

# Knowledge-Based Support For Innovative Design

*Valeri V. Souchkov*

© Valeri V. Souchkov

University of Twente, Enschede, The Netherlands. 1998

---

<i>Chapter 1.</i>	<i>Introduction</i>	<b>5</b>
	Subject of this book	5
	Research Goals	5
	Book Structure	6
	Publications	8
<i>Chapter 2.</i>	<i>AI and Innovative Engineering Design</i>	<b>9</b>
	Introduction	9
	Fundamentals of Innovative Engineering Design	<b>10</b>
	<i>Overall structure of Design Process</i>	<b>11</b>
	<i>Design concepts and detailed designs</i>	<b>13</b>
	<i>Engineering design and natural sciences</i>	<b>14</b>
	Conceptual Design	<b>16</b>
	<i>Phases of Conceptual Design</i>	<b>16</b>
	<i>Design and Creativity: Insight or Science</i>	<b>18</b>
	<i>Systematic versus cognitive approach to conceptual design</i>	<b>20</b>
	AI and innovative engineering design: related research	<b>23</b>
	<i>Role of AI for Design</i>	<b>23</b>
	<i>Knowledge Categories in Design</i>	<b>23</b>
	<i>Models of Conceptual Design</i>	<b>24</b>
	<i>Models of Design Based on Cognitive Methods</i>	<b>27</b>
	<i>Models of Design Based on Design Theories</i>	<b>28</b>
	<i>Discussion</i>	<b>33</b>
<i>Chapter 3.</i>	<i>The Theory of Inventive Problem Solving</i>	<b>35</b>
	Introduction	35
	TRIZ contents	<b>37</b>
	<i>Historical background</i>	<b>37</b>
	<i>Levels of Inventive Solutions</i>	<b>38</b>
	<i>TRIZ Philosophy</i>	<b>39</b>
	<i>TRIZ Structure</i>	<b>40</b>
	<i>Technical system</i>	<b>41</b>
	Theory of Technical System Evolution	<b>42</b>
	<i>Systematic nature of technology evolution</i>	<b>42</b>
	<i>The law of system ideality</i>	<b>43</b>
	<i>TTSE Laws</i>	<b>44</b>
	<i>TTSE Trends</i>	<b>44</b>
	TRIZ Problem solving techniques	<b>47</b>
	<i>Contradiction as the origin of an inventive problem</i>	<b>47</b>
	<i>Types of Contradictions</i>	<b>47</b>
	<i>Principles of Elimination of Engineering Contradictions</i>	<b>48</b>
	<i>Principles for Physical Contradiction Elimination</i>	<b>50</b>
	<i>Substance-Field Modeling</i>	<b>52</b>

---

<i>Pointers to Scientific-Engineering Effects</i>	55
Algorithm of Inventive Problem Solving	57
<i>Problem formulation</i>	58
<i>Problem modeling</i>	59
<i>Problem solving</i>	60
Summary	61

## Chapter 4.

### *TRIZ Critique* 63

Introduction	63
TRIZ contribution to design science	63
<i>TRIZ and German School of Systematic Design</i>	63
<i>TRIZ contribution to science of design</i>	65
Critique of TRIZ	66
<i>Trends of Technology Evolution</i>	66
<i>Principles for contradiction elimination</i>	66
<i>Substance-Field Modelling</i>	67
<i>Pointers to physical effects</i>	68
ARIZ	70
An attempt to formalize TRIZ	71
<i>Mathematical model of design products</i>	71
<i>Conflict Formulation</i>	72
<i>Problem solving</i>	73
<i>Discussion of Glasunov's approach</i>	73
Computer-Aided TRIZ and CAD/CAM software	74
Discussion	74

## Chapter 5.

### *A Knowledge-Centered Model of Systematic Innovative Design* 77

Introduction	77
INDES Project overview	78
<i>INDES Goals</i>	78
<i>Problem area and scope</i>	79
<i>INDES definition of conceptual design</i>	80
<i>INDES limitations</i>	80
Knowledge-centered model of innovative design	80
Knowledge categories for innovative design	82
<i>Deep and shallow knowledge</i>	82
<i>Object and strategic knowledge</i>	85
Object knowledge for innovative design	85
<i>Basic principles for defining and structuring object knowledge</i>	85
<i>Primitive object knowledge</i>	87
<i>Complex object knowledge</i>	91
<i>Object knowledge base</i>	93
Strategic knowledge for conceptual design	93

---

<i>An integrated model of innovative design</i>	<b>94</b>
<i>Meta-level strategy</i>	<b>95</b>
<i>Initial design specifications</i>	<b>95</b>
<i>Problem modelling and problem solving</i>	<b>96</b>
<i>Design concept generation</i>	<b>96</b>
Summary	<b>97</b>

## Chapter 6.

### *INDES: An ontological approach to modeling object knowledge.* **99**

Introduction	99
INDES modeling framework	<b>100</b>
<i>Physical knowledge and System Theory</i>	<b>100</b>
<i>Macroscopic and microscopic observations</i>	<b>102</b>
<i>Requirements for a model of physical phenomena</i>	<b>102</b>
<i>YMIR: a sharable ontology for modeling design knowledge</i>	<b>103</b>
INDES Concepts	<b>104</b>
<i>Physical Components</i>	<b>104</b>
<i>Relations: behaviour and function</i>	<b>105</b>
<i>Structure.</i>	<b>109</b>
<i>Constraints</i>	<b>109</b>
Generic Physical Principle (GPP)	<b>110</b>
<i>Model of generic physical principle</i>	<b>110</b>
<i>Types of Generic Physical Principles</i>	<b>112</b>
<i>Generic Design Concept</i>	<b>114</b>
Examples of modeled Physical Principles	<b>115</b>
<i>Joule Heat</i>	<b>115</b>
<i>Behaviour</i>	<b>118</b>
<i>Form</i>	<b>119</b>
<i>Overall model</i>	<b>120</b>
<i>Thermal Expansion</i>	<b>121</b>
<i>Example of Generic Design Concept</i>	<b>122</b>
Summary and Discussion	<b>124</b>

## Chapter 7.

### *Design Conflict* **125**

Introduction	125
Role of conflicts in problem decomposition	<b>126</b>
Design Product Model	<b>127</b>
Theory of Design Conflict	<b>128</b>
<i>Conflict elimination by innovative redesign</i>	<b>130</b>
Implementation	<b>131</b>
<i>Benefits from the INDES definition of design conflict</i>	<b>132</b>
Summary	<b>133</b>

---

<i>Chapter 8.</i>	<i>Innovative Redesign</i>	<b>135</b>
	Introduction	135
	Role of physical knowledge for design conflict elimination	<b>136</b>
	Innovative Redesign Axioms	<b>138</b>
	Innovative Redesign Principles	<b>139</b>
	<i>Definition of the Innovative Redesign Principle</i>	<b>139</b>
	<i>IRP applicable to non-conflicting situations</i>	<b>141</b>
	<i>IRP for conflict elimination</i>	<b>143</b>
	<i>Mapping between function and GPP</i>	<b>147</b>
	<i>Applicability of Innovative Redesign Principles</i>	<b>147</b>
	General model of innovative redesign	<b>148</b>
	<i>Model of innovative redesign</i>	<b>148</b>
	Case study: Application of IRP	<b>149</b>
	Innovative redesign and ICAID	<b>152</b>
	Summary	<b>152</b>
<i>Chapter 9.</i>	<i>Case Study</i>	<b>155</b>
	Introduction	155
	Problem Description	<b>155</b>
	Problem model	<b>157</b>
	A Design Prototype Model	<b>158</b>
	Solving a specific problem	<b>159</b>
	New design concept generation	<b>163</b>
	Conclusions	<b>164</b>
<i>Chapter 10.</i>	<i>Summary and Conclusions</i>	<b>165</b>
	Summary	165
	Contribution of the research.	<b>166</b>
	Conclusions	<b>167</b>
	Further research	<b>168</b>
	<i>Bibliography</i>	<b>169</b>

---

# *Chapter 1. Introduction*

## **1.1. Subject of this book**

Engineering design is a complex activity which involves different phases: from searching for a problem and formulation of design specifications to completing a set of drawings for manufacturing a product. Today, industries are unable to design products quickly and reliably without computer aid. However, the computer support is available only at later phases of engineering design - after a feasible product concept has been defined.

In contrast to the later design steps which are supported by existing CAD systems rather well, innovative design deals with informal techniques for obtaining new design solutions. The innovative design is a knowledge-intensive process which requires knowledge of diverse domains. However, it is unclear, how to model and represent this knowledge in a uniform and computable way. For this reason, it is claimed to be impossible to automate the innovative design.

This book deals with developing a theoretical background for computer support for the early design phases. The subject of this work is studying how knowledge-based techniques can be used to organize a computer support for innovative design.

The monograph presents results obtained from the study.

## **1.2. Research Goals**

The research entitled INDES (the abbreviation for Invention Designer) was initiated in 1993, after a Knowledge-Based Group of the University of Twente developed YMIR - a general ontology for modelling different design products in uniform and formal way. YMIR proved the thesis that different design products can be modeled and represented similarly.

---

YMIR proposes to regard a design process as a synthesis from primitives. Different primitive components, modeled in YMIR terms can be automatically assembled into a more complex product that match given specifications defining the behavior of the product. The behaviour of the final product can be calculated on the basis of the behaviors of separate components.

However, YMIR has several disadvantages. First, it may not generate new solutions across domains, for instance, combining electrical and mechanical components. Second, it concentrates on an approach to model design products rather than on specific methods of design. Nevertheless, YMIR provided INDES with a basic modeling framework.

Major goal of our research was to study whether a formal framework for innovative engineering design is possible. It was clear in the very beginning that it was not possible to develop the formal framework for designing of all types of products. Therefore, we limited ourselves to those types of products which involve material and energy transformations.

We formulated the following goals of the INDES project:

1. To study and compare known methodologies for engineering design and select those which can serve as the basis for the building the formal methodology of innovative design.
2. Study and compare different approaches to the use of AI and knowledge-based techniques in developing a computer support for innovative design.
3. To extend YMIR with a modeling technique which would make it possible to model design components involving cross-domain relations.
4. To develop the theoretical methodology for designing new products based on the intensive use of knowledge of natural science.
5. To develop a computational model of innovative design.

We did not set up a goal to build a computational model of design which would enable fully automated reasoning. Our study of literature had shown that no techniques were available to achieve this with. Every assumption and dynamic change that can influence the designer's decision can not be taken into account and represented explicitly.

One of the additional goals formulated in the very beginning of the INDES project was to develop and evaluate a knowledge-based system on the basis of the research results. However, due to a complexity of the system to be developed this task has not been completed.

### **1.3. Book Structure**

The monograph is structured as follows:

**Chapter 2** discusses the role of Artificial Intelligence in organizing computer support for Engineering Design. Before we start the discussion, we mention what fundamentals of

---

Engineering Design are, and what the overall structure of design process is. Since we concentrate on modelling a part of the overall design process dealing with innovative solutions, we define a place of conceptual design in the overall design structure. We also mention the roles of natural science and creativity in conceptual design.

Then, we discuss the role of AI in design by comparing two schools of AI in design: systematic and cognitive. We also explain, why a systematic approach is preferable and compare overall costs of the developing AI-based support for design.

**Chapter 3** aims at giving an introduction to the Theory of Inventive Problem Solving (TRIZ). Since the theory is relatively little known outside the former USSR and none of known English-language publications present TRIZ from scientific point of view, we decided to give an extended presentation of TRIZ. We believe it is necessary to present the basic TRIZ concepts for a better understanding of our approach.

**Chapter 4** presents a critique of TRIZ from the AI point of view. Although organized in a systematic way, TRIZ mixes various concepts and its knowledge collections mix various types of representations.

**Chapter 5** explains the INDES approach by introducing a model of knowledge-centered, systematic innovative design. First, we discuss INDES goals, problem area and scope. Based on TRIZ, INDES-based model of design distinguishes two types of innovative design: design as a modification of the existing design product and design from physical principles. We show, how both these types of design are supported with relevant TRIZ tools and how they are interrelated from the point of view of using common knowledge sources for both types of design.

The second part of the chapter explains why we classified all knowledge needed during the design process into two categories: object knowledge and strategic knowledge. We explain why we prefer to model generic knowledge.

In the third part, we focus on INDES strategic knowledge and describe why we selected TRIZ problem solving techniques as a course for strategic knowledge.

**Chapter 6** presents an INDES ontology for object knowledge for innovative design. Our modelling approach is explained in detail. Several examples of modeled physical principles are provided. Besides, we show how to model different technical systems in terms of generic object knowledge.

**Chapter 7** explains what a design conflict is from the INDES point of view and defines the design conflict in terms of energy and material-transforming systems.

**Chapter 8** introduces a model of innovative redesign based on the modification of an existing design product. The innovative redesign is based on the axiomatic approach. We introduce two axioms from which a number of innovative redesign principles are drawn. The applicability of the principles is illustrated by an example.



---

**Chapter 9** describes a case study - an experiment conducted at the industrial company which was supposed to verify the applicability of the framework proposed in the monograph.

**Chapter 10** presents conclusions and thoughts of further research.

The book was written for those who are interested in both engineering design and computer science. However, engineers will probably most interested in chapters 3, 6, 8, 9, and 10.

## 1.4. Publications

Some early ideas on modeling the physical world in terms of object knowledge to be used in knowledge-intensive innovative design were addressed by the author in the *TRIZ Journal* (Sushkov, 1991).

An short overview of TRIZ, TRIZ software support and proposals for developing knowledge-based support for TRIZ-based innovative design was presented in 1993 at *Artificial Intelligence and Creativity* workshop conducted by the American Association of Artificial Intelligence (AAAI) (Mars, Sushkov, and Wognum, 1993).

Early approach to structure, model, and organize physical knowledge which formed the basis of INDES was presented at the *Models and Techniques for reuse of Designs* which was held in conjunction with the 11th European Conference on Artificial Intelligence (ECAI) (Sushkov, 1994).

An extended overview of the Theory of Inventive Problem Solving presented in Chapter 3, and its discussion from the knowledge-based point of view (Chapter 4) were published in *AI in Engineering* magazine (Sushkov, Wognum and Mars. 1995).

Definitions of creative engineering design, a general view of how TRIZ knowledge should be restructured a were presented at the *3rd Int. Roundtable Conference on Computational Models of Creative Design* (Killander and Sushkov, 1995). The article also describes our experience with using TRIZ to solve a real design problem and compares results obtained after a traditional and TRIZ-based innovative designs.

A philosophy behind INDES, a developed approach to modeling physical knowledge in terms of generic components (Chapter 6) were described in detail in the paper presented at the *Artificial Intelligence in Design* conference (Sushkov and Mars, 1996). The paper presents several examples of modeled physical phenomena and our classification of initial design specifications.

General overview of INDES and the obtained results were reported at the *International Conference on Engineering Design* (ICED) (Sushkov and Mars, 1997).

## Chapter 2. *AI and Innovative Engineering Design*

*“Engineering: 1. The application of scientific and mathematical principles to practical ends such as design, construction, and operation of efficient and economical structures, equipment, and systems. 2. The profession of or the work preformed by an engineer.”*

*Webster’s 7th Collegiate Dictionary*

### **2.1. Introduction**

This chapter discusses the role of Artificial Intelligence (AI) in Innovative Engineering Design. First, the place of innovative design in the general process of designing is defined. To outline what is meant with innovative design throughout the book, a number of features distinguishing it from other design activities are described. Furthermore, some successful theories and design methods that innovative design can be based upon are mentioned.

Second, we give a brief overview of how AI helps with developing intelligent knowledge-based support for the innovative design. We compare two AI in Design schools — cognitive and systematic — by studying what AI methods and computational models of innovative design are available today. Towards the end of the chapter, we explain why the systematic approach is preferable.

## 2.2. Fundamentals of Innovative Engineering Design

### 2.1.1 Design process and design products

Engineering design can be regarded as an activity which contributes to society by facilitating the creation of new products to satisfy its needs and aspirations. It is a complex, multidimensional discipline which involves many diverse aspects and knowledge of various natural sciences. Suh, the author of Axiomatic Design approach distinguishes four aspects of the engineering and scientific endeavor involved in the design process:

”...the *problem definition* from a “fuzzy” array of facts and myths into a coherent statement of the question; the *creative process* of devising a proposed physical embodiment of solutions; the *analytical process* of determining whether the proposed solution is correct or rational; and the *ultimate check* of the fidelity of the design product to the original perceived needs” (Suh [1993], p.6).

The design product is the output of the overall engineering design process. Technical entities such as devices, machines, assemblies and individual components are artificial and concrete systems which consist of a totality of organized elements, linked together by relationships caused by their characteristics. The system resulting from the design process is separate from its surroundings while the links with the surroundings define the boundaries of the system. The designed technical systems represent a process within an artificially created environment by which energy, materials and information are routed and transformed. In each elementary transformation, the quantity and quality of factors involved are determined so that the criteria for the precise definition of the task and the evaluation of the system are unambiguous.

Engineering design involves a whole range of different theories and methods (Jones [1981]) applicable at various phases of engineering design aimed at mapping a specified function onto a realisable physical structure of the design product. The designed product is represented as a description of the assembly of various components each of which performs a specific function which is required to perform the overall functionality of the whole design. To design a final product, the designer must make many decisions of various degrees of complexity with respect to what components are needed and how the components should be assembled, arranged and linked to behave correctly. Therefore, we can distinguish two different concepts within engineering: *design product* and *design process*. The design product represents a description of the realisable artifact which is independent of how this description was obtained whereas the design process involves various types of activities aimed at producing the design product.

The design process contains a number of steps needed to create a realisable product description. In general, it can be regarded as a two-step process: first, a decomposition of the specified overall function into more primitive functions and defining what single components are capable of performing those primitive functions and, second - the creation of a new design product by the synthesis process.

With respect to society, the design process forms a loop, since constantly changing societal needs make designers constantly reconsider and improve existing products (Figure 2.1).

