colors. The relevant aspect of this display is its modular design: the display consists of 266 integrated panels [5.21].

Fig. 5.45. Geometrical evolution of flat display screen



A *flat* surface screen may be taken as the initial version.

Perfection of the flat display by complicating the screen shape was caused by the desire to improve its operational capabilities. Subsequently the screen shape changed toward a greater

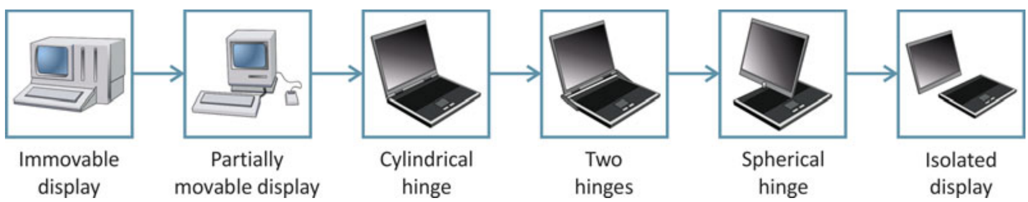
Fig. 5.46. Display as part of a sphere

complexity until it became *cylindrical*. The cylindrical screen allows the image to be observed from all sides. Image is often made moving in a circle for better visualization of information.

The next shape evolution step is a screen shaped into *part of a sphere*. This may be exemplified by a hemispherical display 170 Computer Monitor developed by the Japanese Institute NAIST (Nara Institute of Science and Technology) (Fig. 5.46).

Unlike ordinary flat monitors, such a screen literally surrounds a viewer sitting in front of it. The user does not feel himself to be just a spectator observing the events through a monitor-window, but gets an opportunity to find himself «within» the virtual world. The monitor features a high image resolution [5.22].

Fig. 5.47. Display dynamization



Dynamicity, portability is an important property of any display and a computer combined with it. There are a great number of devices where a computer itself, control tools and a display are rigidly attached to some heavy base. The examples of such a structure are computer terminals, medical diagnostic equipment, control units of processing lines, computerized inquiry desks at airports, that is, all those cases when a *stationary box* with a computer-controlled equipment has a *display*. The example of a stationary display is a refrigerator with Internet connection. Such a display may only be moved together with the refrigerator so it is needless to say about its mobility.

A traditional desktop computer segmented into several independent modules — a system block, monitor, and control tools — has a greater degree of mobility than a stationary display. The computer-forming modules may be easily turned and moved relative to one another. The monitor may be mounted on a special stand in a needed position. However, it is not easy to move the entire system to another place, though the modular construction allows moving it by parts. In this case, mobility is observed at the level of subsystems whereas the system itself is relatively stationary.

Fig. 5.48. A two-hinge display

A notebook is an even more mobile modification of a personal computer. It is not segmented into modules, all its parts form an organic whole. The notebook is light-weight and small-size.

Further dynamization of the notebook is provided by increasing the mobility of its parts, primarily the mobility of the display relative to the system block and keyboard.

For example, the display of an ordinary notebook is attached to the body by a *cylindrical hinge* and can be fixed at any intermediate position while opening or closing.

The mobility of the notebook display may be increased by using *several hinges*.

Fig. 5.49. A display on a spherical hinge

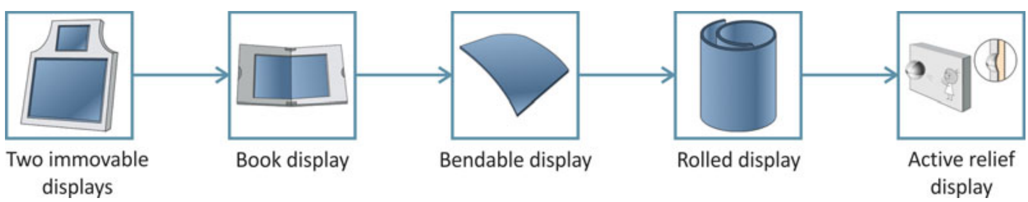
The display suggested by Artem Lebedev's design studio is attached to the notebook housing by means of two cylindrical hinges and an additional frame (Fig. 5.48). Such a display has much more degrees of freedom and may be put with its back panel onto the notebook housing and with its lower edge toward a user, which is convenient, for example, for doing graphics work [5.23].

Fig. 5.50. A display separated from a notebook

The display of Fujitsu's Life Book notebook is attached by means of a spherical hinge and has an even greater degree of mobility (Fig. 5.49). It can rotate through an angle of over 170° in a horizontal plane and can be put onto the notebook housing with its working surface up [5.24].

The maximum mobility of a display relative to a computer with a keyboard may be achieved by providing a possibility of its separation. This is very convenient, for example, for holding presentations for a small number of participants. The link between the computer and the monitor may be both wireless and by means of a thin wire (Fig. 5.50).

Fig. 5.51. Dynamization of a double-screen display



The double-screen display may evolve by increasing the screen mobility. For example, the double screens of the TopHead's display are *rigidly* mounted on a common base, which is not always convenient for users (See Fig. 5.40). The double-screen display's ease of use may be increased by making the screens movable relative to each other. For example, Samsung has developed a 6.6-inch LCD display designed for viewing electronic books. The display consists of two screens connected by a *flexible link*, which makes it look like an ordinary book. Such a display may be opened and closed like a book, which considerably increases its compactness [5.6].

The next step in dynamizing the double-screen display is increasing the degree of mobility of its screens. Matsushita has presented a dynamized double-screen display (Fig. 5.52). *Spherical hinges* attach two separate monitors to a common stand which serves as a system block. Each monitor can turn horizontally through 170 degrees and tilt back and forth within a wide range. Due to the high mobility, the monitors can be set in any position convenient for a user [5.25].

Fig. 5.52. A dynamic double-screen display

Then follows a transition to flexible constructions.

Indeed, making a screen or several screens *flexible* will allow bending them and setting any part of a screen in a desired position (Fig. 5.53). Research works in the field of display creation by non-traditional technologies have been carried out for a long time. A number of companies (Sharp, Toshiba, SEL, Kodak, Sanyo, Seiko and others) have presented prototypes of flexible displays manufactured by different technologies, including the use of organic electro-luminescence.

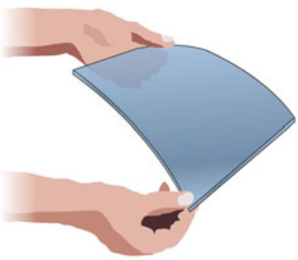


Fig. 5.53. A flexible display

Toshiba Corporation declared about the development of the world's first large flexible LCD display [5.26]. It opened the way to the creation of dynamic bendable screens and, in the long view, foldable displays. The new full-color display has an ultra fine active matrix on a low-temperature polysilicon. The flexible display opens up new vistas to designers. In addition to other merits, flexibility increases the screen stability to impact loads and the new display can be bent in all directions.

This display evolution trend allows resolving the following key contraction in the computer perfection field:

- Microchips are becoming increasingly small so compact computers of miniature size can be produced.
- However, the size of the screen and control tools may only be reduced to a certain limit because of human usability restrictions.

For the computer control tools, this contradiction is resolved by using a flexible, inflatable and virtual keyboard (Fig. 5.54) whereas the screen size remained the main restricting factor of the computer size.

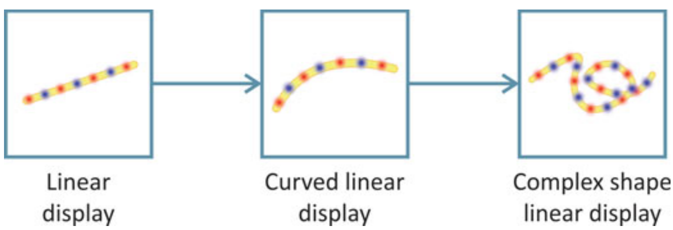
Specialists from Universal Display Corporation resolved this contradiction by creating a *super flexible display* based on thin-film (not more than 1/10,000 of an inch thick) organic light-emitting devices. The company managed to replace the traditional glass substrate with flexible and strong plastic. Such a monitor may be easily rolled into a tube like a sheet of paper. According to the company's representatives, the new technology is very promising because it is capable of solving the main problem of modern pocket computers and third generation mobile phones — that of small screens. When folded, the device with a rolled display looks like a pen. In an operating device, the screen acquires the size which is quite normal for a computer, but very unusual for a mobile phone [5.27].

Combining such a flexible screen and polymer artificial muscles makes it possible to further develop the flexible cinema screen idea obtained by analyzing cinema screen and linear display dynamization patterns. The artificial muscles allow an elastic screen to take the form corresponding to a certain image. The advantage over a flexible cinema screen is that it is unnecessary to project an image onto a screen and simultaneously solve the problem of its uniform sharpness. On the contrary, an image is generated by the screen itself. The only remaining problem is coordinating the change of its surface shape.

Fig. 5.54. A virtual keyboard

Using this idea results in fundamentally new types of displays. For example, creating a remotely controlled computer twin of a man who will participate in negotiations or sending your own copy to your friend as an original present which will look exactly like you and speak with exactly your voice.

Fig. 5.55. Geometrical evolution of the linear display



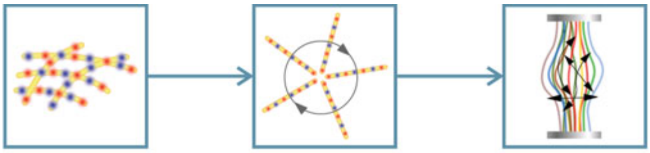
A linear display is a tube made from a conventionally rigid transparent material and enclosing point or extended light sources that can be switched on in a certain order by means of a control unit.

An absolutely rectilinear tubular display is comparatively rare. Generally, even making letters or other simple images requires a more complicated shape. The display evolves towards shape complication.

For linear constructions, transfer form a rectilinear shape to a curved shape suggests bending a straight line *in one direction*. Such a display can be easily put on a flat surface and used, for example, to manufacture advertisements.

Further complication of the linear display shape is provided by bending such a display in *two directions*. This allows full matching of the linear display shape with some *three-dimensional* base.

Fig. 5.56. Dynamization of the linear display



Fabric made of linear displays — Rotating linear displays — Movable object made of linear displays

To provide a high degree of dynamicity, it is better to manufacture linear displays from *elastic materials.*

A flexible linear display may serve as a basis for creating mobile signaling constructions. In the Netherlands, new traffic control systems are being tested where white lines of road marking are replaced with linear fiber-optic displays. It will be possible to change the «intelligent» marking depending on the road situation. Such a system will reduce traffic backups by rapidly changing the number of traffic lanes in accordance with the number of cars moving in different directions.

Flexible linear displays may be used to weave a kind of *canvas*.

Researchers from France Telecom created a prototype of a flexible screen made of multiple thin fiber-light guides which can display immovable or animated images directly on the clothes [5.28]. Flexible screens attached to fabrics may change the notion of fashion and generate a new service industry: everyone will be able to decorate his or her clothing with a drawing or inscription or to replace a pattern with a more suitable one.

The company Luminex has developed a technology for making dresses from such material. Any fabric — from Lycra to wool — is interwoven with superfine plastic optical fibers connected to a power source (Fig. 5.56). As electric current passes through such a light guide, the fabric fibers start glowing brightly forming a magic pattern.

Clothes made by the new technology will be able to serve as a computer display always available to the user. They will receive and display information from computers, mobile phones and other digital devices. In addition to applications in the fashion and entertainment industries, screens made of this fabric will be used by emergency services, furniture manufacturers, in advertising and in many other spheres.

An example of further dynamization of the flexible display is the Hokey Spokes device (Fig. 5.57) designed for enhancing bicyclists' safety. A thin tubular display is attached to the bicycle spokes forming a *rotating luminescent structure*. A small image-controlling computer is installed on the bicycle frame. As the bicycle moves and the wheel rotates, the computer switches on point light sources and produces some text or image on the side surface of the rotating wheel. The developers provided Hokey Spokes with twenty standard patterns and texts [5.29].

Similar illumination can reduce traffic accidents of bicyclists during the hours of darkness. In addition, it is an excellent means of self-expression because everyone can create new patterns and demonstrate them to everybody.

The evolution of the linear display through dynamization looks promising. Let us formulate a concept of a future flexible screen that can be obtained by combining systems having different functions. We *have a flexible screen composed of fibers and capable of taking a desired form. To produce some moving variable-color 3D structure, the flexible screen needs to be supplemented with a drive providing motion of each fiber, a kind of artificial muscle. Two types of NASA-developed artificial muscles are known.*

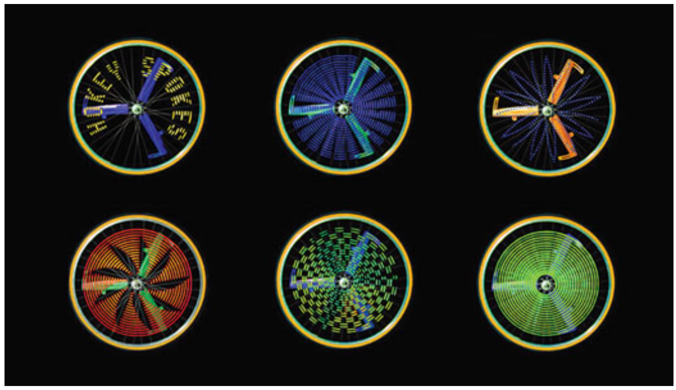


Fig. 5.57. Luminescent patterns on bicycle spokes

The first one is a polymer band composed of chains of carbon, fluorine and oxygen molecules. As an electric impulse passes through this band, the polymer molecules contract or stretch depending on polarity.

The second type of muscles consists of thin plates rolled into cylinders like tobacco leaves in cigars. Positive and negative electric charges are supplied from the cylinder ends making the plates contract. Disconnecting the power supply «relaxes» the cylinders [5.30].

It is the first type of the artificial muscle — polymer band — that suits us best. Such a band may be attached to each fiber composing the flexible screen and the bearing frame may be controlled in accordance with a preset shape. This will allow producing different moving 3D images.

Fig. 5.58. Evolution of the screen surface microrelief

Analyzing different screen versions reveals the pattern of surface micro-relief complication.

For example, the British department of 3M company has developed a liquid-crystal display having parallel vertical protrusions on its surface. The protrusions create a special filter for monitors designed for providing safety of private and confidential information. The secret of the «privacy» filter consists in the unique patented technology of *micro-jalousie* [5.31]. A person working on a computer sees a sharp, undistorted picture on the monitor. Those looking at the screen even at a small angle can see but a black screen.

Рис. 5.59. Real-virtual display (MARS)

The next step is a screen with *surface micro-roughnesses*, which features weak reflection and produces no strain-causing glares.

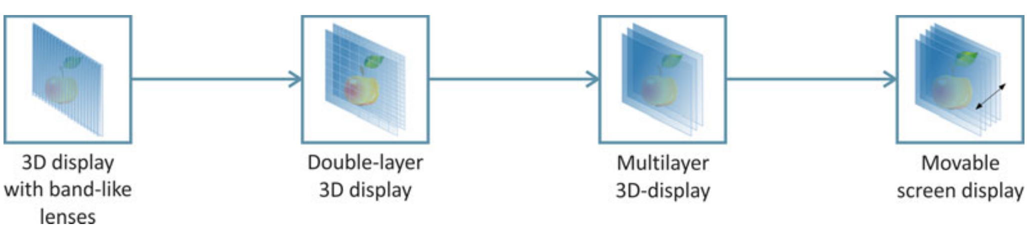
Further evolution of surface micro-relief suggests making the screen surface *active*.

This transformation may be illustrated by the development of the ColumbiaUniversity Computer Graphics and User Interfaces Laboratory. This is an active system of visualization that combines real and virtual images (Fig. 5.59). It can provide a user with comprehensive information about the structure of real world objects. The internal structure visualization system is called the Mobile Augmented Reality System (MARS) and is one of the most advanced technologies in the field of Augmented Reality devices. Using special eye-glasses with a positioning system, a computer keeps track of the user's head position, direction of his look and displays information about the objects at which the user looks [5.32].

A car mechanic will get information about a given engine model; a detective will be able to identify by-passers, which will be helpful in trailing a criminal; architects will be able to see new buildings afield in section and at different angles.

In the subsequent transformation of the screen, the properties of the display's active surface are transferred to an *auxiliary device*. The Japanese Iizuka Denki Kogyo (IDK) developed a liquid-crystal display that may be used to safely display confidential information. Information on such a display is only seen through special glasses. Those without such eye-glasses will only see a glowing white rectangle on the screen [5.33]. IDK achieved this effect by removing the polarizing filter employed in all liquid-crystal displays for cutting off polarized beams passing through a layer of liquid crystal. It is the polarizers that make image visible. To make an enhanced privacy display, the polarizing filters were moved from the screen surface to the lenses of special eye-glasses to be given only to users having the right of access to confidential information.

Fig. 5.60. Mono-bi-poly of 3D display screen layers



The emergence of three-dimensional displays resolves the long-standing conflict which constantly faces painters and filmmakers. The world around us in its essence is big, but until recently we did not have the technical means to convey this to the volumetric image. To enhance the reality illusion, the image should be made 3D. A pseudo 3D image is produced by using a number of special effects, such as the stereo effect.

The artist James Clare found a simpler approach to the problem of creating a 3D image. He invented a 3D display which is a cube composed of a thousand red light diodes [5.34]. The cube is made in the form of a multilayer lattice having light diodes located at its sites. Each light diode may be individually controlled and the cube as a whole may be used as a low-resolution 3D display for demonstrating simple models and images. The cube-display may also be connected to an audio system and used as a color-music device (Fig. 5.61).

However, a computer or TV set display needs a much higher resolution and speed. Therefore, its creators tried to take the construction of traditional displays as a model.

For example, the British department of Sharp developed a liquid-crystal display having vertical parallel *protrusions on the surface* (Fig. 5.62). The protrusions are triangular in section, which allows creating slightly differing images for the left and the right eye of an observer. In the consciousness, the images are combined and a 3D effect is produced. Such an image can only be seen at a certain viewing angle and from certain distances. A small box in the lower angle of the screen shows whether you are at a required point of space: when an observer finds himself in a necessary position, the rectangle acquires a 3D appearance.

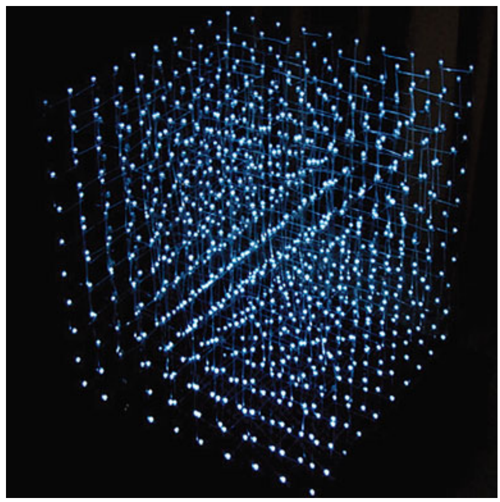


Fig. 5.61. The simplest 3D display

The DVI (Deep Video Imaging) technology uses *two physically separated layers of pixels* creating the impression of true depth (Fig. 5.68). As distinct from ordinary stereoscopic 3D-displays which require use of special eye-glasses, the DVI technology allows producing the 3D effect without using any additional devices [5.35].

Fig. 5.62. Sharp's 3D display

The next step is *using several screens*. This technology is employed by Dynamic Digital Depth (DDD). The DDD's 3D system is composed of a standard flat LCD display, several plasma displays and a long-wave optical filter developed by 4D-vision. The filter splits light into colors: red, blue and green which create a 3D image by deflecting to different directions. This technology is called OpticBOOM 3D [5.36].

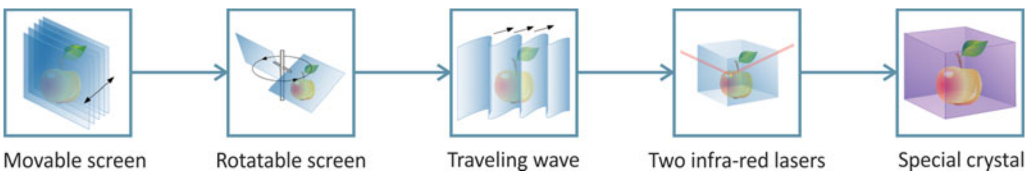
The «Mono-bi-poly» trend is apparent in the evolution of 3D displays. Increasing the number of screens-layers should end with a transition to some new quality, to a higher-level mono-system. For example, a great number of layers may be trimmed into one layer based on a different principle of operation.

The following contradiction occurs while designing multilayer 3D displays:

- If screens-layers are many, the produced image is sufficiently deep and volumetric, but preceding screens overlap following screens, image becomes blurred and the «rainbow effect» arises.
- If the screens-layers are few, the image remains good but insufficiently deep.

This contradiction may be solved by *dynamizing* the screen where an image is formed. Taking one screen and making it *move* rapidly back and forth will produce the impression that the screen is permanently found at each point of the volume. Then a luminous point or several points appearing according to some program can be projected onto the screen, thereby producing an image.

Fig. 5.63. 3D display dynamization



The *reciprocating* motion of the screen is not quite suitable for obtaining a 3D image because cyclical motion will cause huge inertia loads. *Rotating* the screen is apparently more advantageous.

The rotational motion of semitransparent screen developed by Actuality Systems served as a basis for the creation of the first truly 3D display [5.37] which is a transparent ball having 50 cm in diameter having a flat 2D screen inside. The screen rotates at about 10 rpm/sec. Sequential «cuts» of a 3D image are displayed on the screen combining and forming a single picture for an observer. A truly 3D image is produced. It may be observed from different sides as any real object in space (Fig. 5.64).

Such displays will be very useful for air traffic controllers, builders and designers as well as in situations which require visual presentation of experimental results, etc.

This idea may evolve into a flexible screen moved by the action of a traveling wave. In this situation not only a sphere-confined 3-D image may be produced but also a more habitual parallelepiped-confined one.

It is interesting that the 3D display evolution exactly repeats the way of that of the 2D display. Using a perforated rotating disk for image scanning made it possible to produce the first television picture. The color scan of the first color TV set was also mechanical; it was carried out by a rotating transparent three-color light filter.

In the 3D display, this solution is used at a higher level of evolution, i.e. a semitransparent screen is used for image scan in space.

From the evolution view point, scan should move from the macro level to the microlevel — the field level. Accordingly, the next dynamization step is using a field, for example, a light field (Fig. 5.65).

A 3D display on infra-red lasers operates in the following manner. Two invisible laser beams cross in the image-forming area. Their power is selected such that the energy of one beam is insufficient to ionize air or other gas. At the same time, the joint action of the beams at the cross point causes gas ionization and the appearance of a luminous point. Almost the same effect may be produced by *focusing* a single laser beam at a certain point of the image-forming zone. Information about research works dedicated to the creation of displays based on these principles is coming more and more frequently so practical results are expected in the near future [5.38].

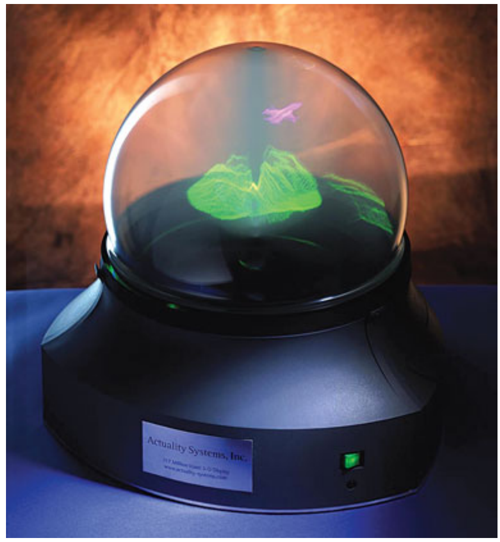


Fig. 5.64. A 3D display Perspecta

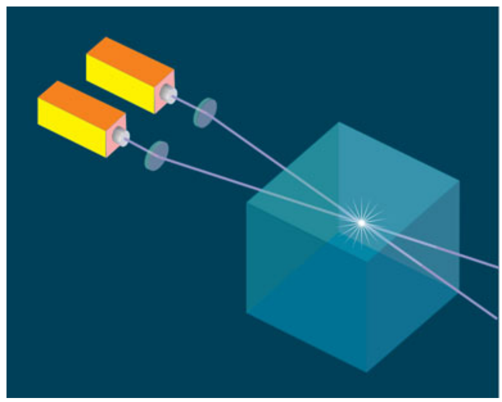


Fig. 5.65.3D image formed by laser beams

An important step in the development of 3D images was made by the research team from the University of California, Los Angeles (UCLA).

The researchers invented a device for controlling the *interaction of molecules* of a transparent crystalline material. The action of an electromagnetic field causes the crystal parts to become lighter, darker or change color practically instantaneously [5.39]. According to the researchers, these properties will allow creating 3D images of practically any shape, including moving images, deep inside the material. In addition, this material will be used in the production of high-speed optical switching circuits and data storage systems. The 3D display based on the new material will be high-speed, because the crystal parts can change brightness, color and contrast within a billionth fraction of a second. This transformation version of the 3D display is the most effective in terms of the technical system evolution laws.

5.7. Segmentation of a display as an object

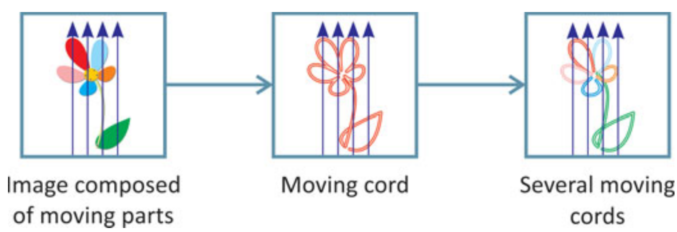
Evolution has considerably exhausted the flat display resources of further improvement. To get new resources, an abrupt qualitative change, a transition to transformation versions fundamentally differing from an ordinary display is needed.

Transformation of substances and objects according to the «*Segmentation*» pattern which now extends the trunk of our Tree offers a large number of new resources. According to this pattern, the display screen will be segmented into separate more or less large parts, then into separate granules and further into smaller particles. Preserving this trend will bring segmentation to the molecular level — the level of liquids and gases. Then follows a transition to the field level where substance interaction is replaced with field interaction. Such segmentation of an object with a desired frequency of transformations may be practically endless until the object turns into vacuum and, finally, into its special modification — a so-called ideal object.

Instead of one working tool — a display screen — segmentation gives many segmented parts, the number of which is permanently growing. This allows use of all evolution patterns, especially the dynamization pattern, providing the mobility of the «segmented display» parts relative to each other, which creates preconditions for increasing the controllability of the produced systems.

The systems obtained using the transformations of the «Segmentation of objects and substances» pattern differ from traditional systems in structure and principle of operation. Sometimes they are exotic versions but determining their position on the Evolution Tree is important for understanding the overall picture of display evolution. At first sight, some display modifications may seem impracticable, but some of their features may be later used for perfection of more realistic constructions.

Fig. 5.66. Image composed of moving elements



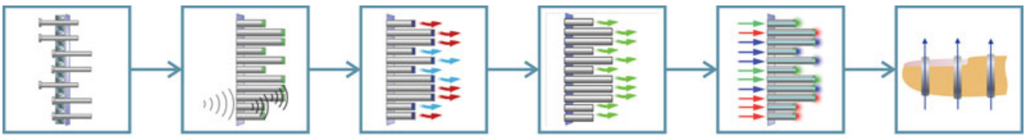
Segmenting an image *into parts* and moving them in a certain manner may produce a moving picture. This is just how the traditional puppet show, the prototype of modern animation, work. The image components may be moved in a number of ways. The puppet show uses strings, fingers and other control components. Another way is controlling the moving components by means of a magnetic field. The Figures' parts may be made from ferromagnetic material (such as plastic or cardboard filled with iron filings) and controlled by electromagnets.

The principle of creating a moving image which employs a magnetic field for moving the Figure parts may also be applied to more dynamic constructions.

For example, a Figure may be made of a *flexible cord* filled with iron filings. With a skillful control, such a Figure looks much more lively and interesting than those made of separate rigid parts. A stage manager and an artist will undoubtedly get new means of expression.

An even more dynamized image may be produced by using not a single cord, but by inventing a system of *several flexible cords* of different colors and different degrees of flexibility. The image presented on such a «multi-cord» display will be even more dynamic and visual.

Fig. 5.67. Needle-type tactile display



Allowing each part of the segmented screen to reciprocate perpendicularly to its surface produces an interesting image-forming principle. Imagine a perforated plate with equal-size stems freely moving through the holes. Applying a hand to the stem ends on the backside will immediately cause an image to appear on the face side.

This funny toy was one of the prototypes of the first TV screen prior to the appearance of CRT. Such a display had the following construction. Stems were actuated by electromagnets. The tail part of each stem was the electromagnet core. Supplying a voltage to the electromagnet pushed the stem out so that it protruded from the plate surface. Changing the polarity returned the stem to its initial place. Controlling the voltage could produce a moving picture (Fig. 5.68).

It must have been a strange TV set: dimply iridescent moving metal bristle creating slowly changing protruding pictures to the accompaniment of the annoying clicks of relay switches. No wonder that immediately after the optical television had been invented, the needle screen was forgotten. It seems not quite just, because there are blind and weak-sighted people in the world, who can perceive visual information through tactile sensations — through finger cushions or skin surface. Such people cannot watch TV or work on an ordinary computer. For such people, a display producing a tactile image would be very useful.

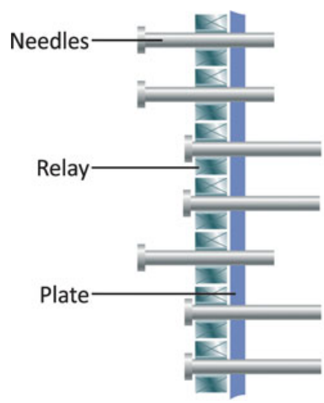


Fig. 5.68. A needle display

The information component of the world is much less accessible to weak-sighted people than to those sighted so each new opportunity to obtain information is very important. One of the disadvantages of the tactile display with relay-actuated needles is the great distance between the neighboring needles, which cannot be reduced because of the necessity to mount control relays. In addition, multiple relays make the display less reliable, more complex and expensive.

Fig. 5.69. Needle displays

The display operation will improve if the needles extend and shorten themselves. This may be achieved by using the piezoelectric effect. Needles or their tail portions can be made from piezoceramic and controlled by means of electric signals. A more advanced solution is actuating the needles by using electroactive polymer developed by NASA for artificial muscles. Such a display may have a sufficient speed of operation to suit modern computers and TV systems. Densely arranged needles will provide a high-resolution relief image and high-quality information transfer. Additional advantages are noiseless operation and a simple design.

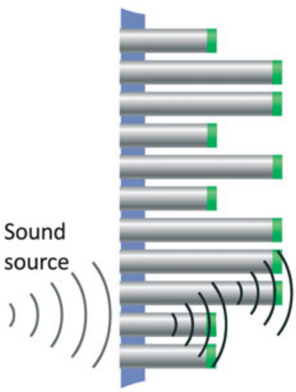


Fig. 5.70. A needle display with local sound

The example of such a display is a computer display for the blind developed by the specialists of the US National Institute of Standards and Technologies (NIST). The monitor surface is provided with a great number of small needles (Fig. 5.69). Each needle can lower and protrude by a millimeter from the array, thereby creating tactile texts and pictures. Properly speaking, it is a «tactile display» that can be used by a blind person for literally touching information, including graphic information [5.40].

What are the ways of improving the tactile display?

Here are some ideas.

This may be done by using TRIZ tools, for example, the MATChEM operator which provides for building improved models of the device by sequentially using the basic fields — mechanical, acoustic, thermal, chemical, electrical and magnetic.

According to the MATChEM operator, acoustic field may be used jointly with a tactile display (Fig. 5.70). Ordinary television and computer displays generally have remote loudspeakers so that sound is always emitted from the same point. It should be taken into account that, as a new sound appears, a sighted person easily determines to which character it belongs just by looking at the screen. A weak-sighted person does not have this opportunity. It therefore looks efficient to supply sound to that part of the screen where a new character or other important information appears. This will facilitate understanding by a weak-sighted person of what is happening on the screen.

Analyzing the expediency of using a thermal field proves a possibility in principle to develop a kind of color tactile display (Fig. 5.71). The fact is that every color is perceived by a man as having a certain degree of heat. There are warm colors such as red and yellow and there are cold colors such as blue and violet. Because a blind person cannot see the color, let us try to «draw» a pseudo-color picture by rendering colors through the temperature of different sections of the screen. For this purpose, let us place a miniature Peltier element at the end of each needle and supply voltage of different polarity to it in accordance with the events on the screen. The needle ends will alternately become hot and cold and, in addition to the screen relief, a blind person will perceive a «color» image of the events occurring on the screen depending on the heating degree.*

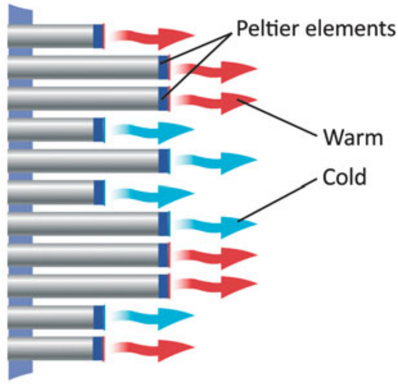


Fig. 5.71. A «color» display for the blind

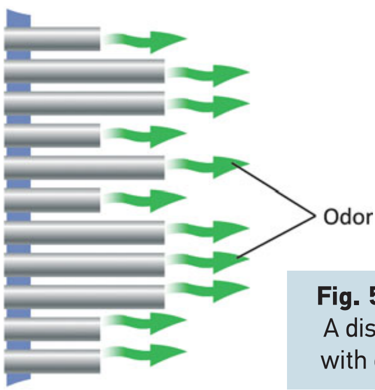


Fig. 5.71. A display with odor

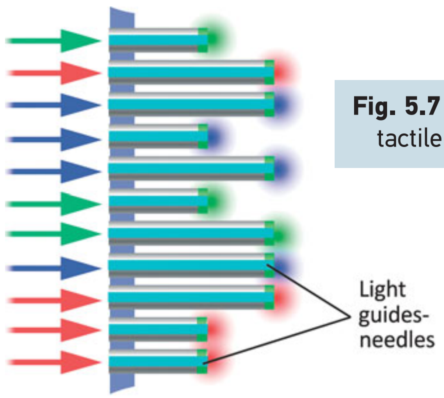


Fig. 5.71. A color tactile display

The next field that can be used in a needle display is chemical. One of such methods of information influence is using chemical substances for combining various odors. In the traditional cinematography, adding an odor to the image is an exotica rather than a need whereas for blind people every information perception channel is greatly valued. Introducing an odor in addition to the image will be particularly effective for the needle display (Fig. 5.69).