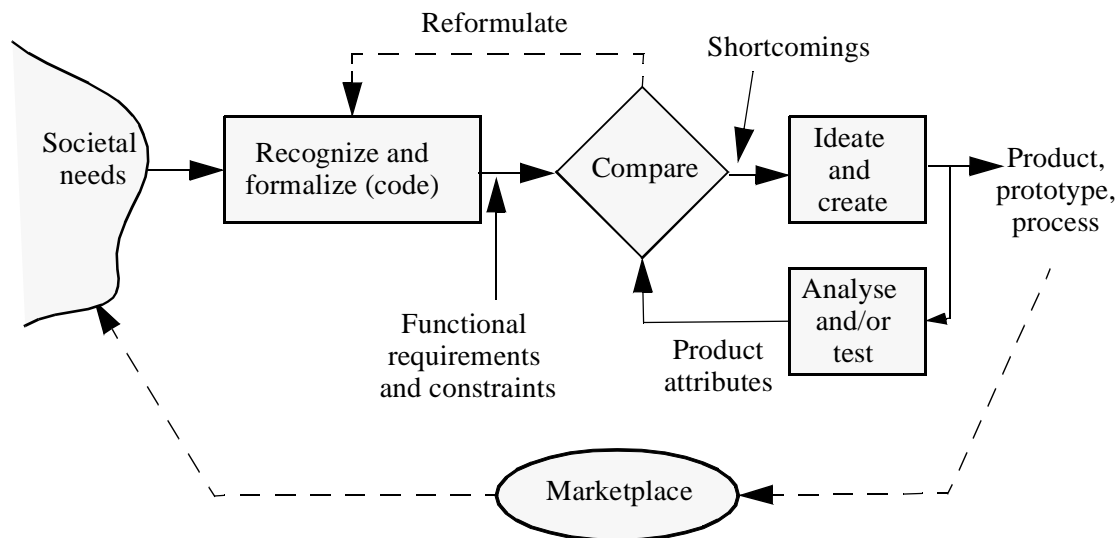


Figure 2.1: The design loop (from [Wilson 1980])

### 2.1.2 Overall structure of Design Process

Before we start a discussion of available design methods and theories, we present an overview of overall engineering design process to define the boundaries of the conceptual design phase and its place in the design process. One of the most consistent and systematic design methodologies that can be found in the literature about engineering design is presented in Beitz & Kuttner [1994]. Despite its orientation to the process of designing mainly mechanical devices, this methodology can be considered as general with respect to all engineering domains dealing with physical entities. Four major phases of engineering design are distinguished:

1. **Phase of Defining Requirements** involves collecting information about market demands and human needs. On the basis of this information, the list of requirements is compiled. Depending on the type of design, the list of requirements might not present information in the designer's language and therefore it might consist of both ambiguously expressed requirements and precise numerical constraints. The requirement list consists of two parts: *demands* and *preferences*. Demands are the

requirements that must be fulfilled in all circumstances while preferences are the requirements of varying significance and should be taken into consideration only if possible.

2. **Conceptual Design** is a part of the design process where a qualitative concept of future design is generated. To generate the design concept, a solution principle(s) that meets the demands part of the requirement list must be found. If a totally new system is designed, several relevant solution principles must be found to fulfil each subfunction if the decomposition of overall function is possible. During conceptual design, it is important to abstract the problem in order to clarify it and to free the designer from fixed ideas and known solutions and to generate new, more effective solutions. Several alternative design concepts can be generated and then they must be checked against the demand part of the requirement list. A choice of relevant design concept can be guided by the preferences part of the requirement list.
3. **Embodiment Design** involves the compilation of techno-economical structure of the design product and making it unambiguous and complete. This phase requires correction and refinement of the design concept generated at the phase of conceptual design by alternating analysis-synthesis procedures. The output of embodiment design is a quantitative model of the product including all necessary individual components.
4. **Detail Design** supplements embodiment design with final specifications on the configuration, arrangement, shape, dimensioning and surface quality of all individual components, by checking all materials, manufacturing feasibilities and costs. A careful check against existing design norms and standards is also performed at this phase. The output of this phase is precise specifications of all aspects of the product including engineering drawings and specification of needed manufacturing facilities.

In many cases it is not necessary to perform all these steps to design products according to a new requirement list. New designs often involve only certain modules or reconfigured previously designed components. In those cases, a new solution principle is not needed and a new design product can be produced by varying dimensions or arrangements within some existing product. This produces three types of engineering designs:

1. **Basic design**, which consists of the compilation of a new solution principle and thus requires all four phases.
2. **Redesign**, which consists of adjusting the embodiment (shape and material) of the existing system while the solution principle remains the same but the system's boundaries might be advanced.
3. **Variant design** (known also as **configuration design**) which involves variation of sizes and arrangements within the boundaries of the existing system.

Basic design is also known as *non-routine design* while redesign and variant design refer to *routine design*.

### 2.2.2 Design concepts and detailed designs

There is a difference between products of routine and non-routine design phases. *Design concepts* are the result of non-routine phases of design while *detailed design* descriptions are produced after a relevant concept meeting crucial designer's requirements has been selected. The major distinction is made by the level of abstraction and generalization of knowledge used to present both design concepts and detailed designs. In both cases, the principal requirement is that the interpretation of the information involved must be adequate in terms of possible physical realization. For this reason, the design concept has to include enough information to verify the physical realisability of the concept.

The detailed design description must specify precisely what the geometry of components should be, what materials are to be used and what the values of the parameters must be at each stage of system operation. These specifications must be defined for *all* system parts. The design concept should not necessarily represent all this information. For instance, any part of a system can be represented as a black-box that hides the possible geometry of components or links between the components. Correspondingly, all parts of the system might be represented in black-box terms. The design concept which passed the verification stage can thus be instantiated into a multitude of detailed designs.

The process of synthesising design concepts consists of assembling the components in such a way that the overall behaviour of the resulting system is correct and fulfils the function required. The synthetical process can be applied if the overall function is represented in such a way that it can be decomposed into more primitive functions and proper primitive components can be found. On the other hand, new requirements are frequently not formulated as precise functions, and might be specified as informal expressions like "to reduce noise produced by a technical device". In this case, multiple interpretations of this expression are possible and it is unclear what starting point can be selected for the design process. Three different starting points can be defined: i) a new device with better noise characteristics should be designed; ii) some part of the device should be modified to produce less noise; iii) something should be changed within the surrounding environment - for instance, to cover the inner surface of a room where the noise-producing device is located with a noise-protecting material. It is clear that any of these problems chosen as a starting point for design will lead to different solution principles.

Another principal disadvantage of the decomposition approach is that a situation can occur when the overall function may not be decomposed into more primitive functions since the relevant physical component that would be capable of performing one of these functions might not be known within the engineering domain given. At this point, a role of knowledge of another domain becomes critical.

### 2.2.3 Engineering design and natural sciences

The field of engineering design, especially its conceptual phase, is closely related to physics and chemistry. In general, it can be said that as knowledge of some scientific topic grows to the point where use of it can be made in everyday life or in industry, the topic passes over to engineering in order to develop and improve things using this information. Exact sciences therefore focus on studying what the properties of the surrounding world are and discovering new facts about materials, energy and information, whilst engineering concentrates on studying how these properties and discoveries can be put to practical use to satisfy the permanently growing needs and demands of society.

Theoretical studies of the physical world and attempts to establish relationships between physical objects can be traced back to ancient Greek philosophers. However, their theories suffered from the lack of experimental research which was crucial to create the prerequisites for regarding the field of engineering as a science instead of magic: experimental research was needed to reveal new properties of previously described physical entities as well as to discover new physical phenomena. For this reason, the best term for those early studies in physics would be “natural philosophy” rather than “natural science”. Further progress in physics, material sciences, applied chemistry and mathematics has led to the appearance of the whole new division of scientific studies known today as engineering.

In spite of the relative success of various design methodologies, one of the major shortcomings of modern engineering is that over the past few decades engineers have come to possess very specific skills and they lack knowledge of other disciplines. Modern engineering is divided into many subdivisions, like mechanical engineering and electrical engineering. However, new product creation is a knowledge-intensive process which requires knowledge of many domains. Consequently, general principles for engineering design should be applicable to any domain. The major role of such general principles is to provide guidelines for using more specific design and scientific knowledge.

In the beginning of the century, the Russian engineer Engelmaier introduced the notion of the so-called “engineering effect” (Engelmaier [1910]). He defined it as “any useful result performed by an engineering system and satisfying the user’s needs”. As seen, this definition is very general and ambiguous. However, it was an attempt to explain the link between physics and engineering. From Engelmaier’s point of view, if the physical effect represents fundamental relations between interacting physical components, then the engineering effect represents the same relations within certain context and focuses on what useful can be obtained within the context. For instance, color change as a result of heating might be essential to represent some engineering effect but not the effect of heating. Furthermore, the engineering effect might involve certain attributes that are not important when describing the physical effect.

From this point of view, the task of the designer performing innovative engineering design phase is twofold:

- Once a new demand has been formulated, the task of the designer is to find what knowledge of the natural sciences could be interpreted as appropriate to meet the demand and later be instantiated into a design product.
- Once new physical properties are studied or new physical phenomena are discovered, the task of the designer might be defined as to analyse how these properties or phenomena can be used to design new useful products.

Extensive utilization of the knowledge of physics helps to make new design products more reliable and simple. Suppose a problem is formulated as: “to prevent an electrical motor from overheating”. One of the known solutions within the electrical domain uses a temperature sensor which reads the current temperature value and an electronic system which switches the power supply off when the threshold value of the temperature is reached. This problem can be solved more easily and the overall design can be made more reliable if the poles of the motor are made of an alloy with a Curie point equal to the required threshold value of the temperature. When the temperature reaches the threshold value, the magnetic properties of the poles change and the motor stops (Petrovich & Tsourikov [1986]). The necessity of introducing a complex and unreliable additional design has disappeared.

The question is whether it is possible to utilize available physical knowledge to design new products systematically. It is obvious that any design product is a physical system which obeys the laws of physics; so why not use physical laws to create new artefacts? The core problem is that physics sees the surrounding world with a different view than engineering -- it studies properties of the world without focusing on what possible applications of discovered facts could be. As noticed by Max Planck:

“...Scientific discovery and scientific knowledge have been achieved only by those who have gone in pursuit of them without any practical purpose whatsoever in view”.

For this reason, encyclopaedia and handbooks do not present physical knowledge in technically applicable way, and attempts to *directly* instantiate physical laws drawn from handbooks into new technical systems are fruitless in most cases.

As an example, let us take the phenomenon of thermal expansion. Despite the fact that it is rather well-known to engineers, most of them will not retrieve the phenomenon from their memory when the problem is given as delivering the particular function “to control displacement of a solid body”. However the function can be fulfilled by subjecting one object to alternating heating or cooling and having another object fixed to the first one. By manipulating the temperature of the first object we can precisely control the displacement of the second one. On the other hand, the displacement, unless regarded at the molecular level, is not a property of the effect of the thermal expansion. To be used for designing, this phenomena can not be analysed only from the point of view of its internal properties and interactions, but from the point of view of those properties and interactions which appear when the phenomenon is used *in combination* with other phenomena.



Another important aspect of studying how physical knowledge can be organized for engineering needs is how physical laws and phenomena are *interpreted*. The role of interpretation is to establish relations between physical phenomena and the multitude of engineering effects that can be obtained on the basis of the phenomena.

## 2.3. Conceptual Design

### 2.3.1 Phases of Conceptual Design

The process of creation consists of mapping a formalized requirement onto some physical structure that is physically realisable. The output of the creation process is a design concept, which can be the concept of a new prototype, new technology or a new product for mass manufacturing. At the phase of conceptual design, it is important to generate a solution principle which would satisfy the demand part of the specification list. This design phase is crucial for the overall design process, since a final detailed design will incorporate all advantages and disadvantages of the principle. An incorrectly chosen principle may cause improper functioning of the system which may result in fatal disaster.

A breakdown structure of the conceptual design phase is shown in Fig. 2.2. It consists of four parts:

- *information definition*, that is the clearance of the requirement list;
- *creation process*, which consists of finding the relevant principles to fulfil each required subfunction;
- *assessment* which means checking the principles found against more specific requirements;
- *decision making* which defines if the design can be built on the basis of the principles or not.

At the early design stages, the list of requirements might not include those specifications which specify detailed physical topology or geometry of the future design artefact. This part of the specifications can be justified only after a relevant physical principle has been found. Quite commonly, the requirement list only includes those key aspects of the design product and its functioning which are independent of its possible physical implementation. In addition, some requirements can restrict the conditions under which the use of certain physical phenomena would be allowed.



### 2.3.2 Design and Creativity: Insight or Science

Conceptual design is often linked to creativity. Like any human activity where new ideas are produced, for instance art or science, conceptual design involves a lot of unorganized and chaotic thinking, and is often identified with insight. So far, neither research in psychology nor neurology has been able to determine what the nature of human creativity is.

As a result, even today, creativity as the capability to produce new useful ideas is often related to exceptional scientists, artists and designers. Although numerous techniques such as brainstorming and synectics are known to activate the search for new ideas, they are still modifications of well-known trial and error methods to create new concepts. However, to solve difficult problems that would require the use of a physical principle that has not been used in engineering before, several thousands trials can be made.

Some schools conclude that creativity is related to analogical reasoning performed by humans (sometimes regarded as intuition). However, the process of how the human mind establishes distant analogies is not yet understood. For this reason, traditional scientific theories assume that new designs or new requirements are the products of a creative process which can only be studied at cognitive level.

With respect to possible constraints on the implementation of creative ideas, three categories of creative processes can be distinguished (Killander & Sushkov [1995]):

1. **Unconstrained creativity.** Unconstrained creativity is not limited by anything (science fiction, art, etc.). Examples are a spacecraft travelling with the speed of light or immediate transportation. These are all the products of unconstrained creativity.
2. **Real-world creativity.** The only limitation for real-world creativity is that its product must be physically realizable regardless of other constraints. For instance, to boil water for coffee, one can use laser beams or nuclear reactions.
3. **Constrained creativity.** A new product must meet numerous constraints identified for a specific situation, such as the following: physical and ergonomic constraints, as well as cost, time and manufacturability constraints.

The most relevant word for referring to verified products of the creative phase of engineering design is "invention". We argue that to successfully perform inventive design, a designer must not be limited to the third category, constrained creativity. Numerous constraints inherent in the design process interfere with the possibilities of obtaining inventions. Moreover, the constraints cause a strong psychological inertia which is very hard to eliminate. A well designed computer-support for creativity should help to overcome this problem and to make the process of inventing good concepts available to any designer regardless of his/her previous creative capabilities.

What does it mean to invent a new artifact from a scientific point of view? If we regard this task as a problem of problem solving, then to search for the solution one must first define the

boundaries of the solution space of all possible solutions, to construct a general theory behind the process of navigating in that space and then to apply the laws of the theory to produce the solution. The most difficult situation occurs when an exact theory for solving a particular problem is not available.

The first results of tackling this problem date back to 300 A.D when, based on the previous works of Euclid and Aristos, Greek mathematician Pappos introduced a concept of heuristics - the science of making discoveries and inventions. A heuristical approach appeared to be very important to develop mathematics -- early mathematicians could not rely on experiential knowledge but on facts, as opposed to physics and chemistry. The further evolution of mathematics from a set of heuristical rules to exact numerical methods have made it possible to solve very complex problems that could not even be approached by the most creative persons a century ago. Therefore, solving differential equations is not the exclusive right of creative persons any more, and the same should have been applicable to solving inventive design problems.

Unfortunately, despite the progress in developing exact sciences, the study of creativity still lacks strong fundamental basis. Researchers mostly focus on studying the psychology of thought instead of studying the products of creative process as would be analogous to mathematical studies. In other words, most of the research concentrates on studying what cognitive process stands behind human creativity, regardless of what the qualities of the output of this process are and if there are general principles for how to obtain this output available.

The situation of understanding a creative process in engineering design changed in 1956 after the Russian engineer Genrich Altshuller published a paper in which he explained the nature of engineering creativity as the ability of a designer to overcome contradictions (Altshuller [1956]). Altshuller showed that the most outstanding inventors managed to eliminate *contradictions* arising between two or more parameters of existing technical systems during attempts to solve the problem in a non-inventive way. Most importantly, he discovered that inventors like Edison and Franklin did not make their inventions in a random way, but utilized certain patterns although the inventors themselves were probably not aware of their existence. Instead, the unconscious use of previous experience stored in an inventor's memory in the form of patterns used to be identified with intuition, or insight. Based on these discoveries, Altshuller developed a systematic approach to solving inventive problems, namely, how to formulate contradictions and to use patterns to eliminate them. The work of Altshuller will be discussed in more detail in Chapter 3.

Summarizing, the cornerstone works published by Engelmaier [1910], Zwicky [1948], Altshuller [1956], Pahl and Beitz [1984] and Suh [1993] created a theoretical basis for making conceptual engineering design a science instead of art. All of them relied on studying factual material rather than examining cognitive activities. These works form the ground for creating a methodology for creative engineering design based on systematic analysis and the search and evaluation of new ideas.

### 2.3.3 Systematic versus cognitive approach to conceptual design

As mentioned above, conceptual engineering design addresses the creative problem solving process. The output of this process, regardless of a specific engineering field, is a conceptual and physically realisable description of a new artefact. Therefore, creativity can be identified with the synthetic process. Unlike exact methods of synthesis known in physics or chemistry, there are no laws or objective principles governing the synthetic process in engineering. For this reason, most novel products are designed in an unorganized manner.

Much work has been done to understand the nature of creativity and to reveal and formulate the objective principles behind human creative process. So far, most of the attempts have been concentrated along two separate directions: first, trials to understand how human mind tackles the problems, and second, creating a design methodology that would make it possible to organize and systematize the design process at early phases. Thus, a clear distinction can be made between *cognitive* and *systematic* design methods.

*Cognitive* methods focus on studying what factors are involved in the creative process. So far, it has been found that the most important among them are abstract reasoning and associative thinking based on previous experience. However, these are difficult to model and understand, so cognitive studies mostly result in developing tools which help the designer to improve his thinking, intuition and inspiration. Among them are brainstorming, synectics, trigger-work technique, attribute-seeking (Harrisberger [1966]) and lateral thinking (de Bono [1992]).

There are also a number of conventional software applications supporting the design process with cognitive methods. They mostly provide an interactive database search for associations or analogies and storing the user's information without its analysis (Holt [1992]). In contrast, systematic approach focuses on studying what specific or general rules and algorithms are applicable to the design process. Among them are such methods as morphological analysis (Zwicky [1948]) and design rules (Gleg [1960], Boothroyd [1987]) as well as numerous attempts to develop general design methodologies (Hill [1970], [Yoshikawa 1985], [Akman et al. 1990], [Tomiyama 1994], [Gero 1994]).

The advantage of the *systematic* approach over the cognitive is that the design process organized in a systematic way is less dependent on human creative capabilities and consists of support for both synthetic and analytic procedures. On the other hand, as the cognitive methods are claimed to be applicable to virtually any design problem, most of above mentioned systematic design methods only apply to specific design tasks and thus may not be generalized over every engineering domain. This makes their use rather restrictive.

The systematic approach can be used as a background for building both better theory behind the innovative design process and knowledge-based systems aiding this process.

Among the most successful systematic approaches to innovative design are:

1. **Design Catalogues** developed by German school of systematic design (Pahl & Beitz [1987]). This method is derived from the early morphological method originated by Swiss astronomer Zwicky. Design Catalogues store information on design primitives and physical principles classified according to specific task that can be solved by the use of one or another principle or design primitive. This information can be reused if a new problem can be identified with a function that has already been delivered by some known physical or design structure (Fig. 2.3). Although design catalogues are a good means for the synthesis of new design concepts, they lack tools for analysing design problems and evaluating proposed concepts.
2. **Axiomatic Design** developed by Suh of the Massachusetts Institute of Technology (Suh [1993]) introduces so-called Design Axioms which can be used to effectively evaluate any design and define tasks for creating new or improving existing designs. The essence of Axiomatic Design is that the functional requirements a design meets should not be dependent on each other. Therefore, the information content of the design should be minimized.
3. **Theory of Inventive Problem Solving** developed by Russian engineer Altshuller (Altshuller [1984]). The theory introduces general principles of design product evolution which are incorporated into so-called innovative design principles. In combination with more specific techniques for problem analysis and formulation they form the basis for systematic problem solving. TIPS also incorporates the technique of synthesis for new design concepts originated by Zwicky and the guidelines for extensive utilization of physical knowledge.

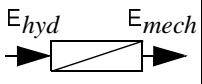
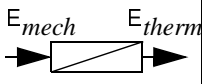
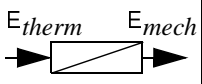
Function	Input	Output	Physical effects			
	Speed	Force Length	Profile lift	Turbulence	.....	Flow resistance
	Force, Speed	Temperature, Quantity of heat	Friction (Coulomb)	1st law	.....	Hysteresis
	Temperature Heat	Force Pressure	Thermal expansion	Steam Pressure	.....	Osmotic pressure

Figure 2.3: A fragment of Design Catalogue (from Pahl & Beitz [1984])

The three design methodologies mentioned above can be regarded as complementary and, when combined together, they form a general methodology for design based on an analysis - synthesis - evaluation cycle. (Figure 2.4).

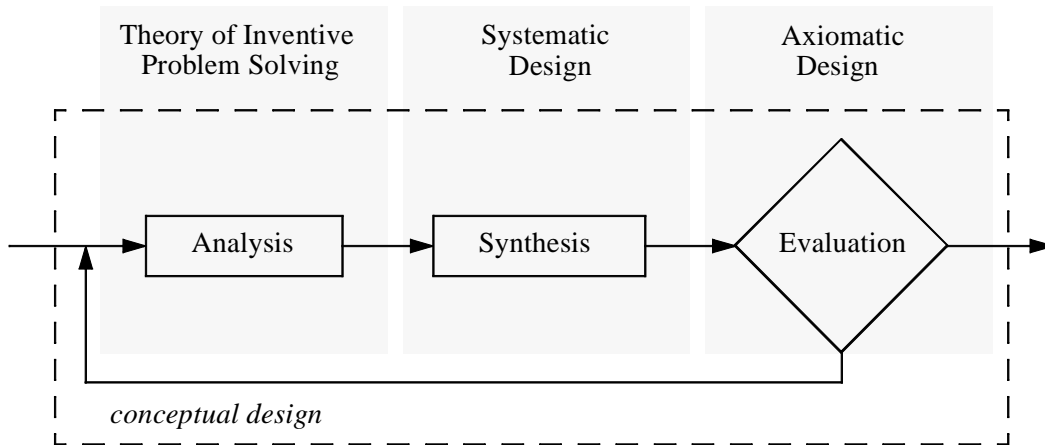


Figure 2.4: Phases of conceptual design with relevant methodological support

In our research, we build a computational model of innovative design which incorporates the basic principles of systematic methods. We have chosen the systematic approach due to two factors:

- this is the only approach which is well understood, and
- it does not depend on the personal abilities of a designer.

In this book, we will label the conceptual phase of design organized in a systematic way as *innovative* rather than *creative*. This difference is important since we do not establish our goal as modeling the cognitive process of creativity. Instead, we aim at developing a formal background which will result in the possibility of systematically creating the same products that could have resulted from creative thinking process. Further in the thesis, we will identify innovative design with non-routine and conceptual designs in order to avoid misunderstandings occurring within the literature on both engineering design and AI.

## 2.4. AI and innovative engineering design: related research

### 2.4.1 Role of AI for Design

While in the previous section we defined what is meant by conceptual engineering design, in this section we will discuss what AI methods are available and how they are used to build knowledge-based systems aimed at the generation of innovative design solutions.

The goal of AI can be defined as building sophisticated, machine-intelligent tools that are capable of performing various types of activities which may only be accomplished under the control of the flexible intelligence of human beings. Modern AI covers different topics of various degrees of complexity — from programs for playing chess to intelligent robots. With respect to design in general, we may also distinguish several different directions where different AI methods and techniques are used. A classical approach within AI states that the design might be modelled as a problem solving process which can be organized as a search within the predefined space of known solutions. The decision of which search space has to be explored is made by the type of initial specifications and constraints given. The more precise and detailed the specifications and constraints are, the smaller is the design search space required to find a solution. Alternatively, when none of previous solutions can meet the requirements given, it is obvious that a search within a larger space is needed. In this situation, two fundamental questions rise: what shall be an extension of the existing space? and how to organize an effective search within the large design space?

One of the important issues which is frequently being discussed during meetings within the AI in Design community is: should an AI system for making new designs be based on existing design theories and methods? Or, should the generation of new design concepts be based on known AI methods, like genetic programming or analogical reasoning. This especially applies to the earlier design phases, when the concept of a design is unknown and therefore can not be found within predefined search space. The question mostly addresses philosophy, since even an ordinary process of creative design which is not supported by intelligent tools can be either based on design methods such as the Theory of Inventive Problem Solving or it can also be conducted in an unorganized manner, for instance, based on associative thinking or brainstorming. An alternative concept regarding all phases of design as a scientific category is presented in detail in Alberts [1991]. The choice of an appropriate tool depends of many factors, whereas the most influential ones are consistency, the availability of solutions and effectiveness.

### 2.4.2 Knowledge Categories in Design

Logan and Smithers distinguish between two types of knowledge required to organize knowledge-based support for innovative design: *domain knowledge* and *design knowledge* (Logan and Smithers [1993]). Domain knowledge expresses facts about design objects, their properties and physical relations underlying the design objects. This knowledge is invariant across designs within a domain (Clibbon & Edmonds [1995]) and forms a so-called *design*



*space*. Design knowledge, in turn, specifies how the design space can be explored and transformed. In other words, it defines a strategy of how to identify problems and how to solve them.

Another important note is that the role of both *experiential* and *deep* knowledge is crucial for innovative design. From our point of view, both types of knowledge are complementary rather than contradictory: using past experience helps to identify problems, the use of deep knowledge helps to synthesise new solutions.

Another important issue is that innovative design as opposed to routine design can not be organized within a single engineering domain. Although new solutions can be obtained by synthesising and instantiating “first principles” derived for each specific domain, their usage strongly restricts the space of possible solutions. It happens since historically, engineering has been divided into many divisions, like electrical engineering, mechanical engineering, optics. As a result, the boundaries between domains and different physical principles the domains based upon restrict knowledge transfer and adaptation from one domain to another.

In our project, we introduce a domain-independent approach to innovative design. To reach this goal we investigate how both design and domain knowledge can be represented in a formal and uniform way to enable domain-independency and sharability of both design and domain knowledge.

### 2.4.3 Models of Conceptual Design

Observing previous works in the AI in Design discipline, it is not difficult to draw the conclusion that AI in Design is a mixture of various scientific and engineering methodologies. As a consequence, it might be the case that the same knowledge concept can be called “theory” in one paper and “model” in another. To avoid this and to bring clarity into further discussion, we would like to define more clearly what is meant by theories, models and methods. There is an approach introducing these definitions and distinctions presented in Smithers [1996]:

1. *Theory* is understood as statements about a particular phenomenon which make no reference to and do not depend upon particular instances of the phenomenon.
2. *Model* is an interpretation of a particular phenomenon which refers to instances or classes of instances of the phenomenon which implies boundary conditions and constraints. Models can be built using empirical knowledge and understanding. Therefore, models are possible *without* theoretical understanding.
3. *Methods* specify actions for a particular kind of designing. They are usually derived from models. However, as opposite to models, if a method does not refer to a theory or a model, it is impossible to estimate its consistency.

Further, we will also discuss which AI techniques are used today to model the process of conceptual design and build intelligent tools based on these models.

From an AI point of view, engineering design can be regarded as a procedure of *mapping the required specifications onto a description of physically realisable artifact* (Tong and Sriram [1992]). A resulting design description has to be correct both in static and dynamic senses, that is, apart from an assembly of physical components, the device should possess the correct behaviour.

Maier distinguishes between three phases which any model of overall design process has to be comprised of (Maier [1990]):

1. *Formulation*: identification and specification of the requirements.
2. *Synthesis*: generation of design alternatives and constraints refinement and justification.
3. *Evaluation*: analysis and evaluation of generated design alternatives according to predefined criteria and given requirements.

The choice of a phase of design also brings substantial differences to models of design. When a routine design requires high computational complexity related to data transformation and implemented logical mechanisms, non-routine design relies heavily on reasoning with heuristic knowledge and operates within an ill-defined problem space. The degree of computational complexity grows when a theory which can define a problem solving methodology is unavailable.

The main difficulty with building knowledge-based systems for innovative design - the lack of *missing knowledge* - is mentioned in Tong & Sriram [1992]: “the missing knowledge might either be knowledge for directly generating new points in design space, or knowledge for directly controlling the design space search”. Two reasons why relevant design knowledge can be missing are mentioned:

1. Most naturally acquirable knowledge might not necessarily be in a directly applicable form. If a designer wants to create a new coffee maker, he can scan various engineering literature and interpret the ideas proposed there with respect to the coffee maker. However, these ideas might have nothing to do with coffee maker design.
2. It is impossible to store the large amount of knowledge that would be necessary to adequately deal with all possible design variations. Presenting various designs at detail level would produce an enormous amount of information even within a single engineering domain.

There is another, very important reason why the same knowledge is difficult to directly reuse at innovative phases of design. The same knowledge might have different interpretations from different views, and it is impossible to predict and model all possible views. For instance, from the view of the transport domain, an air cushion can be interpreted as a way of moving objects; but in microelectronics it can be used to prevent two objects from contacting.

Analysis of current directions within AI in Design shows that the task of AI is seen as not only to acquire, interpret and represent knowledge available in engineering design to enable automated reasoning with design knowledge, but also to study what purely AI methods can be used to organize the design process. To this end, it is important to distinguish between two crucial knowledge levels associated with design:

- *Design theories* defining the basic statements on the nature of the process of designing and designed artifact;
- *Computational models of design*, which bring algorithmic nature into design theories by defining assumptions, constraints, variables, and a set of transformations of initial state into goal state.

One of the sample approaches to building a computational model of creative design is proposed in Gero [1994] and Maher et al.[1996]. In the theory of creative design behind the model suggested, a designing process is viewed first as an exploration of less well-travelled design space, then as a transformation of existing design space. Therefore, the extension of the design space resulting from exploration leads to discovery whereas transformation results in invention. Based on these two modes, a distinction is made between creative and innovative designs: creative design implies the introduction of new design variables into the existing design space, and innovative design is a transformation of a range of values for existing design variables.

In turn, a computational model of design process must define precisely what techniques, operators, sets of variables and types of constraints are needed to implement one or another problem solving methodology with respect to a specific task. Therefore, the same design theory can be represented by a variety of computational models of different degrees of consistency.

According to cognitive and systematic schools of AI, all known computational models of design can be divided into two categories:

- Computational models of design based on a *cognitive approach to design*. Such models are being built on the basis of existing AI tools which simulate the cognitive activity of a human designer. These models are built upon simulating the *process* of creating novel artifacts and refer to the AI process school of thought and known as *process-oriented* models of design.
- Computational models based on *systematic design theories* which are independent of cognitive methods or to integrate both cognitive methods and exact methods. This type of models utilizes the idea that it is more important to focus on how to deal with qualities of products rather than on how these qualities are obtained. These models are known as *product-oriented* models of design.

The advantage of product-oriented design models over process-oriented ones is that they can be checked against physical reality while the others are rather subjective. Nevertheless, it is not possible to model only a process or a product since there might be neither a process without a product nor vice versa. For instance, if a design theory of how to generate new design solutions from physical principles is available, then the theory of how to search for new problems is not known yet. To solve this problem, a cognitive technique, for instance, reasoning by analogy might be integrated with a model built upon reasoning from physical principles. However, availability of results depends on what general reasoning strategy has been chosen. For this reason, we argue that a product-oriented model of design which, of course might incorporate some process-oriented mechanisms, is preferable.

#### 2.4.4 Models of Design Based on Cognitive Methods

One of the first AI methodologies successfully transferred into industry was a problem solving technique for *heuristic search based on logical reasoning*. The famous system DAA dates back to 1985 (Kowalksi [1985]). However the idea of modeling the expert knowledge and reasoning process performed by an inventor failed since the resulting expert systems are very hard to modify when new knowledge emerges. Besides, the framework expert systems are built upon is limited to representing the knowledge of narrow domain. However, as we noted in previous sections, the main distinction between routine and non-routine designs is the degree of utilization of principles behind the specific domain.

Another widespread AI technique for design problem solving is *Case-Based Reasoning* (CBR) (Kolodner [1991], Wills & Kolodner [1994]), which utilizes the idea of the reuse of previous experience. In contrast to traditional expert systems, a CBR system is possible to modify when new knowledge and requirements emerge. The essence of CBR is that it realises a model of reasoning by reusing previous knowledge represented in the form of cases. With CBR, no design theory is needed. CBR organizes the process of innovative design in a manner most human designers do: conscious or unconscious search of their memory for previous problems and solutions that are similar to a given design problem. Computer implementation of CBR consists in a case repository storing previous design cases indexed according to problem specifications. The search for a new solution starts with the identification of what indices of a given problem are identical to the stored ones. Examples of AI systems built on CBR are CADET (Navinchandra [1991]), in which a behavioural description is elaborated to have new indices so that various parts of the behaviour can be identified and CADSYN (Maher et al. [1993]) which decomposes cases into subsystems when no matching case is found. If the solution retrieved does not meet all the requirements given, a case adaptation has to be performed. However, there might be a case when domain knowledge is not enough for meaningful adaptation. As a result, CBR works well whilst no significant adaptation is needed. Therefore, CBR is highly applicable to routine design tasks and less or even non-applicable to solving non-routine design tasks unless a technique for *conceptual adaptation* is known. Although the CBR method is, perhaps the best way available to collect, index and retrieve previous experiential knowledge, for innovative design problems the task of

conceptual adaptation is more difficult and has much higher complexity than the search for a similar case.

Attempts to overcome the limitations of CBR are made by trying to organize design based on **reasoning by analogy**. An analogy in AI is defined as the product of cognitive process which establishes mappings between causal structures of different domains (Keane [1988], Wolverton & Hayes-Roth [1995]). In the approach to design exploration using analogy (Qian & Gero [1996]), it is claimed that in contrast to the literal similarity that is used in CBR, the use of analogy makes it possible to map relational structures and, especially, higher-order relations underlying qualitative casual relations. The disadvantage of using reasoning by analogy is similar to one of CBR: although reasoning about more deep knowledge is possible and more creative solutions can be generated, the lack of a general strategy for limiting the search for a solution appears to be very crucial for this method. Another problem appears when we try to represent knowledge of different domains to organize analogical reasoning: it is unclear in what form all the relations are to be presented and there is no guarantee that some very important chunks of this knowledge will be neglected.

Another rapidly progressing modern approach to designing novel artifacts is based on **genetic programming**. Genetic programming is an analog of the naturally-occurring evolutionary process (Holland [1975]). The genetic programming can be applied to solve various scientific and engineering problems. With respect to design tasks, for instance, in DARWIN (Kruiskamp & Leenaerts [1995]), new topologies of operational amplifiers are designed on the basis of the evolution of an initial population of hand-designed topologies through the operations of crossover and mutation. The fitness of each amplifier is computed using a deviation between the actual behaviour of a circuit and the desired behaviour. Although genetic programming is a promising technique, it is solely based on a predefined set of solutions drawn from a specific engineering domain. As a consequence, the resulting population incorporates the same fundamental principles of the specific domain which the initial population is based upon. To be able to draw and utilize different principles, a system based on a genetic algorithm should have unlimited power and capacity, since successful fundamental mutations take very long time as occurs in nature.