This type of display may also be improved by using an electrical field. The needle screen produces a relief, 3D image. Supplying a light signal of a respective color to the end of each needle is enough for producing a low-relief image. Such a display does not have any additional advantages for weak-sighted people but offers an opportunity for creating a relief display for people having normal sight.

In addition, a large stage opens here for research; it's likely that electric signals of a specific intensity and frequency can act on the skin and provide additional information.

Then, according to the MATChEM operator, follows a magnetic field.

It may be used to resolve the following contradiction. A blind person can promptly catch what is happening on the screen if the screen is small in size and can be fully encompassed by finger-cushions at a time.

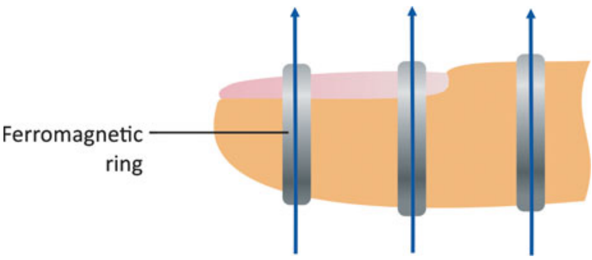
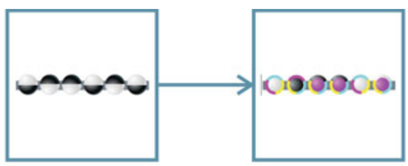


Fig. 5.72. «Magnetic guide»

For better informativity, the screen should be large in size. In this case, however, it is not easy to determine in which part of the screen the events are set and where finger-cushions should be applied.

A solution may be found by building controlled electromagnets in the screen substrate and putting ferromagnetic rings on the user's fingers (Fig. 5.72). Now, switching on the electromagnet in the active zone will «attract» the user's fingers to that zone and allow the user to promptly follow the situation on the screen.

Fig. 5.73. Electronic paper



The electronic paper (Fig. 5.73) developed by Xerox at its Gyricon Media department (Palo Alto, USA) may serve as an example of a segmented display screen. It is a rubber-like material that includes millions of plastic balls, the halves of which are painted *black* and *white*. A sheet of such material contains lubricant in its micropores so that an interlayer is formed around each ball allowing it to rotate due to the action of an electric charge (Fig. 5.75). A wireless transmitter gives a signal to each ball communicating which side — white or black — to show to a viewer. As a result a black-and-white image is produced on the electronic paper surface [5.41].

The principle underlying the creation of electronic paper is very promising so it may be expected that the next step will be increasing at least to five the number of positions at which a ball can stop: the polygraphic trio of purple, yellow and cyan (Fig. 5.76) may come to the aid of the black-and-white. This will allow a higher-quality color image to be produced on the electronic paper.

A powder display as a search object

According to the «Segmentation» pattern, the next transformation version of the system will be fine particles, powder. Making an image from powder is possible using a magnetic field (remember the school experiment with ferromagnetic filings indicating the magnetic field lines). We may suggest that using electrostatics and a large number of electrodes arranged on a flat surface will make particles form a preset image. However, changing such images is problematic.

When filling this transformation version of the «Segmentation» pattern, an interesting situation arose. In addition to the above-described «powder display», no other versions of such displays were found. Then, following the basic Evolution Tree, we built a set of search words according to the «name of the object under analysis + «transformation name» pattern, which, in our case, meant «display» + «powder». As a result, Internet search returned information on the promising Bridgestone's «powder display» [5.42].

The found power display has the following structure: particles of impalpable powder are floating between two panels one of which has a white surface and the other is transparent (Electro Liquid Powder). The size of the particles produced by nanotechnologies is extremely small and the scattered powder allows the white substrate of the rear panel to be well seen (Fig 5.77). Switching on the electrodes placed on the panel causes the power particles to gather on the transparent panel surface and form a distinctly seen black or color point. A similar principle is used in the «electronic paper» (See the previous page); the nano-powder display only differs from it by the design simplicity and high response speed.

According to Bridgestone, the new technology has considerable advantages over the traditional LC technology. Firstly, due to the fast passage of current, the «liquid powder panel is approximately 100 times less inertial that the analogous LC panel. Secondly, to operate, the new display does not require thin-film transistors irreplaceable in traditional liquid crystal displays. And finally, the new material can

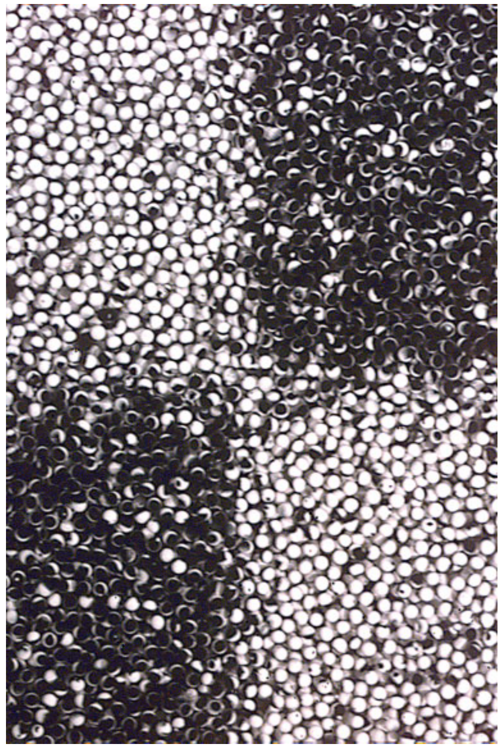


Fig. 5.74. Image composed of balls

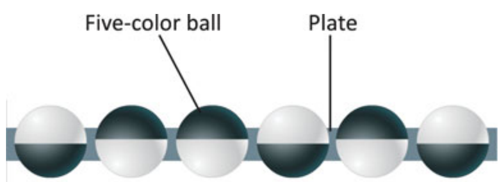


Fig. 5.75. Microball electronic paper

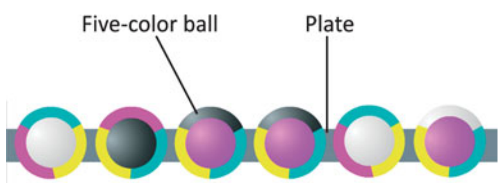


Fig. 5.76. Electronic paper with a color image

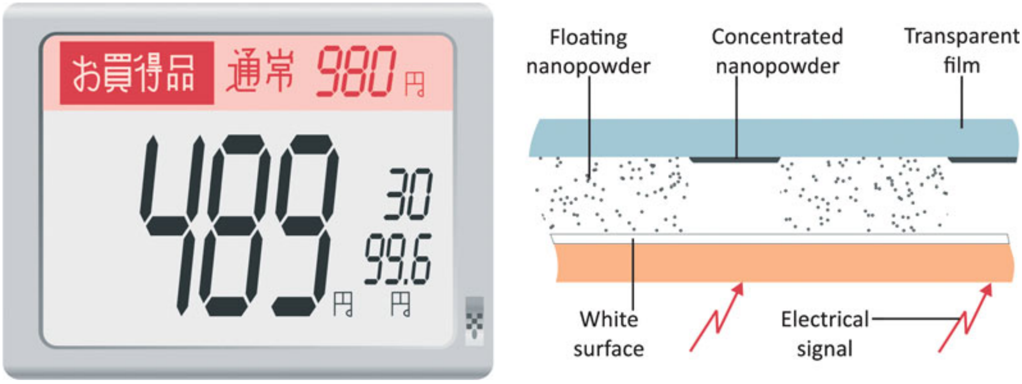
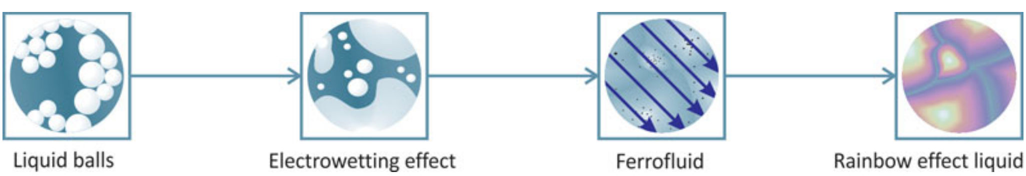


Fig. 5.77. Bridgestone's powder display

reflect up to 45% of incident beams, thereby providing a high image sharpness. Eventually, according to Bridgestone, the new display will consume 5 times less energy and will have a lower manufacturing cost. Bridgestone is starting mass production of nanopowder displays and is planning to use them for electronic price tags organized into a common computer network of a supermarket.

The Internet search experiment illustrates one more advantage of the Evolution Tree. Some of its steps give an impetus to perform a search using key words which are usually difficult to combine. Finding this information without using the seemingly incompatible worlds «display» and «powder» would have been difficult. It is also difficult to find information about the powder display by the company's name, because Bridgestone's main field of activities is not the development of displays but tire manufacture.

Fig. 5.78. Dynamization of the «liquid» display



The «fine particles, powder» step of the «Segmentation of objects and substances» pattern is followed by the «liquid» step. There are two ways of using liquid in a display: the first one is generating an image directly in liquid by locally changing its properties and the second one is using liquid as one of the system's components. A liquid display may be produced by applying a layer of special liquid onto a flat panel and providing control of the optical properties of this liquid.

A simple way of making the liquid layer on the screen surface dynamic, movable is using a liquid having high surface tension and mercury-like properties. Such a liquid easily forms *balls of different sizes*. Controlling these balls will make it possible to produce an image, composed of these balls, on a flat panel.

Philips research laboratory is presently working on the realization of this principle [5.43]. The researchers have created a basically new type of *electronic paper* featuring a high response speed: tiny points arranged on the paper in columns and lines can change color within a hundredth of a second. This is quite enough for demonstrating a video image using arrays of such points. The proposed technology is based on the electro-wetting effect.

Each pixel of such electronic paper consists of *droplets of color ink* on oily base arranged against a white reflecting background. There is a transparent electrode placed on the reverse side of the substrate layer; over the electrode, there is a layer of water-repelling plastic. When in a normal state, the ink droplets spread over the entire cell surface so that the white background is not seen and the sheet looks uniformly black. Supplying an electric potential to the electrode causes the ink to form droplets just as water does on a Teflon frying pan. As a result, the greater part of the background surface becomes visible and its color changes from black to white. Changing the voltage at the electrode can produce all gradations of the grey color. The main advantage of this display consists in that a relatively low potential is needed to «switch» a cell which results in a low summary energy consumption.

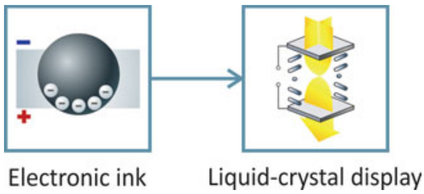
The next version may be creation of a relief image by means of a magnetic field and *ferrofluid*. Controlling the liquid surface shape with a magnetic field makes it possible to generate different images. Making this liquid semitransparent and placing a trans-illuminating flat light source behind it can produce additional light effects, because the liquid layer will be thinner in the lower places of the relief and, respectively, the light will be brighter.

A similar effect is used for producing high-quality multicolor images on large screens of projection TV sets. Such TV sets are installed in the Russian Mission Control Center. Placing a thin oil film carrying DC voltage between a light source and a screen produces a film light modulator. The action of an electron beam causes a charge on the film. It deforms the film surface so that the projected image becomes relief. The light of a powerful lamp is cast on the projected image; the light reflected from the rough surface of the oil film carries the relief image which is focused by a lens and projected onto the screen. A high-quality image is generated [5.44].

Using the rainbow effect seems to be the most promising method of generating an image in liquid and pseudoliquid films. School textbooks usually illustrate it with the rainbows appearing when oil or petrol is spilt in a puddle. A thin film formed on the surface starts to work as a light filter reflecting light within a narrow spectrum range. The wavelength, i.e. the color of this light, depends on the film thickness. The petrol film thickness is not constant so the reflected light color also changes. The same principle underlies the wing coloring of some kinds of butterflies whose colors amaze with their purity and brightness. The same is with soap bubbles colored all colors of the rainbow when their wall becomes very thin.

If we managed to obtain a very thing and strong liquid film on the screen surface and change its thickness pixel by pixel in accordance with the required image, we could speak about a new image-creating principle. It is quite possible that works aimed at using liquid for producing a controlled rainbow effect are already in progress and soon we will hear about some achievements. Anyway, mechanical displays employing the rainbow effect have already appeared. They will be considered below.

Fig. 5.79. Segmentation of image-producing elements



E-Ink Corporation has developed an interesting modification of a display which uses liquid for producing an image. It is called *electronic ink* [5.45]. Electronic ink (Fig. 5.81) consists of multiple transparent microcapsules arranged between two sheets of plastic: front: transparent, and bottom, grey in color. The microcapsules are filled with transparent liquid with the finest balls of white and black color suspended in the liquid. The white balls are positively charged and the black ones are negatively charged. Supplying a positive electric charge to a microcapsule will cause the white particles to concentrate in the upper part of the capsule and a white point will appear on the screen. A negative charge causes the particles to change places and a visible black point will form on the screen. The color display is created in a similar way using balls of three or more colors.

Fig. 5.80. Display on electronic ink

The next transformation step according to the segmentation pattern corresponds to the *liquid-crystal display* which is currently very popular. The liquid-crystal display has the following structure (Fig. 5.82). The LCD screen is an array of small sections, pixels, formed of three parts (subpixels) — red, green and blue. An optical cell controlled by a separate transistor corresponds to each subpixel.

The main component of this cell is a layer of liquid crystal between two transparent panels with intercrossing grooves. On contact with the grooves, the liquid crystal molecules are arranged

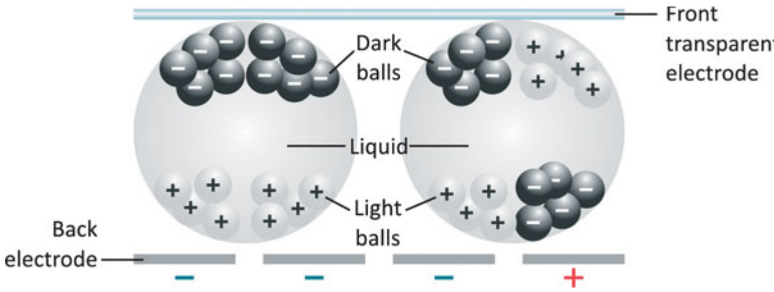


Fig. 5.81. Principle of operation of electronic ink

in a spiral between the panels. At the top and at the bottom, two more polarization filters are added to the transparent filters. The slots of the filters are parallel to the panel grooves. There is a source of distributed light installed behind the panels and polarizers.

The cell operation is based on the ability of the molecular spirals of the liquid crystal to rotate the light polarization plane. When the polarized light penetrates the cell, its polarization plane rotates in accordance with the molecule orientation so that it is turned through 90° at the cell exit. The liquid crystal filter transmits light through the cell. Voltage supplied to each optical cell can vary within a wide range and makes the liquid crystal in each cell turn through different angled, which determines the amount of light passing through the cell up to its full absorption. In this case a subpixel becomes black. At the rear, behind the «sandwich» of the panels and polarizers, a distributed source of light and color filters (red, blue or green) are set for each cell. This makes it possible to reproduce the image on the screen of any color, controllably altering the brightness of the cells.

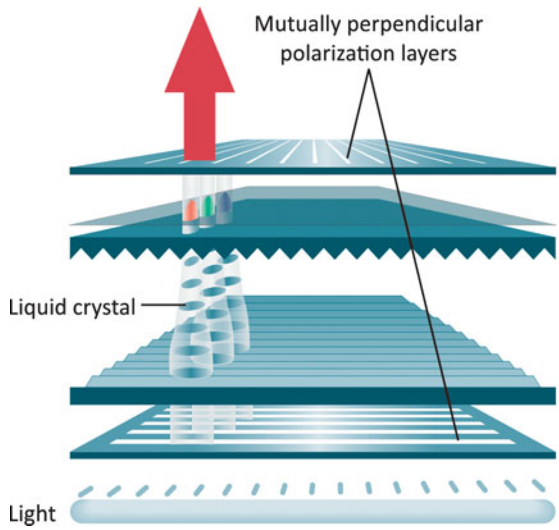
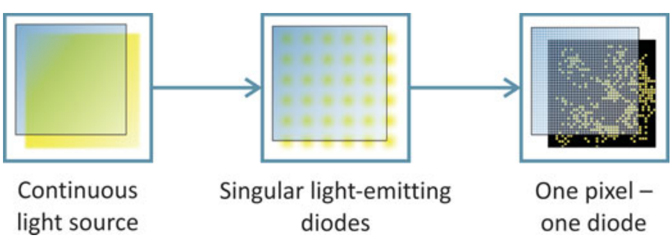


Fig. 5.82. Liquid crystal display

Because each pixel is formed of three subpixels — red, green and blue — it becomes possible to reproduce any color on the screen by changing the brightness of obtained colors.

Fig. 5.83. Segmentation of the distributed light source



A series of tubular luminescent lamps and a reflector-scatterer are used as a light source in the liquid-crystal display. The light source emits light uniformly and illuminates all pixels indiscriminately, irrespective of whether each given pixel should be dark or light while generating an image on the screen. This wastes energy. At the same time, there is no need to illuminate dark portions of the screen. This disadvantage was taken into account while designing the DHR display developed by Sunnybrook Technologies jointly with the researchers from the University of British Columbia and the University of York [5.46].

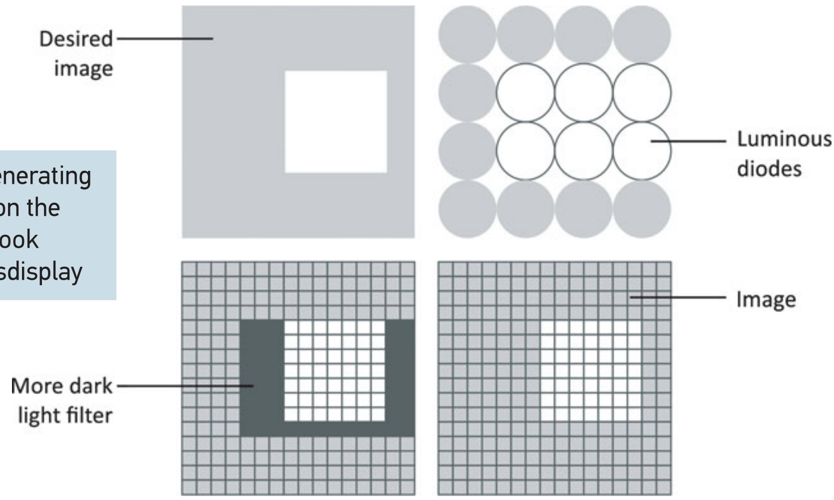
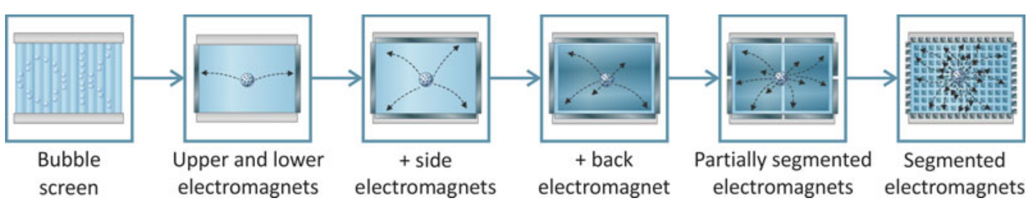


Fig. 5.84. Generating an image on the Sunnybrook Technologiesdisplay

Instead of a continuous light source used in ordinary LCD, the DHR display comprises a panel having a grating of extremely powerful white light-emitting diodes created on the basis of organic light-emitting substances (OLED) (Fig. 5.84). Due to this, the light flux is concentrated at necessary places and a very bright image is created. As for the dark parts of the screen, light-emitting diodes are disconnected there, which gives a greater depth of the black color. As a result, the brightness range in different sections of the screen is very wide. The DHR display can generate an image that is 30 times brighter and 10 times darker than that produced by ordinary displays.

The next step may be additional segmentation of the light source to the «one pixel — one light source» level. In this case, the coordination of the light flux with the display image considerably increases, which substantially reduces energy consumption of the display.

Fig.5.85. The bubble display



According to the «Segmentation of objects and substances» pattern, liquid is replaced with foam — a mixture of liquid and gas or liquids of different density.

A group of scientists from Pittsburg who developed *the bubble display* called it the Information Percolator [5.47]. This display consists of 32 transparent tubes arranged in a row in a transparent box filled with liquid, such as water. The tubes are open at the top while at the bottom each of them is connected to a pump. Pixels are air bubbles supplied under control of a microcontroller connected to a computer. The computer coordinates the operation of individual pumps, thereby forming an image from bubbles rising in the tubes (Fig. 5.86). The information presented by the bubble may be different. The authors suggest several application fields for their brainchild: a clock-face, a signaling device, in advertising. Finally, the bubble display may serve as a piece of furniture, a kind of original fountain.

Fig. 5.86. The bubble display

The bubble display's functionality is strongly limited because bubbles move in the transparent tubes vertically up due to the action of the buoyancy force. The analysis of the Tree branches proves that increase in the bubble motion controllability may be achieved by using a magnetic field jointly with ferromagnetic particles, by analogy with the «Segmented monolith» and «Fine particles» transformations.

This bubble display modification can look as follows: not air but a viscous liquid having a lower density than water, a suspension of *ferromagnetic particles*, for example, in oil is supplied into the transparent tubes. Arranging electromagnets at the top and at the bottom of the display will allow the speed of rise of the ferromagnetic suspension balls to be controlled.

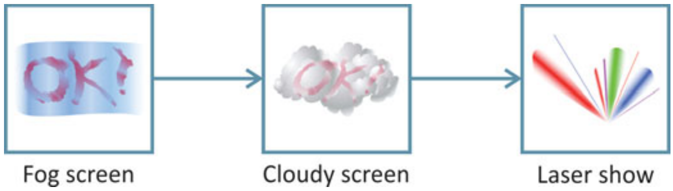
The control of the suspension ball motion may be enhanced by introducing *two side electromagnets* in addition to the upper and lower ones. In this case, the transparent tubes may be removed from the system and the suspension balls will be able to deflect from the vertical and move along a complex trajectory.

Introducing *one more electromagnet* and arranging it on the back wall of the water-filled box can stop the suspension balls at any point of the screen and press them against the back wall of the box, make them blur and increase in size.

The ball control effectiveness may be increased by segmenting the electromagnets, for example, by segmenting each of them *into two parts.* In this case, the electromagnets may be switched in turn, which provides a fuller compliance of the suspension ball motion with the designer's idea.

The final step of this pattern suggests segmentation of all the electromagnets into parts approximately equal in size to the suspension balls (Fig. 5.103). In this case, the efficiency of the ball motion control will be high and the system itself will be coordinated well enough.

Fig. 5.87. Geometrical evolution of the «foggy screen»



Original types of screens may be produced by making substance molecules more mobile, dynamic. For example, the mobility of liquid molecules increases gradually due to segmentation to fog or vapor which is a mixture of air with liquid molecules.

A «*foggy screen*» sounds strange but it is a real development (Fig. 5.88) of researchers from the Technological Institute of Tampere (Finland) [5.48]. Such a screen (they call it virtual) is created using a bar of carbonic acid (dry ice) placed in a special chamber at the lower portion of the screen. A compressor pumps air into the chamber. The air catches up the evaporated carbonic acid and pushes it through a special filter which equalizes the pressure and rate differences and produces a uniform laminar flow of white fog having a sufficient density to reflect light. This is just the fog on which any image may be projected. On this screen, a semitransparent picture literally hangs in the air.

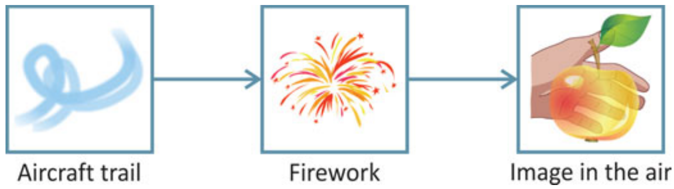
The developers think that their invention may be used for creating billboards and «virtual rooms» where the visitor will be able to pass through the walls of fog.

Fig.5.88. Foggy screen

The flat foggy screen may be followed by a relief screen. It is not easy to obtain a *relief surface* matching an image superimposed on it. One of the solutions is placing a projector at a large distance from a screen. This may be illustrated by the attempts to show films by projecting an image on the clouds by means of a powerful projector situated on the earth. In addition, the flat foggy screen may be used to create a *3D image*. The researchers from the Technological Institute of Tampere also work on this project. Using several computer-controlled compressor guns, they intend to create 3D virtual images from fog. It would be appropriate to remind here about the already existing 3D images projected on fog. We mean popular laser shows where interesting effects are produced due to a well-thought-out arrangement of lasers and generation of clouds from fog in preset places, even over the audience. Combining these two methods allow stage managers to achieve terrific effects.

Further segmentation of substance suggests transfer from vapor, fog to gas. The most easily available gas is air. This transformation includes all cases of generating an image directly from air. Most events observed by a man occur in air; therefore, air may be considered a universal display.

Fig. 5.89. Image activation in a gas medium



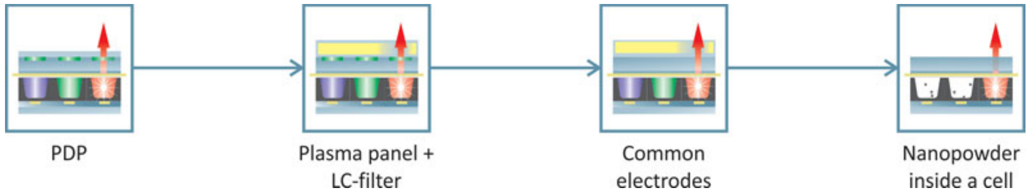
Contrails made by the exhaust of the aircraft engine or trails of special color smokes used during air shows may be considered the first step of this pattern. Using colored smoke, experienced pilots can depict a flag, some inscription or other objects in the sky.

The next case of forming an image in the air is *fireworks*. Multicolored fires that draw intricate patterns in the sky may be used both for entertainment and for advertising. Every summer, a firework festival (Hanabi) is held in Japan. The best-known masters make a performance of this event, special rockets and fires are manufactured to burst into millions of lights in the dark summer sky, forming an amazing spectacle.

Fig. 5.90. Heliodisplay developed by IO2 Technology

The above «displays» have low information content and offer few opportunities for transmitting meaningful images. It is desirable to have an «air display» capable of generating a complex 3D *image simply «suspended in the air».* Such a device seems to have been created. The American company IO2 Technology formed by the graduates of the Massachusetts Technological Institute has publicly presented Heliodisplay (Fig. 5.90). A 3D image appears directly in the air over a small box. An operator places his hand into a luminous image and easily makes it move. The «screen» is created by using the local temperature difference between air layers that causes instantaneous condensation of water vapors in the air. The device transforms air within a very limited volume without adding or extracting anything out of it. In addition, the image is interactive; therefore, a hand or a finger may be used as a computer mouse cursor. The image is flat but seems 3D from a small distance, though physically it has no depth. «The viewing angle» is from 75 to 150 degrees. No special eye-glasses or screens are needed to see the image [5.49].

Fig. 5.91. Combining a plasma panel with a different type of display



The next step of the «Segmentation of objects and substances» pattern is plasma, the level at which substance consists of parts of atoms. The example of this transformation is the *plasma display.*

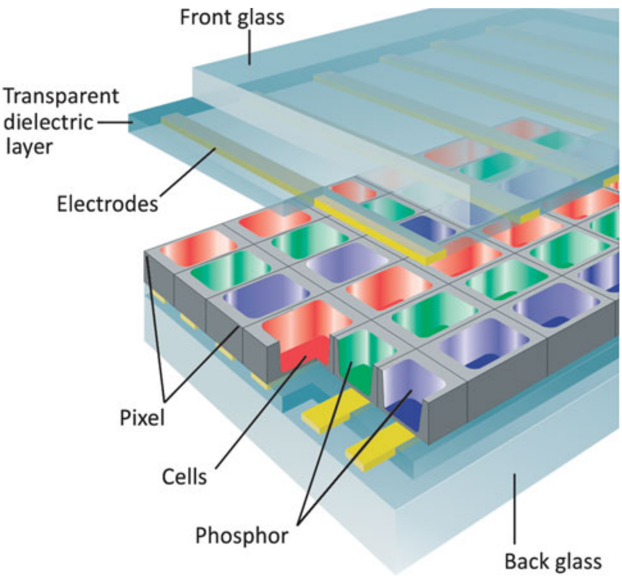


Fig. 5.92. The plasma display structure

The plasma display has the following structure (Fig. 5.92): the back wall of a panel filled with inert gas is provided with multiple miniature depressions, elementary cells. Each pixel of the display combines three elementary cells (subpixels), the walls of which are lined on the inside with phosphors of the primary colors — red, green and blue. Voltage is supplied to two transparent electrodes placed over a cell. A plasma discharge occurs between the transparent electrodes. Ultraviolet emitted in the process makes the phosphor of a given cell glow with different intensity. This is achieved by different glow duration of a corresponding cell of the plasma panel: the brightest elements glow during the entire impulse whereas in the darkest places elements are not activated at all. As a result, each pixel gives multiple shades of color and brightness, which provides high-quality color rendition.

One of the plasma display disadvantages is enhanced refection of incident light. Due to this, the panel looks insufficiently black at dark image spots. To obtain a brightness-coordinated image, it is necessary to increase the glowing brightness at light spots, which leads to energy overconsumption. There arises a contradiction: on one hand, the cell walls should properly reflect light to increase phosphor radiation, and on the other hand, they should not reflect light well to avoid causing reflection of the light incident on the screen surface.

This contradiction can be solved by using an *additional non-transparent light filter* capable of closing an idle cell (Fig.5. 93). Analyzing the display modification by means of the Evolution

Tree proves that such filters are used in «transilluminated» displays, i.e. in LCD and Bridgestone's powder displays. Thus, the combined system looks like a plasma display having a non-transparent nanopowder filter attached at the top (with the coordination of pixels). However, such a system is uneconomical because to function practically it needs two displays.

To increase efficiency, the combined system should be trimmed. The first step may be *removing the light filterelectrodes* and using the plasma panel electrodes to control the opening and closing of transparent windows in the powder light filter. To realize this idea, it is necessary to change the parameters of the electrodes and to solve some of additional problems. Realization of such transformation will reduce the combined system cost.

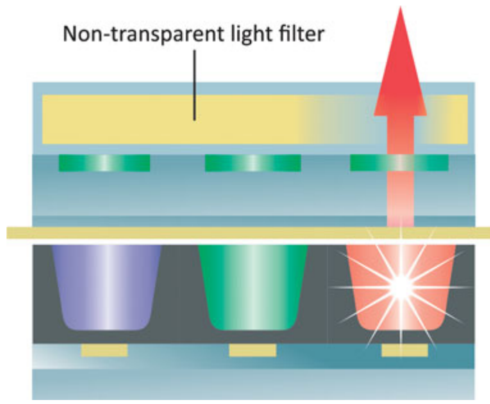


Fig. 5.93. Combination a plasma display panel with a light filter

We have two structurally similar partially trimmed systems operating in parallel, i.e. the plasma panel is a closed box the internal structure of which is filled with inert gas. The powder light filter, in its turn, is also a closed powder-filled box. To continue trimming, we can combine these two boxes and place *powder directly into the plasma panel*. When the electrodes are switched on, the powder opens the window directly over an operating pixel; when the electrodes are switched off, the window is closed. Practical realization of this idea is likely to meet with difficulties, but the produced trimmed system will combine the advantages of two different types of displays.

Fig. 5.94. Trimming a light image

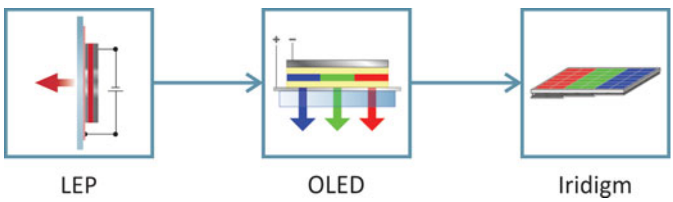
Modern displays (in the popular sense) are electronic devices operating on the field transformation principle. Generally, an electric field turns into a light field by means of varyingly complicated transformations. Therefore, it is difficult to give fully correct examples of this transformation and to build evolution patterns. So let us agree to refer to those transformation examples where the image is formed with a minimal transformation of fields.

One of the examples of field level transformations may be a shadow play where image in generated by introducing moving Figures between a light source and a screen. Cinematography is the example of trimming such a system because it generates image not by using real Figures but by placing alternating photographic images, optical copies printed on a transparent film, between a light source and a screen.

Removing a film from the display projection system makes it possible to obtain an even more trimmed system — a laser projector where an image on a screen is drawn by a laser beam. In this situation, a material image carrier is not needed, but its electronic model. A promising development is an extension piece for the Mini-beamer mobile phone developed by SIEMENS. It is a tiny, really pocket-size projector designed for holding presentations at any place [5.50].

Next within the «Trimming» pattern comes the *projection TV set* where all the units are combined in one housing and the image is projected onto a screen by means of a system of mirrors.

Fig. 5.95. Trimming of light-emitting elements



The design of displays with direct energy transformation allows electric energy transformation into visible radiated or reflected light.

The example is the display manufactured by *LEP-technology* (LEP — light-emitting polymer), the patent for which is owned by Cambridge Display Technology (Fig. 5. 96).

Polymers capable of emitting three colors are deposited onto a flexible transparent coating. Electrodes are attached to two sides of the polymer layer. When a weak current is supplied to the electrodes, the layer of light-emitting polymer emits a light spot of a specific color and such microspots are used to form an image. The display structure is simple, but the manufacturing process of the needed materials, especially of the *light-emitting polymers* themselves, is complicated and expensive.

The LEP display has some advantages over the liquid-crystal display: it is compact, simple and efficient. No glass vessel for liquid crystals, backlighting system, polarizers or color filters are needed. Polymer pixels emit light uniformly, which provides a 170° viewing angle.

Jet deposition of polymers makes feasible production of panels of the most different formats [5.51].

Then follow displays based on *organic light-emitting diodes* (OLED). They are very similar to LEP displays: a cellular structure from a thin film of organic substance with insignificant admixtures of fluorescent molecules is placed between two conductors (transparent and metal). Each pixel includes three subpixels of the primary colors. A potential difference is supplied to the conductors: electrons and «holes» recombine in the emission layer, emitting photons.

The fluorescent admixtures in the organic substance serve to produce specific colors of a light spot (Fig. 5.97). The OLED display is thin and light weight and can be easily made on glass but not on a flexible film. For example, Universal Display Corporation has announced creation of thin-film light-emitting devices not more than 1/10,000 of an inch thick. Organic light-emitting displays look very promising. They have some advantages: low energy consumption, enhanced image sharpness, high-speed response sufficient for demonstrating video films [5.52].

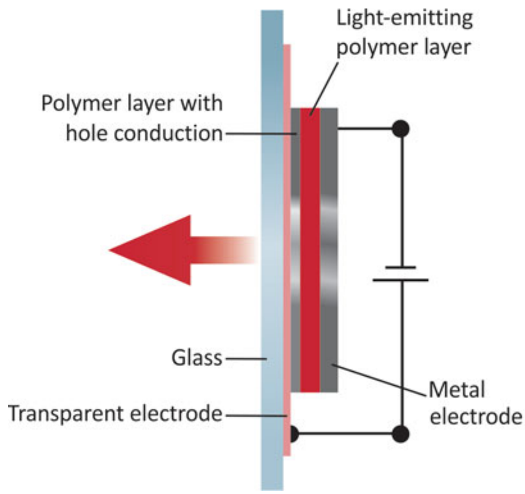


Fig. 5.96. The display on light-emitting polymers (LEP technology)

The engineering company Iridigm formed in 1996 by two graduates of the Massachusetts Technological Institutes suggested fundamentally new displays capable of competing with those existing. The new technology is based on the principle of lightwave interference which provides the rainbow coloring of soap bubbles (Fig. 5.98). To form an image, such displays employ three types of elements — according to the number of the primary colors. The elements forming the red, green and blue colors differ in the size of the clearance between the glass substrate and the metal membrane placed behind it. Each element may be in two states. Until the membrane is at a certain distance from the glass, an element reflects light of a fixed wavelength. Supplying an unlike charge to the electrode and the membrane will attract the membrane to the glass while the element will cease reflecting the light causing the cell will become white.

The main advantage of such a display is that each color cell has a built-in memory: while the membrane and the film exchange charges, the cell remains in the same state. If no changes occur on the screen, the display may consume no energy at all [5.53].

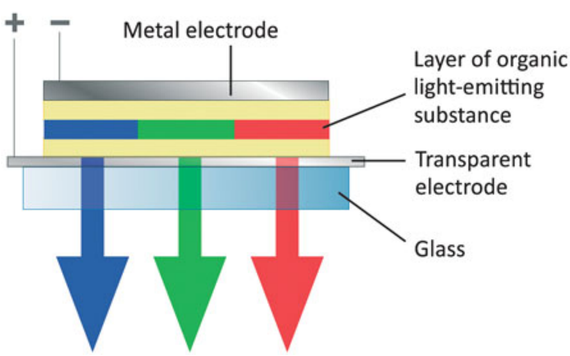


Fig. 5.97. The display based on organic light-emitting diodes

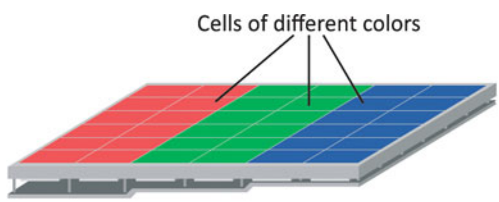
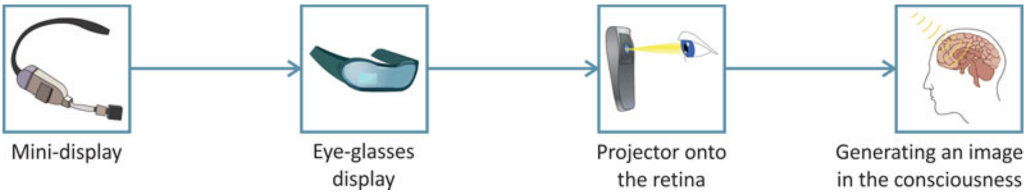


Fig. 5.98. The interference display structure

Fig. 5.99. Trimming the display to the ideal



A display may be considered ideal if it does not exist in the usual form but its function is performed. Such displays resolve one of the basic contradictions of modern electronics which have reached the state when computer devices can be miniature enough but the size of input-output devices must be convenient for a person.

Portable displays may be considered close to the ideal. First of all, they are mini-displays placed in close vicinity to eyes. The development of Mitsubishi Electric [3.63] is the example of such a display. This display stands out against the background of compact-size analogs presented on the market due to its compact size that allows a user to see not only a picture on the screen but, at the same time, the surrounding world. This became possible due to a considerable reduction in the light scattering coefficient and, as a result, the display size restriction (to 4 mm vertically and 7 mm horizontally). The image on the screen is equivalent to a 10-inch display viewed from a distance of 50 cm and the device weight is 20 g. The mini-display is planned to be used jointly with video players, pocket computers and mobile phones [5.54].

The *eye glasses display* is more ideal than the portable display because one device performs the functions of two different systems. For Formula-1 racing drivers, the engineering department of BMW has developed a miniature 6x7 mm display built in a helmet. Information needed during a race is shown on the screen of a mini-display placed directly on the protective glass of the helmet, in front of the driver's eyes. Because the image is translucent, the pilot's attention is not distracted from the situation on the race track [5.55].

Even more ideal is the display developed by Microvision. It projects a high-quality color *image directly onto the retina* [3.65]. The device is based on a micro-mechanical scanner — a small vibrating mirror controlled by an electronic circuit. Using a light beam composed of the beams of three color light-emitting diodes, a micro-mirror «draws» through a lens a picture on the retina. With this virtual screen, a viewer, using a miniature device, can see an image of a standard computer size [5.56].

The next step toward the display ideality may be a display causing visual images *directly in the human consciousness.* This method, described repeatedly in fairytales and science fiction, only seems absurd at first sight: do not we see very vivid dreams sometimes? We may draw a conclusion from this fact that causing some visual image in the human consciousness is not difficult but it is difficult to control this process.

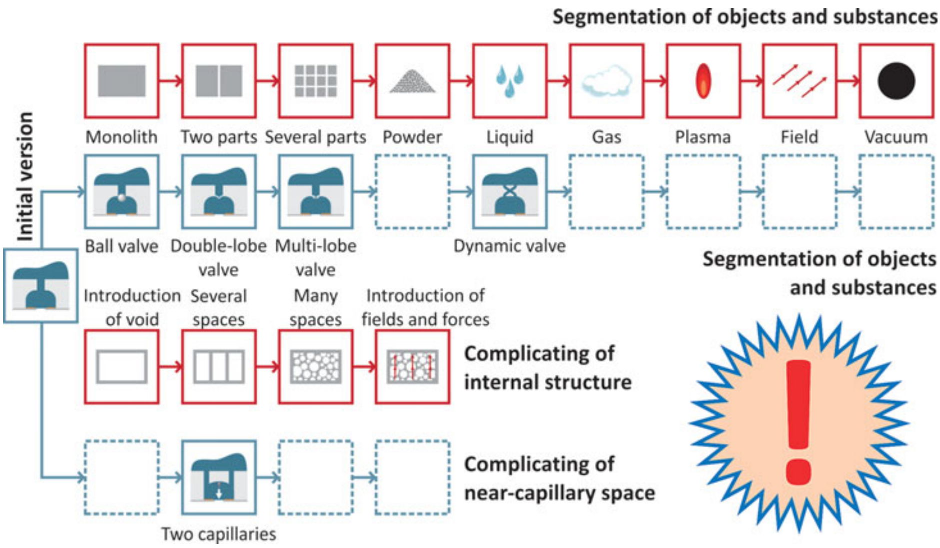
Currently, first attempts to create such a «display» are being made. For example, the Japanese Takara Co has announced the creation of a device that allows simulating a dream to the client's desire (Fig.5.123). The client must look attentively at the image of what he wants to dream and then to write down comments to the picture he likes. Then the «dream studio» will prepare an individual dream to his order [5.57]. Using voice and music, light and odors, the «dream machine» will control the dreams for several hours, until the client is woken by soft light and music. Ideally, he must remember the minutest particulars of his dreams. According to the company's employees, the device still needs improvement because it does not always succeed in generating the desired dream. However, the tendency towards the development of devices for transmitting an image directly into the consciousness is obvious.

Developers of new principles of visual information transfer will probably focus on the controllability problem. Finding a reliable method of generating a preset picture directly in the user's mind will become a breakthrough because such a display is ideal from the viewpoint of technological evolution.

Chapter 6

Application of Evolution Tree

Use of the Evolution Trees offers significant advantages at all information handling stages. This concerns both search and analysis of retrieved information as well as production of new concepts and technical solutions while circumventing patent sand forecasting technical system evolution.



6.1. Basic principles of work with the Evolution Tree

The Evolution Tree can be used to solve various problems while working on inventive projects. The purpose of any inventive project — problem solving, system evolution forecast or patent circumvention — is obtaining new, previously unknown information. The sequence of steps aimed at obtaining new information can be represented by the algorithm illustrated in Fig. 6.1 diagram.

After formulating a problem, patent and other sources are searched for relevant information. This information is mostly dedicated to the implementation versions of a problem technical system which needs improvement. In the course of searching, all system functioning nuances and all possible improvement directions are taken into account. This stage deals with the information field extension and the accumulation of knowledge about a prototype and design problem.

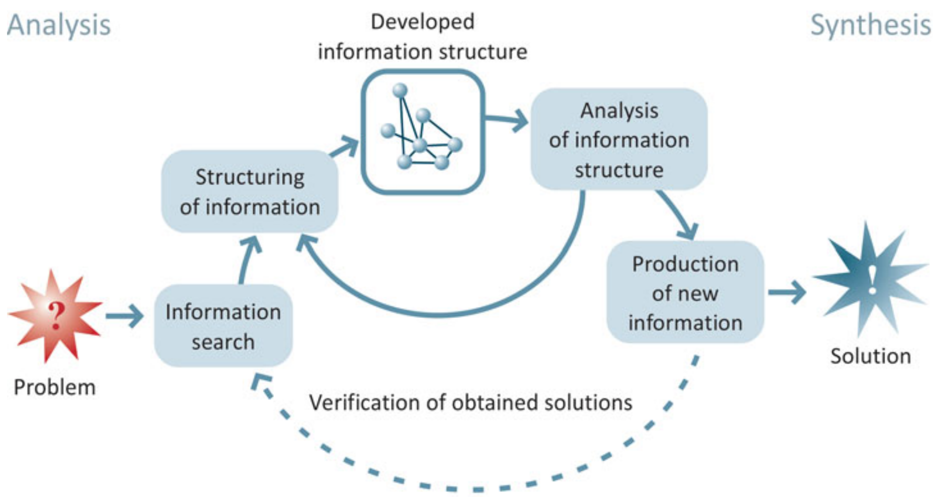


Fig. 6.1. Sequence of operations during information search

Further, unstructured information should be organized into a convenient structure for analysis. The higher the quality of the information content of a structure under analysis and the more logical its classification, the better the analysis. This will ensure successful performance of the second information processing stage — synthesis, building of new concepts. At this stage, we are primarily interested in those concepts and technical solutions which resolve contradictions available in a system and allow prototype improvement. Generation of new ideas and concepts is the key moment of the work with the Evolution Tree. This stage includes *synthesis* of new knowledge about a system under investigation. All previous actions are analytical in character and preparative for the production of new information.

Then the obtained new technical solutions are checked for novelty because some of these ideas will inevitably turn out to be known. Consequently, a new search cycle in the patent collection should be conducted using specified parameters to make sure that the obtained idea is really new and then the information structure should be specified.

Paradoxical as it sounds, in analyzing a prognostic information field, obtaining a correct data organization structure is more important than obtaining some new technical solutions. An analogy with beekeeping is appropriate here. Experienced beekeepers say it is not production of honey but production of bees that is the most important thing in their business while honey is a by-product of this process. The same is applicable to our situation: the main purpose is obtaining a full and effective information structure while new prognostic solutions — our honey — are formed as if of their own accord when the development of the information structure oversteps a certain line and the structure becomes sufficiently full and logical.

Objective and logical structuring of information units by means of the Evolution Tree considerably simplifies solver's actions at each information-processing stage. The Evolution Tree is a kind of multidimensional matrix showing which cells are filled and what versions of a technical system should be sought. If no missing information units have been found, we can start their construction.

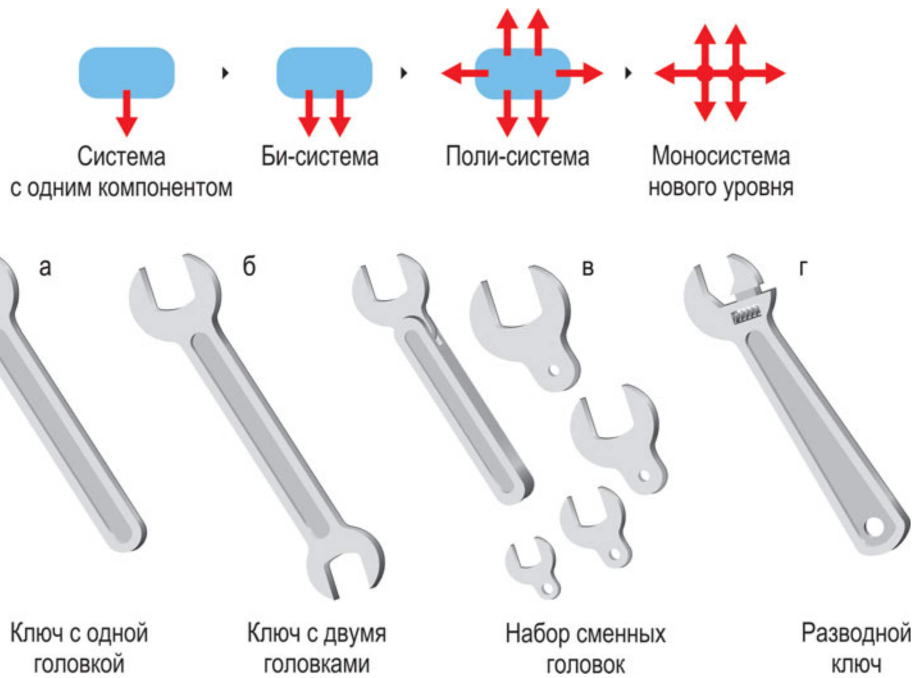


Fig. 6.2. Comparing the Evolution Trees of a basic and real object

The Evolution Tree is composed of technical system evolution patterns. Each pattern is represented at two levels — generalized which gives abstract descriptions of system versions and specific where examples of such transformation versions for a real system are shown. To analyze the Evolution Tree of a specific technical system for fullness, it is necessary to compare the evolution patterns of this Tree with the abstract evolution patterns of the basic Evolution Tree (Fig. 6.2.). Basic evolution patterns can be compared to a template, matrix. The matrix «cells» correspond to the key transformations of the patterns. These «cells» should be filled with problem solving versions. The cells which remain unfilled will suggest new solution versions.

Placing near each other an evolution pattern of a specific object and an abstract evolution pattern describing the same transformation may cause two situations.

1. *Some transformation versions (See 6.3, a) are omitted on the specific pattern.* After revealing the unfilled boxes of the specific Evolution Tree, it is necessary to make a repeated search by key words. If the search does not give any results, there is a chance that it where to find unknown, unpatented versions of the object of interest. Such gaps are the most convenient places for attacking a competing patent and searching for prognostic solutions (for more detail, see sections 6.4 and 6.5). In addition, unknown versions of a system under investigation — existing or future ones — are hiding there.
2. *Revealing the uncompleted evolution patterns of the specific Tree is equally important.* It often happens that the end parts of the patterns are not filled (Fig. 6.3, b). The most promising transformation versions which could become a basis for new, more perfect technical solutions, are there.

The proposed method of generating new information is based on that the Evolution Tree shows omitted versions of a system. In addition, it is possible to obtain a description of an omitted version built by analogy with a corresponding system version available in the basic pattern. Indeed, we already know the name of the system, the function it performs, the abstract description of its omitted version. In addition, due to the analogy with the basic pattern of the Evolution Tree we understand the evolution logic of this system. We can also see the examples of other systems from the analogous pattern of the Evolution Tree and make an attempt to transform our system by analogy.

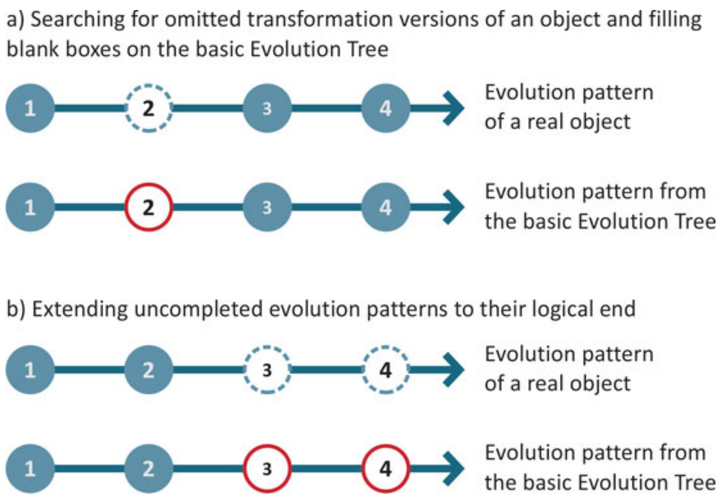


Fig. 6.3. Comparing the basic Evolution Tree and the Evolution Tree of a real object

It may be said that comparing the Evolution Tree of a real object with the basic Evolution Tree gives a kind of «prompting machine» or «generator of concepts», which allows obtaining «portraits» of omitted transformation versions of a real system. With such portraits-descriptions, it is much easier to find the way to real constructions while generating new technical solutions.

Let's use the example of analyzing line «Segmentation» of variations of an aircraft propulsion unit» to demonstrate.

Section 3.1.4 described the «Segmentation of the aircraft propulsion unit» evolution pattern (see Fig.3.9). The pattern is full enough, but some transformation versions described in the basic evolution pattern «Segmentation of objects» turned out to be uncovered [6.1.].

These versions are: segmentation to granules, segmentation to liquid and segmentation to vacuum.

We can perform a shortened analysis of the obtained through example and try to fill in blank boxes. Let us summarize the basic pattern steps and transformations obtained for the unit of propulsion in a table.

In principle, we may try to fill in the basic pattern steps straight away by generating new concepts.

Table 6.1

	Transformations from the basic evolution pattern «Segmentation of objects and substances»	Evolution pattern of the aircraft's unit of propulsion
1	Solid object	Single-blade propeller
2	Object segmented into two parts	Double-blade propeller
3	Object segmented into several parts	Multi-blade propeller
4	Object segmented into many parts	Double-row propeller
5	Granules	. . .
6	Liquid	. . .
7	Gas	Reactive jet
8	Plasma	Ion flow
9	Field	Photon flux
10	Vacuum	. . .
11	Ideal object	Glider's "unit of propulsion" (air)

For example, «Granules»

How can granules be used for creating a propellant impulse? The novel character Captain Vrungel accelerated his yacht during races by firing corks from bottles of champagne placed on the stern. A similar effect may be produced by shooting from on-board armament for accelerating or decelerating an aircraft (Fig. 6.4).

The recoil force of an aircraft gun achieves several tons (for example, the recoil of each of two Volkov-Yartsev 23 mm cannons mounted on IL-2 attack aircraft was 4600 kg). A backward-firing cannon placed in the tail is a kind of «granular» engine that gives an additional acceleration impulse to the aircraft. At the same time, active use of the forward-firing on-board armament will allow decelerating the aircraft, for example, during an emergency landing. Of course, this is hypothetical

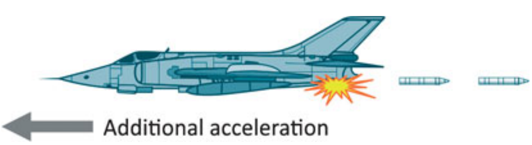


Fig. 6.4. Firing tail cannons accelerates the aircraft

reasoning but important is the principle of using small-size objects for accelerating or decelerating an aircraft.

The next blank box of the table is «Liquid»

Here we can see a possibility to use liquid together with an aircraft propeller. A small amount of water splashed behind the rotating propeller will mix with air, thereby forming fog, and the propeller thrust increases abruptly because the rejected air becomes denser. Naturally, it is too expensive to carry water all the time, but this effect may be used during a short time, for example, for accelerating an aircraft during take-off (Fig. 6.5).

The effect of creating a support for an air jet is employed on aircraft carriers where enormous efforts are made to reduce the takeoff acceleration distance. The main starting pulse is given by a steam catapult, but the aircraft engines must help it. To create a high-power starting pulse behind an aircraft ready to start, an inclined steel sheet is mounted. A reactive jet hits this sheet, is decelerated and creates a support for gas molecules escaping from the nozzle. The engine thrust increases abruptly and the aircraft actively starts takeoff acceleration (Fig. 6.6).

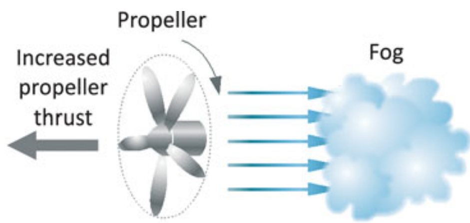


Fig. 6.5. Using water fog for increasing the propeller thrust

How can «Vacuum» be used to propel an aircraft?

Most likely, we can speak here about something tractive, arranged in front of an aircraft. For example, fuel is combusted ahead of a rocket. A low-pressure area is created which abruptly reduces air resistance (Fig 6.7).

Fig. 6.6. An aircraft is taking off from an aircraft carrier (there is an impingement plate seen behind the aircraft)

Fig. 6.7. A rocket flying in a vacuum cavity

This simple method of generating new prognostic solutions does not always work. Attempts to immediately find a good technical solution covering a blank box of the Tree often fail as is the case with the «Segmentation to granules» version of the aircraft propulsive unit evolution pattern. It is unlikely that somebody will fire all guns to decelerate or accelerate an aircraft. However, as seen from the following example, such an express analysis can prove very useful in practical problem solving.

Let us consider the Evolution Tree application at each information-processing stage in more detail.

6.2. Search by marking the information field

With traditional information search methods, interaction of a designer with the information collection may be shown as in (Fig. 6.8). Search is obviously most effective when the area the person searches perfectly overlaps the area where the information exists. In actual practice, full overlapping of these areas is unlikely: relevant information may be randomly «dispersed» throughout the entire information field.

The traditional technical and patent information search system has a serious disadvantage: search generally starts randomly, from an arbitrary benchmark. Then, as information is obtained, new possibilities for correcting the direction and widening the search field are opened to a designer. In other words, initially, the information field looks homogeneous and the information units it contains seem equivalent. And only after the information field structure has been thoroughly studied, the most promising zones accumulating the most valuable information are identified.

Such introduction of query areas into the area of needed information may be called *frontal*. This method is effective where automatic search programs are used. These programs may rapidly «turn over» the entire search field and procure all available information which is close to a preset topic. However, this leads to the information noise problem because much information unrelated to the query is retrieved per one unit of relevant information. Of course, search systems are being perfected, the ‘noise» is being reduced, but the ideal is still far away.

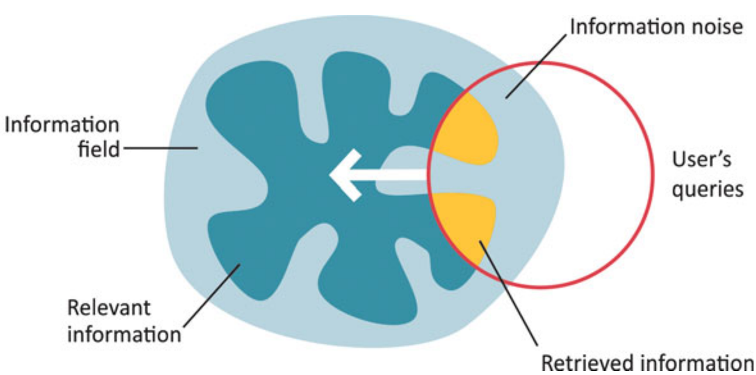


Fig. 6.8. Frontal information search

We think that the information search effectiveness may be considerably increased by determining in advance the areas where the needed information is accumulated. It would be useful to have a kind of map indicating the points of the most probable «occurrence» of data needed to designers [6.1]. In other words, instead of the frontal «attack» on the information field and consistent «grinding» of all information, search starts from pre-marked key areas*. In patent search, the task is facilitated by that technical and patent information is more strictly structured than other information areas.

Using the basic Evolution Tree makes it possible to determine the areas of the needed information concentration. These areas are found in those places where the basic versions of the system under analysis are situated and they are specified as its versions suggested by the basic Evolution Tree are found. Such structuring also allows determining «entrance points», i.e. the most typical versions of system transformations suitable for starting a search. The search starts just from these points, the «area of the known» widens, gradually occupying more and more sections of the information field. Links are established between individual search zones, some of the zones merge. Based on the already studied key areas, a kind of logical skeleton of the information body is formed.

After finding (or generating) a new information unit (a version of a system under study), we can easily determine its place in the Evolution Tree structure. This is often accompanied by the appearance of new entrance points, which allows the search direction to be corrected. As distinct from *frontal search*, this type of search may be called structural (Fig. 6.9). Frontal search is easier to perform because it does not require any thinking operations: just search information units one by one and see whether the found one is similar to that sought. No wonder that such a routine search is performed by machines. Structural search requires serious thinking effort, but the result is incomparably more effective. This approach does not exclude use of computer search; on the contrary, it improves this search.

The basic Evolution Tree allows a designer to obtain sets of key words for search. Each such set comprises two parts — the name of an object and the name of transformation. When added to the object name, a characteristic definition describing the essence of the technical system transformation in the basic Evolution Tree can significantly simplify search for necessary information.

Such form of a query corresponds to the organization principle of patent search systems where it is necessary to indicate the object name and give definition of its important characteristic property.

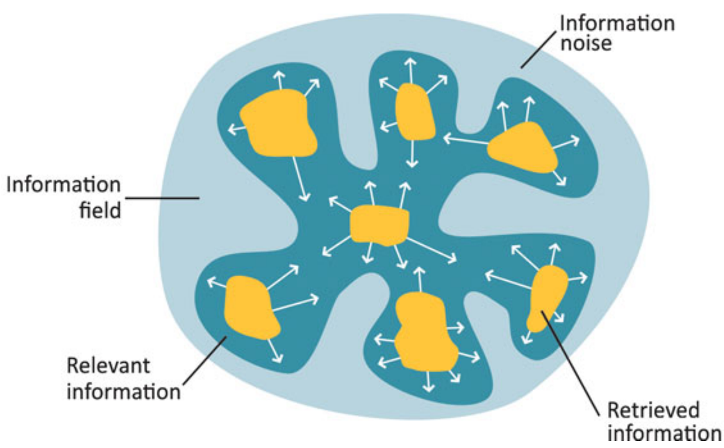


Fig. 6.9. Information search with a pre-marking of the information field. The arrows show the search directions from the information concentration centers

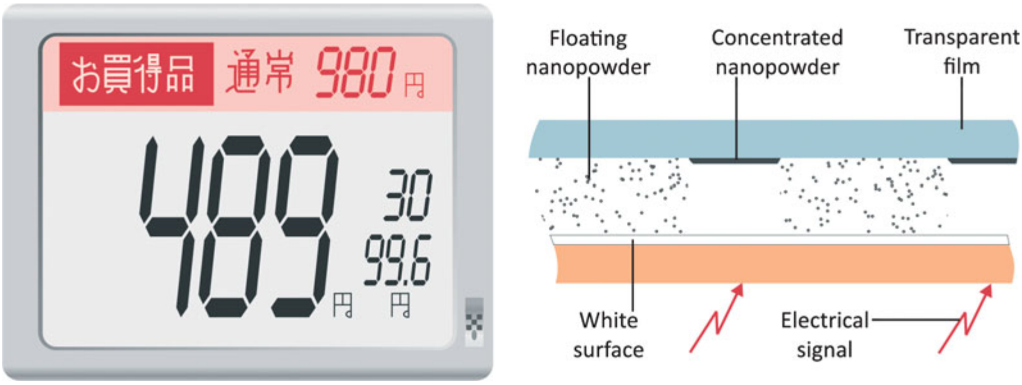


Fig. 6.10. Nanopowder Bridgestone display

Let's return to the example of «Bridgestone display».

The search for information necessary for building the Display Evolution Tree did not give any versions for the «Segmentation of objects to powder» transformation. No display modifications made of fine particles, granules or powder were revealed. The transformation name gave us the first key word — «Powder» (characteristic property), the word «Display» (the name of the object being transformed) became the second key word. These words were used as a query and the search system retrieved a large volume of information about Bridgestone powder displays (Fig. 10).

Finding these displays by random search would be difficult because Bridgestone's principle area of activities is tire manufacture.

It is important to note that the structuring of information into the Evolution Tree considerably facilitates search for additional modifications of a system under analysis. The search does not proceed easily until some limit is achieved; found information units often duplicate each other. However, after some «critical mass» has been achieved, the general structure of the Tree becomes clear and each piece of information occupies its own place on the Tree branches surprisingly quickly and precisely. The search addressness improves and the search itself becomes simpler.

6.3. Using the Evolution Tree for problem solving

6.3.1. Expanding the spectrum of preliminary concepts

While solving inventive problems, there often occurs a situation when we obtain many preliminary concepts but neither of them suits us. Besides, a customer wants to have not only a solution to his problem but also the analysis of possible methods for transforming his technical system or technology. Using the Evolution Tree would be helpful in this case. The purpose of using the Evolution Tree for problem solving is maximally expanding the spectrum of proposed concepts and providing material for the analysis of alternative ways of solving a given problem. This prevents us from leaving out of account some problem solving possibilities and helps to find optimal solutions for specific problem conditions.

Identifying an evolution pattern means searching for regularities within a set of concepts. To properly determine regularities, it is necessary to identify a key transformation to which each solution version corresponds and then to check the bulk of solutions for the presence of akin transformations. Identifying at least one akin transformation allows speaking about some regularity. If no akin transformations have been found, a single transformation should be related to one of the basic evolution patterns. One should remember that each version of a system can correspond to different transformations at the same time. That is, the same concept can participate in several patterns. Naturally, this adds ambiguity to a solving process. The ambiguity, however, provides a considerable advantage because it allows the number of investigated patterns to be expanded even if the number of preliminary concepts is small. Thus, should any doubts occur as to which pattern the concept belongs, it is better not to cut off possibilities but, on the contrary, to look for solution versions in both patterns.

Of importance is the solver's ability to consider the same solution version from different viewpoints. One should remember that any solution is normally a combination of transformations. Almost each solution version can be included in different patterns. This widens the field of search for new versions.

It would be practical to keep to the following sequence of operations:

*Preparatory stage.*Obtaining preliminary solution concepts.

Solving a technical problem by any methods and algorithms and generating a number of solution concepts.

*Step 1.*Selecting an evolution pattern of a basic Evolution Tree.

Identifying evolution patterns into which the obtained concepts can be organized.

Step 2. Determining the place occupied by the concepts in evolution patterns.

Arranging the concepts according to corresponding transformation versions of patterns.

*Step 3.*Determining the transformation potential.

Studying the recommendations of transformations for which no solutions have yet been found.

*Step 4.*Generating alternative concepts.

Filling the gaps by generating concepts according to the recommended transformations.

*Step 5.*Improving the concepts.

Developing and intensifying the concepts so as to obtain promising solutions.

*Concluding stage.*Selecting a concept for implementing.

Analyzing the obtained concepts, selecting the most promising ones and building the final solution of the problem on the basis of the selected concepts.

Analyzing a fragment of the Evolution Tree of a capillary for a jet printer chamber (real problem):

The jet printer operates in the following manner: a miniature ink-filled chamber is equipped with an activator for ejecting ink droplets. It works on the principle of an ordinary piston and may also be piezoelectric or steam-driven. A steam-driven activator is more ideal than a mechanical one because it has no moving parts and employs a vapor bubble, generated by a micro-heater located in the chamber, as a «piston» (Fig. 6.11).

Voltage supplied to the heater instantaneously boils the ink around it; the forming steam forces a portion of ink through a nozzle down onto the paper. Then the chamber gets filled with ink again through the upper capillary and becomes ready for a new «shot».

The main problem of the jet printer is its comparatively low productivity. This happens because some time is needed to fill the chamber with ink before ejecting a next portion of ink.

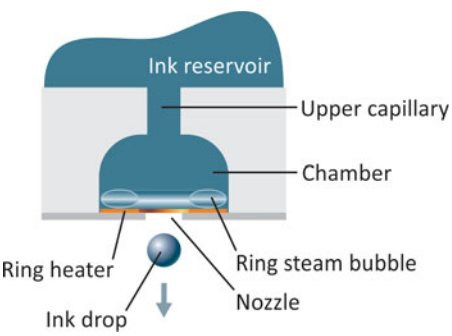


Fig. 6.11. A jet printer chamber with a steam-driven activator

An obvious way to solving this problem is increasing the capillary diameter. This, however, can only be done to a certain limit. The problem is that while «shooting» the ink is not only ejected from the chamber through the nozzle, but also rises into the capillary. The back current of the ink not only weakens the energy of an ejected droplet but also creates a counterflow in the capillary, thereby reducing the chamber filling rate. The following contradictory requirements are made of the capillary (Fig 6.12).