The general disadvantage of the above mentioned AI techniques is that while a limited set of innovative design solutions within known design space is possible to generate, the solutions based on principles unknown in this particular engineering domain might not be found. To solve this problem, better understanding of the nature of the innovative design process is needed. The goal of understanding the nature of innovative design is twofold: first we want to know if general principles behind the innovative design process exist and, if they do, how to apply the principles to generate new design solutions.

#### **2.4.5 Models of Design Based on Design Theories**

In Smithers [1996], a dramatic trend within the AI in Design community is outlined: “a survey of AI in Design research shows not just a lack of development of usable theory (or theo-

ries) of design process, it also demonstrates widespread ignorance and neglect of related and relevant work on the fundamental nature of design process by researchers in the Design Research Community". This is a very important issue, and it addresses not only the question of the place of design theories in AI. The best design theory, by definition, is the theory which is based on a set of axioms that would make it independent of the cognitive abilities of a human designer. Examples of systems built upon CBR or genetic programming show that, without such a theory, the size of the knowledge base must be enormous in order to incorporate all existing designs; the reasoning methods available are not effective enough to generate the expected solutions, and there is no guarantee that the feasible solution(s) will be found at all.

As follows from the analysis of failures of using pure AI techniques to build AID systems for solving non-routine design tasks, it becomes obvious that there are two main problems to be tackled:

1. How to *model* and *represent* the existing design knowledge of many domains to make it available at different steps of domain-independent reasoning process.
2. What *reasoning strategy* should be used to support reasoning with this knowledge to generate and evaluate innovative design solutions in different domains.

In order to answer these questions, many research efforts have been undertaken lately. First, it is clear that it would not be possible to model and represent all available design knowledge in a uniform way and independently of a particular task. Therefore, there is the need to select what knowledge categories are to be chosen and how they fit the general reasoning methodology. Second, the reasoning strategy must be able to deal with ill-defined problem formulations and less travelled design search space.

One of the solutions to the problem of dealing with high complexity of engineering knowledge is seen as abstracting from specific engineering data to the level of physical knowledge. Any engineering system can be modelled as a set of physical phenomena occurring within the system and interacting with each other. Therefore, the most difficult part of the reasoning process concerning innovative modification of the system can be performed using abstract knowledge. Apart from that, the advantage of such a modeling framework is that physical principles can be organized and accessed in a systematic way while it is unclear how to access detailed engineering knowledge, especially when cross-domain search is involved. From this point of view, a conceptual engineering design can be modelled as a two-level knowledge translation process: first, finding what physical principle(s) best matches the requirements and constraints given, then - instantiation of the selected principle(s) into specific design description. The reasoning strategy should be based on systematic search and evaluation of the solutions.

Attempts to build a theory of reasoning with abstract physical knowledge, but to prevent this knowledge from possible incorrect interpretation led to appearance of *naive physics* (Forbus [1984], Akman & ten Hagen [1989]). The basic assumption behind this approach is that the

highly complicated world of physical systems should be regarded from of viewpoint of what the qualities of the systems are, rather than what their quantities are. The dynamic aspects of a system's behaviour are expressed through so-called *qualitative simulation* based on qualitative approximation of differential equations constituting relations between system parameters. An example of integration of qualitative reasoning and Case-Based Reasoning is shown in Sycara & Navinchandra [1989]. However, modeling the qualitative behaviour of physical systems seems to be more suited for tasks of explanation and generation within AI whereas quantitative models are more applicable for mathematical simulation. This is an argument which is difficult to discuss unless AI systems for innovative design synthesis based on the concepts of naive physics are available and, therefore, might be studied, compared and evaluated. However, naive physics has given impact to modeling physical knowledge for several AI in Design applications which will be discussed below.

One of the well-known models of innovative design based on the utilization of physical knowledge is *designing from first principles* proposed by Williams [1990]. This approach is based on the assumption that engineering devices are constructed by focusing on qualitative differences between how alternative devices work. The design process is represented as *reasoning from first principles*. It is also assumed that every specific domain has its own set of first principles, and new devices topologies can be composed by generating new behaviours by imagining all possible interactions produced by components and connections. The basic limitation of this approach is that it regards innovative design as a modification of design space based on transformations which are allowed by the principles previously incorporated in a particular engineering domain. As a consequence, utilization of the principles of other domains does not seem to be possible.

More recent research efforts have concentrated on tackling this dilemma by extending the paradigm of designing from first principles to *designing from physical principles*. In Taura et al. [1994], a framework for constructing the natural law database is presented which studies the method for representing physical knowledge based on object-oriented paradigm. The goal of creating the database is to establish a relation between past technological inventions and the physical principles the inventions are based upon. Nevertheless, the selected modeling framework does not enable the automatic synthesis of new design concepts although it might be useful for assisting purposes.

A further extension of the paradigm of designing from physical knowledge is *designing from generic knowledge*. One of the first design-oriented AI systems based on the reasoning from generic knowledge was EDISON (Dyer et al. [1986]). The system was intended for the generation of novel mechanical devices and consisted of so-called conceptual primitives representing models of physical phenomena. However, these primitives were rather intuitive.

To enable cross-domain reasoning, all available physical laws and phenomena are to be modelled in a way which would make it possible to combine the principles selected from different physical domains without violating the law of energy preservation. As a fundamental background allowing for such modelling, a theory of *Bond Graphs* (Karnopp et al. [1990]) is used. The Bond Graphs theory was derived from more general System Theory (Shearer et

al.[1969]) and regards the world of physical phenomena as a set of operations on energy transformations and storage. This makes it possible to represent various physical effects in a uniform way. Once a particular physical phenomenon is described as a specific transformation of the energy flow of a certain type, it can be connected to another phenomena which is able to deal with the same type of the energy flow (Zaripov [1988]). The material information is not included in Bond Graphs structures.

To enable mappings between real word problems and predefined models of physical effects, a problem given is identified with a physical function which is to be delivered by a resulting concept as well as with a set of constraints indicating physical limitations. The goal function of the device and intermediate functions defining connectivity of physical phenomena are modelled in terms of elementary energy transformations. Applying the law of energy preservation, it is possible to synthesise a physical structure containing a set of effects linked together which would deliver the required function.

An example of the research using Bond Graphs theory as a modelling framework is shown in Malmqvist [1993]. The system for automated generation of design alternatives discussed in this work is based on Hubka's theory of engineering design (Hubka & Eder [1988]) which regards machines as transformation systems. Given a physical function, it is possible to synthesise several alternative bond graphs of the process structure which can later be translated into detailed design description.

The principal disadvantage of modeling frameworks which use Bond Graphs is that they are limited to modeling energy-transforming systems. For this reason, information and material transformations can only modelled if they are carried by an energy flow. As a consequence, a large number of innovative problems dealing with material transformations may not be supported by the Bond Graphs approach unless these transformations are not represented in terms of energy transformations. Another problem has to do with the evaluation of generated design alternatives: due to a lack of material information, it is impossible to evaluate the quality of the solutions at the phase of concept generation with respect to material implementations.

An attempt to tackle the above mentioned disadvantages was undertaken in building Knowledge Intensive Engineering Framework (KIEF) (Ishi et al. [1993], [1995]) which enables designers to synthetically build models of mechanisms with a description of physical laws that are essential components to compose design object models. Early versions of KIEF included a knowledge base of physical phenomena modelled on the basis of the Qualitative Process Theory. However, to eliminate the limitations of qualitative modeling, KIEF was extended with external design object models such as Bond Graph models and models for processing quantitative information.

Another approach worth mentioning is based on a combination of product-based design derived from the German school for systematic design and Suh's *General Principles of Design* which provides guidelines for improving design process (Schmekel [1992]). The General Principles of Design are formulated on the basis of the so-called *The Independence*



*Axiom* which states that “a design which maintains the independence of functional requirements is better than a design which does not” (Suh [1993]). Therefore, by analysing what parameters of the existing design negatively influence each other, it is possible to evaluate what design has a lower information content or what parameters have to be decoupled to achieve the better design. However, implementation of this methodology can probably give best results at the phase of evaluation of design alternatives when tools specifying how the functional requirements can be decoupled are not available.

Recently, the **ontological approach** has gained popularity for developing modeling frameworks. (Gruber [1993]). An ontology enables one to make commitments about various types of knowledge, represent knowledge in a uniform way and make the knowledge sharable between various application domains. As a consequence, knowledge represented within one ontology can easily communicate with knowledge of another ontology after having agreed on ontological commitments. This point is crucial for innovative design since it deals with very different types of knowledge. Another aspect — sharability of knowledge — plays an increasingly important role in establishing a reasoning process when the same knowledge has to be shared between different domains.

Among recent examples of building ontologies that would be applicable to engineering design are MathEng (Gruber & Olsen [1994], an ontology for representing mathematical information about technical systems, SHADE (McGuire [1993]) which enables knowledge sharability for collaborative engineering, and PhysSys ontology for physical systems (Borst et al. [1995]), to name a few.

An example of an ontology which introduces a framework for modeling both domain and design knowledge is YMIR (Alberts [1993]). YMIR is based on a formal approach to design drawn from System Theory and captures many advanced techniques for knowledge modelling and reasoning. As opposed to systems which produce recommendations in terms of general physical principles or ready-to-use specific design descriptions, YMIR has the ability to introduce intermediate levels of abstraction between a generic design principle based on a certain physical phenomenon and the design instances of this principle. To instantiate generic knowledge into design descriptions, a set of ontological concepts and reasoning strategy were developed to enable automated translation between levels. Since, as a modelling framework YMIR is the best suited to the goals of this thesis, it will be discussed in more details in Chapter 4.

A proof that innovative design can be very effectively organized in a systematic manner was developed in the Russian school of inventive design. Several computer systems were built on the basis of the Theory of Inventive Problem Solving, among them are INVENTION MACHINE (Tsourikov [1993]), EDISON (Glasunov [1993]). In reality, these systems are still far beyond the utilization of AI methods for representing knowledge and automated reasoning. However, the theoretical background behind these systems is based on integration of designing from physical principles and methods for engineering conflict elimination which provide guidelines for clearing specifications. Therefore, a procedure for the generation of

new design concepts can be performed by searching for the available physical principles guided and constrained by formulated contradictions.

### 2.4.6 Discussion

In this chapter we did not attempt to provide a comprehensive coverage of all existing AI in Design tools and methods. The goal was to indicate, highlight and compare research projects which so far have been promising and relevant to our area of study.

The overview of AI-based approaches to automate innovative engineering design presented in previous sections shows that there is no an agreement between researchers of what innovative design is either in AI or in Engineering Design. First, there is no clear understanding of what the fundamental background of innovative design is - is it an activity which can be governed by theoretical statements or is it a kind of activity solely based on human cognitive capabilities? Until now, it has been difficult to accept or reject arguments of either side.

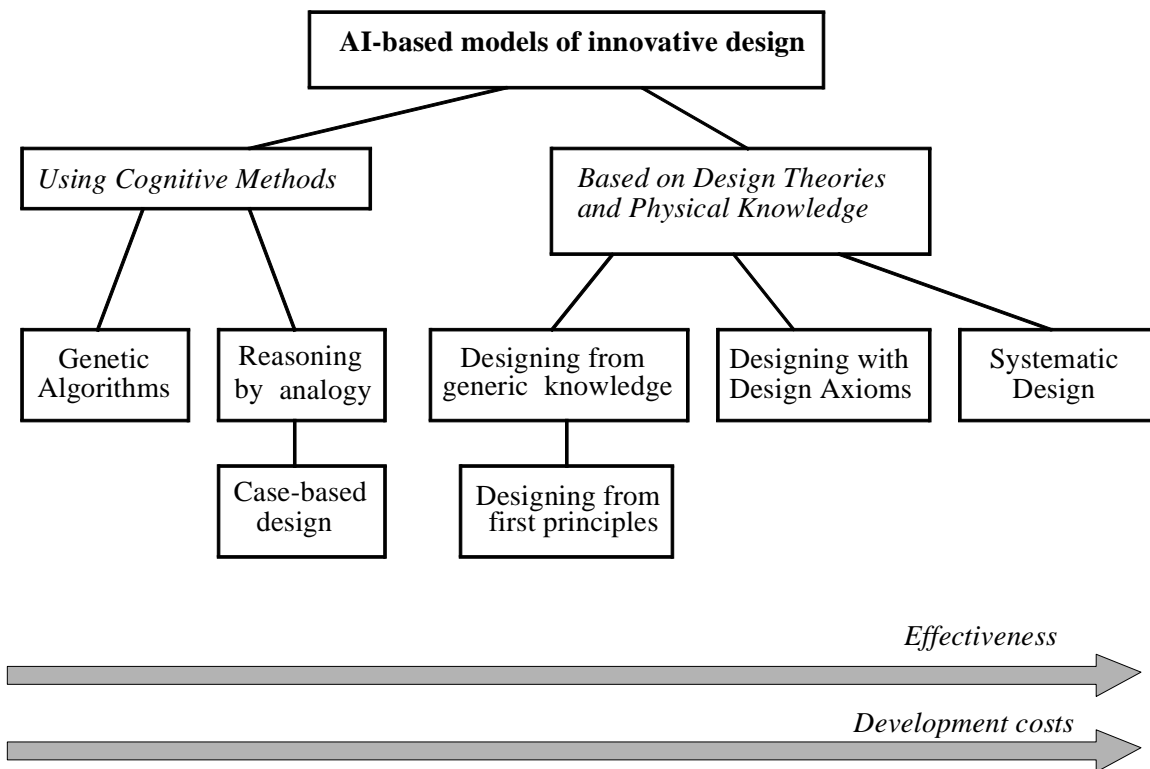


Figure 2.5: AI-based models of innovative design

Sometimes, the argument rises that conceptual design can not be modelled as a search process. The reason is that an output of conceptual design is new knowledge that has not been known before and, therefore, may not be found among available solutions.

From our point of view, innovative design can be organized as a search process, since the resulting knowledge is actually not a new knowledge regarded at the high level of generalization. A new design product description is an *instance* of some known fundamental knowledge after a novel *interpretation* of this fundamental knowledge in terms of a specific situation has been made. We argue, that the search might be organized at the level of fundamental knowledge and new design concepts can be represented in terms of fundamental physical knowledge.

Figure 2.5 presents a classification tree of models of innovative design based on the existing mainstream within AI in Design research according to their efficiency. The efficiency rate is an integrated factor which comprises such characteristics of models as the size of search space, applicability in different domains, availability of solutions, quality of solutions generated, and a capability to evaluate solutions.

As seen from the picture, the most promising are models built upon general theories of design. To be successfully implemented as computer tools they must be able to deal with various types of knowledge representation and reasoning methods. This, however, causes growth of development costs.

Another important conclusion is that any computational model of innovative designs would involve reasoning about a large diversity of knowledge. This knowledge can be of various degree of abstraction: it might be facts on physical properties of specific materials as well as general principles for design. To be successfully applied, it should be represented in a sharable way. As a possible means to achieve the sharability, we use an ontological approach to modeling knowledge.

## Chapter 3. *The Theory of Inventive Problem Solving*

*“I invent nothing. I rediscover”.*

*Auguste Rodin (1840-1917)*

*“Keep on the lookout for novel ideas that others have used successfully. Your idea has to be original only in its adaptation to the problem you’re working on”.*

*Tomas Edison (1847-1931)*

### **3.1. Introduction**

Until now, designers and engineers did not have access to a systematic approach to solving problems arising in the predevelopment phase of design. As a consequence, a lack of general methodology for solving inventive problems in a systematic way did not allow computer scientists to develop a computable model that could be implemented as a computer program.

The previously little known, the Theory of Inventive Problem Solving (TRIZ<sup>1</sup>) originated by Russian engineer and scientist Genrich Altshuller proved that a systematic approach to the inventive process is possible. A major conclusion of Altshuller's studies was that inventions were not a result of unorganized thinking, but instead the products of objective laws and trends of technology evolution. However, these laws have been unconsciously utilized by inventors. TRIZ made these trends explicit.

Comprehensive studies of the patent collections<sup>2</sup> following this discovery resulted in two more findings. First, Altshuller shows that an inventive solution results from *elimination* of a *contradiction* which is caused by attempts to improve preceding design products (Altshuller [1956]). Attempts to compromise without eliminating the contradiction do not allow a designer to achieve the desired degree of improvement. The second conclusion is that the majority of the patented inventions complies with a relatively small set of basic principles for eliminating the contradictions.

Based on these findings, Altshuller and his research associates have developed a scientifically-based problem solving methodology which codifies numerous inventive principles and incorporates the laws of engineering system evolution (Altshuller [1984], Arciszewski [1988], Fey et al. [1994], Linde et al. [1994], Sushkov et al. [1995]).

Assuming that there is a lack of literature available in English that presents TRIZ from a scientific point of view, we decided to briefly discuss the major aspects of TRIZ fundamentals and problem-solving techniques in the book<sup>3</sup>.

This chapter is divided into three parts. First, we present TRIZ historical background and philosophy. Second, TRIZ problem solving techniques based on a systematic approach to inventive design are discussed. To provide better understanding of TRIZ concepts, they are illustrated by examples of inventive solutions drawn from different engineering domains. In the last part, we present ARIZ, an integrated problem-solving technique which aims at solving the most difficult inventive problems.

We should however note, that our presentation of TRIZ slightly differs from TRIZ presented in Altshuller [1984] due to two factors: i) a direct translation of TRIZ terms into English made in the book appeared to be misleading and we had to improve translations; and ii) our study resulted in a restructuring of some TRIZ parts, namely, the theory of technology evolution.

- 
1. There are two abbreviations available in the literature on the Theory of Inventive Problem Solving: TRIZ (English abbreviation) and TIPS (Russian abbreviation). There is still no common agreement what abbreviation to use. To avoid confusion, we will further use the abbreviation TRIZ since it is more widely used.
  2. It is claimed that over 40 years TRIZ researchers studied more than 1.500.000 patent descriptions drawn from different areas of technology.
  3. Recently, there has been much discussion about what TRIZ is between various schools of design. Definitely, TRIZ is not an exact theory from a scientific point of view, since most of TRIZ conclusions were obtained by abduction. We would prefer to label TRIZ collections of problem solving techniques as a *set of recommended practises based on empirical rules of design*. However, to avoid confusion with existing literature, we will use the abbreviation *TRIZ* throughout the book.

## 3.2. TRIZ contents

In this section, we present a brief overview of TRIZ historical background, TRIZ philosophy and the structure of modern TRIZ.

### 3.2.1 Historical background

TRIZ was originated by Altshuller in 1946. The first problem solving technique labelled *inventive principles* resulted from studying more than 40.000 patent descriptions. It indicates what general principles are applicable to solving various types of contradictions occurring in different engineering domains. Currently, the collection consists of 40 inventive principles.

Further, TRIZ evolution resulted in the creation of a more precise technique for inventive problem solving - a collection of Inventive Standards, also known as the rules for substance-field modification. An Inventive Standard enables a designer to represent a part of a product where a problem occurs in terms of the specific components the design product consists of and modify the product.

The third TRIZ major component is Pointer to physical, chemical and geometrical effects. They provide a mapping between a specific function a new design product has to deliver and descriptions of physical, chemical and geometrical phenomena that are capable of delivering the required function.

Every TRIZ problem solving technique can be used independently. All three problem solving techniques are presented in section 3.4.

Further studies of patent collections revealed the trends of technology evolution (section 3.3) which form a theoretical background of modern TRIZ.

Apart from specific problem-solving techniques, Altshuller developed an integrated framework which enables consistent use of different TRIZ problem solving techniques. The framework was labelled *Algorithm for Inventive Problem Solving (ARIZ)*<sup>1</sup>. The first version of ARIZ was developed in 1961 (Altshuller [1969]). Since, the ARIZ has been largely extended and modified. The latest Altshuller version of ARIZ<sup>2</sup> is ARIZ-85B. ARIZ-85B is presented in detail in section 3.5.

Lately, TRIZ has proven itself as a very strong tool to solve inventive problems and invent new products. In 1984, about 300 TRIZ research and training centres existed in the USSR. In the end of nineties, TRIZ became known outside Russia and the former Soviet states.

Today, TRIZ continues to develop and there are also attempts to integrate TRIZ with other methods for design. For instance, one research direction investigates the applicability of TRIZ for supporting Suh's General Principles of Design (Nordlund [1994]). Another direc-

---

1. Similarly to TRIZ, ARIZ is also known as AIPS.

2. Other versions of ARIZ developed by different authors are also known. However, it is the professional judgment of the authors that ARIZ-85B is the most consistent version of the algorithm.

tion studies how TRIZ can be combined with other methods for product conception, such as Quality Function Deployment (Verduyn & Wu [1995]).

### 3.2.2 Levels of Inventive Solutions

To provide better understanding of the origin of inventions, Altshuller developed a new classification of design solutions. He divided all technical solutions into five large groups. The groups differ in the way the solutions were obtained:

1. **Routine Solution** does not require that contradictions be solved, since the needed solution can be found within the same engineering domain. Standard solution results from a redesign of some of the existing components of a system.

#### *Example 3.1.*

To enforce a building, its walls might be made more thick. □

2. **Change of a System:** pertains to more difficult problems which contain contradictions. A contradiction arises when mutually exclusive requirements should be placed on the same system. However, the contradiction might be eliminated by methods known within the discipline.

#### *Example 3.2.*

Foam instead of water can be used to intensify putting out fire. □

3. **Innovation:** a solution to a problem is outside of the engineering discipline, but *available* in another field of engineering. Innovation results in replacement of a system which causes the contradiction.

#### *Example 3.3.*

To measure a thickness of a thin conductive wire in microelectronics, it was proposed to use a known method in the wire manufacturing industry for measuring the diameter of microwire using the physical effects of Corona discharge. □

4. **Invention:** the solution is not available within existing engineering fields, but it can be found among physical or chemical phenomena. As a result of invention, a new technical system is synthesized.

#### *Example 3.4.*

Invention of a vehicle using air cushion. □

5. **Pioneering invention:** the solution is based on recently discovered scientific phenomena. Pioneering inventions launch new engineering disciplines.

**Example 3.5.**

Invention of radio or photography. □

Using TRIZ, it is possible to successfully tackle engineering problems of the second, third and fourth groups: those problems which can not be solved without elimination of contradictions. Instead of a chaotic search for the solutions proposed by cognitive methods, such as brainstorming or synectics, TRIZ organizes the search in a systematic way.

### 3.2.3 TRIZ Philosophy

More than 40 years of studying patents in different areas of engineering resulted in several important discoveries which form the TRIZ philosophy:

1. Every engineering system evolves according to regularities which are general for all engineering domains. These regularities can be studied and used for innovative and inventive problem solving, as well as for forecasting the further evolution of any engineering system in design terms.
2. Engineering systems, like social systems, evolve through the elimination of various kinds of contradictions. The principles for eliminating the contradictions are universal for all engineering domains.
3. An inventive problem can be represented as a contradiction between new requirements and an engineering system which is no longer capable of meeting the requirements. Finding an inventive solution to the problem means to eliminate the contradiction under the condition that a compromise is not allowed.
4. Frequently, when searching for the inventive solution to a problem formulated as a contradiction, there is the need to use physical knowledge unknown to the domain engineer. To organize and direct the search for appropriate physical knowledge, pointers to physical phenomena should be used. In the pointers, the physical phenomena are structured according to technical functions that can be achieved on the basis of the phenomena.

Classical TRIZ which will be discussed below consists of several problem modeling and problem-solving techniques. It introduces a uniform way of modeling inventive problems by representing them in terms of contradictions and generic principles for resolving the contradictions. A comprehensive study of patent collections undertaken by TRIZ researchers and thorough tests of TRIZ within industries have proven the fact that if a new problem is represented in terms of a contradiction, then it can be solved by applying the relevant TRIZ principle. The principle must indicate how to eliminate the same kind of contradiction encountered



in some engineering domain before. However, the most important achievement in TRIZ has been the formulation of general problem solving principles covering virtually all possible types of innovative and inventive problems.

### 3.2.4 TRIZ Structure

The structure of TRIZ is shown in Figure 3.1. It includes the following parts:

1. **Laws of engineering system evolution.** This part of TRIZ studies and formulates general trends of engineering system evolution.
2. **Problem solving techniques.** The techniques aim at building a problem model and producing recommendations on how to solve the problem. Among them are:
  - 2.1. Principles for the elimination of engineering conflicts;
  - 2.2. Principles for substance-field modeling for solving inventive problems through representing them in terms of substance-field interactions and applying generic patterns of interaction transformations.
  - 2.3. Pointers to scientific-engineering effects. This part of TRIZ focuses on studying how to use the knowledge of exact sciences in the inventive process.
  - 2.4. Algorithm of Inventive Problem Solving - an integrated technique aimed at solving most difficult inventive problems.
3. **Collections of advanced patents.** This part contains patent descriptions drawn from various engineering domains. The patents are structured according to inventive principles used to solve one or another type of contradictions. The patents can be used as analogous design cases illustrating the applicability of the principles and making a problem-solving process easier.
4. **Functional and Value Analysis.** It is a modified version of traditional Value-Engineering Analysis proposed by Miles [1972] with the focus on functional analysis of technical systems.

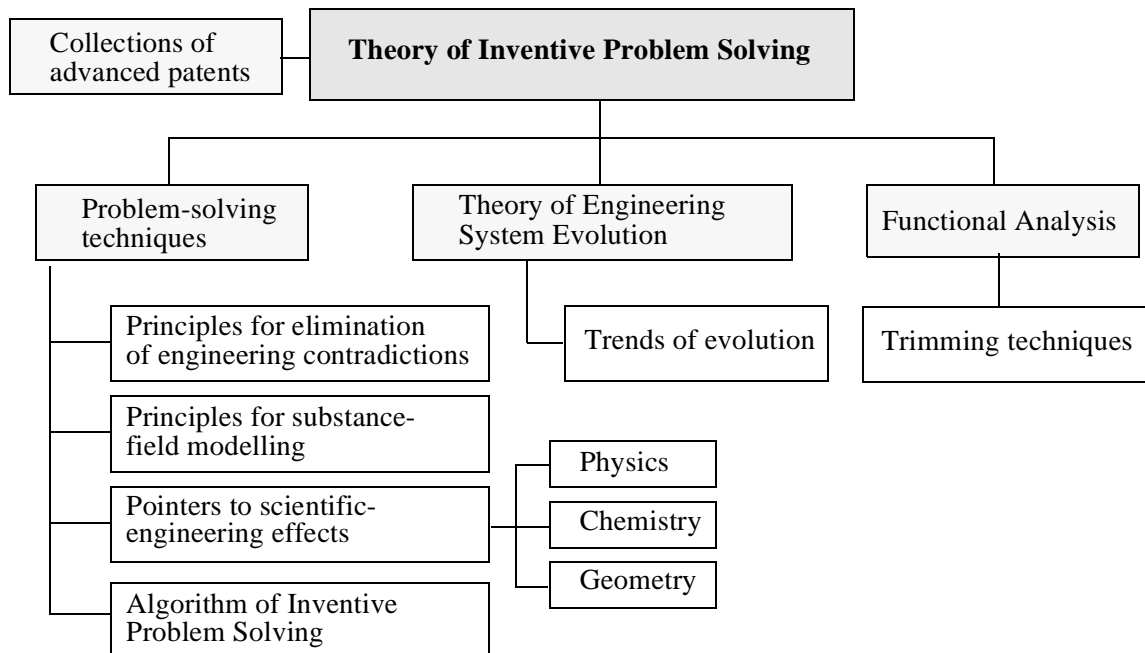


Figure 3.6: Structure of the Theory of Inventive Problem Solving

### 3.2.5 Technical system

TRIZ operates with a number of concepts which are used for problem formulation and solving. Since TRIZ was developed to support solving problems dealing with physical systems, its concepts can be regarded as generic with respect to any design product involving material-energy transformations.

To systematize knowledge about the inventive process, and to make reading the next sections more clear, we should define the major TRIZ concepts:

#### *Definition 3.1.*

**Technical system<sup>1</sup>:** any artificially created system which consists of interacting material components and energy flows. The technical system possesses a deterministic behaviour in order to deliver a given functionality. Any design product can thus be regarded as a technical system if it involves transformations of materials or energy flows no matter how many components are involved.

1. In some other publications on TRIZ, the term “engineering system” is used.

It is important to note that most of the concepts presented above are defined in TRIZ in a qualitative manner. In the next sections, the major TRIZ problem-solving techniques based on these concepts are discussed in details.

### 3.3. Theory of Technical System Evolution

The theoretical background behind TRIZ is formed by the *Theory of Technical System Evolution (TTSE)*. Our opinion is that an understanding of the systematic nature of technology evolution is very important in building a computer support for design. Respectively, in this section, we present trends of technology evolution as well as more specific evolution patterns. Since we do not aim at comprehensive coverage of TRIZ in the book, we will only present the basic concepts.

#### 3.3.1 Systematic nature of technology evolution

A long-term study of patent collections and hundreds of bibliographical sources presenting the history of technological development resulted in the conclusion that the process of technological evolution is not chaotic even if a designer is not aware of it. The development of technology correlates with the evolution of societal needs which, in turn, obey more general laws of nature. Therefore, the process of evolution of technical systems can be studied to reveal what regularities underlie it.

A further research in TRIZ was aimed at revealing and studying basic laws and trends of technical system evolution (Altshuller [1984], Salamatov [1991]). The laws and the trends indicate general regularities of evolution which can be applied to each branch situation. TTSE states that the evolution of any technical system obeys these regularities independently of the domain the system belongs to. As a consequence, the applicability of the laws and trends is not restricted to a single domain, they are valid for the every engineering discipline.

In TRIZ literature, no separation is made between the laws and the trends. Nevertheless, by looking at TTSE laws as proposed, for instance, in Salamatov [1991], it becomes obvious that there is a clear distinction between:

- *the laws of evolution* which denote general conditions for the creation and development of technical systems.
- *the trends of evolution* which denote what particular phases of evolution a system passes. However, not every system passes all phases of evolution indicated in the trends - some of the phases could be missed. Therefore, it would not be correct to refer to them as to laws.

### 3.3.2 The law of system ideality

The primary law of the theory of technical system evolution is the *law of system ideality*. It states that during evolution over the time, any system tends to increase a ratio between the overall performance of the system and the expenditures necessary to provide the required performance.

The law indicates a principal design requirement which every designer has to keep in mind while creating a new engineering system: a system being designed must be able to deliver the best performance possible, whereas the expenditures required to provide the system life-cycle should be minimized. The expenditures in this definition are all types of energy, material and informational resources required to deliver the given functionality and meet all other requirements.

Formally, the end point of ideality growth for system  $\Phi$  can be expressed as:

$$I = \lim_{\substack{E \rightarrow 0 \\ P \rightarrow \infty}} \Phi_p(E)$$

where  $I$  is a system ideality and  $E$  are the expenditures involved to achieve a system's performance  $P$ . In turn, two particular types of achieving the ideality are possible:

$$I_1 = \lim_{\substack{E \rightarrow 0 \\ P = \text{const}}} \Phi_p(E)$$

that is, expenditures decrease without performance change and

$$I_2 = \lim_{\substack{E = \text{const} \\ P \rightarrow \infty}} \Phi_p(E)$$

that can be interpreted as performance growth without increasing the expenditures.

Strictly speaking, an ideal system is a philosophical category since it may not exist. Nevertheless, it is useful to use this notion in engineering just like it is used for modeling purposes in exact sciences. For instance, idealized models in physics or chemistry help to create and describe the existing and possible physical systems.

#### **Example 3.6.**

When developing a portable radio station for mountain-climbers the problem of ensuring temperature stabilization of the oscillator's quartz crystal arose. Instead of using a con-

ventional thermostabilizer which is too heavy and needs a special power source, it was suggested to fix the crystal on the mountain-climber's body. Hence, the function of temperature stabilization is fulfilled with no special design at all. □

### 3.3.3 TTSE Laws

TTSE presents four laws of engineering systems evolution including the law of system ideality growth. Other TTSE laws are:

1. **Law of system completeness:** any technical system must have a complete material-energy structure to deliver the required function.
2. **Law of energy conductivity:** a necessary condition of the functioning of any system is providing energy flows through all parts of the system.
3. **Law of irregularity of system's parts evolution:** the more complex a system becomes during the evolution the more irregularly its parts evolve. As a result, further development of the system becomes more difficult due to contradictions arising between system's parts.

### 3.3.4 TTSE Trends

If the laws denote primary conditions and observations for the creation and development of *any* technical systems, the trends define how the system's physical structure changes during the system's evolution. They include generic patterns which present detailed information on those changes. Among TTSE trends are:

1. **Trend of increasing a number of material-energy interactions:** any system tends to increase the degree of interacting material-energy components to provide better performance and controllability.
2. **Trend of frequency and form adjustment:** during evolution, any system tends to adjust frequencies and forms of interacting components.
3. **Trend of dynamics growth:** any system tends to increase the degree of freedom of its movable parts by transition to a more flexible physical structure.
4. **Trend of transition to microsystem:** any system tends to replace a physical principle behind its components delivering main function with a new physical principle utilizing properties of more fragmented materials, microparticles or physical fields.
5. **Trend of transition to macrosystem:** a system which has approached its limits of evolution can evolve further through merging with other systems (that produces a new function); or it can be eliminated if its function might be delivered by other systems.

Knowledge of the trend helps to estimate what phases of evolution have been passed by a system. For instance, the trend of increasing the number of material-energy interactions indicates the early phase of system evolution, whereas the law of transition to microsystem indicates that the system has approached its final phase and soon will be replaced with another system based on another physical principle.

***Example 3.7.***

The law of transition to microlevel states that in the final stage of the system's development, a basic physical effect responsible for performing some system function is replaced with a new effect capable of performing the same function by using field interactions instead of mechanical actions. For instance, instead of mechanical cutting of materials a laser beam can be used. □

***Example 3.8.***

Two solutions can be mentioned to illustrate the law of dynamics growth. The first solution is taken from the aircraft industry: a nozzle of a jet engine is made as a telescopic pipe. While operating, the nozzle is pulled out to its fullest extent, and pulled in for transportation (U.S. Patent 3561679). Another solution is taken from the optical industry: a mirror with changeable geometry. A sectioned pneumatic chamber is placed behind the mirror made of flexible material. The curvature of the mirror surface can be changed by varying pressure in different sections of the chamber. □

The significance of the trend of technical systems evolution for engineering design is that they can be used to estimate what phases of the evolution a system has passed. As a consequence, it is possible to foresee what changes the system will experience. In reality, such evaluation is not easy to perform for a whole system due to the law of irregularity of a system's parts evolution. For this reason, a model of a system can be decomposed into a number of subsystems according to their functions and the physical principles used to deliver those functions.

***Example 3.9.***

Evolution of a mirror in a way of increasing the degree of fragmentation according to the law of transition to microlevel is shown in Figure 3.2. □

Knowledge of both the laws and the trends is very important to avoid developing the system in a wrong direction. There are many examples in history of technology showing how violations of the laws led to making wrong decisions and the design of non-competitive products.

Since the modelling of the laws of evolution is not a part of our research, we omit a full description of them. More information on the laws and patterns of evolution can be found in Altshuller [1984], Altshuller et al. [1989] and Salamatov [1991]. However, further in the the-

sis, some trends will be used as a basis for modelling innovative and inventive designs and we will explain them in more details when needed.

The system of the trends of evolution can be regarded as an independent technique for problem solving. However, with little experience in TRIZ, its applicability is rather difficult. For this reason, they were smoothly incorporated into TRIZ problem solving techniques that makes their use more convenient. An overview of TRIZ problem solving techniques with some examples is given in the next section.