- *if the capillary diameter is large, the ink easily enters the printer chamber but at a working pulse, the ink is ejected back into the capillary;*
- *If the capillary diameter is small, the back current of the ink is insignificant, but the chamber gets filled slowly.*

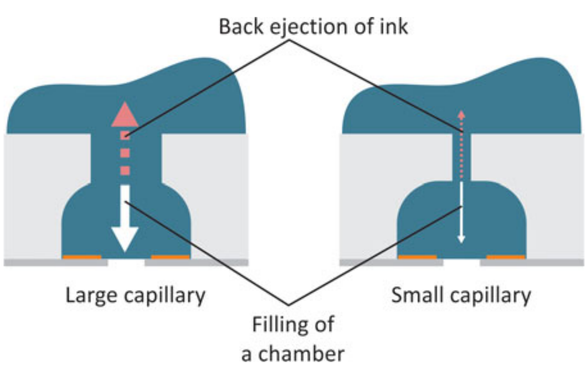


Fig. 6.12. The scheme of contradictory requirements of the capillary size

While solving this contradiction, several known solutions were found and new concepts were formulated.

Concept 1. Making two side capillaries instead of one central capillary. In this case, the back current of fluid caused by the pulse will first hit the chamber dome and partially lose its energy (Fig. 6.13, a).

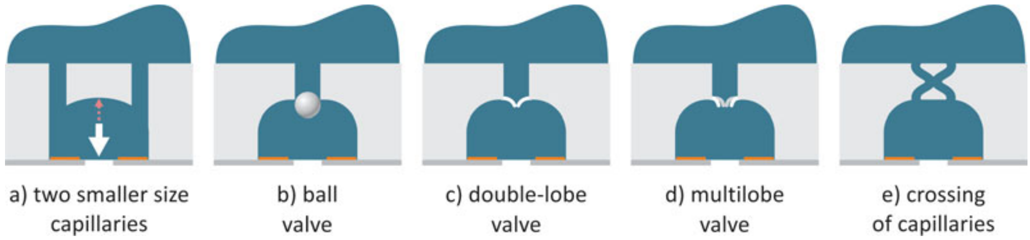


Fig. 6.13. Concepts 1–5

Concept 2. Placing a freely floating microball into the chamber. During ink ejection, the ball will block the entrance to the capillary, thereby preventing the back current of the ink (Fig. 6.13, b).

Concept 3. Using a double lobe valve at the capillary entrance (Fig. 6.13, c) instead of a microball.

Concept 4. Mounting a multilobe valve at the capillary entrance (Fig. 6.13, d).

Concept 5. Changing the capillary shape from straight to zigzagging with the crossing of capillaries if the chamber is fed by means of two capillaries (Fig. 6.13, e). At the point of crossing, the reverse flows of ink ejection will collide mutually extinguishing each other.

It should be noted that each of the proposed concepts has disadvantages either with respect to prevention of the back current of ink or with respect to technological effectiveness.

To obtain additional concepts suiting the customer, an express analysis of the available information was carried out. The generated solutions were organized into an Evolution Tree

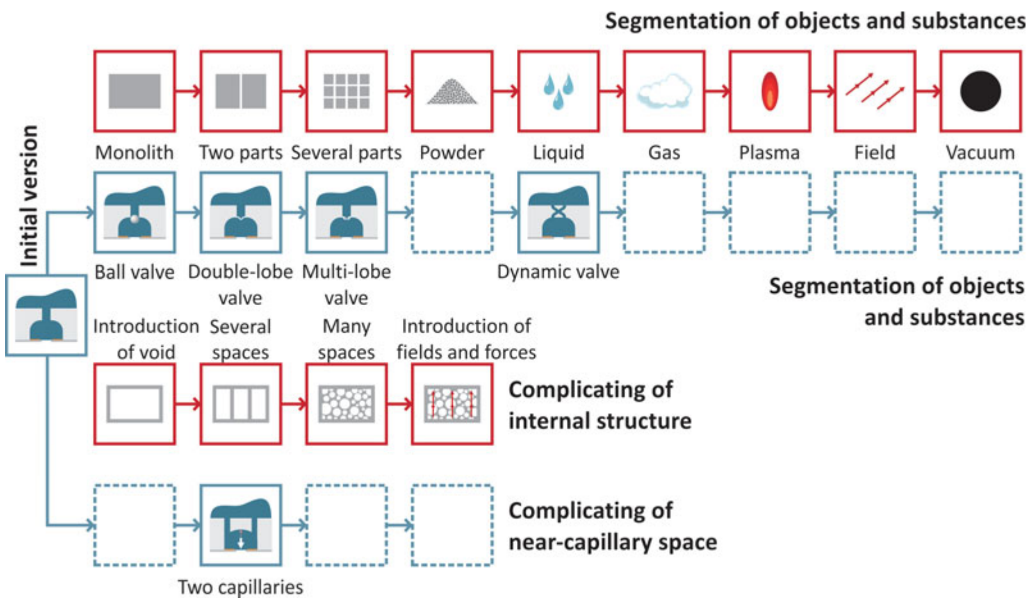


Fig. 6.14. A fragment of the capillary Evolution Tree

and arranged on corresponding evolution patterns (Fig. 6.14). There are two such patterns: — «Segmentation of the valve blocking element» and «Complicating the near-capillary space». Solution versions were found for the omitted steps.

Let's consider «Segmentation of the valve blocking element» pattern.

In the pattern «Segmentation» were unfilled embodiments: the locking member of a powder or granules, gas, plasma, field components and vacuum.

In the pattern «The increasing complexity of space» the following transformations were available: the introduction of emptiness, emptiness division into multiple volumes (micro-capillary structure) and the introduction of fields and forces to the area around the capillary.

To find solutions, unfilled transformation versions through the «Segmentation» pattern from «Powder» to higher transformations were analyzed. Results of the analysis also included a version of «fluid» because it is extremely important in this object setting where one of the main components is liquid ink.

The following solution concepts were found.

The «Liquid» step

Liquid is one of the main resources of this system. That is why two new concepts were obtained in addition to the already available one:

Concept 6. Generating in an ink reservoir ultrasonic vibrations having frequency equal to that of the droplets ejected onto the paper may increase pressure in the chamber during a pulse and reduce it during filling. Then a counter-pressure occurring during a pulse will prevent back current of the ink through the capillary and facilitate its passage through the nozzle.

Concept 7. Weakening the counterflow using liquid (ink in a reservoir) may be achieved by creating increased pressure in the reservoir.

The «Gas» step

Concept 8. Mounting a micro-heater in the capillary and switching it on simultaneously with a chamber heater. Then a steam plug will be formed in the capillary. It will block the capillary for a fraction of a second, thereby preventing the back current of the ink.

The «Field» step

Concept 9. Using ink having electro-rheological properties. Supplying an electric voltage will «freeze» the ink in the capillary for an instant, thereby preventing its back ejection.

Through the «Complication of space around capillary» three transformations are free, but we have only considered the last two: micro-capillary structure and the introduction of fields and forces. We obtained the following concepts:

The «Segmentation of space into multiple volumes» pattern

Concept 10. Segmenting of space into multiple volumes, we will obtain a capillary-porous structure where the motion trajectory of ink micro flows will be very complicated (Fig. 6.15).

A high-speed back current formed at the pulsed ejection of the ink gets into randomly arranged capillaries, changes its direction many times and is well decelerated. Because the chamber-filling rate of the ink is much lower than that of the back ejection, practically no deceleration of the flow occurs.

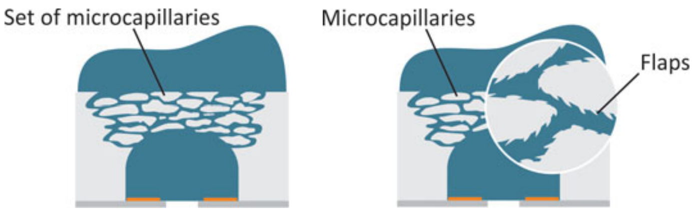


Fig. 6.15. A capillary-porous structure of ink supply

Using a micro porous material for manufacturing the jet printer head is adaptable to streamline production and may be realized without any difficulties.

The «Introduction of field and forces» step

Concept 11. We are primarily interested in the forces directed against the back current ejection. Accordingly, it may be proposed to provide the capillary surface with flaps similar to those available in blood vessels and facilitating blood circulation. Such flaps do not prevent filling of the chamber, but can prevent the ink from getting into the reservoir.

This technical solution combines the advantages of a micro-capillary structure and a return valve. However, considerable difficulties are expected in practical realization of this device.

Thus, using the Evolution Tree for solving the jet printer chamber problem resulted in six sensible concepts. The final solution — making the printer head from a capillary-porous material — was based on one of those concepts which were not found at the problem solving stage but were worked out just by analyzing the information obtained in the course of problem solving.

Special methods may be used to search for new effective solutions in the course of the Evolution Tree analysis. Some of them are described in the following section.

6.3.2. Structural analogy method

A special case of technical system improvement is transformation by analogy. Our method of *structural analogy* [6.2] proved highly efficient in work with Evolution Trees. Let us consider it in the first instance. The simplest and fastest problem-solving method is solving by analogy: find a corresponding analog for each situation.

Analogy is correspondence between objects and processes affording ground for the transfer of information, characterizing one object, to another object having similar essential properties.

It was the analogy with cobweb stretched between branches that suggested Samuel Brown the original construction of a cable-stayed bridge (Fig. 6.16).

Fig. 6.16. Cobweb and a cable-stayed bridge

Studying the structure of tall plants or bones of living organisms allowed engineers to better understand the construction principles of high towers. Today a 1300 m high city tower is under construction in China. It will house 100,000 dwellers. The unique construction will be able to withstand fire, flood, earth quake and hurricane [6.3]. The cypress — the tree capable of withstanding the strongest winds — was selected as the analog to the tower (Fig. 6.17).

A cypress crown consists of a great number of small segments which allow wind to pass through. Its root system, buried only 50 cm below the surface, extremely branches out and is similar to a sponge in structure. As the tree grows, the root system is actively complicated and expanded so that overturning or stubbing this tree is very difficult. The root system of the cypress city is buried in an artificial flat island surrounded by a lake. Such a support construction is designed for cushioning shocks of earthquakes.

The organization of the body of a quick and agile box-fish suggested to Mercedes engineers the optimal shape of the «Bionic Car» body (Fig. 6.18). The selected prototype is very interesting. As distinct from strong high-speed fish such as sharks, or, for example, dolphins, the box-fish has minimum energy expenditure exactly when moving at low speeds and is very maneuverable despite its weak fins and tail. In addition, the fish shape really resembles a box so placing items in the car is very convenient. The wind shape factor of the Bionic Car is 0.19 (for comparison, for most modern cars this factor equals 0.28-0.35). It is an excellent indicator for the car meeting all safety and convenience standards for ordinary city cars [6.4]

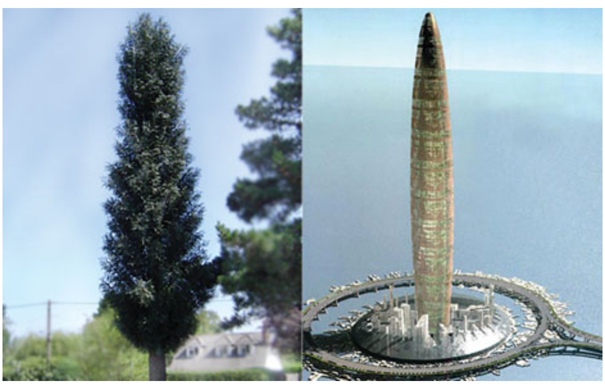


Fig. 6.17. The project of tallest building is the analog of a cypress

Fig. 6.18. A Box fish, Bionic Car

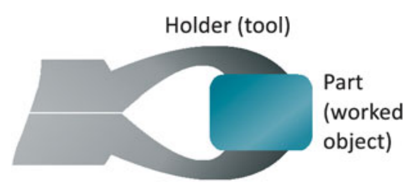


Fig. 6.19. An example of performing the «Holding an object» function

A particular case of analogy is the so-called *structural similarity*. Structural similarity may be presented as the development of the focal object method created by E. Kuntze and C. Whiting, which provides for a transfer of the properties of randomly chosen objects to a technical system being perfected. In the process, there may occur unexpected combinations which are then developed by way of free associations.

The main idea of structural similarity is it is always more convenient to compare similar assemblies which present certain structures than individual objects. It is most convenient to consider structures composed of two components: a tool and an object* worked by this tool.

A holder is a tool and a part is an object. They perform the function «Holding an object». Such a structure may be part of many technical systems (Fig. 6.19).

Often you can see that some structural groups making part of different systems are often similar to each other. This is especially obvious when they participate in the performance of similar functions. At the same time, the designation of the systems incorporating such components may be different.

The function of an automobile is «Carrying loads and passengers». This system comprises, among other things, a two-component structure — a fuel tank and fuel within the fuel tank. The components of this structure perform the auxiliary function «Holding liquid» being constituent parts of other automobile components.

An automatic coffee machine is designed «To make and sell coffee». A paper cup which, properly speaking, contains coffee, is part of this technical system. The cup and coffee are a two-component structure performing the function «To hold liquid».

The coffee-filled cup and the fuel-filled tank perform similar functions and have similar structures which may be determined as a «Shell and Filler» (Fig. 6.20).

New versions of making a two-component structure may be found by analogy

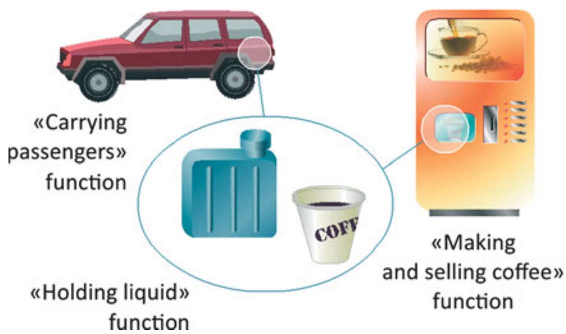


Fig. 6.20. A similar function of the components of different systems

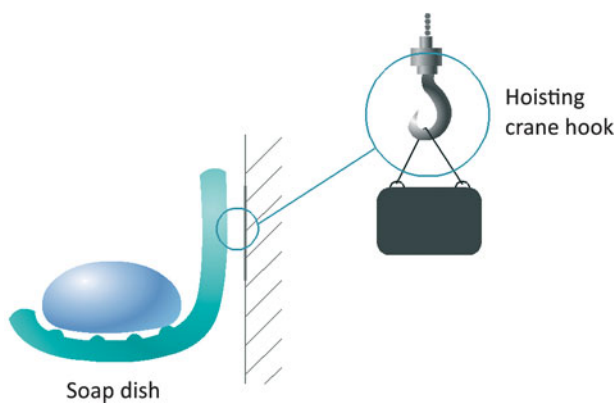


Fig. 6.21. Analogous structures: «soap dish + wall» and «hook + load»

*) Structural similarity can be represented as the development of the method of focal objects, developed by E. Kunze and C. Whiting and providing for the transfer of the properties of randomly selected objects perfected technical system. This may result in unexpected combinations that develop further through free association.

with a similar structure. For example, searching for versions of mounting a detachable soap dish to a wall may be based on the analogy with a load-handling device which consists of a holding component and an attached load (Fig. 6.22).

The simplified evolution pattern for the load-handling device will look as follows:

1. *hook,*
2. *tongs,*
3. *grab having elastic fingers,*
4. *vacuum gripper,*
5. *electromagnetic gripper.*

Accordingly, the soap tray mounting modifications may look as follows:

1. *mounting a soap dish by nails or screw nails*
2. *mounting the back wall of a soap dish in a clamp*
3. *mounting a soap dish in a spring clamp*
4. *mounting a soap dish by suction cups*
5. *mounting a soap dish magnetically.*

Using the analogy method resulted in a number of concepts for a technical system having quite a different designation compared to the initial system.

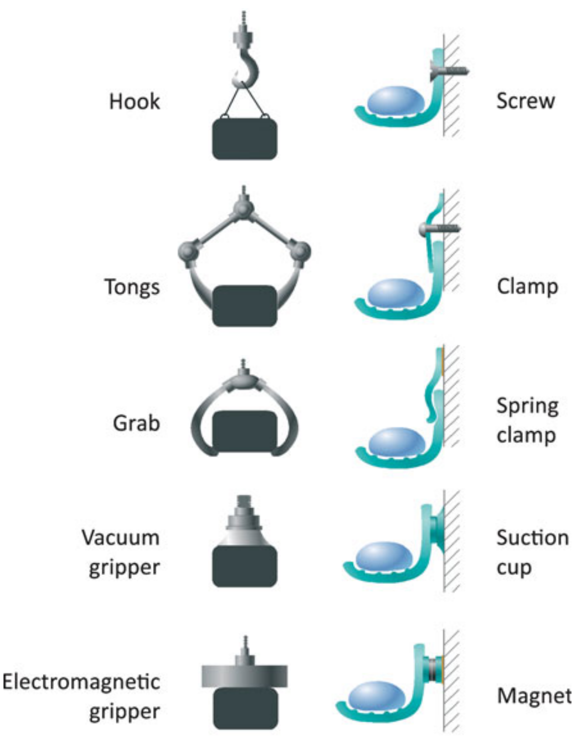


Fig. 6.22. Soap dish mounting by analogy with a load-handling device

To make sure that a search for new concepts by structural analogy is possible, we analyzed a number of structural diagrams composed of interacting «tools» and «work objects»: «wheel and road», «ruler and part being measured», «shaft and bearing», «hinged joint», «shell and filler».

Over 40 objects were studied for the «shell and filler» structural pair, including an automobile tire, a vacuum flask, a parachute canopy, a boat, a motor vehicle (both the boat and the motor vehicles may be considered as a shell (body) with filler (passengers), an air-balloon, a light bulb, a submarine, a fuel tank, footwear and many others.

We selected the *automobile tire* as a basic analog for comparing the transformations of objects under study, because the tire itself is a well-studied technical system. The tire was also convenient to use as a basic system because we had already constructed the tire Evolution Tree. We selected several tire modifications from this tree, all of which were united by the same purpose — increasing the motor vehicle safety.

Here are all these modifications arranged into an evolution pattern:

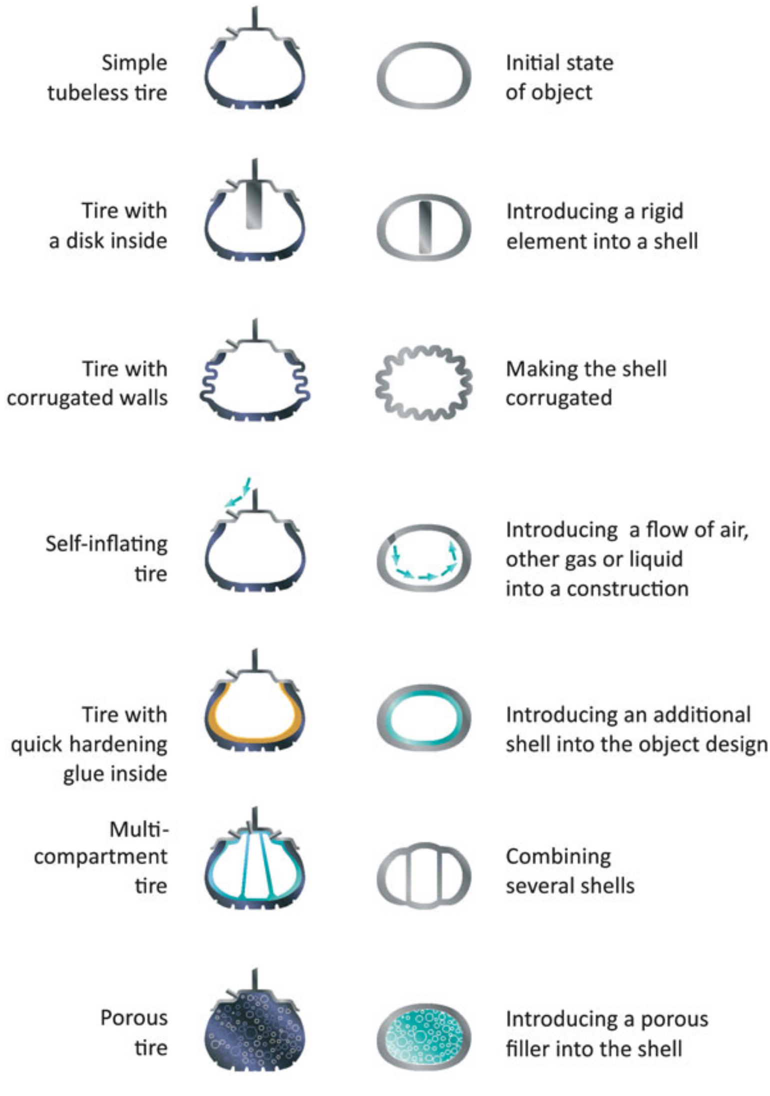


Fig. 6.23. The automobile tire evolution pattern and transformation concepts

1. A simple tubeless tire. The most vulnerable modification. On puncture, air escapes through the hole and the tire loses its carrying capacity. In addition, the flat tire may lie alongside of the rim causing the wheel to veer.

2. A tire having a disk inside. A strong disk mounted within the tire does not allow it to collapse in the event of an abrupt puncture or breakage. This allows an automobile to stay on the road and prevents an accident.

3. A tire with corrugated walls. It often happens that after getting a puncture, the tire collapses in a random manner, not necessarily directly under the rim. If it collapses sideways, a lateral force occurs and the car can diverge from the road. The tire with corrugated walls is free from this defect. After getting a puncture, such a tire collapses along the corrugations and its tread finds itself directly under the wheel rim.

4. A self-inflating tire is another method of dealing with punctures. After puncturing and pressure loss within a tire, a compressor is switched on. It supplies air into the tire and maintains pressure at a necessary level. Special automobiles equipped with such a system can normally go after the tire has been shot through with a bullet.

5. A tire containing a quick-hardening glue. Some liquid glue placed into a tire during mounting will be distributed over the entire inside surfaces. In the event of a puncture, the glue penetrates the hole together with the escaping air, thickens and closes the hole.

6. A multi-compartment tire. There are several unconnected torus-like rubber compartments arranged within a multi-compartment tire. During inflation, air is uniformly supplied into all the compartments. In the event of a puncture, only one compartment is damaged and the remaining two are enough for getting to a repair shop.

7. A porous tire. A porous tire is the most reliable wheel. This wheel was invented at the beginning of the last century for a gun carriage, but did not find application in automobiles because it was too rigid when made of porous rubber. About a hundred years were needed for this system to come alive at a new evolution loop. This wheel made of special porous plastics excellently suited automobiles trying to set records by cracking the sound barrier on the ground.

Let us formulate the concepts corresponding to these transformations (Fig. 6.27):

- *«A simple tubeless tire >The initial state of the object.*
- *«A tire having a disk inside» > Introducing a rigid component into shell.*
- *«A tire with corrugated walls» > Making the shell corrugated*
- *«A self-inflating tire» > Introducing a flow of air, other gas or liquid into the object design.*
- *«A tire having a quick-hardening glue inside» > Introducing an additional shell into the object structure.*
- *«A multi-compartment tire» > Combining several shells.*
- *«A porous tire» > Introducing a porous filler into a shell.*

Let us consider several objects of a similar structure, to which the above described transformations were applied: *a vacuum bottle, a parachute canopy, a boat, and a car.*

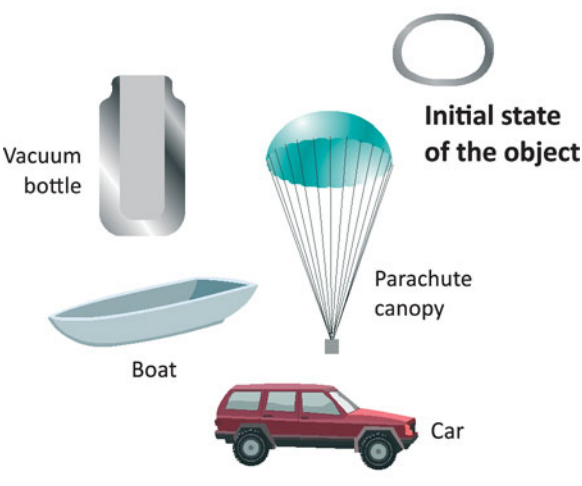


Fig. 6.24. The initial state of the system under analysis

1. The initial state of the object (Fig. 6.24)

A vacuum bottle. A double hermetically closed vacuum bottle forms a vessel that may be filled with liquid. Air is evacuated from the space between the bottle walls. This maximally reduces heat loss so that liquid in such a bottle maintains the initial temperature for a long time. The vacuum bottle is a shell and the liquid it contains is filler.

A parachute canopy. An air-filled shell made of durable cloth and shaped into a semi-sphere is attached to a man by means of shroud lines and a suspension system and provides his low-speed descent.

A boat. A boat, ship may also be considered as a shell containing various mechanisms, loads, crew, etc.

An automobile. An automobile is a rather complicated system. But in a generalized sense its outer part, the body, is a shell and everything that is inside is filler.

2. Introducing a rigid element into a shell (Fig. 6.25)

A vacuum bottle with an additional stiffening rib. One or several rings tightly put on or stuck to the outer surface of the vacuum bottle will increase its strength.

A canopy with an inserted metal spring element. A spring made of elastic wire is inserted in a small pilot parachute which, while releasing, extracts the main canopy from a bag. When collapsed, such a spring resembles a chamomile flower; while releasing, it turns the pilot's parachute into a small ball.

A load-carrying frame inside a boat (canoe). The tourist boat — canoe — is very popular. The main requirement of the canoe is small weight and high strength. In addition, the canoe must be easily packed into a handy pack unit. Therefore, its construction is executed in such a manner that the boat strength is mostly provided by a light and strong folded frame with water-proof cloth stretched on it.

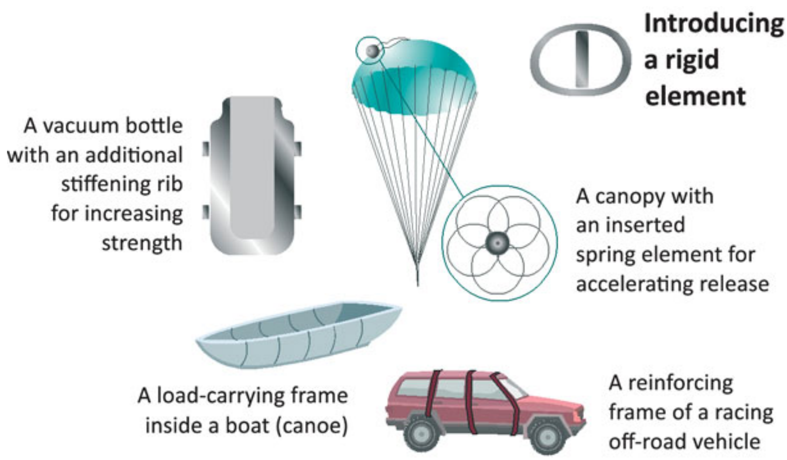


Fig. 6.25. Systems with an inserted rigid element

The reinforcing frame of a car body. A sports car for off-road races must have a strong body. This requirement is because such a car has more opportunities to overturn than an ordinary one. Increasing the body metal thickness is unpractical, therefore, an additional reinforcing frame — a spatial construction of strong pipes — is mounted on the car body. The power frame can be built in the body walls directly in the process of manufacturing, as is done in many off-road motor vehicles.

3. Making a shell corrugated (Fig. 6.26)

A vacuum bottle with corrugated walls. There is one more method for increasing the vacuum bottle strength — making it corrugated. Corrugations are a kind of stiffening fins standing up to deformation.

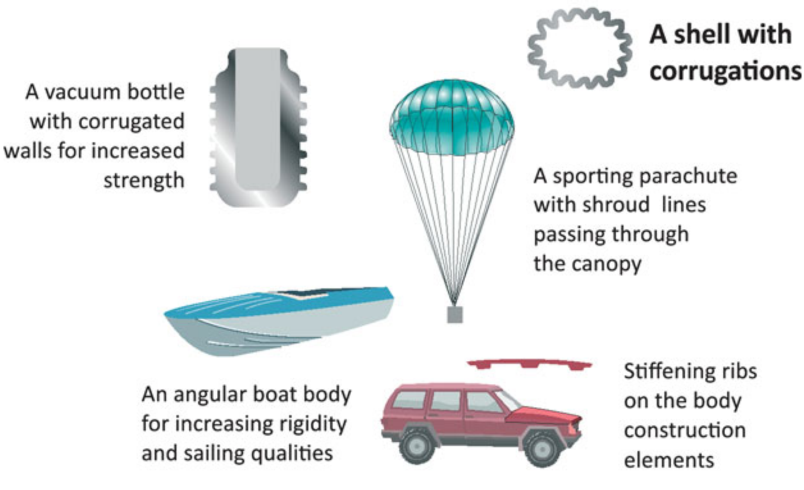


Fig. 6.26. Systems with a corrugated surface

A sporting parachute. The shroud lines of this parachute are attached not to the canopy edge, as in a military parachute, but pass through the canopy toward the apex dividing the canopy into radial segments. This is done for regulating air escape from under the canopy and providing a more stable descent.

An angular boat body. The roundish shape of the boat body does not provide sufficient strength and motion stability. That is why the body is provided with longitudinal protrusions and hollows. These unusual «corrugations» may perform different functions — from increasing the body rigidity and strength to improving the speed and sailing qualities.

Corrugations on the components of the car body construction. The car body is a good example of using corrugations or folds for increasing the shell strength. For example, hoods of many automobiles are not smooth but have Figured folds which make them more rigid.

4. Introducing a flow of air, other gas or liquid into the object construction (Fig. 6.27)

Evaporation cooling of a vacuum bottle. Wetting a surface with water and then blowing it with an air flow reduces the surface temperature due to the evaporation cooling effect. This effect may be used for additional cooling of a vacuum bottle.

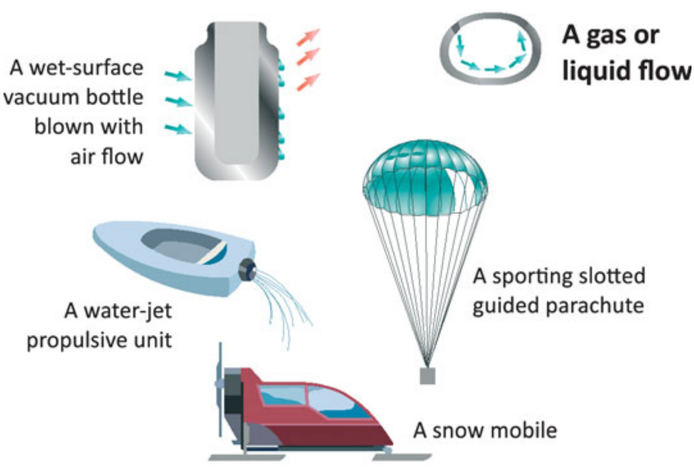


Fig. 6.27. Systems with air or water flows

A sporting slotted guided parachute. An air flow may be used to control a sporting parachute. Slots in the rear part of the parachute canopy permit part of air to escape in a horizontal direction. The parachute starts moving ahead. Special side slots serve to turn the canopy.

A water-jet propulsive unit. It is easy to give an example of how a liquid flow is used to move a boat: any propelling screw pushes off the water creating a dynamic water jet behind the boat.

A snowmobile. An aeromobile — an automobile provided with an air screw — moves similarly to a motorboat. Such a construction is widely used for moving in winter, using skies. The ski version of the aeromobile is called a snowmobile.

5. Introducing an additional shell into a construction (Fig. 6.28)

A vacuum bottle having an additional body. To reduce heat loss, a vacuum bottle may be enclosed in an additional heat-insulating container.

A gliding wing parachute. Such a parachute has a doubled surface, the upper and lower sheets of which are connected by means of profiled nervures. When opening such a parachute, an oncoming air stream inflates cavities in the canopy so that the canopy becomes wing-shaped.

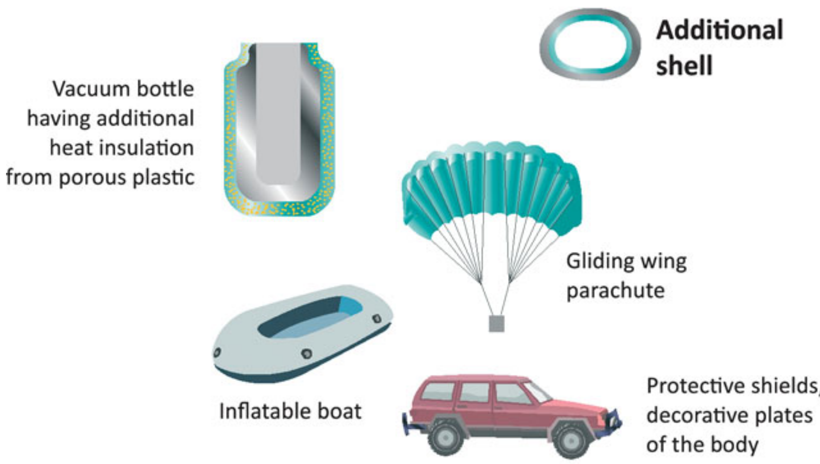


Fig. 6.28. Double-shell systems

An inflatable boat. One more construction which is popular among fishermen and tourists is an inflatable boat. It is a hermetically closed doubled profiled shell, which takes the form of a boat upon inflation. When deflated, the boat may be packed in a knapsack.

Protective shields, decorative plates on a car body. Decorative and protective plates, shields for protecting individual units may be considered as an additional shell of a car body.

6. Combining several shells (Fig. 6.29)

A combined vacuum bottle for different dishes. A vacuum bottle used for different dishes may serve

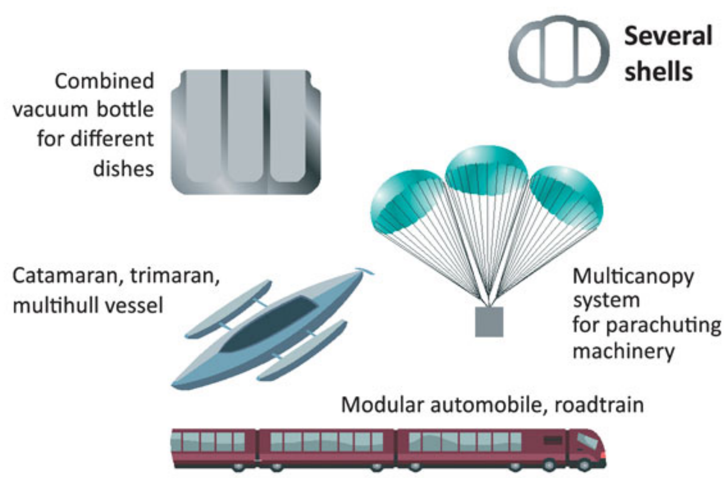


Fig. 6.29. Several shells in one system

as an example of using several shells for a vacuum bottle. It is a body containing several bottles. Each bottle may have different contents.

A multi-canopy system used for landing equipment. Parachutes may be used to land not only people but also heavy equipment. To provide an acceptable lowering speed of a multi-ton machine, a huge parachute canopy is needed. However, the larger the parachute canopy, the slower it opens. That is why multi-canopy systems are used in parachute delivery of machines.

Fig. 6.30. LeTourneau road train

A catamaran, trimaran, multihull vessel. Using several interconnected hulls instead of one hull in a submarine will abruptly increase the stability and usable area of the vessel. The most popular construction — a catamaran — has two adjacently installed hulls. Also known are trimarans and multi-hull vessels capable of withstanding a storm.

A modular automobile, roadtrain. Because the road width is limited, it is impossible to place several automobiles in a row. That is why a «multibody» automobile is a number of sequentially arranged modules — a roadtrain. Interesting is the railless construction of LeTourneau train built in 1940 (Fig. 6.35). It was 141 m long and comprised 11 sequentially connected modules [6.5].

7. Introducing a porous filler into a shell (Fig. 6.31)

A heat-retaining container. A vacuumized bottle retains heat for a very long time. However, often it is only needed to keep liquid hot for a short time. In that case, a foam container with a foam lid and a usual reservoir placed inside may be used.

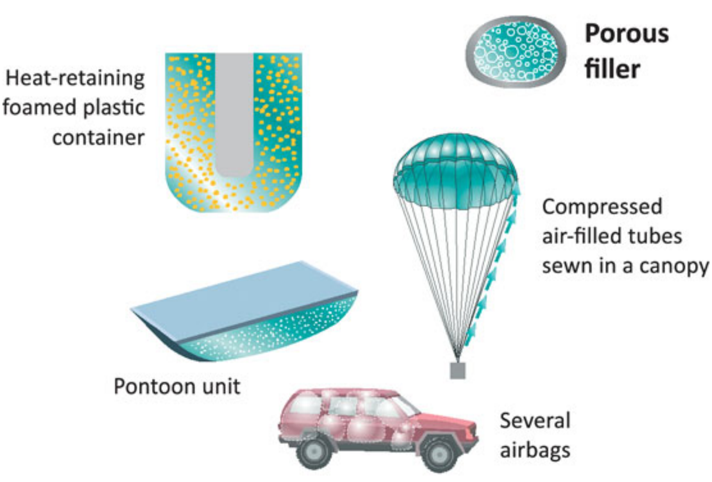


Fig. 6.31. Use of porous filler

A quick-to-open canopy. Sometimes, it is necessary that a parachute canopy open very quickly. One of the methods for accelerating its opening consists in sowing a light tube in the canopy along its perimeter and providing the parachutist with a spray can containing compressed air. As the parachute deploys, air is supplied into the tube turning it into a rigid ring that instantaneously stretches the canopy.

A pontoon unit. The most reliable vessel is a strong shell filled with foamed plastic. Unfortunately, such a vessel has no room inside for machines and the crew. But, for example, pontoon units used in the army for assembling floating bridges are coposed just of a metal shell filled with foamed plastic.

Several car airbags. Porous plastics are widely employed in the automobile construction for noise suppression and filling closed cavities. In addition, airbags used in cars may be considered as a kind of porous filler. Another technical solution is filling a passenger compartment with inflatable elastic balls upon a collision.

Let us accumulate all the transformation versions in a table (Fig. 6.32). As we see, well-grounded versions of similar transformations were found for each technical system.

We see that we have managed to find valid versions of similar transformations for each structurally similar technical system (tire, parachute, boat, car). The fact is that pairs of structurally similar components have similar resources which may be used for their transformation. As systems having similar structural elements evolve, similar transformations repeat in one form or another. During the evolution of a system having similar structural groups, transformations will repeat in one form or another. Consequently, such an approach can be used to search for transformation possibilities for technical systems under consideration.

Fig. 6.32. The spreadsheet of transformations of different systems

The evolution of any systems having similar structural groups will be accompanied by transformations of the same kind. For example, the display Evolution Tree proved that transition of different types of displays to the microlevel is observed in several cases [6.6]. Considering the four basic types of display — a cinematograph, a black-and-white TV set, a color TV set, and a 3D TV set — reveals the following situation (Fig. 6.33).

Cinematograph

The operating principle of a cinematograph consists in mechanically pulling a semi-transparent film in front of a light source directed at a screen. It does not matter for the cinematograph whether a black-and-white or a color film is loaded. For the cinematograph, transition to a microlevel means giving up the mechanical method of changing the film frames. In a laser projector, the image is drawn on a screen by three laser beams — red, green and blue. Special mirrors that deflect the laser beams move luminescent dots horizontally and vertically on the screen.

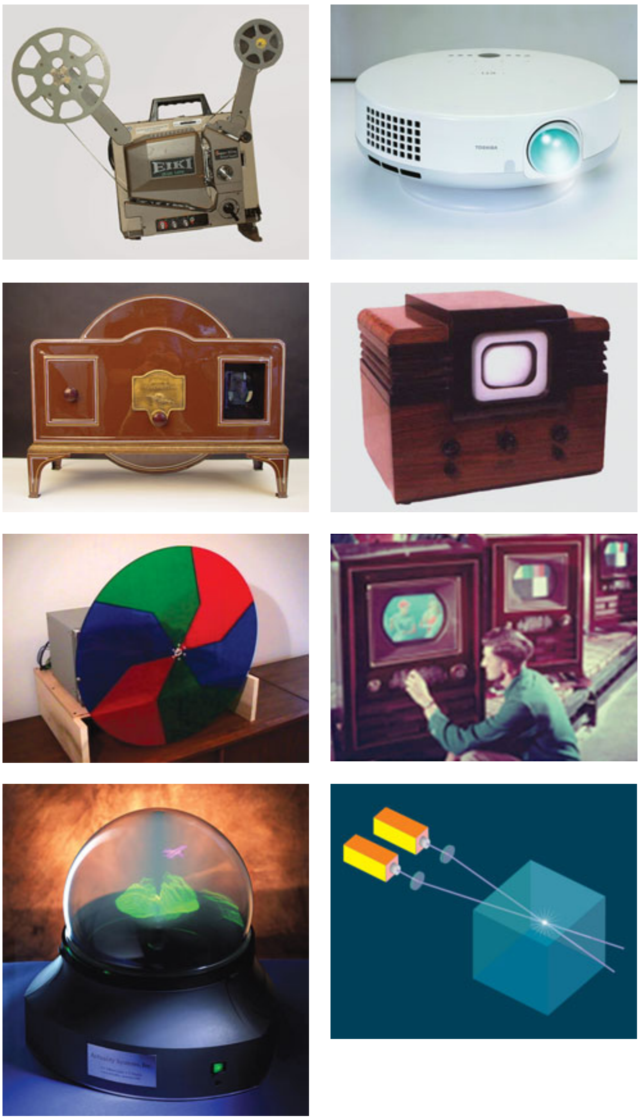


Fig. 6.33. Evolution of display

Black-and-white TV set

By analogy with a cinematograph, first TV sets scanned image mechanically by means of the Nipkov disk. The rotating disk had some helically arranged holes through which dot-pixels were projected onto the screen, forming a primitive image.

This system passed to the microlevel after the invention of a CRT TV set. In a cathode-ray tube, image is scanned by magnetic deflection of an electron beam generated by an electron gun. The electron beam activates phosphor-coated pixels placed within the tube on its front panel. These pixels draw a luminescent image.

Color TV set

By analogy with a cinematograph and black-and-white TV set, the same rotating disk formed by four sectoral light filters (red, blue, green and yellow) was initially used. When rotating, the disk covered the screen of an ordinary black-and-white TV set with its edge so that all the four colors sequentially passed in front of the screen. The black-and-white display worked according to a special program and four times faster than usual. When the red light filter passed before the screen, the display showed a picture for the red color, the green filter passing before the screen highlighted dots which had to become green, etc. One complete turn of the disk produced four images which merged into one color image for a viewer. The color mechanical scanning TV set Zenith was produced and sold in the USA.

Transition to the microlevel for the color TV set occurred very quickly and mechanical scanning of color was replaced with electronic scanning. Color TV sets got the same cathode-ray tube as black-and-white TV sets, the only difference being use of three electron guns activating phosphor of a certain color — red, green or blue — instead of one.

3D display

The sample principle — rotational motion of a semitransparent screen — was used by Actuality Systems for creating the first really 3D display. It is a transparent ball having 50 cm in diameter with a flat 2D screen inside, rotating about a vertical axis at about 10 revolutions per second. Sequential «cuts» of a 3D image are displayed on this screen so that a viewer sees a single picture as a result. A real 3D image is produced which is observable from different sides as a real object in space. This display with a rotating screen differs from multilayer displays where a pseudo-3D image is formed only visible from a certain position.

The predicted evolution trend of a 3D display is transition to the microlevel. One of possible solutions is producing luminescent dots in crystals or gaseous medium by the action of invisible infrared radiation. Beams produced by two infrared semiconductor lasers cross at a predetermined point of space filled with active gas. The summarized energy of the two beams is sufficient for producing a luminescent dot. Controlling the beam deflection, for example, by means of movable mirrors can build a 3D image from luminescent dots throughout the entire gas volume.

And now let us discuss the use of the Evolution Tree for express-analysis of a technical system to be improved. From the viewpoint of using the structural analogy method, the Evolution Tree gives a considerable advantage over other information organization methods because it is a huge data base of analog systems. Solutions can be found directly by analogy with better studied systems.

In p. 6.1., we have shown that new promising versions of a system can be found by comparing the Evolution Tree of a real object with a basic Evolution Tree formed by generalized evolution patterns. Comparing the transformations of an object being improved with the Evolution Tree of another, better-studied object can intensify this process and make it more vivid. It is common knowledge that the major part of funds allotted for scientific research

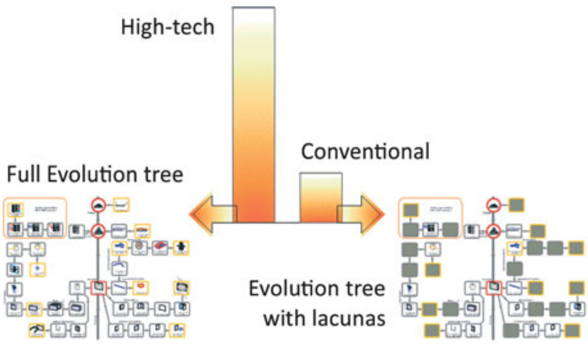


Fig. 6.34. Comparing the Evolution Tree for an ordinary and advanced technical system

are concentrated in the military, space and other science-intensive branches of industry. Building an Evolution Tree for similar systems — one belonging to household appliances and the other to space branch — will prove that the second tree is far richer (Fig. 6.34). Accordingly, it is necessary to seek analogs of a technical system being improved in advanced fields of knowledge and compare this system with a structurally similar system from other fields of technology.

In other words, if you are going to improve a frying pan (*a shell heated on one side*), it is wise to analyze the implementation versions of a propulsive nozzle wall which has little structural distinction from an ordinary frying pan. That is, to analyze a system to be improved, it is necessary to select a structurally similar system from a well-studied field of knowledge, build the Evolution Tree of that system and see if you can find a prompt for your problem by comparing the analogous system implementation versions with those of the system under analysis.

Let us consider an example of solving a real problem by the structural analogy method.

Example: Forming a sheet from molten aluminum

To produce a sheet, molten aluminum is fed between two rollers rotating toward each other. It is squashed into a thin strip, cools down and exits the rollers in the form of a finished sheet (Fig 6.35). Engineers are permanently preoccupied with improving the aluminum sheet quality, primarily, with obtaining a smooth-surface sheet free of cavities and scores. The conflict responsible for surface roughness occurs between the sheet itself and the roller surface. The fact is that semi-solidified metal has a good adhesion with the roller surface and its particles break away from the sheet and remain on the roller surface. The process is further aggravated as the particles stuck to the roller form cavities on the sheet surface, thereby lowering its quality.

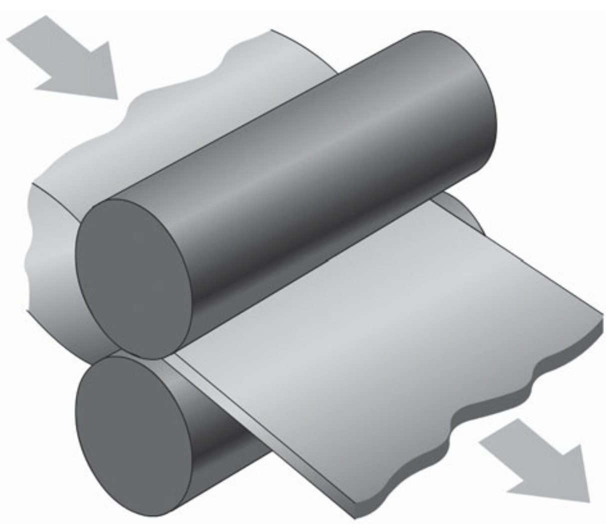


Fig. 6.35. A device for forming an aluminum sheet from melt

What is to be done?

It is obvious that rolling an aluminum sheet which is accompanied by simultaneous solidification of metal is a very complicated and difficult-to-understand process. Let us approach the problem from a different side and formulate it in the following way. We have a smooth roller surface which makes contact with solidifying aluminum. How can the roller be transformed to improve the interaction?

How is surface transformed in other technical systems?

To answer this question, let us analyze the display Evolution Tree (see p. 5). The fact is that the display surface and surface state is one of the most important display characteristics, so examples will be in abundance.

What do we see?

A smooth surface of a cinema screen is coated with micro-pyramids (see p. 5), the surface of TV displays becomes rough and is coated with strip micro-lenses, micro-shutters, etc. the micro-roughness size being not arbitrary but strictly prescribed.

In this connection, there occurs a prompting question: what will happen if the roller surface is transformed by analogy with the screen surface — coated with micro-protrusions or a network of shallow marks? Will it be better or worse than using a smooth roller?

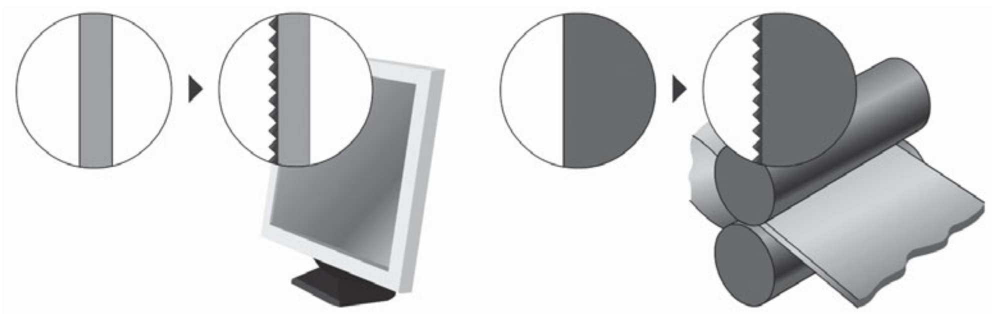


Fig. 6.36. Transition from a smooth surface to a rough surface can prove useful for various technical systems

Now, a possibility of changing the state of the roller surface is obvious to the work team. This immediately causes a discussion which results in something like «paradoxical as it may seem, a very smooth roller produces a higher adhesion to an aluminum sheet than a rough one.» The problem is that well-polished steel surfaces stick together very strongly because adhesion forces are replaced here with Van der Waals forces, forces of molecular interaction.

The answer is as follows: providing the rollers with roughnesses sufficient for excluding intermolecular interaction but not reducing the sheet surface quality is a promising development trend for this technology. Further work should be aimed at experimentally determining the optimal parameters of micro-roughnesses.

6.4. New possibilities for circumventing competing patents

6.4.1. Main approaches to patent competition

Patent struggles are the quintessence of competitive struggles. Successful competitive struggles are primarily achieved by companies with aggressive inventive ambitions. In addition to the profit gained from licensing, this ensures better investment policy management, effective and timely response to the market requirements.

Any strong business idea is based on an effective technical solution, a concept of a new competitive product which is usually protected by a patent. A car, fiber-tipped pen, TV-set, sewing machine — all of them are examples of commercially successful technical solutions.

Practically everything may be patented — new information, new ideas and concepts obtained in the course design work, technological processes, control algorithms, etc. Each patent can become a sticking point in a conflict between a patent holder and manufacturing companies. Here the interests of many competing companies come into conflict with each other and lead to real patent battles. Each company wants to be the first, tries to circumvent patents of other companies and protect its own product.

Harley-Davidson, an American manufacturer of motorcycles, faced a problem of style duplicating. Japanese companies started manufacturing motorcycles that were similar in appearance to the popular «Harley».

To prevent the spread of those machines, Harley-Davidson filed an application for the firm basso sound produced by Harley motorcycle engines (high-speed engines used in Japanese motorcycles produced a high «sporting» sound during operation). The United States Patent and Trademark Office has approved the application and confirmed that it is this sound characteristic of these particular motorcycles that distinguishes these machines from other types of motorcycles. The patenting process drags on too long but the company has already benefited from this application by using it as an advertising campaign.

One of the episodes of the never-ceasing patent battles is taking place in Russia

The Oktyabrsky District Court of Petersburg has sustained the case of the Gatchina resident Viktor Petrov against the Russian branch of Siemens concerning the infringement of copyright. The court ruled unlawful the use of «smileys» (the plaintiff's invention) in Siemens cellphones. Viktor Petrov patented the popular faces as early as 2004. The invention was called «The way of a directed regulation of human psycho-emotional state».

A new patent war is about to break out among the computer-manufacturing companies. Microsoft has patented the «the double click» of a computer mouse. Patent No. 6727730 was issued on April 27, 2004. The computer mouse double click has been used for a long time in all computer programs. The current initiative of the company can touch practically all computer makers. So far, Microsoft has not declared how it is going to use this unique invention and what fee it is going to demand from computer mouse manufacturers. One thing is clear: if the patent is not avoided, the prices for computer mice will go up and the company's revenues will grow.

Microsoft is famous for its aggressive patent policy. For example, patent No. 6 754 472 states that Microsoft owns the right for the methods and apparatuses for energy and information transfer by using a human body. Microsoft is planning to attach various electronic devices of the pager of mobile phone type to a human body for receiving information about a man's condition and transferring data about the environment conditions to the carriers of these devices. According to Microsoft, the human body is but a part of the intellectual property object created by the company and subject to licensing.

Some of the world's largest IT firms, such as Microsoft, HP, Dell, and Apple intend to cancel the patent for the method of organizing a wireless network which belongs to the Australian company «Scientific and Industrial Research Organization» (CSIRO). The development of wireless networks started only five years ago whereas the patent was obtained as early as 1996. SCIRO foresaw that this invention would become a very valuable piece of intellectual property and did not miss out to patent it. Today, this patent may become a sword of Damocles hanging over the well-being of many large manufacturers because it covers practically everything that is connected in one way or another with the organization of such networks as well as devices capable of connecting to these networks. SCIRO is already making efforts by offering reasonable licensing conditions to Buffalo Technology. The license cost may amount to several billion dollars.

Of course, it often happened that the introduction of a new invention lagged behind. In that situation, there was no question about the protection of author's rights because all reasonable terms of patent protection elapsed. The greatest share of profit was usually received by businessmen who were the first to appreciate a promising idea, organize the patent protection of this idea, production and sale of a profitable product. Inventors themselves often could not use the fruits of their labor and received recognition too late. In such a situation, a businessman did not need to think about buying a license, he only had to think about organizing production and promotion of a cheap high-quality product.

In 1888, the engineer J. Loud got a patent on a ball unit for marking boxes and sacks with sugar. He could hardly imagine that several decades later his invention, introduced practically without alterations, would bring huge profit to the enterprising businessman M. Reynolds. Just like the owners of parachute-manufacturing companies did not pay anything to Leonardo da Vinci who proposed the scheme of the first «device for safe descending from height».

However, the situation has cardinally changed. The fact is that the time of bringing an idea to mass production has reduced abruptly. This is due to the improvement of both technological equipment and calculation, experimental and design methods. Wide use of robots and computerized equipment allows redistribution of force. The number of those employed directly in production is permanently reducing whereas the number of people employed in the production process maintenance, including creation of new products, is growing. The work of designers has become much more effective, too. A designer using Pro Engineering software can do the work done by an entire design office some 50 years ago.

All this resulted in a considerable acceleration of the innovation process. Today, mass production of patent-covered developments starts very quickly, well before expiry of the term of a patent (which usually is 15 — 20 years). Under such conditions, a businessman cannot just take an idea and produce a product he likes, because a company holding a corresponding patent or the inventor himself lays a well-grounded claim against him. The availability or absence of a patent is becoming one of the very important factors which determines the success of any business.

Fair information competition

A good technical solution, an idea of an absolutely new product is a great piece of luck for any large company: it is an excellent opportunity to strengthen its market position and gain a good profit. However, inventing a basically new product is not easy. On the one hand, the basic human needs are not many. On the other hand, there still exist technological limitations even though technological advance is gradually widening its frames. As a result, the «innovation field» is not limitless and it is often necessary to develop new modifications of products already available on the market.

When a company wishes to enter some market segment, the selected niche of consumer demand often turns out to be occupied by the product of other companies. Winning the competition is only possible by convincing consumers that your product is the best. This is the function of advertising. The other way is creating a really better product and offering it to the market at a lower price. This is not easy because developing the product and organizing its mass production need huge investments. In addition, promising directions are often occupied by competitors and protected with patents. Taking out a license is also rather expensive.

You may try to copy a product manufactured by a competing company, develop its manufacturing technology and start production. This, however, is difficult within the limits of fair competition, without using illicit methods because companies generally take serious measures to protect their intellectual property. One of the possible methods of legally obtaining information about a competitor's product is the *reverse technical analysis* [6.7]. When a device produced by a company is already in open sale, competitors can buy and examine it. Thoroughly doing this work may help understand how this device operates, what it is made of, what technology is used and use this information to organize production.

After taking a decision to produce a similar or analogous product, two important problems need to be solved: development and protection of the technological process for this product, product adaptation to the manufacturing capabilities of the company and its protection. A thorough analysis provides much information and allows the company to develop its own technology. Full doubling of the technological process used by the competitor makes no sense because every company has its own set of machine tools and processing lines. In any event, it is necessary either to adapt the existing technology to the manufacturing capabilities of the company or to create a new, more effective technology.

Production processes are seldom protected by patents. Companies merely keep them secret. If another company develops and use an identical or very similar process, this will not entail any sanctions. Generally, protection of a commercial secret concerning a production process is the responsibility of the company itself. The law only prohibits use of such methods as industrial espionage, bribery of employees, etc.

As distinct from the *production process*, the *product* produced by a company practically always has patent protection. Protection may cover the design of the product and the design of its parts, principle of operation, operating rules of a device, materials of its components, etc. It is here that the key sphere of confrontation of the competing companies' interests, the patent struggle arena is located. Some companies try to maximally protect their developments while their competitors tend to attack or circumvent these patents and to obtain their own patent for the manufactured product.

Juridical and inventive patent-circumventing methods

To circumvent a competing patent, just like in case of obtaining a patent for a new invention, joint efforts of two key Figures are required. They are an inventor who offers some idea and a patent agent who prepares an application for obtaining a patent in accordance with the patent law and files it with a patent office. They are valued specialists at any company. It is their work that sometimes results in a great success. It is their mistakes that can cost dearly to the company.

If both an inventor and a patent agent act each in his own direction while circumventing a patent, two main patent-circumventing methods may be identified.

The first and most important one is *juridical*.

This method implies use of imperfections of the patent law and mistakes in the invention description in a patent. The object of invention itself is not changed at all in this case. Using the juridical method of circumventing a patent, a patent agent tries to protect the object of a necessary invention without changing it at all. Presenting the situation in the form of a diagram (Fig. 6.37) will show that all actions are performed in accordance with the lower portion of the diagram i.e. the hero in the juridical method of circumventing a patent is the patent agent. The inventor in this case performs an auxiliary function and checks that changing the patent formulation does not cause loss of the essence of the invention from the technical point of view.

To juridically circumvent a patent, it is necessary to find in a competing patent incompatibility of features of a real object, method, substance with the text of the claims and description. Then it is necessary to describe the existing object of invention using other terms and reformulating the description according to certain rules, the content of which is know-how of each experienced patent lawyer. The main landmark in this work is the acting patent law. This work often results in a possibility either to litigate the patent by proving that it protects a previously known invention or to compose a new application and get an alternative patent for the existing invention.

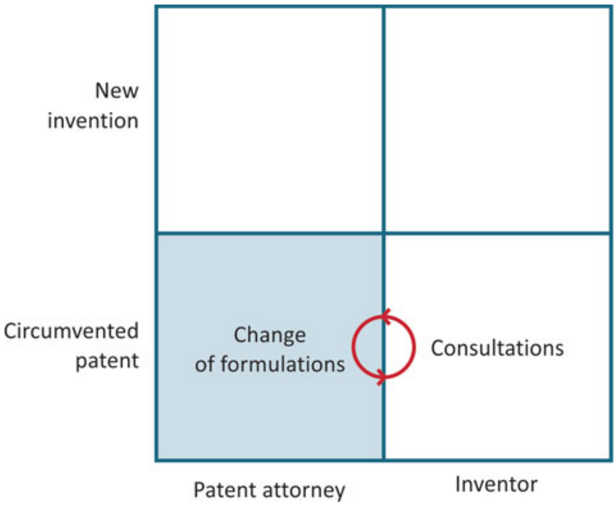


Fig. 6.37. Juridical method of circumventing a patent

With certain skills and qualification of a patent agent, the juridical method may prove very effective. An important resource is the errors committed while preparing the patent being circumvented. Making up claims and a description reliably covering the obtained solution requires thoughtful and thorough work. We may suggest that this work was not done carefully while preparing a competing patent. The preparation and examination of an application, as well as obtaining a patent are very complicated and costly procedures that involve tens of specialists. A fault at one of its stages may lead to insufficient protection of an invention so that the obtained patent will be litigated in future. In addition, no absolutely perfect patent law exists and the levels of patent protection vary to a considerable extent in different countries, which sometimes leads to curiosities.

One fine day, the patent agent John Keogh from Melbourne became world famous. Using the imperfection of the Australian patent system, he put an end to the disputes about where the wheel was invented — in ancient Asia or in ancient Africa — and literally managed to obtain a patent for the wheel. According to the document issued by the Australian Patent Office in 2001, Keogh patented a «circular transportation facilitation device» [6.8].

Circumvention of patents by using tricks of law is a very exciting and interesting topic, but it is beyond the scope of our book. Every patent agent has lots of his own methods and approaches to circumvention of competing patents. Sometimes, they easily manage to do this, but usually it requires a serious and thoughtful work. There are also patents which are very difficult to circumvent using the juridical method.

Example: Legend about Singer's patent

According to the patent legend, Singer, the creator of the first truly practical model of a sewing machine, patented his invention ingeniously. The formulation was exclusively simple: «The distinctive feature of Singer's sewing machine is that the hole on the needle is located closer to the pointed end» (Fig. 7.2). This patent is practically impossible to circumvent by the juridical method.

In those days, this patent was considered an ideal solution making impossible further perfection of the principle of operation of sewing machines because the needle patented by Singer was present in any operable system, leaving other inventors only an opportunity to introduce minor alterations such as a foot drive.

The second method of obtaining a new patent is the *inventive method.*

This method implies more or less significant transformation of the device or production process structure which is the object of invention. In this case, the main role belongs to the inventor. With this method, all actions are located in the upper portion of the diagram (Fig. 6.39).

It is the most efficient way out of a problem situation: finding a better solution concept than that of the competitors and patenting this concept, i.e. a new solution to a technical problem, a significant change in the object of invention is made. This may be achieved by:

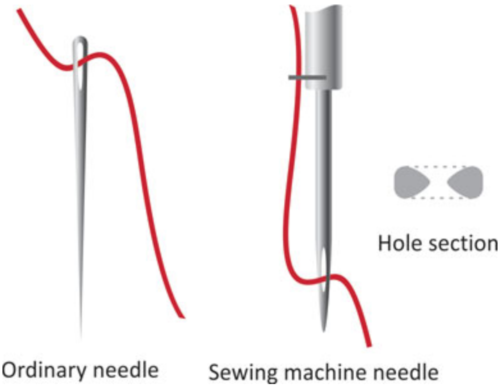


Fig. 6.38. An ordinary needle and the needle patented by Singer

- analyzing the available technical solution and determining its disadvantages,
- trying to remove these disadvantages and determining the reasons why it is impossible to do,
- removing these reasons and obtaining a new technical solution.

This is the method used by many inventors and it often proves very effective. A solution concept may be obtained by using all TRIZ tools designed for effective problem solving. In addition, the juridical part of the work — preparing an application and obtaining a patent — becomes much simpler.

However, it happens often enough that the version patented by a competing company fully satisfies us and making any significant change is inexpedient. The following contradiction occurs:

- it is necessary to change the object of invention so as to obtain an alternative patent «overlapping» the competing patent;
- changes should not concern the principle of operation of a technical system which is the object of invention.

This contradiction can be resolved by combining the juridical and the inventive method into one *«juridical-inventive»* method, the essence of which may be formulated as a *«change without*

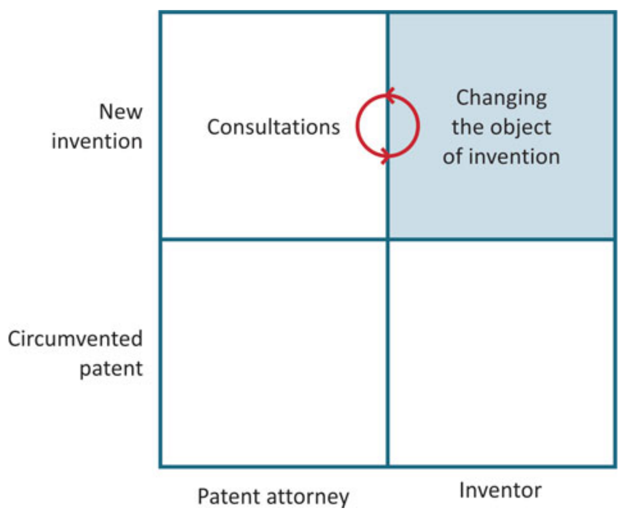


Fig. 6.39. The inventive method of circumvention of patents

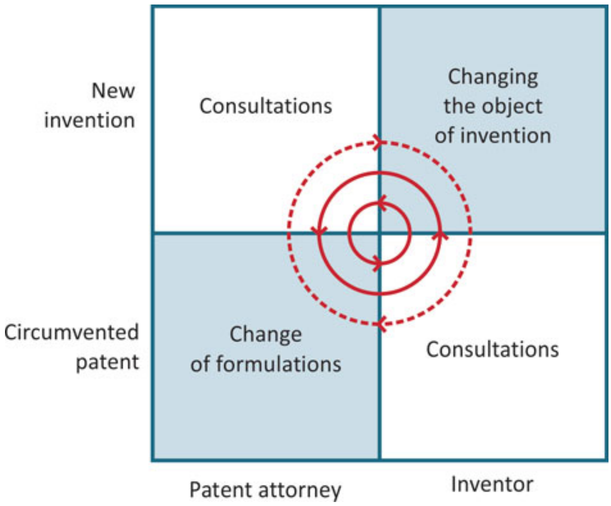


Fig. 6.40. «Juridical-inventive» method of circumvention of patents

changing». This means that a patent agent and an inventor are equivalent, work jointly and their efforts are focused in the front portion of the diagram (Fig. 6.40).

It should be noted that the composition and structure of any system usually suggest some variability; therefore, some alterations can always be made in the system's components. The presence of such insignificantly differing components within a system does not change its principle of operation because the object of invention undergoes minimal changes from the technological viewpoint. This, however, may turn out to be sufficient for obtaining a juridically significant distinction of a certain patent feature, which, in its turn, may give a patent agent additional opportunities in legal protection of an alternative patent.

To solve such problems, it is practical to use the Evolution Tree — an organized set of evolution patterns of the system described in a patent being circumvented. The more significant alternative solutions we obtain, the easier to orient in this field and to find more effective solutions suitable for circumvention of patents. Because the evolution patterns are based on objective laws of technology evolution, we can expect that all the basic transformation versions of a system can be described by a selected patent feature. Then it only remains to find technologically satisfactory unfilled boxes in the given tree and start attacking a competing patent.

Properly speaking, you are dealing here with a patent information handling process described in Section 6.1. At the input of the process, there will be a prototype to be transformed, which means that it is necessary either to improve the prototype itself or to find an analogous equivalent solution permitting patent protection of this prototype. At the output, you must obtain a series of prototype transformation concepts allowing you to compose a patentable alternative version of the given system. Such an approach offers new opportunities in circumvention of patents. In addition, study of possible changes in system's components and patent features of the invention may also be used for effective patent protection of your company's developments. To this end, it is necessary to perform a reverse operation — to consider the technical solution, patented by your company, as a competing patent and to conduct a search for basic alternative versions of the components of the system. Building several new models of systems that include changed components will reveal possible ways of circumventing your patent in the future. The alternative models of a system should be patented using a so-called patent umbrella which closes the door to the main possibilities of circumventing your patents by competing companies.

What actions should be performed to reveal the alternative versions of a competing invention?

Preparatory stage. It is necessary to determine the prototype, that is, the patent to be circumvented.

*Step 1.*Determining the function, composition and structure of the technical system covered by this patent.