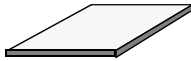
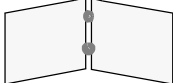
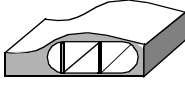

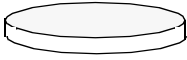


Phase of evolution	Examples	
Solid object ↓		A traditional mirror
Object with joints ↓		A mirror with joints
Flexible object ↓		A pneumatic chamber behind a flexible mirror surface
Powder ↓		A sphere covered with powdered mirror particles
Liquid ↓		Liquid mirror (optics, telescopes)
Gas ↓		Gaseous mirror (plasma mirror)
Field ↓		Use of electromagnetic field instead of visible light

Figure 3.7: Evolution of the traditional mirror according to increasing the degree of fragmentation (the law of transition to microlevel)

## 3.4. TRIZ Problem solving techniques

### 3.4.1 Contradiction as the origin of an inventive problem

A goal of the design process is to map a function onto a physical principle that would be capable of performing the function. But what can be done in a situation when an exact function to be performed is not available? Or, none of previous solutions does not meet the new specifications? For instance, a new requirement might be formulated as "to reduce the noise produced during the manufacturing process". Without knowing a design concept apriori, it is not possible to represent the problem in terms of functional specifications that can be mapped onto some known physical principle. Consequently, innovative design is difficult to perform due to the uncertainty of how an original problem can be translated into the functional specifications.

From this point of view, the most important TRIZ achievement, which perhaps will impact not only engineering science but other types of human creative activities as well, was that TRIZ revealed the common cause of innovative design problems: *contradictions* (Altshuller [1956]). A contradiction arises from mutually exclusive demands that may be placed on the same system where compromising does not produce the required result. Instead of solving innovative problems ad-hoc, TRIZ introduces principles for the formulation and elimination of the contradictions in the systematic way. A systematic approach to eliminating contradictions follows from the systematic nature of technology evolution: each specific pattern of the trend of evolution eliminates a certain type of contradiction arising in the existing technical system (see Chapter 3).

#### *Example 3.10.*

The pattern of transition from a flexible structure to powder (see Figure 3.2) eliminates a contradiction between a necessity for the same object to be solid and liquid. The powder is solid whereas under certain conditions it possesses the properties of liquid (for instance, it can "flow").□

TRIZ states that to obtain an inventive solution, a contradiction must be eliminated whereas compromise or optimization methods are not allowed.

### 3.4.2 Types of Contradictions

A problem that requires an inventive solution arises when a new demand is applied to an existing system and a method for directly meeting the demand can not be found among all previous design cases available in a given engineering discipline. This situation indicates that the problem contains a contradiction that should be eliminated by innovative redesign.



Altshuller proposed to formulate inventive problems in terms of contradictions with respect to already existing design. The existing design can be labelled “a prototype” which has to be improved through redesigning.

Two types of contradictions are known in TRIZ: *engineering* and *physical*. The engineering contradiction arises when it is required to improve some feature of the existing prototype but all solutions known within the domain do not produce the required result or their use would cause a negative effect. The impossibility to improve one parameter and to prevent another important parameter from deterioration is the main feature which separates inventive problems from problems that can be solved by a procedure of routine design.

***Example 3.11.***

Engineering contradiction: To achieve a positive effect “to protect a radiotelescope from lightning”, a lightning rod can be installed near the radiotelescope. However, it would cause the negative effect “the rod screens radiowaves”.□

More precise form of contradictions arising in engineering is a physical contradiction. A physical contradiction indicates that part of a design prototype should have two mutually exclusive values of the same physical parameter at the same time. To eliminate the physical contradiction, a change to the existing physical principle(s) the prototype is comprised of is needed. Formulation of the physical contradiction makes it possible to identify precisely what part of the prototype is responsible for producing the negative effect. It is important to note that if the given problem may not be solved by using the principles for the elimination of engineering contradictions, such a situation indicates that the physical contradiction is present in a system.

***Example 3.12.***

Physical contradiction: the lightning rod installed near the radiotelescope must be conductive to protect against lightning and must not be conductive to prevent from screening radiowaves.□

### **3.4.3 Principles of Elimination of Engineering Contradictions**

The first TRIZ problem solving technique was a collection of inventive principles aimed at eliminating typical contradictions. The inventive principles are heuristic principles based on the accumulated and generalized previous experience of inventors. Due to a high degree of generalization, the inventive principles are available in the form independent of any particular engineering domain. The use of this experience is organized in the systematic way according to which type of engineering contradiction is present in a problem.

To make the inventive principles applicable in a systematic way, Altshuller formulated 39 generalized engineering parameters, like “the weight of a movable object”, “speed”. These are generalizations of diverse specific technical parameters.

A new problem can be solved by the use of a proper inventive principle, after the problem has been formulated as an engineering contradiction in terms of predefined generalized parameters: “a generalized parameter to be improved versus a generalized parameter which deteriorates”.

An inventive principle provides a guideline indicating in what way to solve a problem without causing negative effect. The principle itself doesn't give a solution to the problem, it recommends a method for eliminating a certain type of engineering contradiction.

***Example 3.13.***

The weight of a short steel pipe is small enough and does not hinder the movement of the pipe inside a kiln during thermal processing. However, to process a long pipe is more difficult: its large weight makes the transportation difficult. In this situation, a contradiction arises between the parameters “length of the movable object” and “weight of movable object”. One of the inventive principles suggests the use of pneumatic and hydraulic structures to eliminate this kind of contradiction. One of the known solutions to the problem is to create an air cushion in the kiln, which provides the required movement of long pipes. □

Examples of other inventive principles are:

***Variability Principle:***

- Characteristics of the object (or external environment) should change so as to be optimal at each stage of operation.
- The object is to be divided into parts capable of movement relative to each other.
- If the object as a whole is immobile, to make it mobile or movable.

***Segmentation Principle:***

- To divide the object into independent parts.
- To make the object such that it could be easily taken apart.
- To increase the degree of the object's fragmentation (segmentation).

At the moment, 40 inventive principles aimed at resolving contradictions between generalized parameters are known.

The inventive principles can be used in a systematic way by accessing the principles through indices in a matrix. Along the vertical axis of this matrix the generalized parameters which

have to be improved are specified. Along the horizontal axis the parameters which deteriorate as a result of improvement are specified. These parameters can be looked up along the vertical and horizontal axes and the matrix suggests up to four principles that can be used to solve the contradiction.

<i>Parameters</i>	<i>what deteriorates as a result of improvement</i>				
	<i>Speed</i>	<i>Force</i>	<i>Stress</i>	<i>.....</i>	<i>Stability</i>
<i>Speed</i>		13,28,15,19	6,18,38,40	.....	28,33,1
<i>Force</i>	13,28,15		18, 21,11	.....	35,10,21
<i>Stress</i>	6, 35,36	36,35,21		.....	35, 2,40
<i>.....</i>	.....	....	....		.....
<i>Stability</i>	33,28	10,35,21	2,35,40	.....	

Figure 3.8: A matrix of principles for engineering contradiction elimination. Numbers indicate what principles have to be used: 1 - Fragmentation; 2 - Removing; 10 - Preliminary action; 13 - Other way round; etc.

Selected principles are ordered according to their applicability. The principle that will most likely solve the problem is given first (Figure 3.3).

### 3.4.4 Principles for Physical Contradiction Elimination

Originally, inventive principles were developed for the elimination of engineering contradictions. However, both quality and validity of the resultant formulation strongly depend on the personal views of a designer. For this reason, the method of constructing engineering contradictions is ambiguous and it is unclear how the engineering contradictions can be modelled in a formal way.

As said above, an advanced form of contradictions is physical contradictions. To model an inventive problem as a physical contradiction, one has to identify first what physical object of a prototypical design must have two conflicting properties.

#### *Example 3.14.*

Suppose, we need a roof in a house to protect the house from rain and we do not need the roof when the weather is good.

To solve problems containing physical contradictions, the principles for physical contradiction elimination are used. Among them are:

- *Separation of conflicting properties in time.*

***Example 3.14 (continued).***

A roof is made foldable. It can be folded when the weather is good and unfolded when the weather is bad.

- *Separation of conflicting properties in space.*

***Example 3.14 (continued).***

The roof is made of the porous material which stops the rain drops but passes the air and sunlight.

- *Separation of conflicting properties in physical structure.*

***Example 3.14 (continued).***

The roof is made of a material which becomes solid when interacting with water and porous when water is out.

- *Separation of conflicting properties at microlevel.*

***Example 3.14 (continued).***

The roof is not needed when a heat flow over the house makes the rain drops evaporate.

- *Separation of conflicting properties at macrolevel.*

***Example 3.14 (continued).***

The roof is not needed if something else in the surrounding environment protects the living space. For instance, the whole town can be under a glass cover.□

As seen, the principles of elimination of physical contradictions are very general. Consequently, they are very hard to use without additional support by more specific knowledge. The output of using principles of physical contradiction elimination might not be directly in the form of a design concept or in the form of any specific physical function to perform.

This problem might be tackled by structuring inventive principles according to more general principles for eliminating physical contradictions. In Litvin and Guerassimov [1991] it is

shown that all the inventive principles can be classified according to the principles of physical contradiction elimination. For instance, the inventive principles of variability and periodic action are used to eliminate contradicting requirements in time; whereas the principle of segmentation eliminates contradictions in space. Therefore, the principles for physical contradiction elimination form five large classes of inventive principles. Each type of a physical contradiction is related to a certain inventive principle that contains all necessary knowledge to eliminate the contradiction.

### 3.4.5 Substance-Field Modeling

Another TRIZ problem solving technique is a collection of so-called “Inventive Standards”. Inventive Standards are drawn from the fact that most inventions refer to conceptual modification of *physical* systems. This means that there should be some common problem-solving method applicable to the whole group of similar inventive problems. If problems from different domains result in identical physical models, this means that the problems are similar. As a consequence, they can be solved by applying the same method.

Inventive Standards operate with physical entities a part of the design where a problem arose consists of. This makes the Inventive Standards more specific than the inventive principles.

To model the physical structure of a system which causes the contradiction, the so-called *substance-field modeling*<sup>1</sup> is used. The basic idea behind substance-field modeling is that any part of an engineering system can be represented as a set of substance components and field interactions between the components. The problem is indicated as an undesirable, insufficient or missing interaction between two components. To obtain a solution to the problem means that the given physical structure which contains the undesirable or missing interaction has to be transformed into a structure in which the desired interaction is achieved. Inventive Standards are the rules which indicate what patterns are to be used to transform a given substance-field model.

A substance-field model is an abstract model of the system part where the undesired interaction occurs. Substance components that have complex physical structure can be generalized and modelled in a black-box manner. Boundaries of the system are usually defined by two interacting substance components and a field providing an energy flow between the components. The problem is then specified in terms of the physical attributes of the components or the field to be changed.

The left part of an Inventive Standard specifies conditions of a problem: what restrictions on the introduction of additional components are and what type the substance-field model is. The right part specifies how the model should be transformed to eliminate the undesired interaction. The Inventive Standard itself doesn't specify exactly what particular substances and fields are to be introduced. It presents a general pattern indicating how they should be introduced into the system in terms of abstract physical components.

---

1. In other sources, substance-field modeling can be referred as “S-Field modeling” or “su-fi modeling”.

**Example 3.15.**

Inventive Standard 1-2-1: If a useful and a harmful effect appear between two substances in a substance-field model, with no need to maintain direct contact between the substances, the problem is to be solved by introducing a new substance between them (Figure 3.4). □

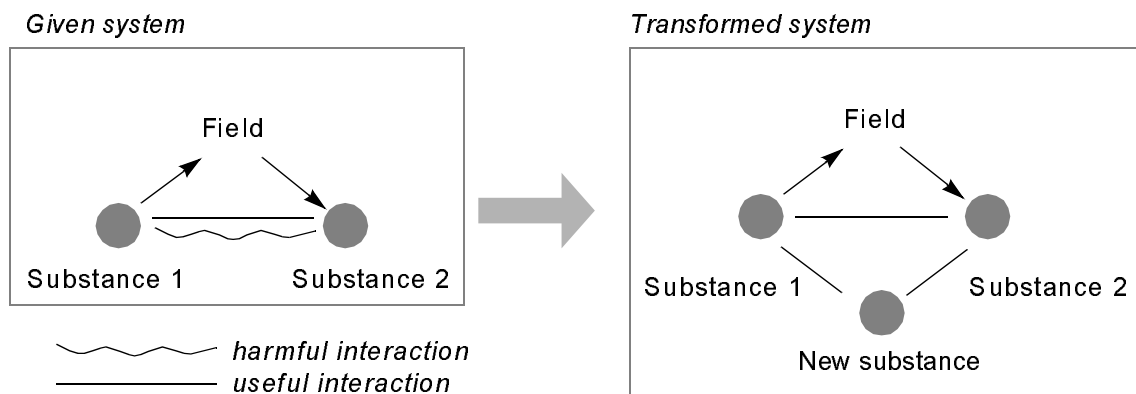


Figure 3.9: Graphic representation of the principle of substance-field transformation

**Example 3.16.**

A system for melting glass ampules filled with a liquid medicine consists of a nozzle producing a flame which melts the ampule's tip, and a container in which the ampules are stored. A problem arises because of the difficulty to support the necessary intensity of the flame: when the flame becomes larger than required, it overheats the medicine.

The substance-field model of the problem is depicted in Figure 3.5 (right side). The arrows show the direction of actions provided by the fields.

This problem can be solved by applying the Inventive Standard 1-2-1 (Example 3.15). One of possible solutions is to fill the container with an incombustible liquid, e.g., water. The liquid protects the lower part of the ampule containing the medicine whereas the ampule's tip is to be kept above the liquid surface. In this case, the need to control the flame disappears. □

We believe that there are two general principles which underlay all possible types of substance-field transformations:

- The value of some physical parameter of a substance or its physical state can be altered by subjecting the substance to a physical field which is capable of changing the value of the parameter or the object's physical state.
- The value of a parameter of a physical field in the required place of a system can be altered by a substance component introduced into this place and capable of changing the field's parameter.

These principles are implicitly incorporated into most of the Inventive Standards.

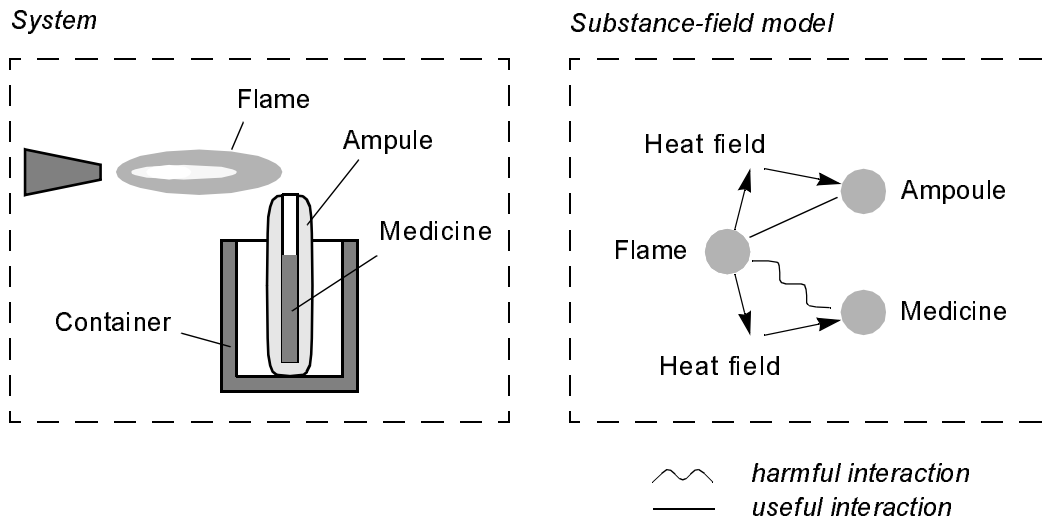


Figure 3.10: Substance-field model of the problem with melting ampoules.

In TRIZ, 76 Inventive Standards are available. However, apart from the Inventive Standards operating with substance-field models, the system also contains the Inventive Standards drawn from the trends of an engineering system evolution directly applicable for problem solving. A problem with these Inventive Standards is that they are very informal, so their practical use is quite difficult.

**Example 3.17.**

Inventive Standard 3-2-1: The effectiveness of a substance-field system can be enhanced by matching (or mismatching) the frequency of a field's action with the natural frequency of a product (or a tool).□

### 3.4.6 Pointers to Scientific-Engineering Effects

Another problem-solving technique suggested by TRIZ consists of the collection of so-called *scientific-engineering effects*. While principles for contradiction elimination and Inventive Standards do not produce recommendation in terms of what physical substances or fields should be used, scientific-engineering effects provide the mapping between technical functions and known natural phenomena.

Studies of the patent collections indicated that the best and, as a consequence, more ideal inventive solutions are obtained by utilizing natural phenomena previously unused in engineering domain. Knowledge of natural phenomena often makes it possible to avoid the development of complex and unreliable designs. For instance, instead of a mechanical design including many parts for the precise displacement of an object for a short distance, it is possible to apply the effect of thermal expansion to control the displacement.

Finding relevant natural phenomena that would be capable of meeting a new design requirement is one of the most important tasks in the early phases of design. However, it is nearly impossible to directly use formulations of natural phenomena in the form in which they are given in handbooks on physics or chemistry. The descriptions of natural phenomena available in handbooks or encyclopaedia yield information on the properties of the phenomena from a scientific point of view, and it is unclear how these properties can be used to deliver specific technical functions.

TRIZ collection of scientific-engineering effects (Gorin [1973]) is designated to bridge a gap between engineering and science. Thus, the definition “scientific-engineering effect” implies that each natural phenomenon in this collection is identified with a multitude of various technical functions that might be achieved on the basis of the phenomenon.