Step 2. Determining the patent features which need to be changed.

Step 3. Carrying out a patent search, finding the main alternative versions of this system.

Step 4. Building an Evolution Tree for the system under analysis.

Step 5. Identifying transformation versions which are not covered by patents by comparing the basic and specific Evolution Trees (See Section 6.1).

Step 6. Assessing the possibilities of using these versions in the system, selecting the most suitable ones.

Step 7. Proposing of technical solutions based on these versions.

*Concluding stage.*Proving of legal protection of these solutions.

Example: possible circumvention of Singer's patent. Let us go back to the legend about Singer. From the technological viewpoint, circumvention of Singer's patent (Fig. 6.41) could look like as follows: the sewing machine design wherein the needle moving up and down is replaced with a hook swinging about some center of rotation (Fig. 6.41, b). This modification was proposed by the French tailor Timonie. The hook has some advantages over the needle, the most important being a simpler drive.

One more version concerns the needle itself, or better to say, the hole for thread. The hole may be located on the needle axis in such a manner that the thread enters at the top and exits at the bottom (Fig. 7.5, c). In addition, the generally accepted shape of the hole may be replaced with an inverted L-shape where the thread enters a hole at the side and exits at the bottom.

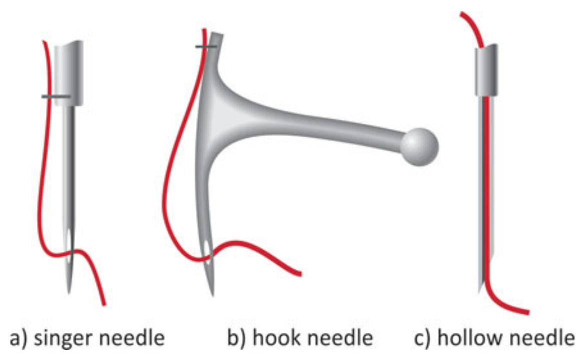


Fig. 6.41. Alternative modifications of the sewing machine needle

6.4.2. Using Evolution Tree for circumventing a real patent

Let us give a real example of how the proposed methods work. The author's goal was finding the circumvention versions for a patent on a device for supply of conditioner solution into the tank of a vertical drum type washing machine (Fig. 6.42). The housing of such a washing

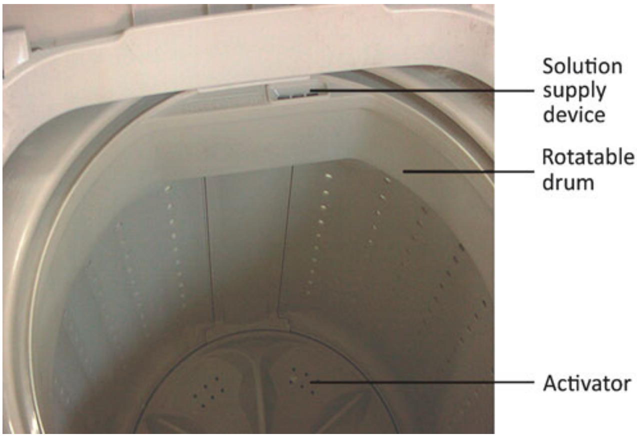


Fig. 6.42. A vertical drum type washing machine, top view

machine contains an open top drum that is rotated by an electric motor. At the drum bottom, there is a rotatable blade-type activator for washing and rinsing. In addition, the drum itself can also rotate quickly for removing water from the clothes.

The considered construction belongs to the economy class, therefore its manufacturing cost should be minimal. Such machines have a strict technological cycle. It includes one washing cycle and two rinsing cycles (Fig. 6.43). For presetting the required intensity of these processes, the user can only change the duration of each cycle and the activator motion character. After every washing and rinsing, clothes are wrung out and water is discharged from the washing machine due to the fast rotation of the drum together with the activator.

Before washing, detergent powder is poured into the washing machine tank and a conditioner solution that must get into the tank during the last rinsing cycle is poured into a special vessel located on the drum. The conditioner solution supply system is located directly on the drum, in its upper portion.

The design of the device for conditioner solution supply into the washing machine drum is shown in Figure 6.44.

There are several recesses molded within the hollow drum: 1, 2, 3 and 4. Pairs 1-2 and 3-4 have common walls and are located so that recesses 1 and 3 are positioned closer to the inner wall of the drum while recesses 2 and 4 are situated closer to the outer wall. Recesses 1 and 3 have a greater depth that recesses 2 and 4. In addition, recesses 2 and 3 are connected by an inclined channel. Over recess 1, there is an opening for pouring a solution. The bottom of recess 4 has an opening for discharging the solution into the tank. The device employs the centrifugal force that occurs during fast rotation of the drum.

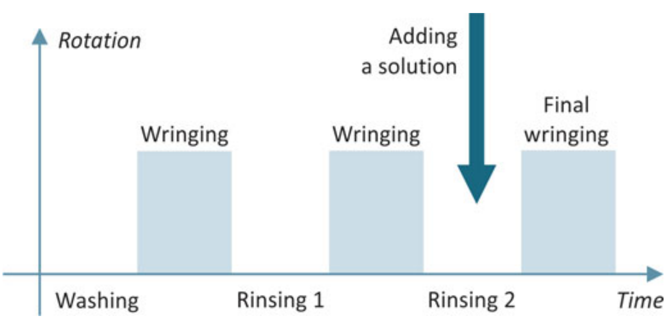


Fig. 6.43. The washing-machine operating cycle

The device works in the following manner.

The solution poured in before washing first gets into recess 1. During washing, when the drum is motionless and only the activator rotates, the solution remains in this recess.

When the drum starts to rotate for wringing out the clothes, the centrifugal force makes the solution flow into a shallower recess 2. During the drum rotation, the solution remains pressed against the wall in the second recess.

Then follows rinsing, the drum stops and the solution flows through the channel from recess 2 into recess 3.

Further the drum starts rotating quickly again to rinse the clothes. The solution flows into recess 4 and remains pressed against the wall until the rotation is finished.

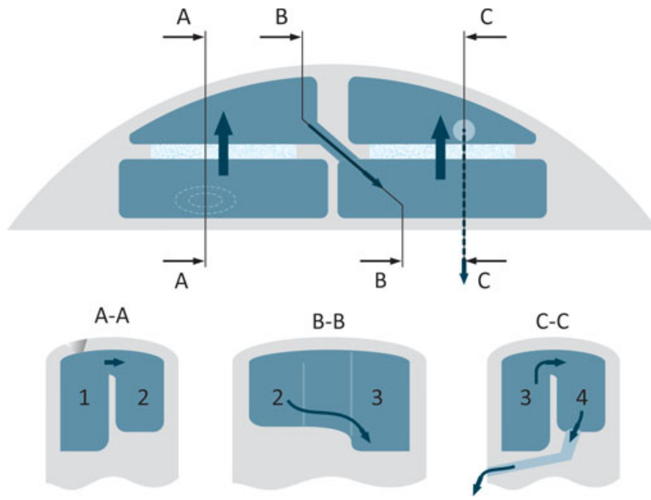


Fig. 6.44. A device for supply of conditioner solution

The drum stops for the second rinse, the solution flows down to the bottom of recess 4 and then flows through the channel into the tank.

Worth noting is the extreme simplicity and functionality of this device. It has no moving parts, valves, energy supply; it is only necessary to pour in a solution before washing. For the given conditions, this system may be considered close to ideal. Unfortunately, this «ideal system» had one serious disadvantage: it had been patented by competitors. Purchasing a license seemed the only way out.

Then the question of circumventing the patent came up. The problem was that the simple and effective prototype selected by designers fully suited them and changing the principle of operation was undesirable. We had a classic version of a patent obstacle: the system was fully satisfactory but it could not be produced and sold for legal reasons.

Before making a move towards circumventing the patent, it was necessary to select a patent feature to be analyzed using the Evolution Tree. For this purpose, a simplified model of the device for conditioner solution supply was built.

The simplest model will be a device comprising three chambers connected with each other by inclined channels (Fig. 6.45). One of the chambers (A) is filled with liquid and located closer to the center of rotation, the other one (B) is empty and is located at a greater distance from the center of rotation. The third chamber A1 is located similarly to the first one and connected with the chamber B by a channel.

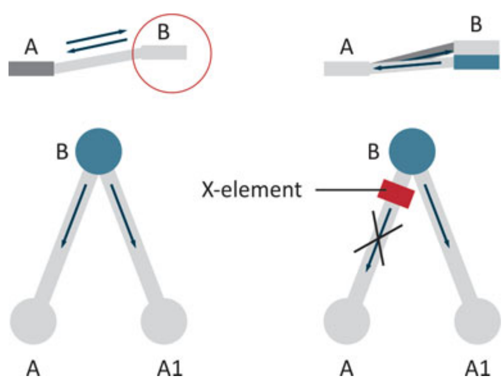


Fig. 6.45. A 3-D elementary module

It is apparent that at fast rotation and due to the action of the centrifugal force, the liquid starts flowing into the second chamber and remains there until the rotation ceases. Then the liquid starts flowing through the inclined channels into the first and third chambers.

To prevent this, the channel inlet should be located in the upper portion of the second chamber and the outlet in its bottom portion. Then the liquid will not be able to go back during rotation and will flow into the next chamber. This suits us excellently because it is necessary to direct the liquid to the next chamber without permitting it to go back.

Thus, to move the liquid by one step corresponding to the fast rotation of the drum, it is necessary to have a module comprising three chambers arranged in a certain manner. Connecting several such modules in series may produce a device providing a step-by-step motion of the solution from the first chamber to the subsequent level A chambers in some direction (Fig. 6.46).

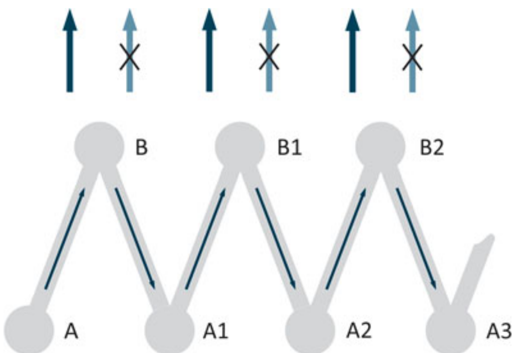


Fig. 6.46. A counter of fast revolutions of the washing machine drum

This model allows better understanding of the details of the device operation and determining the patent features to be analyzed.

Limitations imposed on the change of this technical solution are very severe. It is not allowed to complicate the key diagram of the device, introduce valves or gates controlled by the washing-machine computer. In addition, the chamber parameters (the tilt angles and surface properties of the walls, size and depth of the chambers and channels) are well optimized. Changing these parameters can negatively affect the device operation. It would seem that the drum rotation frequency may be used as the liquid movement control parameter. However, the drum rotates with the maximum speed during wringing and increasing that speed turned out to be impossible.

Thus, there only remains one patent feature that can be changed: *the layout of the three-chamber modules.*

In the existing design, three-chamber modules are arranged in levels under the upper surface of the drum in the rotation direction. We are dealing with the linear structure transformations that may be described by the Evolution Tree using the following evolution patterns:

- *Coordination of the chamber arrangement with the drum rotation direction;*
- *Geometrical evolution of a linear structure.*

To transform this device modification, we can try to use the basic Evolution Tree fragment that includes these evolution patterns (Fig. 6.47). The generalized transformation versions are arranged in the following manner.

The evolution pattern «Coordinating of the chamber arrangement with the drum rotation direction»:

- *in parallel with the rotation direction, in one direction;*
- *in parallel with the rotation direction, in different directions;*
- *perpendicularly to the rotation direction, toward the drum center;*
- *perpendicularly to the rotation direction, away from the drum center;*
- *perpendicularly to the rotation direction, down along the drum generatrix;*
- *perpendicularly to the rotation direction, up along the drum generatrix;*
- *at an angle to the rotation direction.*

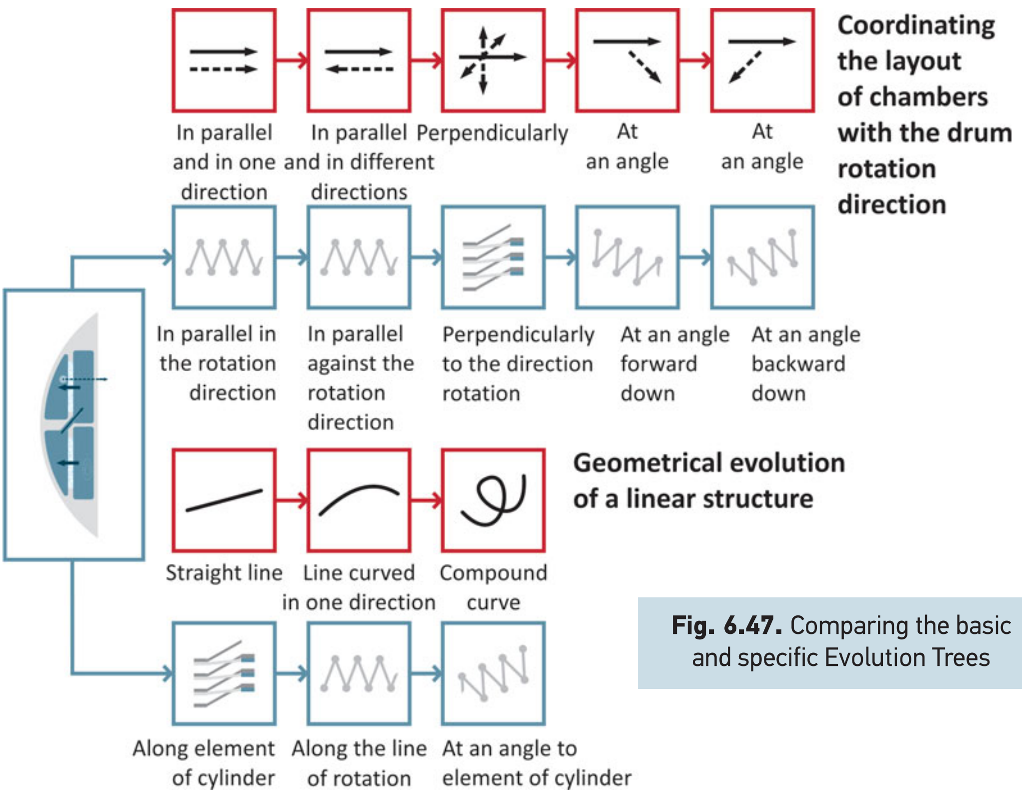


Fig. 6.47. Comparing the basic and specific Evolution Trees

The evolution pattern «Geometrical evolution of a linear structure»:

- *straight line;*
- *line curved in one direction;*
- *compound line.*

The Evolution Tree fragment for a real system — a solution supply device — will look as follows.

The evolution pattern «Coordinating of the chamber arrangement with the drum rotation direction»:

- in parallel with the rotation direction:
  - — in the rotation direction
  - — against the rotation direction

- perpendicularly to the rotation direction:
  — top down;
  — bottom-up;
  — inside the drum;
  — outside the drum;
- at an angle to the rotation direction:
  — forward down;
  — forward up;
  — backward down;
  — backward up.

The evolution pattern «Geometrical evolution of a linear structure»:

- straight line:
  — along the cylinder generatrix;
- line curved in one direction:
  — along the rotation line;
- compound line:
  — at an angle to the cylinder generatrix.

After considering possible module layout versions, we can draw a conclusion that there is some choice (patented layout charts of three-chamber modules are marked with red circles in Fig. 6.47 and the remaining ones may be used to build alternative technical solutions).

Placing the modules toward and away from the drum's center of rotation should be immediately excluded because it will be necessary to make the drum wall much thicker for placing the device composed in this manner. The variant of arranging the modules at an angle can also prove difficult to manufacture.

Placing the chamber against the rotation direction is equivalent to the initial patent. This version may be considered as a possible way of circumventing the patent, but requires advice of a patent agent concerning the scope of claim of the initial patent. It is legally easier to protect the top down arrangement of the modules along the cylinder generatrix and top down arrangement at an angle to the generatrix — backward and forward.

Technologically simpler is the top down module layout along the cylinder generatrix. We choose it as the alternative to the initial patent (Fig. 6.48).

What can the structure of a device with a vertical arrangement of modules be like?

Chambers for solution are mounted on two levels. The internal and external chambers are arranged in pairs and have a common wall. Over upper internal chamber 1, there is an opening for pouring a solution. External chamber 2 has an inclined bottom. The lower portion of the chamber has a slit. A diverter made as a continuation of the slit directs the solution into internal chamber 3 of the lower level. The bottom of lower external chamber 4 has an opening for solution discharge into the tank (Fig. 6.50).

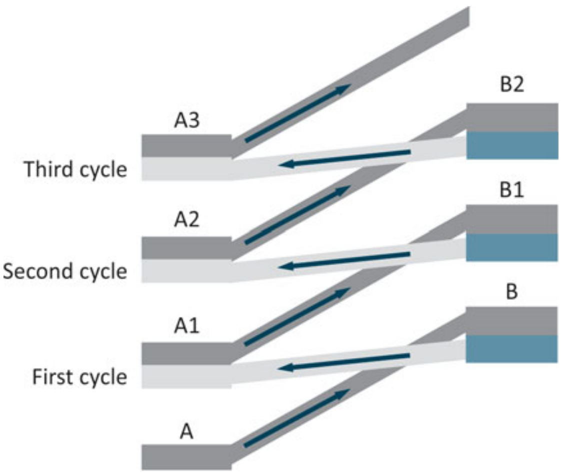


Fig. 6.48. Vertical arrangement of modules

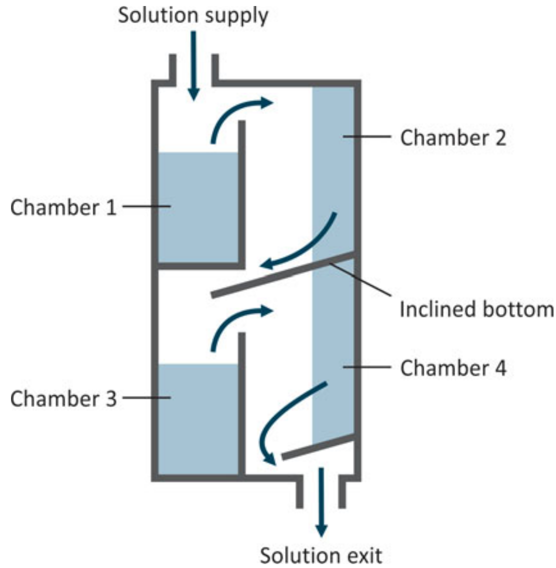


Fig. 6.49. Design layout of the new version of the device

This alternative device works in the following manner (Fig. 6.50).

- Solution is poured into chamber 1 and kept there.
- The drum starts rotating to remove water. Centrifugal force makes the solution flows into chamber 2 where it remains pressed against the wall.
- The drum stops for rinsing and the solution flows on the inclined bottom and the diverter from the recess of chamber 2 into chamber 3.
- Then the drum starts rotating quickly for wringing the clothes. The solution flows into recess 4 and remains pressed against the wall until the rotations ceases.
- The drum stops for the second rinsing cycle and the solution flows to the bottom of recess 4 and through the channel into the tank.

Thus, using the Evolution Tree, we have obtained alternative technical solutions which offer new opportunities for circumventing the existing patent. For introducing into production,

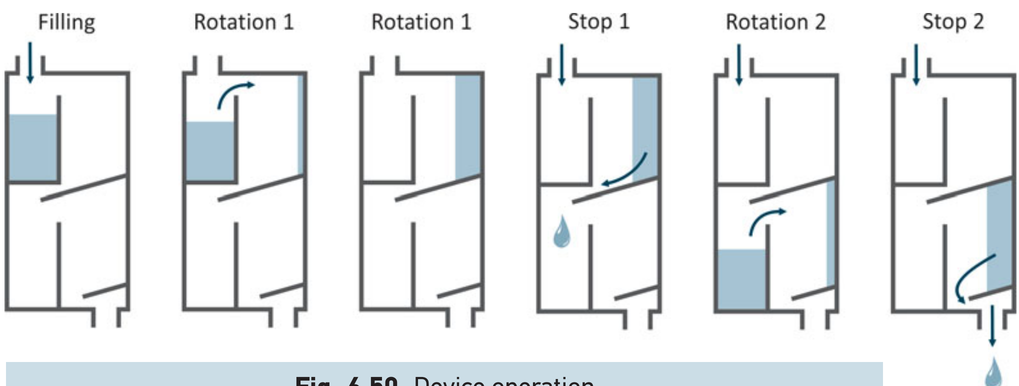


Fig. 6.50. Device operation

we have selected one, the most technologically advantageous version and proposed it for patenting. It should be noted that most of the described versions are operable and effective. They can also be successfully patented.

6.5. Effective forecast of systems

6.5.1. Using Evolution Trees in forecast

A man permanently forecasts his actions. It may be the simplest forecast, for example, when crossing a street one assesses the most important factors: whether the green light is on, whether the cars have stopped, whether the road is slippery and forecasts a successful crossing of the street.

It can also be a complicated long-term prediction when events of a very remote future are to be planned. Then we have to take into account and analyze a lot of factors, build models of possible courses of events and take decisions based on the analysis of these models. Some people do it better and they are considered to be successful; others are not so good at forecasting and events «fall upon» them all of a sudden without leaving them time to respond adequately to them.

Of particular importance is accurate forecast in the innovation activities of a company. To create a promising product, it is necessary to maximally accurately forecast the development of analogous products, technical systems. If an idea of a new competitive product has been found, the company will be able to hold its advanced position among competitors. Any engineer, say nothing of an inventor, needs to know the forecast methods and be able to see several steps forward, because no high-quality creative work is possible without such an ability. Science and technology cannot be developed without attempts to dip into the future, without clear understanding what problems are the most urgent today and how they can be solved tomorrow, on which directions the maximum efforts should be focused. The entire variety of forecast methods may be reduced to two main ones [6.9]: *qualitative and quantitative*.

The *quantitative* forecast method has been considered next door to the only one possible for a long time. It is based on the extrapolation of the already known trends and models. The advantages of the extrapolation method are the well-studied models being used and possibility (as a rule) of quantitative assessment.

For example, to forecast how much fuel a medium car will consume per 100 km

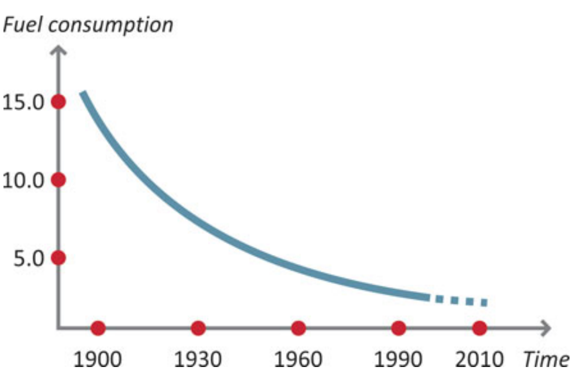


Fig. 6.51. Forecasting by single extrapolation

in 2010, we can trace the tendency to increase (or decrease) in this characteristic for a long period of time, present the collected information in the form of a diagram (Fig. 7.15) and extend the fuel consumption curve to the date of interest. According to the quantitative forecast, the fuel consumption of a car in 2010 will be 2 to 5 liters. The probability frames may be narrowed considerably by turning to experts, but their opinions may be subjective.

The quantitative method may be considered more or less satisfactory for a near-future forecast of some characteristics of technical systems: production volume, displacement tonnage, aircraft velocity, etc. This method works well within the framework of one fundamentally differing system evolution stage.

For example, in 1965 Gordon Moore made a forecast according to which the density of elements on integrated circuits had to double every 1.5 years. This forecast has been valid for 35 years so far but, according to many experts, it may soon become invalid. The problem is that silicon microcircuits have practically reached their potential and new technologies have emerged capable of essentially changing the entire situation. Even Moore himself admits that his forecast can lose effect after 2017.

It is very important to discern a qualitative transition from one type of technical system to another. The technology history knows many cases of underestimation of the importance of qualitative transitions. The textbook example is the statement made at the end of the 19th century. The forecast of London transport system development was as follows. The number of carriages was expected to grow manifold so that all the streets would be covered with horse manure up to the first floor windows. The qualitative transition — the invention of a car — that was not taken into account while history turned that forecast into a joke.

Among such mistakes include forward-looking statements, and the following:

- «Heavier-than-air flying machines are impossible!»

  (Lord Kelvin, President of the Royal Society of England, 1895)
- «640 Kilobyte ought to be enough for anybody.»

  (Bill Gates, 1981)
- «Computers in the future may weigh no more than 1.5 tons.»

  (Popular Mechanics Magazine, 1949)
- «Drilling for oil? Do you mean drilling into ground for oil? You are crazy!»

  (The principle of the drilling company whom the famous geologist Edwin Drake wanted to attract to his project).

The quantitative forecast does not satisfy the requirements of the inventive practice, because qualitative transformations are the result of accumulating quantitative transformations of technical systems. This is just what an inventor is interested in: how will a forecasted technical system change? Which subsystems will undergo maximum changes? What contradictions can arise as a result of these changes and what problems will they have to solve?

A *qualitative* forecast should primarily give an idea of possible fundamental changes in a system, forecast the approach and essence of a qualitative leap in its evolution. However, it is also important that the forecast be built not only on the subjective experience of experts but also on objective criteria. In this connection, forecasting based on objective laws of technology evolution used in TRIZ would look advantageous. Such a forecast is more correct because it is based on a complex analysis of specific system evolution from the standpoint of its conformity with the evolution laws of technical systems.

A traditional forecast based on experts' evaluations predicts change of some mechanisms of a system but does not often determine how this change will be achieved. A forecast made by using TRIZ offers specific technical solutions, allows formulating an integral and relevant conceptual model of a new generation product or process. Thus, a prognostic solution may be obtained by using all the TRIZ tools; moreover, any solution obtained at the invention level is prognostic.

A patent application is neither more nor less than a description of a new system modification having advantages over its prototype. When filing an application or developing a new solution, it is not clear whether the given system will take this way of evolution or the obtained patent will remain unclaimed and will be but an attempt of failed forecast. Accordingly, exclusively important is increasing the validity of forecasts in the evolution of technical systems [6.10].

Forecasting the evolution of some system eventually means building a temporal sequence of its modifications. If we managed to do this along the whole length of the time axis — both in the past and in the future — the entire evolution of the system would be spread before the eyes. Placing the already existing versions of a system on the time axis is simple enough while understanding which of its versions will come tomorrow, in a week or in a month suggests some degree of probability.

Suppose, the versions of the system under consideration are arranged according to the creation time (Fig. 6.52). Moving along the time axis will inevitably bring us to the day when we do not know what system's version is going to be next. This is similar to going through train cars. When

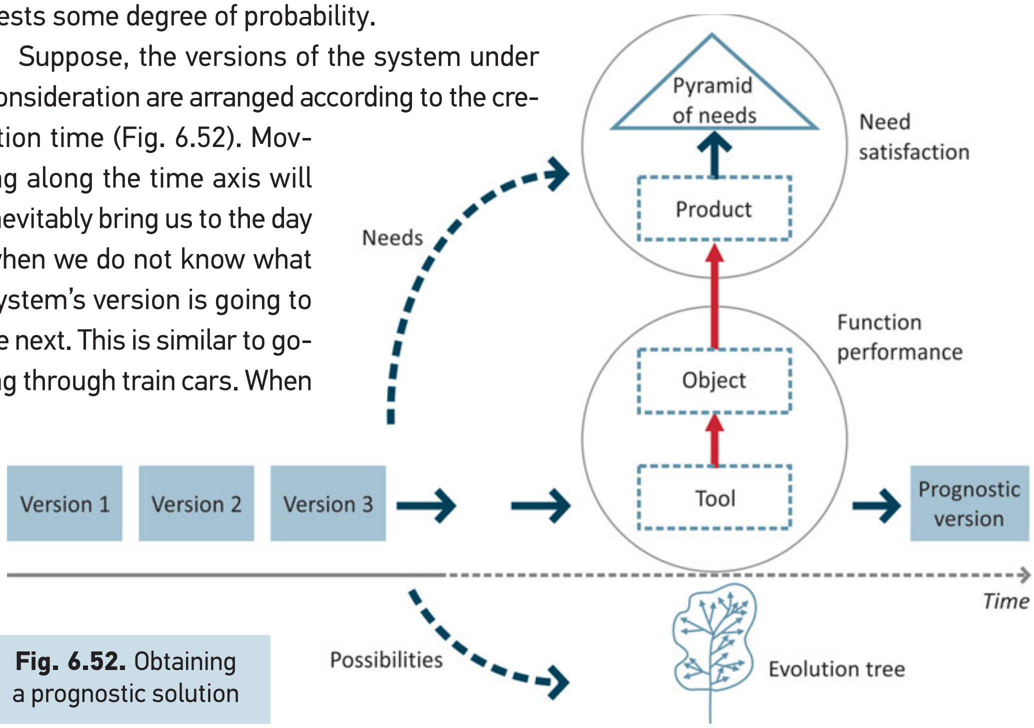


Fig. 6.52. Obtaining a prognostic solution

the doors between the cars are open, we easily go from one car to another and can examine the structure of each of them in detail. But suddenly we come to a locked door (today) and do not know what is behind this door — another car or something else.

To understand this, it is necessary to leave the frameworks of the temporal evolution model of the system and analyze its possible versions by using some other models. Such models may be some structure describing human needs or the Evolution Tree visualizing the system's versions. Using these models can satisfy these needs. A successful forecast may be defined as finding the optimal balance between human needs and technological capabilities of satisfying these needs.

When forecasting its evolution, any system may be presented as a means of satisfying some needs. The thought that a man needs not objects but *functions* that can be performed by means of these objects and *products* obtained as a result of performing these functions has become very popular in these latter decades [4.46]. We buy not goods but the ability of these goods to provide some property, give a product that satisfies some of our needs. The goods itself is but a means for obtaining such a product.

There exist many motivation theories that analyze and structure human needs making a man act in this way or another. The authors of many theories — McClelland, Herzberg, Maslow and others — proceed from the premise that a man has needs when he feels the lack of something in a physiological or psychological aspect. Accordingly, needs are also different: primary (physiological) and secondary (psychological).

Very popular is Maslow's classification of needs [6.11] illustrated by the so-called Maslow's pyramid and comprising:

- self-actualization needs,
- esteem, social recognition and social status needs,
- social needs,
- safety needs,
- physiological needs.

The pyramid is built in accordance with the hierarchy of needs. Maslow thinks that a man will primarily tend to satisfy the most important, vital needs — physiological, as well as safety needs. As soon as he succeeds in satisfying this type of needs, they cease to be the motivator giving way to the needs that are next in importance, for example, social needs and social recognition. The top of the pyramid is occupied by the need for self-actualization.

All human needs are satisfied by using some objects that are formed as the product of operation of some systems (see Fig. 6.52). For example, the need for communication is satisfied, among other methods, by using a sheet of paper with letters written on it.

As follows from the diagram given in Fig. 6.52, a system will be effective and competitive if it finds itself at the meeting point of two vectors. One of the vectors symbolizes users' needs while the other one the action of the system that produces the product satisfying these needs. The more ideal the system and the less expensive and of a better quality the product it produces, the fuller will be the satisfaction of the users' needs within a given market niche.

Accordingly, the system improvement should be directed toward the cheapening of the product, improving its quality and consumer properties.

The minimum composition of a system producing some product is a combination of a working tool and a work object. Of course, to provide their interaction, it is necessary to introduce some additional parts into a real system, but its function is performed by these elements. In forecasting, it is necessary to understand how these elements should change in the future. A new version of a system may be obtained by replacing one of these components.

To make a forecast, it is necessary to imagine the main promising embodiments of this system. An object to be worked to satisfy some need has been already determined. Hence a forecast will consist in determining which of the most promising versions of the working tool will replace the existing ones.

According to the diagram presented in Fig. 6.52 the following algorithm for obtaining promising prognostic modifications of a system is visible.

- determining the user's needs,
- determining the work object of a system,
- determining the product to be obtained by the action on the work object,
- transforming the work object (if necessary),
- determining the working tool of a system,
- transforming the working tool.

For effective transformation of the working tool and work object, it is necessary to have a sufficient number of their versions in view, ideally all the basic embodiments of these components of the system.

6.5.2. Determining the main versions of system's components

As was mentioned above, an advanced method of search for new modifications of the system's components is qualitative and technological extrapolation.

As distinct from quantitative extrapolation which analyzes a certain quantitative index, this deals with the evolution of the system's composition and structure. The evolution of a selected object is usually studied from the very first to the most advanced versions, then the revealed trend is extended in accordance with one of the evolution laws of technical systems by building a certain evolution pattern of its components. Further a prognostic solution is built, that is, a probable version of the object under examination is determined.

The disadvantage of this method is the narrow field of view, impossibility to visualize all basic transformation versions of a component (Fig. 6.53, a). This looks like a strictly determined route to be followed by wildlife area visitors rather than a free walk in a recreation park.

We believe that it would be useful to have not one variant of such a «route» but an entire map of possible ways, so that we can see all the possibilities simultaneously (Fig. 6.53, b). Such a map is neither more nor less than a field of search for a solution to a prognostic problem. The efficiency of the prognosticator's work grows abruptly if he has a structured and

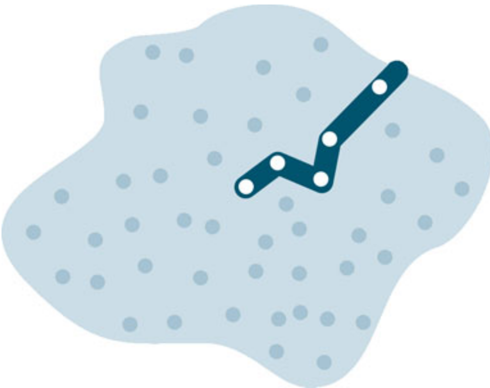


Fig.6.53, a. Forecasting by extrapolating from a single evolution pattern

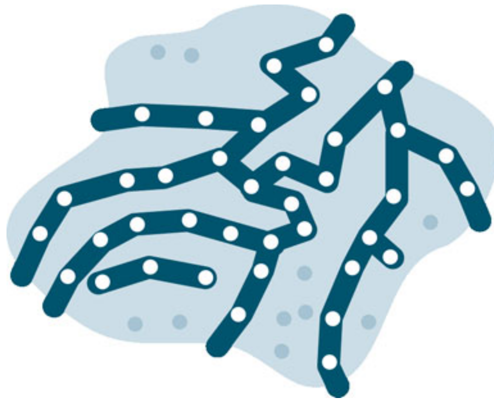


Fig.6.53, b. Forecasting by extrapolating from a set of evolution patterns

visualized collection of information at hand. The Evolution Tree suits excellently the role of such a map because it presents, in a regular order, all the existing versions of an object under examination. In addition, significant information on the absent, probable versions of the object may be obtained.