The search for effect is possible through indicating a design requirement to be met by a new design. These requirements are represented in the collection of the effects in terms of predefined technical functions. Each technical function indicates an action to be provided by a system to be designed, for instance, “to move a loose body” or “to change density”.

Each technical function refers to a list of several possible effects to use, which makes it possible to apply even well-known effects in non-ordinary situations. The organization of the pointer to physical effects is shown in Figure 3.6.

#### ***Example 3.18.***

In the TRIZ pointer to physical effects, the function “to achieve a precise displacement” refers to the physical effect of magnetostriction. The applicability of this effect for performing the function is illustrated by solving the problem of achieving a precise distance between a magnetic head and the recording surface in a tape recorder. One end of the magnetostrictive rod is fixed in a rigid position and the opposite end is fixed to the head. A magnetic field is applied to compress the rod exactly to the required distance between the head and the recording surface (SU A.c. 517 927).□



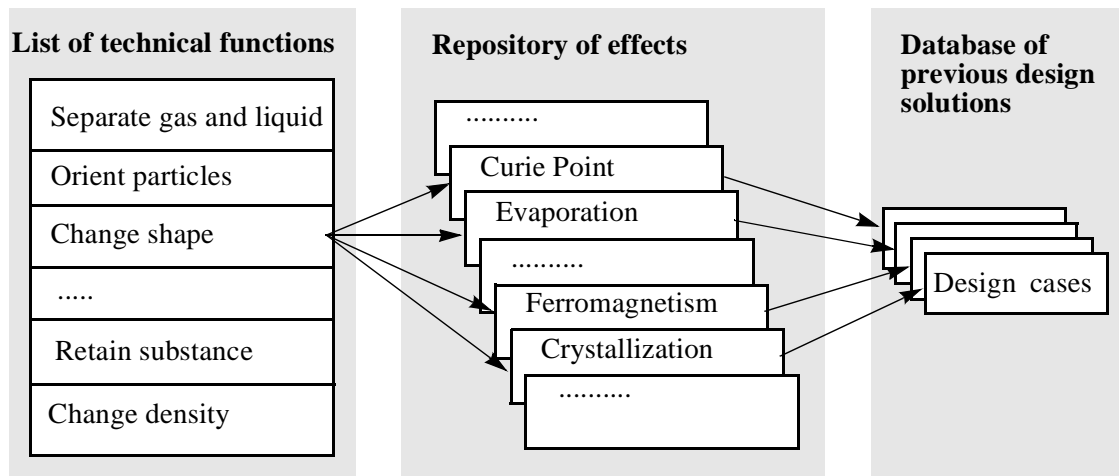


Figure 3.11: Organization of the pointer to scientific-engineering effects

The way of organizing the TRIZ collection of scientific-engineering effects is similar to the approach used by the German School of Design for developing the Design Catalogues (Pahl & Beitz [1984]). However, there are two distinctions between the TRIZ pointers to the effects and the Design Catalogues: first, TRIZ understands a technical function in the broader sense, and any physical function is regarded as generic with respect to producing a multitude of technical functions; second, TRIZ contains unique collections of geometrical and chemical effects which relate chemical (Salamatov [1988]) and geometrical (Vikentiev & Efremov [1989]) knowledge with technical functions. The collections of geometrical and chemical effects are organized in the same way as the collection of physical effects.

### Example 3.19.

Many of the technical functions can be performed by using various geometrical shapes. For instance, the function “to obtain different properties on one side of an object” can be provided with the use of a Möbius tape. The Möbius tape is a twisted tape, glued in a ring. While rotating, it has a one-sided surface, whereas actually it has a two-sided surface. By covering both surfaces of the sheet with different materials, it is possible to obtain two different properties on one side of the sheet. □

### 3.5. Algorithm of Inventive Problem Solving.

After applying TRIZ problem solving techniques, there might be a situation when the inventive principles, inventive standards or the pointer to scientific-engineering effects do not eliminate the formulated contradiction. This situation indicates that the actual origin of the problem has not been extracted during the problem analysis phase. Although the three problem solving techniques discussed above are rather powerful and may be successfully used, they have the following disadvantages:

1. The techniques operate with fuzzy information about an engineering system and its surrounding environment. They are not supported by tools for verifying a problem statement. As a consequence, a model of a problem is constructed ad hoc and may include incorrectly chosen components.
2. An ill-defined inventive situation can be decomposed into various engineering contradictions and substance-field models. There is no tool available in inventive principles or Inventive Standards to evaluate what model to choose from a number of alternatives.
3. The TRIZ problem-solving techniques do not remove psychological inertia in full: specific engineering terms shackle an inventor to ingrained concepts about technical objects, thus significantly complicating the search for new solutions.

Inventive principles and Inventive Standards are supposed to conduct superficial analysis of the problem. In many situations, the most difficult inventive problems are featured by a physical contradiction: the same component of a technical system has to have mutually exclusive values of the same physical parameter or should be in different physical states. In contrast to procedures for formulating the engineering contradictions and substance-field modelling, the physical contradiction often may not be directly extracted from the initial problem description. To find what physical contradictions are present in the system, a more sophisticated problem analysis is needed. To translate the initial problem statement into the physical contradiction, an integrated technique called *Algorithm of Inventive Problem Solving (ARIZ)* was developed. ARIZ focuses on modeling a problem as a physical contradiction and its elimination.

ARIZ systematizes the use of various types of TRIZ modelling techniques. Following the ARIZ steps, a designer is able to gradually refine its problem model starting with an engineering contradiction and translating it into a physical contradiction.

ARIZ-85B (Altshuller [1986]) consists of several parts aimed at supporting all phases of the problem solving process with TRIZ: from the initial problem statement to the qualitative evaluation of the obtained solution. In addition, ARIZ includes a large collection of previous design cases which are various industrial solutions accumulated during the long-term development and use of TRIZ. The collection of design cases has illustrative aid and is used to

help with a better understanding of TRIZ techniques. Many of the cases store detailed information on how solutions were obtained.

The repository of design cases can be viewed as a storage of inventive experience with tackling difficult situations occurring in different steps of problem solving. This information is used to guide the process of solving a new problem.

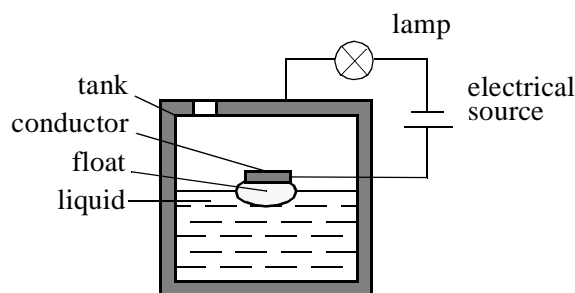
In the following subsections we will describe the ARIZ structure and a case study illustrating problem solving by ARIZ. To make understanding of the basic ARIZ concepts more clear, we will structure a problem solving process by ARIZ into three steps: i) problem formulation; ii) problem modeling and iii) problem solving.

### 3.5.1 Problem formulation

The first step of ARIZ is the analysis of an inventive situation and decomposition of the initial situation into manageable subproblems. A selection of a problem to solve from the subproblems is made according to the following principle: a subproblem that involves minimal modifications to obtain a solution should be selected.

#### *Example 3.20.*

Suppose the following situation: to indicate the level of a combustible liquid in a tank, a float detector is used. When the level of the liquid reaches the top of the tank, a conductor placed on the float and the inner surface of the tank's top close a circuit and a signal is generated (Figure 3.7). However, when there is a small gap between the float's conductor and the tank's inner surface, a spark may jump across them, that can lead to inflammation of the liquid. In this case, an inventive situation takes place. It can be decomposed into a set of subproblems: to prevent the spark jumping, to isolate the gap from the liquid, to replace an electrical system with some other system, etc.



*Figure 3.12: A problem of spark jumping*

As said above, ARIZ recommends to solve a mini-problem: a solution should be found with minimal modifications of the existing system. The mini-problem is formulated in the following way: “An engineering system for providing float detection includes: the float, the combustible liquid, the tank, the conductor, the conductors, an electrical source, electrical current and a lamp. It is necessary to prevent the spark jumping with minimal system modifications.”

### 3.5.2 Problem modeling

After the mini-problem has been formulated, it is represented in terms of an engineering contradiction. In ARIZ, the formulation of an engineering contradiction slightly differs from the formulation used in the inventive principles (see Section 3.5). The ARIZ formulation of the engineering contradiction consists of two parts. Both parts indicate opposite situations of the contradiction and relevant positive and negative effects produced in both situations.

#### *Example 3.20 (continued).*

The engineering contradiction corresponding to the mini-problem formulated in the previous section:

*Contradiction A:* An electrical current must be within the system to provide indication of the liquid level, but the liquid can explode as a result of a spark jumping;

*Contradiction B:* Without a current, the spark does not appear, but the level detection is not provided.

After this step has been completed, ARIZ suggests the construction of a substance-field model of the problem and to attempt to find a solution based on Inventive Standards. However, even if a solution is found, one needs to continue the work with ARIZ because there is no guarantee that the best possible solution has been found.

The next step is an analysis of the available substance-field and other resources. Cheap available resources should be used as much as possible to solve the problem with the lowest possible cost. At this step, the so-called *Ideal Result* is formulated. The Ideal Result displays the ultimate goal of problem solving: the desired result should be obtained without introducing new components into the existing system. This also means that new components that need to be introduced to obtain a solution should utilize or be derived from available resources.

The next step is define an *operation zone* and *operation time*. Operation zone defines a part of space where the negative effect takes place. Operational time specifies at what time interval(s) the negative effect occurs.

After a comprehensive analysis of the problem has been completed, the engineering contradiction is transformed into a physical contradiction. There are two types of physical contradictions: a *macrolevel contradiction* and a *microlevel contradiction*. The macrolevel contradiction is represented in terms of exclusive demands on the physical state of some

component in the system. The microlevel contradiction specifies where and when the particles of a substance should provide a required useful effect and must not cause the negative effect.

**Example 3.20 (continued).**

An operation zone includes the upper surface of the float, an internal surface of the tank where the contact is closed, and a distance between them when the spark jumps. An operation time includes the time interval for the spark to jump.

*The macrolevel contradiction:* an electrical current should be in the system to provide detection and should not be in the system to prevent the spark from jumping.

*The microlevel contradiction:* electrons must flow when both conductors are closed and should not flow when they are opened regardless of the distance between the contacts.

### 3.5.3 Problem solving

The problem solving phase consists of two parts:

1. Decide on *what* modification of the design prototype has to be done to eliminate the physical contradiction.
2. Decide on *how* this modification can be achieved by using a natural phenomenon that will make the solution physically realisable.

These parts are probably the most difficult from the point of view of the designer. The macro- and microlevel contradictions are the most precise problem formulations which result from accurately performing previous ARIZ steps. At this phase, principles for physical contradiction elimination are used. They provide guidelines to specify a new function that has to be realised to obtain the solution and what physical effect can provide this function.

**Example 3.20 (continued).**

One of the principles for the elimination of physical contradictions states that the contradiction formulated at the previous step can be resolved by the “division of contradictory properties in time”. Thus, we obtain a new formulation of the problem: “the electrical current has only to be generated when the contacts are closed, and must not be generated when the contacts are opened.”

To solve the problem in such a formulation, after using the pointer to physical effects it was suggested to use the Seebeck effect: appearance of electromotive force in contacting heterogeneous metals with different values of thermopowers. One of the possible solutions is based on this effect: the tank and the conductor placed on the float are made of different metals (Figure 3.8). The metal which the tank is made of has the value of thermopower  $S_1$  and the metal which the float conductor is made of has the value of thermopower  $S_2$  respectively. When contacting, they form a cold junction of a

thermocouple, whereas another junction is formed by a new conductor made from the same material as the float's conductor and placed outside the tank. This new conductor is supplied with a heating source. Since a current can only be generated in a closed circuit, in this case the physical contradiction is eliminated: the current only arises when contacts are closed (SU Author Certificate 904 532). □

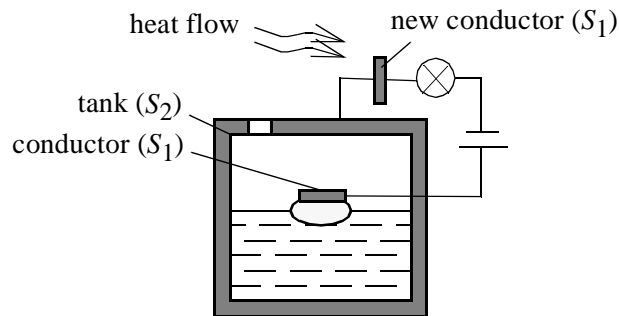


Figure 3.13: Solution to the problem of preventing a spark jumping

If the solution can not be found after the final phase has been accomplished, ARIZ recommends to go back to the phase of problem formulation and conduct the problem analysis more carefully or to select another subproblem. ARIZ also includes special rules to help with managing this situation. The overall process of problem solving with ARIZ is depicted in Figure 3.9.

### 3.6. Summary

We have presented a brief introduction to and overview of major components of the Theory of Inventive Problem Solving. Our goal was to demonstrate that a systematic and knowledge-based approach to solving inventive problems is possible. TRIZ problem-solving techniques and collections of scientific knowledge can be regarded as knowledge sources for building an AI-based system for conceptual design.

On the other hand, TRIZ concepts and techniques are still far from being formal enough to build a computational model of design upon. To understand how TRIZ knowledge can be structured, modeled and represented in an AI system, we have to formulate and solve a number of problems. In the next chapter, we perform a critical overview of TRIZ from the knowledge-based point of view and formulate these problems.

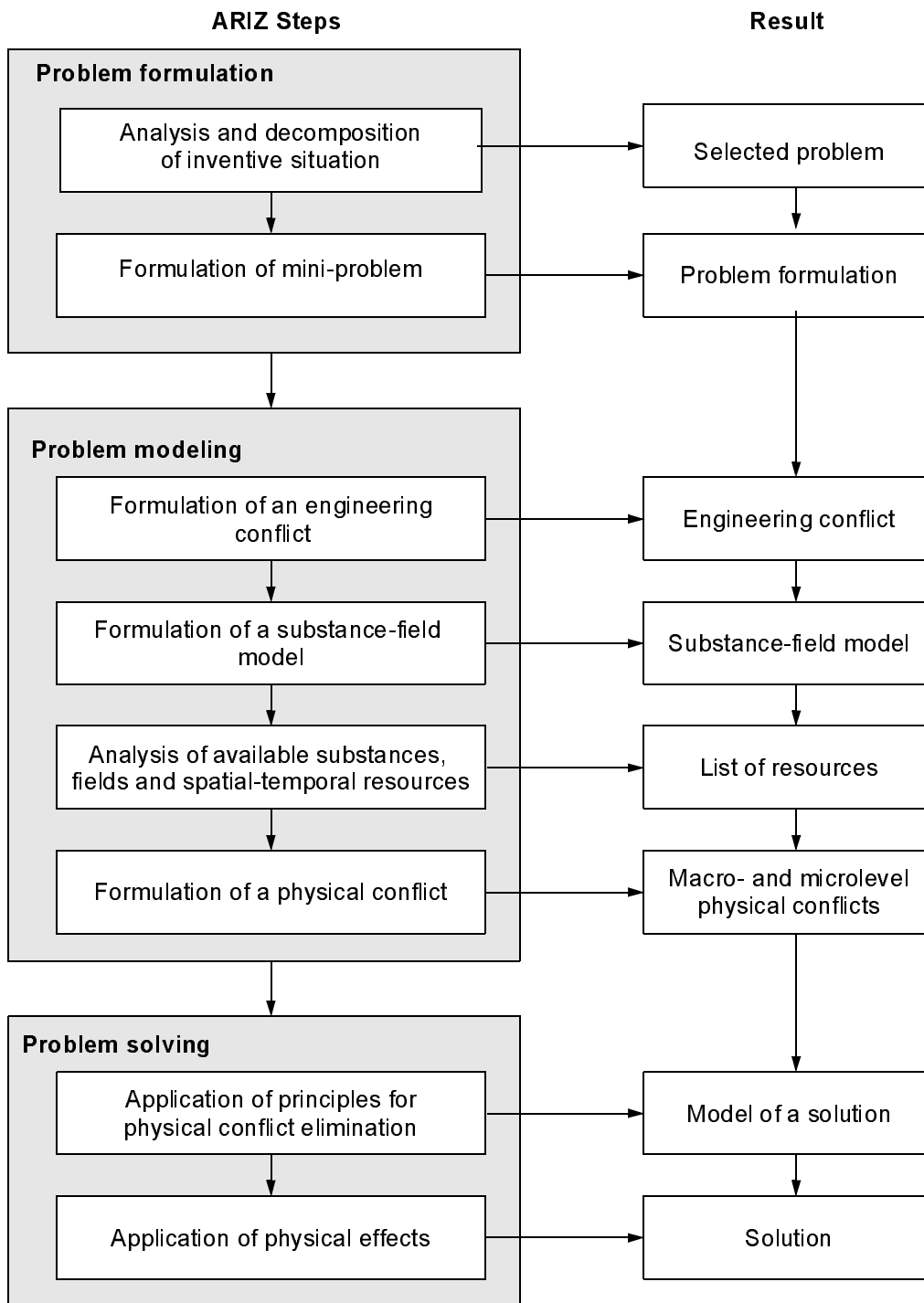


Figure 3.14: Algorithm of Inventive Problem Solving

## Chapter 4. TRIZ Critique

### 4.1. Introduction

In the previous chapter, we presented an introduction into the Theory of Inventive Problem Solving (TRIZ). A goal of this chapter is to study what problems will be faced during the development of computer aid for TRIZ-based innovative design. To do this, we perform an analysis of TRIZ problem solving techniques from the knowledge-based point of view.