The mechanism used in determining the promising versions of an object under analysis is described in section 6.1. New versions of the object may be found by comparing our real Evolution Tree with the evolution patterns constituting the basic Evolution Tree [6.12]. Properly speaking, the marking-out of the information field and presentation of the «portraits» of still nonexistent system's versions take place. This allows clearly seeing which components can be used in the system and focusing on selecting the best ones. In this case, the objective approach to building the Evolution Tree allows us to not only focus on ordinary, trivial versions, but to produce new, most unexpected ones.

In addition to the basic Evolution Tree, other information structures may be used as «landmarks» for the perfection of a real technical object. Much information on a system under examination may be provided by the structural analogy method considered in p. 6.3.2. The method consists in the use of the closest analogies between the system components, which occur when these components have so-called structural similarity. With such an approach, we can analyze a selected structure and search for its transformation versions using analogy with the transformations of another, better developed structure.

In the world of technology, systems have different degrees of development. For example, the best specialists, scientific potential and huge financial resources are employed in the automotive industry or creation of space technologies. Corresponding technical systems and their components are maximally developed. Studying the evolution history of such systems allows building very detailed evolution patterns and then trees of their components. Just like basic Evolution Trees, these will not have any omitted steps. At the same time, specific Trees of well-developed technical objects and systems have a great advantage over basic Trees as a check point for comparison. The fact is that all transformations presented in specific Trees are more visual and informative than those of the abstract transformations of a basic Tree.

Alongside well-developed fields of technology, there exist systems that have always involved less scientific and financial resources. As a result, the evolution patterns of their components have holes. It is not always possible to extend such a pattern and to make a well-grounded forecast. A very effective forecasting method is comparing two Evolution Trees of structurally similar systems — a forecasted and a well-developed one. More simply, to invent a new frying pan, it would be useful to search for analogy in the design of a propulsive nozzle. In both cases, we have a shell heated on one side, which affords ground for comparing these two systems.

Comparing the Evolution Trees of two real systems not only gives new, interesting versions of a system under consideration. Analyzing a well-developed system gives a clear idea of what it was changing for in the process of evolution and what contradictions were resolved as a result.

A forecast is always an equation in many unknowns. Using Evolution Trees for visualizing probable system versions makes it possible to determine some of these unknowns. This eventually simplifies the forecasting procedure and makes a forecast fuller and more precise. In a «pre-marked» information field, a researcher sees a clear picture, representing all basic system versions. With this information, he can fully focus on solving a reverse problem and analyzing the system versions trying to find an answer to the following questions:

- What one or another version of the system may be needed for?
- Which parameters of a new system will change and how will they change?
- What advantages does any version have?
- What disadvantages does it have?
- How will the property of the product produced by the system change?
- Will the new version be more expensive or less expensive?
- How will the system's supersystem change?

Thus, a normal research work is meant here: study of the models of a new hypothetical system.

6.5.3. Examples of prognostic solutions from the display Evolution Tree

Building and analyzing the display Evolution Tree resulted in a number of prognostic solutions some of which are given below. Worth noting is the following fact: the display Evolution Tree was built as far back as 2001-2002. And what about prognostic solutions? Have any of them come true at least in a small measure? Have displays evolved in the directions indicated in the Tree?

Certainly, the display Evolution Tree shown in this book is illustrative to some extent and based on open sources. To make a more accurate forecast, much research into closed information sources is needed as well as the analysis of special collections. Such work provides a strong result but no company will agree to publish the results of some real prognostic research work because it has a great advantage over a competitor.

So let us have a look at our simple Evolution Tree of a display which in fact only differs from a big prognostic map by the study depth.

Building and analyzing the display Evolution Trees resulted in a series of prognostic solutions, some of which are given below.

A) Forecast. Increasing the informativity of displays for the blind and weak-sighted people (2001)

Need — obtaining maximum information.
Work object — sense organs (excluding sense of sight).
Product — sense organs subjected to information action.
Working tool — a needle display.

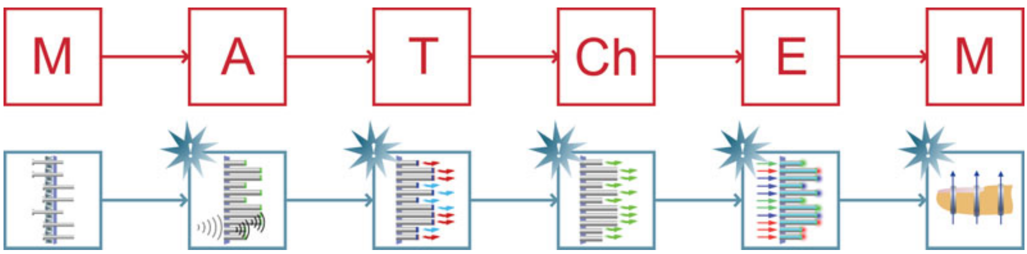


Fig. 6.54. Analyzing the needle display

The version available on the Evolution Tree is a mechanical needle display. Transforming it by means of the MATChEM operator (see p. 3.3.2) resulted in the following prognostic versions (Fig. 6.54):

- Acoustic field — display with local supply of sound directly through the needles.
- Thermal field — «color» display for the blind where a color picture is formed by heating and cooling the needle ends.
- Chemical field — an odor-emitting display
- Electric field — a color raised display with electric discharges on the needles.
- Magnetic field — a device directing user's fingers to a necessary place on the screen.

A) Implementation (2005)

As far as the general trend is concerned, the evolution of displays for the blind is directed toward providing a complex action on the skin for information transmission. For example, in 2006, the Forehead Retina System was developed which employs electrical pulses of a certain frequency for image transmission. The research results proved that pulses of different frequencies are perceived as different colors.

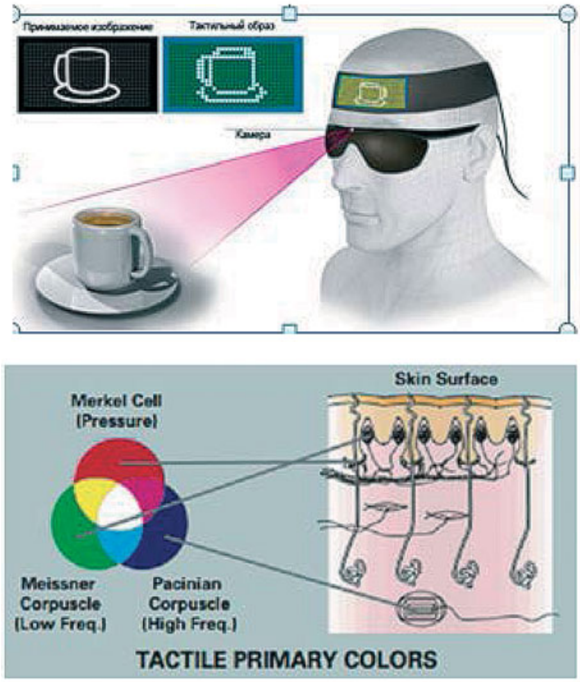


Fig. 6.55. Producing an image by the Forehead Retina system technology

B) Forecast. Displays forming a real movable copy of a presented object (2002)

Need — producing a maximally real image.

Work object — sense organs (first of all vision).

Product — sense organs receiving a realistic image.

Working tool — a display.

Analyzing the «Dynamization of double screen display» pattern (p. 5.5) and comparing it with the basic «Dynamization» pattern proves that the specific pattern is uncompleted, it lacks two transformation versions — «Use of field» and «System having segmented parts» (Fig. 6.56).

Using a force field, for example, of a field of active polymers jointly with a flexible screen allows production of a maximally realistic 3D image. Such a display is an executive mechanism, robot that may be similar to any object or man. Such a device may also be obtained by analyzing the display segmentation evolution patterns, for example, by forming an image from «display fibers».

B) Implementation...

Strange to say but we have not found any information about such devices in technical literature whereas the idea of such a machine, for instance, a robot having a face shaped into

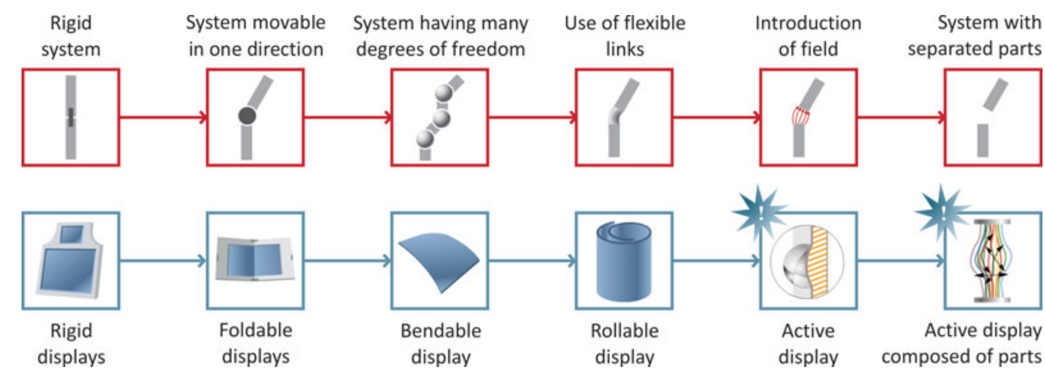


Fig. 6.56. Analyzing a dynamic display

a flexible display, is obvious. This robot could have a considerable advantage over its faceless «fellows» by demonstrating an absolutely human (or non-human) complexion, play of colors, etc. I believe we will soon hear about another break-through in this direction.

C) Forecast. Dynamic screen display (2001)

One of the ways of adding vividness and visuality to an image is transition to volume. Yet 3D displays are rather expensive and imperfect, while a 3D image can be imitated on a flat screen.

Need — obtaining a visual pseudo-3D image
Processed object — sense organs (primarily, vision).
Product — sense organs which receive a realistic image.
Tool — display.

Analyzing the evolution pattern «Double-screen display dynamization» (see the last but one step of the pattern in Fig. 6.56) and comparing it with the basic «Dynamization» pattern shows that the display evolution pattern misses the «Using a field» version. Here a display version having an active bas-relief may be located, having the surface shape that can change with the image. Following the pattern logic, it can be represented by a flexible twistable screen the shape of which is changed by a magnetic field (Fig. 6.57).

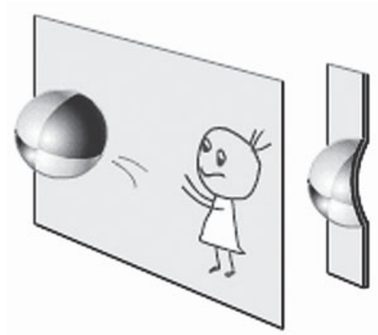


Fig. 6.57. Display having a dynamic micro-relief

C) Implementation (2006)

In 2006 and 2007, there appeared information in the press about displays having active microrelief (Fig. 6.58). One of them was invented in Japan by Prof. Yoichiro Kavaguchi. In the Gemotion display, image is projected onto a flexible screen. The screen shape is changed by as few as 72 pneumo-cylinders so it is too early to speak about transmitting some sensible information.

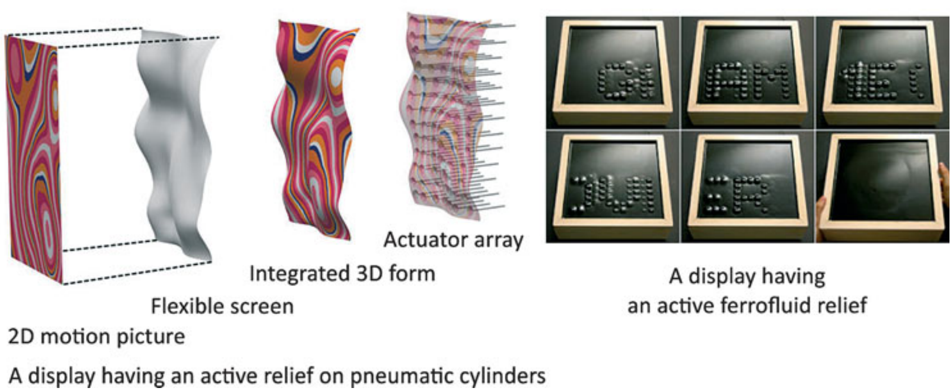


Fig. 6.58. Displays with a dynamic microrelief of a screen

The second display is a tray filled with ferrofluid. The action of magnetic field changes the ferrofluid surface shape.

Further evolution of active micro-relief displays will proceed in accordance with the diagram given in Fig. 5.17: surface segmentation into increasingly small pixels, coordination of their position, shape, size, etc, additional dynamization of each screen portion and provision of good controllability. This may result in the creation of an operable screen, for example, for advertising applications.

D) Forecast: Notebook with a detachable display (2001)

Need — obtaining a maximally movable notebook display.
Processed object — sense organs (primarily vision).
Product — sense organs which receive a realistic image.
Tool — display.

Analyzing the «Display dynamization» pattern (p. 5.4.6) and comparing it with the basic «Dynamization» patterns shows that the given pattern is not complete, it misses the last transformation — «System with segmented parts» (Fig. 6.59).

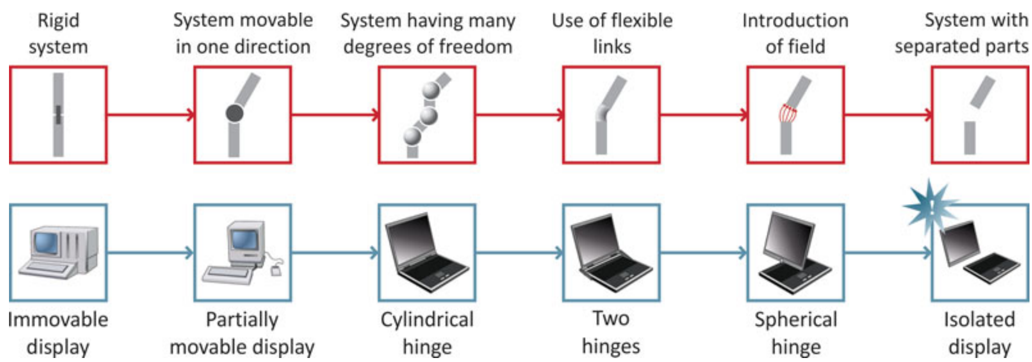


Fig. 6.59. Dynamic display analysis

Fig. 6.60. Notebook display detachable from the keyboard

This step offers the following prompt: making the notebook display detachable from the keyboard. Such a display can be hung on a wall and used for mini-presentation or be placed in any other convenient place.

D) Implementation (2006)

In 2006, Samsung announced the creation of a new product: Samsung M70 notebook having a screen that could be detached from the keyboard and work by itself (Fig. 6.60).

E) Forecast: Ideal display (2001)

An ideal display is a method of inducing visual sensations directly in the consciousness (Fig. 6.61). The entire evolution history of displays and microelectronic devices is the tendency toward minimization, idealization. The top of the Evolution Tree is occupied by the most ideal displays. However, we would like to make them even more ideal.

Need — obtaining visual and other information without any devices.

Work object — consciousness.

Product — brain that received visual information.

Working tool — display.

Comparing the «Trimming a display» pattern with the basic «Trimming» pattern shows that uncovered is the final display transformation version. The essence of this transformation in the basic pattern is the use of the ideal object. What display will be the most ideal?

It is apparently a display that is absent while its function is performed, i.e. the most suitable will be a display version that induces an image directly in the user's mind, bypassing eyes and other sense organs (section 3.4.7).

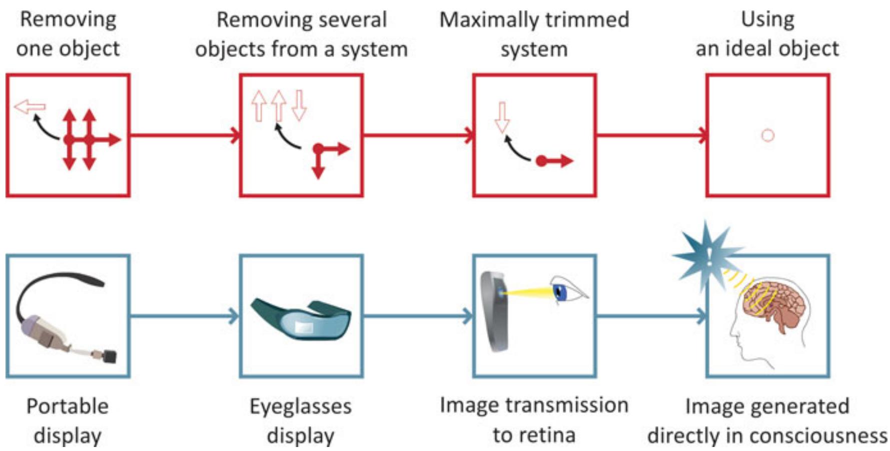


Fig. 6.61. Analyzing the ideal display

E) Implementation (2003)

Many companies felt interest in the creation of an ideal display, for example, Takara Co, which developed a «dream machine» described in p. 5. However, all those efforts were not too serious until Sony reported that it worked on calling forth visual and other images directly in the neuronic chains of a man.

In 2003 and 2004, the company received two patents on this method. Being familiar with the system of patenting novelties at Japanese companies, one may safely say that the works on the new technology are at full speed.

It should be noted that prognostic versions of a system often exist in some form but specialists may have no full understanding that it is just where the key evolution direction of the system is found. The analysis carried out by using the Evolution Tree allows better understanding of the system evolution logic and forecasting which versions exactly (existing or newly created) are the most promising from the technological evolution viewpoint.

6.5.3. Computer approach to the analysis of complicated systems

The effectiveness of prognostic analysis by means of the Evolution Trees can be enhanced by using a computer.

The Evolution Tree only describes transformations of one component whereas function performance requires at least two components, i.e. two Trees should be used for correctly describing the interaction of two components. The forecast effectiveness is considerably enhanced by taking into account the transformation of the surrounding where the action is performed, i.e. one more Tree is needed. Hence to describe performance versions of only one function, it is necessary to build at least three Trees and examine them in the aggregate changing each time one or several components — participants of the operational zone.

At the early stage of the method development, we used the Evolution Tree for specific projects in the following manner: the tree structure was traced on huge sheets of paper. Individual steps were drawn and described on small sheets which were then glued down in a preset order along guiding lines. As a result, we had huge «sheets» about the size and complexity of real trees. That looked very spectacular, but bringing such a poster to some other place, for example, to a presentation hall, was a real adventure. It was also difficult to record all the obtained function performance versions, because it was necessary to describe all combinations and draw by hand produced system models.

It is quite natural that from considerations of visualization, efficiency, mobility and information conveyance, there arose a desire to computerize this process. Based on the Evolution Trees, we have developed an idea and prototype of a computer program that helps a designer quickly obtain, in a visual form, a large number of qualitatively different methods of performing different functions. Each method may be used to obtain an elementary solution concept while a combination of the methods can give a complex concept of building a technical system being designed. After producing a number of concepts, a designer can start analyzing them and selecting the most promising ones according to the recommended criteria.

The main part of the program (Fig. 6.62) is its data base which comprises a large number of images of object transformation versions and textual descriptions of these versions, taken from the Basic Evolution Tree. To the user's command, these versions can be displayed in the operation window where the elements performing the function under examination are «assembled». Transformation versions are selected by means of a special navigator built on the basis of the Basic Evolution Tree [4.49].

As a result, a designer obtains a number of various concepts including those already existing — either realized or patented — as well as new ones, synthesized by the designer himself. Because designers are usually well aware of the situation in their sphere of activities, it is not difficult to mark the existing solutions on the Evolution Tree of the element under consideration.

The program is aimed at acquainting a designer with all basic versions of performing the needed function so that he could select the most suitable one from all known realization methods and use it in the system. Omitted steps and combinations of existing solutions form a search field of new concepts.

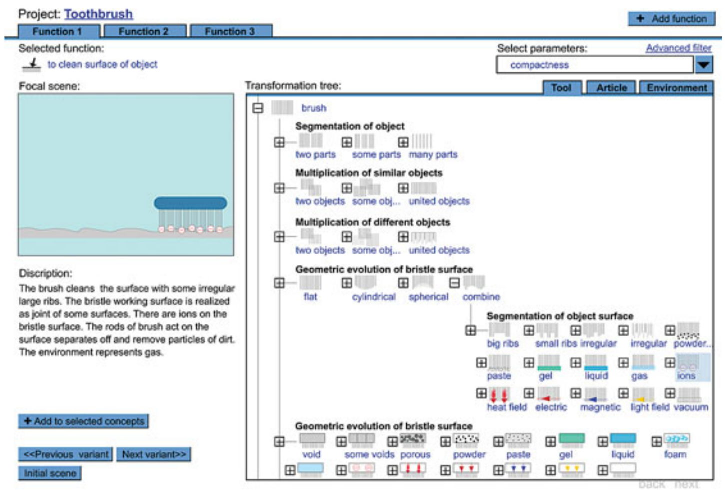


Fig. 6.62. Fragment of the Evolution Tree for the tool «Brush»

The program suggests automatic selection of concepts according to the parameters being improved or, in case of independent selection of concepts, gives certain recommendations. Proceeding from the analysis purposes, available resources and operational parameters to be improved, the designer chooses the most promising concepts.

A possibility is provided to build functional chains while analyzing complicated systems (Fig. 6.63).

At present, there exists a well-developed prototype of this program. The program like Concept Generator can become a useful supplement to any automated programming system such as AutoCAD, Pro Engineering or any other tool of this sort.

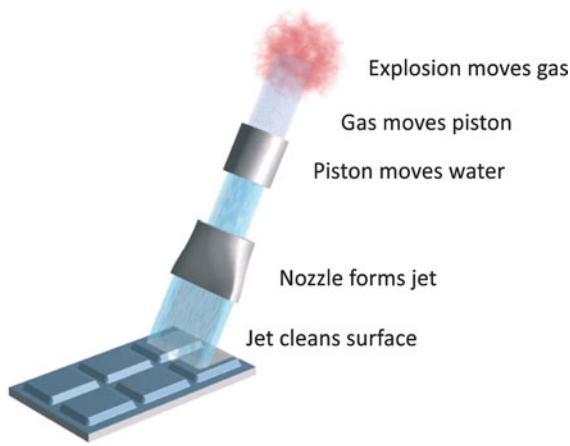


Fig. 6.63. Building a functional chain

The Evolution Tree based approach allows a designer to create a virtual technical world — the world of function-performing models which can be used for building any technical system.

Author’s afterword

Our world is becoming increasingly dynamic and is changing visibly. Unit recently such things as cellular communication, integrated circuit, nano-technologies and Internet were unknown to us. Now we have to learn to live in the permanently changing world and, consequently, to solve a large number of permanently arising problems. It is difficult to understand what challenges the life will through out tomorrow and what problems we will have to solve in near future, but we need to know this.

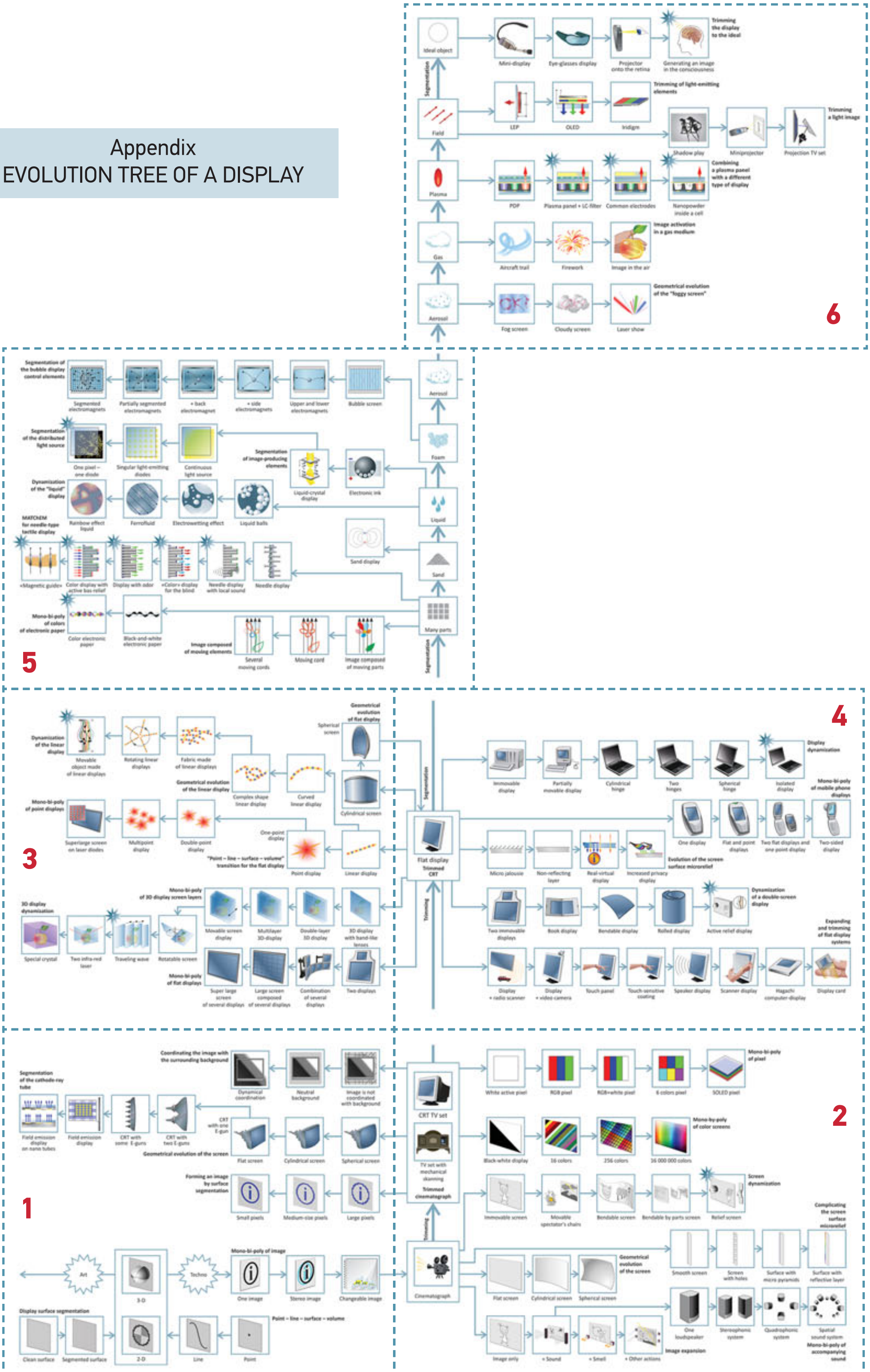
The Theory of Inventive Problem Solving is currently the most promising tool for improving machinery, removing technical, social and marketing contradictions. Many researchers work on making this methodology even more effective. One of the possible TRIZ development directions is using Evolution Trees for work with clusters of information units. The Evolution Tree based on our experience in practical problem solving is an excellent visualizer of possible versions of a system, which shows an engineer the direction of the system’s evolution and available improvement reserves.

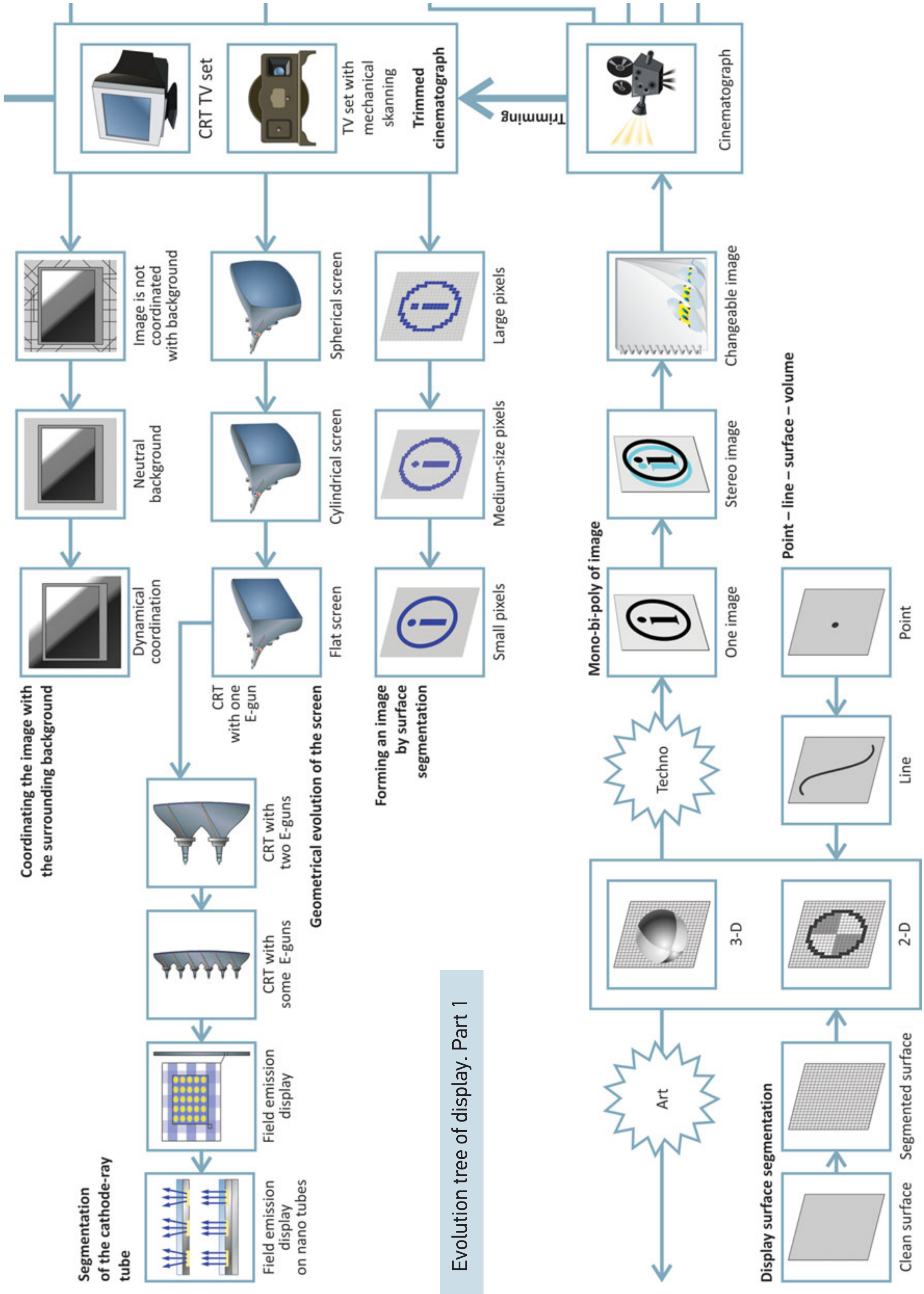
I hope that this book will prove interesting and useful, help acquire new skills for use in problem solving, and allow better understanding of the essence of the process of work on a problem.

I invite all readers to visit our web site «Generator» (http://www.gnrtr.com) and take part in a discussion about the book on the site’s forum. The author is ready to answer all questions pertaining to the practical solving of inventive problems, system evolution forecast and circumvention of patents.

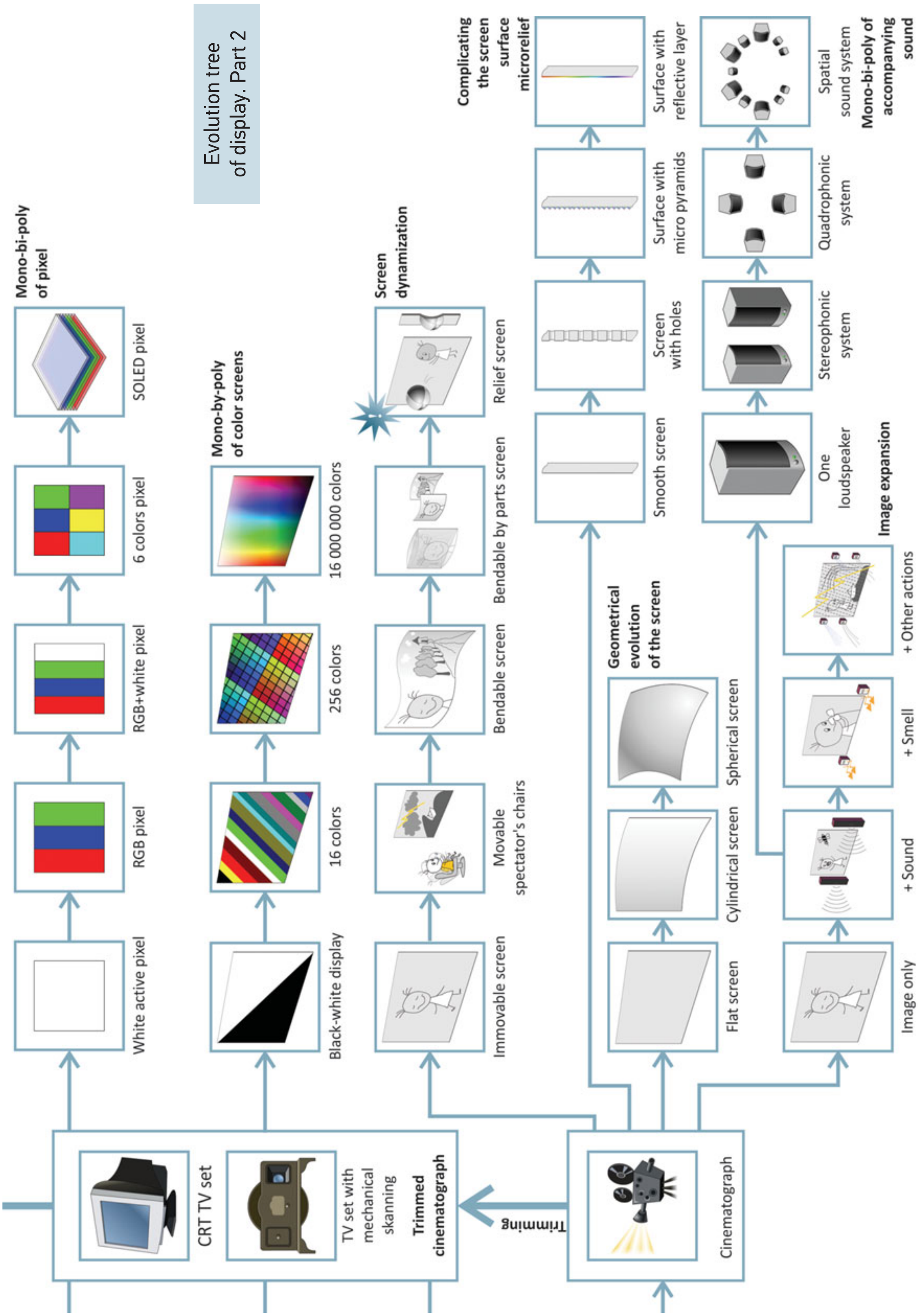
I hope wholeheartedly for the continuation of the dialog.

Appendix
EVOLUTION TREE OF A DISPLAY

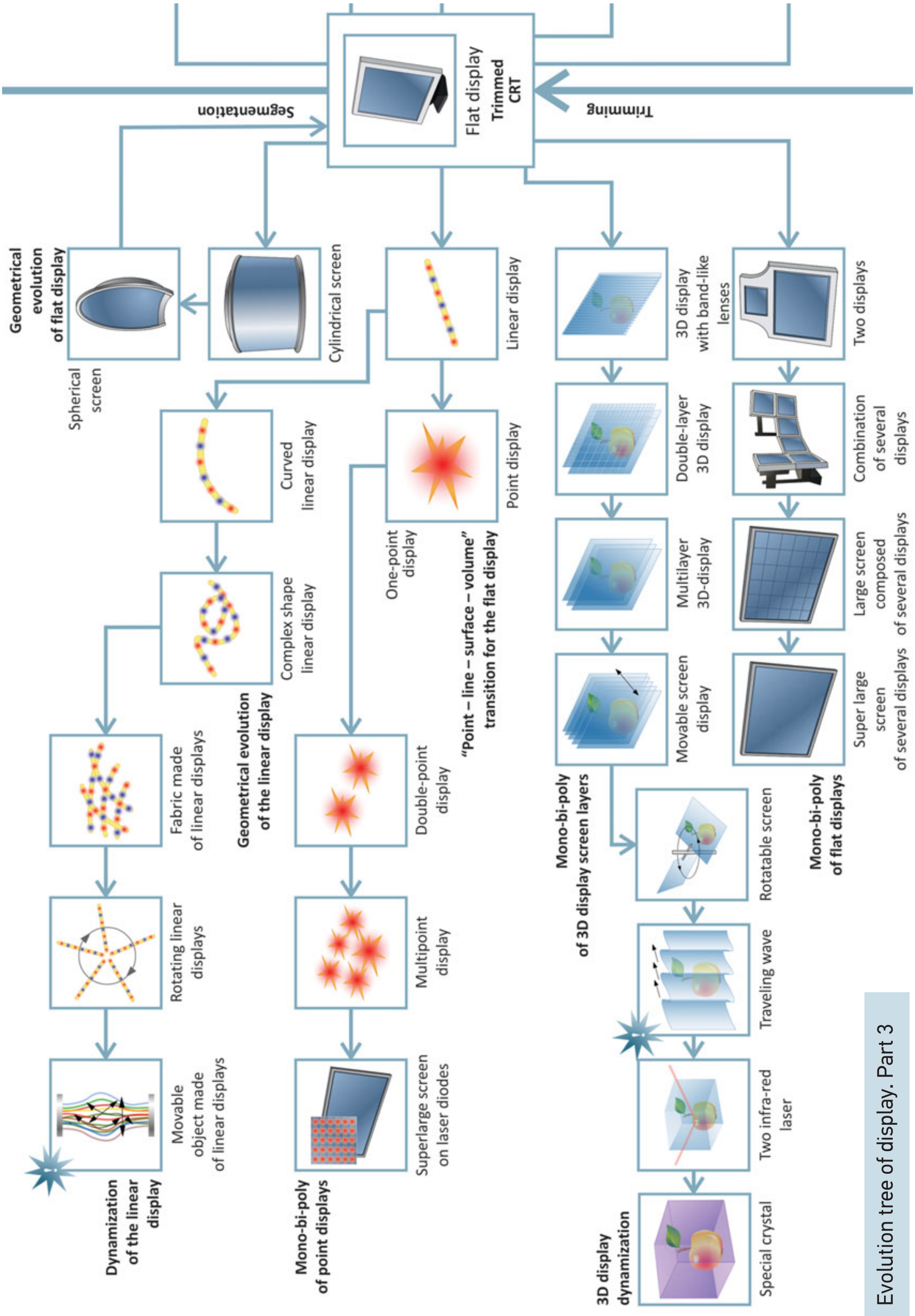




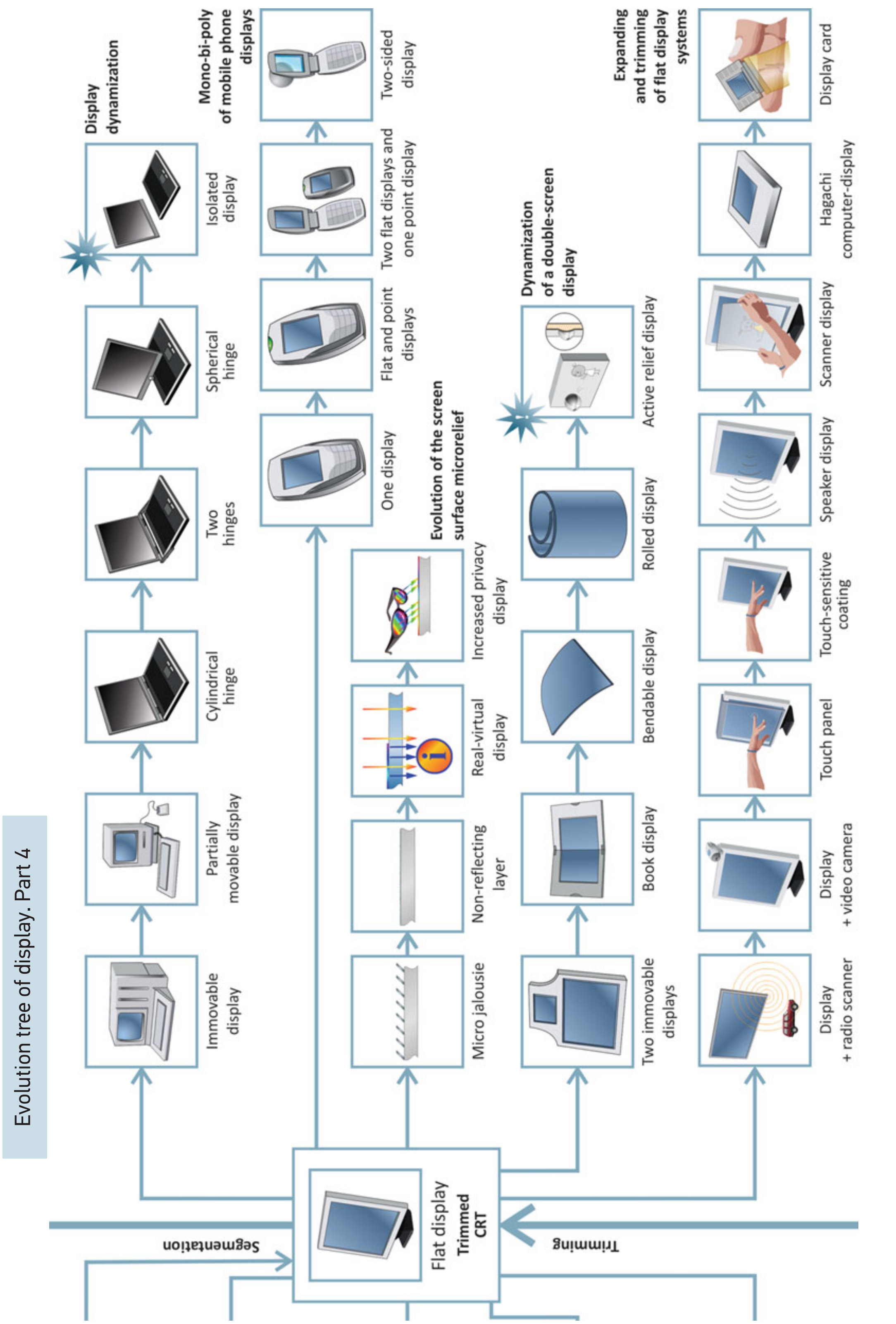
Evolution tree of display. Part 1



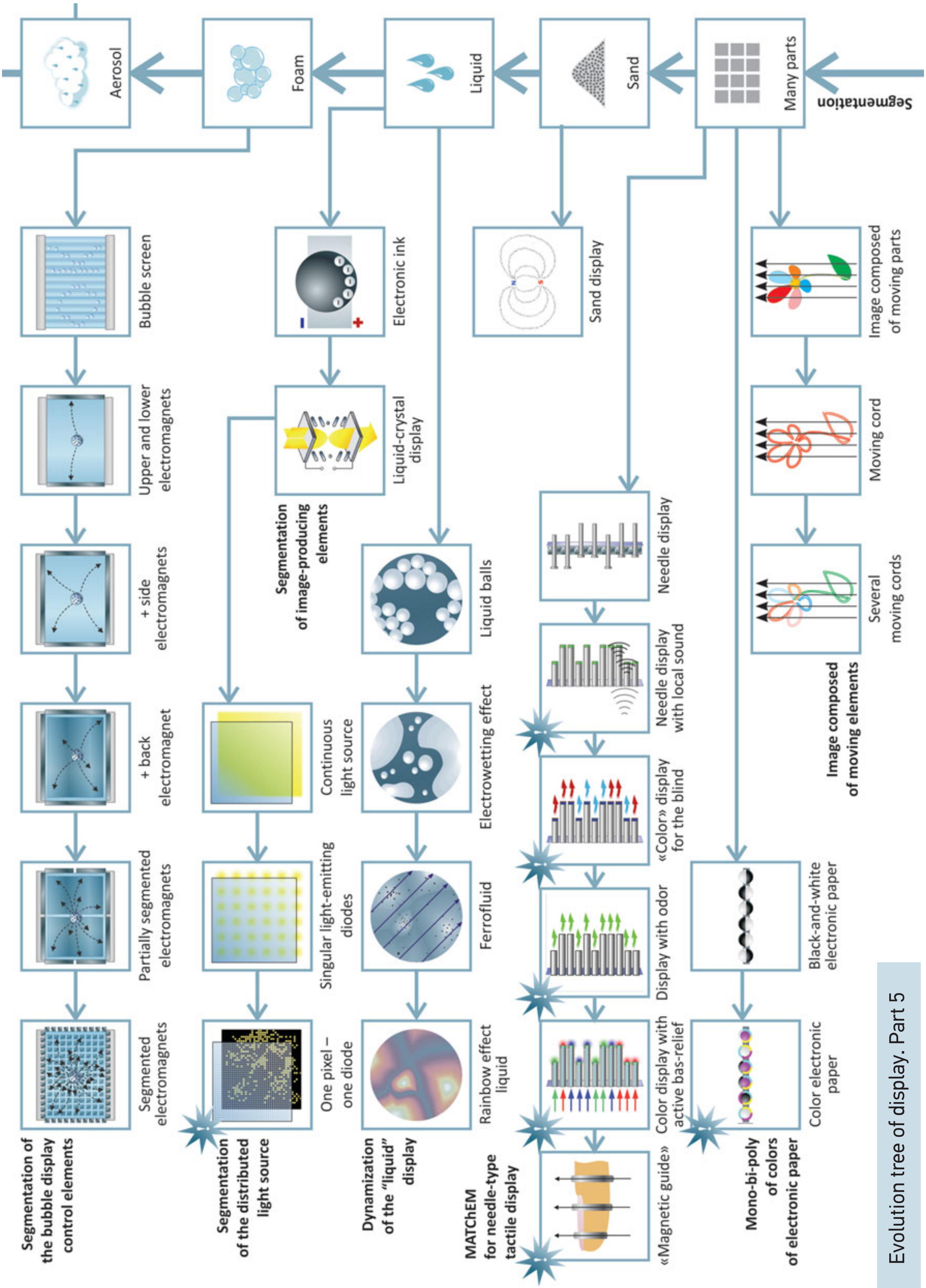
Evolution tree of display. Part 2
Mono-bi-poly of pixel
White active pixel
RGB pixel
RGB+white pixel
6 colors pixel
SOLED pixel
Mono-by-poly of color screens
Black-white display
16 colors
256 colors
16 000 000 colors
Screen dynamization
Immovable screen
Movable spectator's chairs
Bendable screen
Bendable by parts screen
Relief screen
Complicating the screen surface microrelief
Smooth screen
Screen with holes
Surface with micro pyramids
Surface with reflective layer
Geometrical evolution of the screen
Flat screen
Cylindrical screen
Spherical screen
Image expansion
Image only
+ Sound
+ Smell
+ Other actions
One loudspeaker
Stereophonic system
Quadrophonic system
Spatial sound system
Mono-bi-poly of accompanying sound
CRT TV set
TV set with mechanical skanning
Trimmed cinematograph
Trimming
Cinematograph



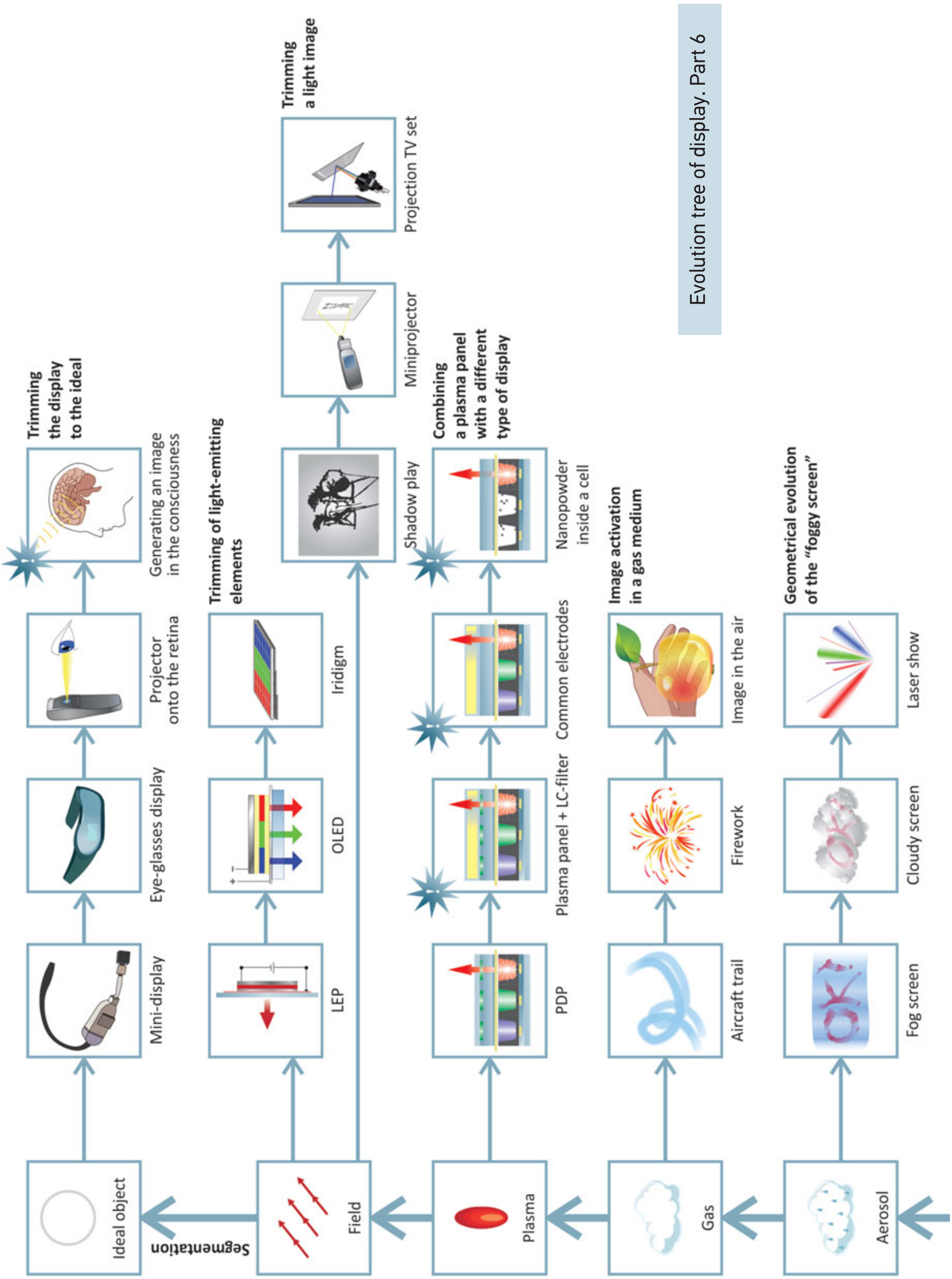
Evolution tree of display. Part 3



Evolution tree of display. Part 4
Segmentation
Trimming
Flat display
Trimmed CRT
Display dynamization
Immovable display
Partially movable display
Cylindrical hinge
Two hinges
Spherical hinge
Isolated display
Mono-bi-poly of mobile phone displays
One display
Flat and point displays
Two flat displays and one point display
Two-sided display
Evolution of the screen surface microrelief
Micro jalousie
Non-reflecting layer
Real-virtual display
Increased privacy display
Dynamization of a double-screen display
Two immovable displays
Book display
Bendable display
Rolled display
Active relief display
Expanding and trimming of flat display systems
Display + radio scanner
Display + video camera
Touch panel
Touch-sensitive coating
Speaker display
Scanner display
Hagachi computer-display
Display card



Evolution tree of display. Part 5


Evolution tree of display. Part 6
Ideal object
Segmentation
Field
Plasma
Gas
Aerosol
Trimming the display to the ideal
Mini-display
Eye-glasses display
Projector onto the retina
Generating an image in the consciousness
Trimming of light-emitting elements
LEP
OLED
Iridigm
Shadow play
Miniprojector
Projection TV set
Trimming a light image
Combining a plasma panel with a different type of display
PDP
Plasma panel + LC-filter
Common electrodes
Nanopowder inside a cell
Image activation in a gas medium
Aircraft trail
Firework
Image in the air
Geometrical evolution of the "foggy screen"
Fog screen
Cloudy screen
Laser show

Dear reader,
you can download
the poster in high resolution
(60x90 cm) here:

www.target-invention.com

Literature

1.1. Severinets G. A., Novitskaya E. L. Bronze Horseman. // Web site Generator. http://www.gnrtr.com/problems/ru/p09. html /

1.2. Altshuller G. S. Basics of invention. — Voronezh: Central Black Earth Publishing, 1964.

1.3. Altshuller G. S., Shapiro R. B. On the psychology inventive creativity. // Questions of psychology. — 1956.— № 6.— S. 37–49.

1.4. Altshuller G. S. How to learn to invent. — Tambov: Tambov Book Publishing House, 1961.

1.5. Altshuller G. S. Algorithm invention. — Moscow: Moscow Worker, 1973.

1.6. Altshuller G. S. Creativity as an exact science. Theory of Inventive Problem Solving. — Moscow: Sov. radio, 1979.

1.7. Kudryavtsev A. V. Review of the methods of creating new technical solutions. — Moscow: USSR State Committee for Inventions and Discoveries 1988.

1.8. Altshuller G. S., etc. Search for New Ideas: from Insight to Technology. — Chisinau: Kartia Moldoveniaske. 1989.

1.9. FSA Handbook / Ed. M. G. Karpunin and B. I. Maydanchik. — Moscow: Finance and Statistics, 1988.

1.10. Devoino I., Skuratovich A. Idealization of Engineering Systems through Trimming: Engineering Systems as Technological Proceeding of TRIZ Future. 2002.— Strasbourg, 2002.

1.11. Gerasimov V. M., Litvin S. S. Accounting zakonomernostey razvitiya technology during the activity-based costing processes. // Web site OTSM-TRIZ. http://www.trizminsk. Org/e/216001.htm/

1.12. Skuratovich A. I. Typical objects super system for technical objects: Abstracts of the articles of the international scientific-practical seminar «Project IM-96». — Minsk, 1996.

1.13. Litvin S. S., Gerasimov V. G. Development of alternative systems by combining them in a super-system. // TRIZ. — 1990.

1.14. Altshuller G. S. Algorithm invention. — Moscow: Moscow Worker, 1973.— P. 221–224.

1.15. Altshuller G. S. Typical methods of eliminating technical contradictions. // Website Altshuller Foundation. http://www.altshuller.ru/

1.16. Altshuller G. S. On formalization of the process of drawing up a physical contradiction. // Website Altshuller Foundation. http://www.altshuller.ru/

1.17. Strizhakov A. V., Martirosov A. L., Kubarev A. E. Descriptive Geometry . — M.: Phoenix 2004.

1.18. Gorin Yu. Pointer application of physical effects and phenomena . // Science and Technology. — 1981–1982.

1.19. Hutsol A. F. Ranque effect: Methodological notes — http://ufn.ioc. ac.rU/archive/russian/lndex97.html#rel_4_1 99 7_180 /

1.20. Jean-Louis Naudin. The Coanda Effect. Created on September 26th, 1999 — JLN Labs — Last update April 4.2006.http://jnaudin.free.fr/html/coanda.htm /

1.21. Require an airfield, amphibious, high-efficiency, hi-tech aircraft novogo tipa «EKIP». // Website Concern EKIP. http://www.ekip-aviation-concern.com/rus-b/O.shtml/
1.22. Liquid armor. Popular Mechanics. http://www. popularmechanics.ru/part/?articleid=451&rubricid=7/
1.23. Regenerating polymer // Web site Generator. http://www. gnrtr. com/solutions/ru/s049. html/
1.24. Miracles nanotech. Part Two: // nanonadezhdy Website Membrane. http://www.membrana.ru/articles/technic/2002/01/31/181300.html/
1.25. Nanotechnology allowed to create a metallic rubber // Website Membrane, http://www.membrana.ru/lenta/?3515/
1.26. Timokhov V. I. Pro biological effects: To help teachers of biology / VI Timokhov. — Gomel 1993.
1.27. Altshuller G. S. process of inventive problem solving: the basic stages and mechanisms. — 1975. http://www.altshuller.ru/triz1.asp
1.28. Daring Formulas of Creativity: A Collection. — Petrozavodsk: Karelia, 1987.
1.29. Thread in the Labyrinth: A Collection. — Petrozavodsk: Karelia, 1988.
1.30. Rules of the game without rules: Collection. — Petrozavodsk: Karelia, 1989.
1.31. How to become a heretic: Collection. — Petrozavodsk: Karelia, 1990.
1.32. Chance for Adventure: Collector. — Petrozavodsk: Karelia, 1991.
1.33. Karpukhina S. I . The protection of intellectual property and patenting. — MA: Center for Economics and Marketing. 2002.

2.1. Altshuller G. S., Vertkin I. M. Llines increase «emptiness». http://www.altshuller. ru/triz/zrts5.asp/
2.2. Petrov V. M. laws of systems development // Website Trizland.http://www.trizland.ru/index. php/
2.3. P. Bauer. Aircraft innovative schemes / P. Bauer; per. from English. — Moscow: Mir, 1991.
2.4. Problems of creation of perspective aerospace / ed. V. G. Dmitriev. — Federal State Unitary Enterprise Central Aero-Hydrodynamic Institute. prof. N. E. Zhukovsky, 2005.
2.5. NASA experiencing first-transformer plane // Website Reporter. http://www.korrespondent.net/main/73346/
2.6. Parenchik G. I. Terminological aspects of TRIZ. Technical creativity: the theory, methodology, practice // Website methodologist, www.metodolog.ru/00702/00702.html /
2.7. Patent Law of the Russian Federation of September 23, 1992 // website of Rospatent. — Http://www.fips.ru/npdoc/LAW/PATlawOO.HTM/
2.8. Meleshenko S. Machinery and regularities of its development. — AL: Lenizdat, 1970.
2.9. Polovinkin A. I. Structure of laws and the development of technology. — Volgograd: FS. 1985.
2.10. Polovinkin A. I. Automation search design. — Moscow: Radio and Communications, 1981.
2.11. Goldovskiy B. I., Vainerman M. I. Rational Creative. — M.: River transport, 1990.

3.1. Shpakovsky N. A. Space mirror / NA Shpakovsky // Web site Generator. — http://www.gnrtr.com/
3.2. Podkatilin A. V. Do not throw money to problem: Effective Solutions. — M.: TRIZ-profi. 2005.
3.3. Shpakovsky N. A. One of the lines of propeller aircraft / NA Shpakovsky // Web site Generator. — Http://www.gnrtr.com/
3.4. Kizevich G. Principles and Theory of creativity every day. — Moscow: Williams 2004.
3.5. Belsky, I., «Improve your Thinking: Substance-Field Analysis», TRIZ4U, Melbourne, ISBN 978-0-9803293-0-8, 2007.

4.1. Kuznetsov S. V., Titov V. V. Classification: system-morphological approach /. — M.: Inventa, 1998.
4.2. Kudryavtsev A. V. Brilliant master Raymond Lully. // Website methodology 2006 // http://www.metodolog.ru/00110/00110.html /
4.3. Zwicky F. Discovery, Invention, Research through the Morphological Approach. — Toronto: The Macmillian Company, 1969.
4.4. Salamatov Y. P. System of laws of the art. // Chance for Adventure: Collector. — Petrozavodsk: Karelia, 1991.
4.5. Goldovskiy B. I. TRIZ in a rapidly changing world. On the different approaches to the subject of TRIZ. // TRIZ. — TZ — 1.92.— No. 3.
4.6. Polovinkin A. I. Fundamentals of engineering creativity: Proc. allowance for students of technical colleges. — M: Mechanical engineering, 1988.
4.7. Korolev V. A. Structure of needs. // Website Encyclopedia of TRIZ. http://triz.port5.com/works/ws2.html/
4.8. Morozov A. D. Introduction to the theory of fractals: Contemporary Mathematics. // Izhevsk: Institute of Computer Science, 2004

5.1. Cinema History // Website Sovkino. http://sovkino.sky.ru/history.htm /
5.2. Smell-o-vision // Website Retro future. http: // www.retrofuture.com/smell-o-vision.html/
5.3. Prokhorov A. Computers acquire odor // http://www.compress.ru/Archive/CP/2000/3/31/
5.4. From the «electric eye» to A. Stoletov modern TV // Website Radio Museum, http://radiomuseum.ur.ru/index7.html/
5.5. Halfin A. M. Mechanical & Electronic TV. — Moscow: State Publishing House, 1937. Web site of Samsung Electronics. http://www.samsung.com/
5.6. MPC vs. RGB: color TV reinvented // Website Membrane, http://www.membrana. ru/articles/technic/2004/09/02/214800.html /
5.7. The Web site of Universal Display Corporation // http://www.universaldisplay.com/
5.9. Ambilight. // https://ru.wikipedia.org/wiki/Ambilight
5.10. Company Website Candescent Technologies Corporation // http://www.candescent.com/
5.11. Sviridenko Yu. TV nanotube from the company Applied Nanotech. // Website Science RC. http://www.sciencerf.ru/
5.12. Company Website NaturalNano // www. naturalnano. com /
5.13. Bezhanov S. Sponsored yesterday watching what muzykuslushayut passing motorists. // Website Computerra. http://www.compulenta.ru/2002/11/27/35753/
5.14. Dalton A. V., Ellis C. S. Sensor-Enhanced Energy Management for mobile Device // http://www.cs.duke.edu/~angela/faceoffreport.pdf /
5.15. Paramonov O. To touch as possible! // http://www.computerra.ru/focus/new/27376/
5.16. Company Website Fujitsu // http://www.fujitsu.com/
5.17. Talking screen // Website Toaster. //http://www.toster.ru/653/
5.18. Display a surprise Toaster website, http://www.toster.ru/877/
5.19. Are the Japanese preparing «killers» PDA? // CNews web site.ru. http://www.cnews.ru/
5:20. Students invented a television wine bottle // Website Membrane. http://www.membrana. ru/lenta/?1794 /

5.21. Company Website TopHead // http://www.tophead.com/usa/products/products.htm /
5.22. Company Website Seamless Display // http://www.seamlessdisplay.com/
5.23. Website Nara Institute of Science and Technology (NAIST) // http://www.naist.jp/index_en.html/
5.24. Website Lebedev's Studio // http: // www.artlebedev.ru/portfolio/xox/
5.25. Company Website Panasonic // http://panasonic.biz/
5.26. Company Website Toshiba // www.toshiba.com/
5.27. The Web site of France Telecom // http://www.francetelecom.com/en/
5.28. Company Website Hokey Spokes // http://www.hokeyspokes.com/
5.29. Worldwide Electroactive Polymer Actuators // http: // ndeaa. jpl.nasa.gov/nasa-nde/lommas/eap/EAP-Be6.htm /
5.30. 3M (tm) Computer Privacy Filters // http://cms.3m.com/cms/US/en/2-22/kzRziFW/view.jhtml /
5.31. Steven K. Feiner MARS — Mobile Augmented Reality Systems. http://www1.cs.columbia.edu/graphics/projects/mars/mars.html /
5.32. Company Website IDK Tech // http://www.idk-tech.com/index_e.html/
5.33. Kartashov I. Cubic LED display. // http://www.rol.ru/news/it/news/04/09/01_009.htm/
5.34. Company Website Deep Video Imaging //http://www.deepvideo.com/
5.35. Company Website DDD // http://www. ddd. com/
5.36. Company Website Actuality System // http://www.actuality-systems.com/
5.37. Knigin S. 3D displays . — Part 1-3 //web site 3D News, http://www.3dnews.ru/display/3dd/
5.38. After 10 years, will be a three-dimensional crystal TV // Website Membrane // http://www.membrana.ru/articles/technic/2002/04/02/233200.html /
5.39. The blind will be able to «feel» Internet // Website CNews. ru. http://www.cnews.ru/
5.40. Electronic Reusable Paper // http://www2.parc.com/hsl/projects/gyricon /
5.41. Bridgestone company's specialists presented the thin display on the basis of the powdered material, created using nanotechnology // Website ITware. http://itware.com.ua/news/2002/03/29/Bridgestone.html/
5.42. Kartashov I. Color electronic ink for demonstration of motion pictures. // http://science.compulenta.ru/42184/
5.43. Bugrov V. A. Basics kinotelevizionny technology. — Moscow: Progress, 1964.
5.44. Company Website E-lnk. Technology // http://www.eink.com/technology/howitworks.html/
5.45. Alizar A. Created display brightness of the sun. // Web site 3D News, http://cnews.ru/reviews/articles/index.shtml?2004/03/09/156122 /
5.46. Bubble Imaging Technology: all the time in magic bubbles // Website Membrane. http://www.membrana.ru/articles/business/2002/08/16/194200.html /
5.47. The Web site of The FogScreen // http://www.fogscreen.com/
5.48. Company Website IQ Technology // http://www.io2technology.com/
5.49. Web site of the company Siemens // http://www.siemens.com/index.jsp/
5.50. The Web site of Cambridge Display Technology // http: //www.cdtltd.co.uk/
5.51. Company Website Qualcomm // http: //www. qualcomm. com/qmt /
5.52. Williams M. Mitsubishi Crafts Tiny LCD. Small display unit could make head-mounted displays comfortable, feasible, http://www. pcworld. eom/news/article/0,aid,83024,tk,dn020602X,00.asp/
5.53. BMW Develops Head-Up Display For Ralf Schumacher's F1 Helmet. Website Automotive Intelligence. http://www.autointell.com/News-2002/October-2002/October-2002-1/October-02-02-p8.htm /

5.54. Company Website MicroVision Technologies // http://www. mvtscopes. com/
5.55. In Japan, invented the machine that controls the dreams // Web site Grani. ru. http://www. grani.ru/Techno/m.56891.html /

6.1. Shpakovsky N. A. From the screw! // Effective solutions: Collection. — M.: TRIZ-profi. 2005.
6.2. Titov V. V. System-morphological approach in engineering, science, social / V. Titov // http://www. metodolog.ru/0003/00039.html /
6.3. Novitskaya E. L. Transformation of structurally similar elements of technical system. // http://www.gnrtr.com/tools/ru/a09.html /
6.4. Bionic tower — vertical city — data. http://www. torrebionica. com/bvs-english/bvs-english. htm /
6.5. Bionic Mercedes: Fish on Wheels boasts streamlined // Website Membrane. http://www.membrana.ru/articles/technic/2005/06/09/155400. html /
6.6. Giants LeTourneau: 174-meter road train consists of 11 trailers // Website Membrane. http://www. membrana.ru/articles/technic/2003/08/05/160000. html /
6.7. Shpakovsky N. A. Transition to the micro level — one of the main display evolution trends / NA Shpakovsky // Website methodology.http://metodolog.rU/00492/00492.html#1/
6.8. Maggs P. B., Sergeev A. P. Intellectual Property. — Moscow: Jurist, 2000.
6.9. Anisimov V. A wheel was invented in Australia in the XXI century. // Website Netoskop // http://www. netoscope.ru/news/2001/07/04/2795.html /
6.10. The development of functional thinking. // Website TRIZ School. http://www.natm. u/triz/function/fun_01.htm/
6.11. Maslow A. Motivation and Personality. — SPb.: Eurasia 1999.
6.12. Shpakovsky N. A., Chuksin P. I., Novitskaya E. L. Tool for generating and selecting concepts on the basis of trends of engineering systems. // Web-site Generator. http://www.gnrtr.com/ tools/ru/a05.html/
6.13. Shpakovsky N. A., Novitskaya E. L. Concepts Generator. // Website Generator, http://www. gnrtr. com/powers/ru/soft01.html /