First, we compare TRIZ and the German School of Systematic Design and summarize the contribution of TRIZ into design science. Second, we discuss TRIZ problem solving techniques from the position of developing a knowledge-based system (KBS). We also show why a computational model of innovative design can not build upon existing TRIZ techniques. Third, we discuss advantages and disadvantages of a formal theory of innovative design which is closely related to our study and was developed in Russia.

We also show the difference between available TRIZ-based software packages and CAD/CAM systems.

At the end of the chapter, we summarize general problems which have to be solved in order to organize knowledge-based support for innovative engineering design.

### 4.2. TRIZ contribution to design science

#### 4.2.1 TRIZ and German School of Systematic Design

Since TRIZ belongs to the category of systematic methods, it would be useful to compare TRIZ with some other systematic design method. Malmqwist et al. [1996] performed a com-



parative study of classical TRIZ<sup>1</sup> and Systematic Approach of Pahl and Beitz (SAPB) which is claimed to be the most internationally accepted systematic design methodology. The results of this study are demonstrated in Table 4.1.

**TABLE 4.1.** Comparison between classical TRIZ and SABP (adapted from Malmqwist et al. [1996])

Aspect	Classical TRIZ	SABP
Scope	Inventive phase of design, Difficult problems, Component design	Entire Design Process, Simple and difficult problems, System Design
Task clarification	Trends of Technology Evolution	General procedure
Problem formulation	Identified contradiction	Abstraction of an essential problem
Systematic methods	Matrix of inventive principles Functions coupled to physical effects Principles for Substance-Field transformation	Functions coupled to physical effects and machine components
Creative methods	Not included	Brainstorming, Synectics, 6-3-5 method, Delphi method.
Function vocabulary	30 standard functions	5 generally valid functions
Solution space	Restricted to best known solutions	All possible solutions considered
Product models	Substance-field model	Design Specifications, Function structure, Concept, Component structure
Principles	Principles for contradiction elimination Principles for substance-field modification	Rules, principles and guidelines for all phases of the design process
Knowledge base	Physical, chemical and geometrical effects Inventive Principles Trends of Technology Evolution Selected Patents	Physical effects Design Catalogues Engineering knowledge
Evaluation	Not available	Use-value analysis VDI 2225
Learning time	Long	Short
Computer support	Commercial	Research prototypes

1. A number of TRIZ versions are known which slightly differ from the original TRIZ proposed by Altshuller. For instance, Altshuller's TRIZ couples physical effects with 30 general functions while the Invention Machine software package specifies 264 standard functions. Further in the text, we will label Altshuller's version of TRIZ as "classical TRIZ".

As seen from the table, TRIZ primarily focuses on the conceptual phase of design while SAPB aims at supporting every step of the overall design process. A support for conceptual design by SAPB is organized through: i) the use of cognitive methods for activating thinking process and ii) browsing through a very general description of physical effects. Apart from coupling physical principles with generally valid functions, SAPB does not contain systematic means to solve particular inventive problems. In contrast to TRIZ, SAPB is unable to deal with ill-defined inventive problems.

However, SAPB contains a procedure for evaluating new design solutions while TRIZ does not. Design catalogues are also of a great help in finding solutions when the functionality of a solution has been clarified.

As a conclusion, TRIZ and SAPB may be regarded as complementary methods rather than alternative. TRIZ can be used for generating novel design concepts and clarifying functionality of the concepts while SAPB can be used for mapping specific functions onto solutions available in Design Catalogues.

#### **4.2.2 TRIZ contribution to science of design**

Since TRIZ is mostly comprised of a number of heuristics, it is difficult to evaluate from the point of view of exact sciences. Unlike fundamental sciences, TRIZ is not based on the axiomatic approach and does not include formal means for problem solving and verification of results. Instead, its techniques resulted from a comprehensive study of previous engineering experience that does not guarantee that the techniques will be applicable to *every* situation that may occur when designing new products. No proof of absolute applicability is possible due to the heuristic nature of TRIZ.

On the other hand, TRIZ discovered a number of principles and introduced new concepts which, although have not been formalized yet, have proven their applicability for solving practical engineering problems and considerably accelerating the process of new product development. This fact should not be neglected when studying TRIZ. Many years of experience with using TRIZ indicated that the discovered patterns and principles can be successfully applied to solve virtually any inventive problem (Altshuller [1988]). For this reason, TRIZ rapidly became a part of the engineering curriculum worldwide as a general methodology for conceptual design of new products (Ouellette [1996]) and developing new technologies. An example of using TRIZ for developing a new technology for direct manufacturing of metal parts is presented by Apelskog Killander [1996].

In summary, major contributions of TRIZ to engineering are:

- TRIZ discovered a systematic nature of technology evolution and described a number of domain-independent evolution trends.
- TRIZ introduced a new classification of design solutions.

- TRIZ proposes to regard a contradiction as a cause of inventive problems and states that inventions results from eliminating contradictions.
- A set of basic principles for contradiction elimination was proposed.
- Access to the basic principles was organized in a systematic way.
- TRIZ proposed to model design products in terms of substance-field interactions and apply generic patterns to transform the physical structure of products.
- TRIZ proposed a novel way to couple physical principles and technical functions.

We believe that TRIZ concepts such as contradictions and principles for substance-field modelling might be used for developing a fundamental science of innovative engineering design.

### **4.3. Critique of TRIZ**

In this section, we perform a brief critical overview of TRIZ problem solving techniques presented in Chapter 3 from the knowledge-based point of view.

#### **4.3.1 Trends of Technology Evolution**

Although the Theory of Technical System Evolution (TTSE) is claimed to be of a large importance for evaluating existing design solutions and predicting future solutions, it is unclear how to use it. Many TRIZ schools refer to significant difficulties in teaching TTSE to students because of the generic nature of evolution patterns and the lack of a product model TTSE would be able to deal with. For these reasons, we do not see a possibility of representing TTSE in a formal and structured way and will not use it in our further study.

#### **4.3.2 Principles for contradiction elimination**

As shown in Chapter 3, one of the most important TRIZ achievements which brings a better understanding of innovative engineering design is that TRIZ regards the process of technology evolution as the successive elimination of various types of contradictions. A contradiction arises when the designer's wish to improve the performance of an existing system by redesigning the system results in a deterioration of another system's parameter.

To eliminate the contradiction and prevent other system parameters from deterioration, TRIZ recommends that a problem be formulated in terms of a pair of conflicting parameters and eliminate the conflict using inventive principles. For instance, the contradiction between the weight of a moving object and its speed is eliminated by using the "Counterweight Principle" which states that "to compensate for the weight of the object it should be merged with other objects that have lifting forces". Examples of other inventive principles are "to increase a

degree of object's fragmentation", "to place one object inside another", etc. (see Appendix A). As seen, TRIZ principles are formulated in a very general manner and do not indicate what should be done with respect to a particular problem.

Summarizing, there are four shortcomings of the existing approach to solving inventive problems by formulating and eliminating contradictions:

1. No formal definition of a contradiction is available in TRIZ.
2. TRIZ does not provide exact recommendations on how to formulate contradictions with respect to a particular problem. As a result, a contradiction is constructed *ad-hoc* since no analysis of a prototypical design is performed.
3. To identify an inventive principle which has to be used for solving a problem represented as a specific conflict, the conflict has to be reformulated in terms of generalized engineering parameters. However, this can only be done intuitively since no translation technique is available in TRIZ.
4. Inventive principles do not propose a solution to a given problem. They only recommend a method which was used to solve a similar contradiction before.

For these reasons, the problem solving process turns out to be a process of numerous iterations until an appropriate inventive principle is identified. Furthermore, no quantitative reasoning is supported by TRIZ thus making it impossible to evaluate feasibility of solutions.

To eliminate these shortcomings, a formal definition of the contradiction concept is necessary. It should allow the designer to define contradictions in terms of specific parameters after an appropriate analysis of product models has been performed. This will make it possible to derive contradictions from product models. However, this can be only be done after agreeing upon how to model diverse design products.

### 4.3.3 Substance-Field Modelling

Inventive Standards (Section 3.7) deal with more specific problem formulations presented in terms of interactions between substances and fields. Altshuller points out that Inventive Standards are an advanced form of inventive principles since they relate conflicts and physical effects. However, this relation remains implicit since no contradiction is identified when a problem is modelled as a set of interacting physical objects.

The next disadvantage is that it is unclear which components of a system should be included in the model and how modelled components are related except interaction via physical fields. To explain what he means by field, Altshuller writes (Altshuller [1984], p. 53):

“In technology, the term field is used in a broader sense: there is space to each point of which a certain vector or scalar magnitude stands in relation. Such fields are often linked with vector or scalar bearer-substances, the temperature field, the

field of centrifugal forces, for instance. We shall use the term field in a very broad sense, and together with the legitimate physical fields regard all possible kinds of “technical” fields, - heat, mechanical, acoustic, etc. as such”.

From the point of view of exact science, this definition is ambiguous. It allows different designers to give different interpretations of the term “field” when describing the same physical process occurring in a system. TRIZ permits the same component to be regarded as a substance in one situation and as a field in another (for instance, it allows the water flow to be presented as a “field”).

Due to this ambiguity, inadequate interpretations of the same Inventive Standards are possible. Experience drawn by the author from teaching TRIZ to various categories of students shows that the students with different backgrounds tend to interpret the same Inventive Standard differently. As a result, an attempt to use the Inventive Standard to solve the same problem by two students does not guarantee that both students will come up with similar solutions.

Problems with organizing a computer-based reasoning with Inventive Standards are similar to the problems mentioned when discussing inventive principles. We need a set of modelling concepts for representing diverse types of product information in uniform way, and we need to redefine Inventive Standards in terms of such concepts. This will relate inventive principles and Inventive Standards.

#### **4.3.4 Pointers to physical effects**

The basic assumption behind organizing a design process as synthesis from physical knowledge is that any part of a design product can be decomposed into a set of causally related fundamental physical phenomena. This makes it possible to reveal what particular functions can be delivered by one or another phenomenon and group those phenomena that deliver the same function. The Pointer to Physical Effects (PPE) available in TRIZ utilizes this approach to provide a designer with a structured collection of physical phenomena coupled with technical functions (see chapter 3).

The principal claim of PPE over other approaches to using physical knowledge in innovative design is that the PPE set of functions was drawn from patent descriptions. The basic principle behind the PPE is that the more general the function is, the larger number of effects can be identified with the function. However, the high degree of generalization is not always useful: the function “to move a solid body” refers to a large number of physical phenomena since constraints are not supported by the PPE (Figure 4.1). This causes a new problem: how to restrict a selected set of phenomena to those that are applicable to the situation given? Obviously, the high degree of generalization may result in a loss of the informational contents of a specific problem.

Another problem is that due to the heuristic origin of PPE, in some cases, PPE does not map a function onto an effect which directly performs the function. In these cases, the function is

mapped onto an effect that produces conditions needed to achieve the required result in combination with other effects. For instance, the function “to move a solid body” refers to the effect of thermal expansion. It is clear, that the effect of thermal expansion can move the body only in combination with the effect of inertia. Thus, the required result can be obtained by causal chaining several effects. However, the procedure of chaining is not supported by PPE. As a consequence, the use of proposed effects might be misleading.

With respect to the development of computer tools on the basis of PPE, the major drawback of PPE is a lack of uniform representation of physical knowledge. Different effects are described with different degrees of details. No relationships between physical phenomena and their possible design implementations are available in PPE.

In summary, there are the following shortcomings of PPE:

1. A solution to a problem is only proposed as a general description of a physical effect without any indication of how an implementation based on these effects will look like.
2. There is no means to generate solutions by combining several physical effects.
3. A function is often mapped onto an effect that does not produce the needed function itself but provides conditions to activate some other effect which fulfils the function.
4. Evaluation of the applicability of proposed effects is left to the responsibility of a human designer. Specific design cases provided in PPE are too domain-specific and do not always help in establishing an analogy with a given situation.

We believe that deep physical knowledge can be modelled in the same way as design product knowledge. From this point of view, there should not be any difference between conceptual modeling of specific design products and physical phenomena.

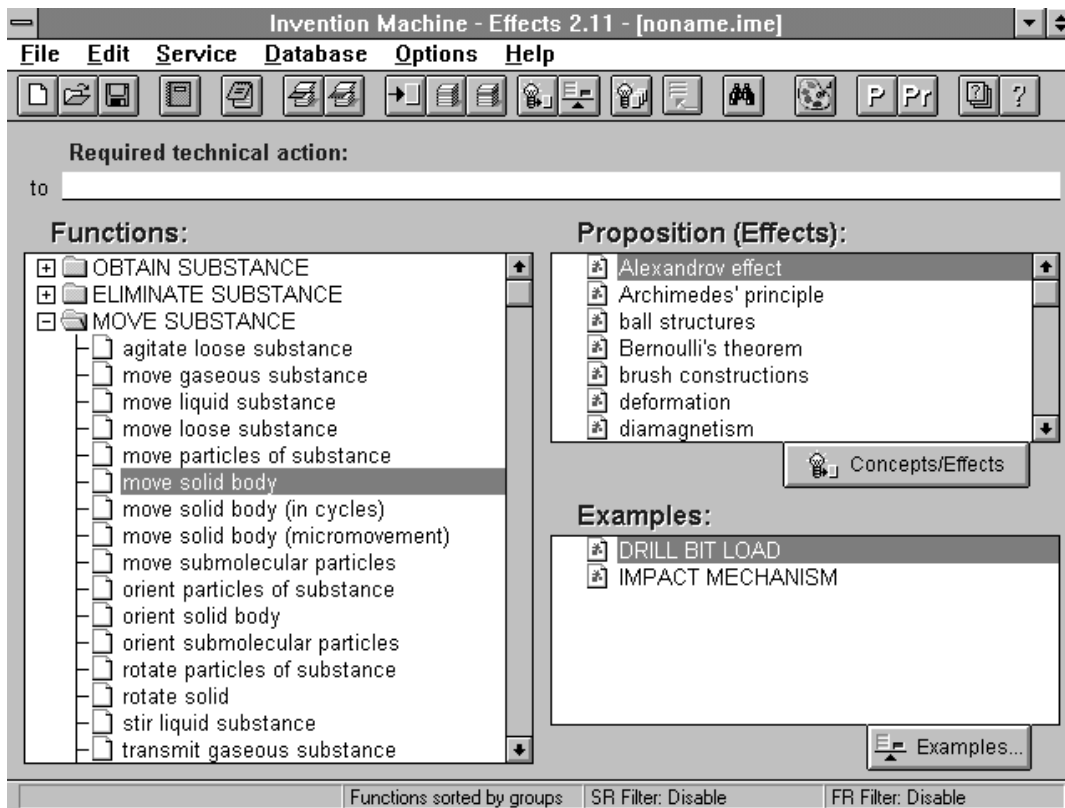


Figure 4.15: Screenshot of Invention Machine's Pointer to Physical Effects

### 4.3.5 ARIZ

Created to solve most difficult inventive problems, the Algorithm of Inventive Problem Solving (ARIZ, Section 3.9) attempts to tackle a problem by identifying *physical conflicts* in problems (Section 3.6.1). A physical conflict seems to be more accurate form of problem formulation than engineering contradiction since it exactly indicates what part of a design product causes a problem.

The use of ARIZ is, however, strictly human oriented. A major drawback of ARIZ is its inconsistency: information obtained at earlier steps of analysis might be inadequate for later steps which are aimed at concept synthesis. Besides, different ARIZ parts operate with different problem models (such as contradiction, function, ideal final result, substance-field model), and it is unclear how all these models are related. It is also unclear how to translate one model into another. Working with ARIZ requires extensive engineering background and TRIZ experience. As a result, the quality of obtained solutions relies heavily on the previous scientific and engineering knowledge of a designer.

As pointed out in Dikker et al. [1992], this happens because relations between Inventive Principles, Inventive Standards and physical principles are unclear. Each technique operates with its own problem model that makes it impossible to use the techniques consistently.

## 4.4. An attempt to formalize TRIZ

The work related to our study which is aimed at building a formal theory of design based on TRIZ is presented in Glasunov [1990]. To formalize concepts introduced by Altshuller, it studies the applicability of mathematical modelling of design products. Since INDES also aims at building a formal framework for modelling TRIZ knowledge, we decided to make a reader familiar with this approach.

### 4.4.1 Mathematical model of design products

Glasunov proposed to draw potential conflicts from a mathematical model of a design product. The mathematical model of the design product is defined as a set of so-called *Complex System Features (CSF)*. In turn, a CSF is an integrated parameter which is defined on the basis of more specific design parameters which involve the same type of variable. Examples of CSF are overall mass, system reliability, overall dimensions, energy consumption and so forth.

A value of each CSF is calculated as a sum of values of specific parameters describing different system components:

$$F = \sum_{i=1}^n P_i$$

A design product is presented as an ordered tuple of CSFs:  $TS = \langle F_1, \dots, F_n \rangle$ . In turn, each CSF is identified with an ordered tuple of particular parameters:  $F_i = \langle P^i_1, \dots, P^i_m \rangle$

It is obvious that in any physical design the same parameter might be involved in two different CSFs. Therefore, two sets may intersect so that  $F_i \cap F_j = F_{ij}$ . This results in the set of common parameters  $F_{ij} = \langle P^{ij}_1, \dots, P^{ij}_m \rangle$

If the same element  $P^{ij}$  is used as a parameter of two CSFs  $F_i$  and  $F_j$ , such CSFs are labelled *linked*. Two linked CSFs possess a general property: any change of values of variables within  $F_i$  results in change of the value of  $F_j$ .



## 4.4.2 Conflict Formulation

With respect to the desired improvement of the system, two categories of CSF are distinguished: positive ( $F^+$ ) and negative ( $F^-$ ). A CSF is positive when its value should increase (e.g. reliability). In contrast, a CFS is regarded as negative when its value should decrease (e.g., energy consumption).

### Example 4.1.

The parameter *mass* of a hammer might be defined as a variable involved into both CSF<sup>+</sup> *Force* and CSF<sup>-</sup> *Energy Consumption*. Therefore, an attempt to increase the force through increasing the mass will lead to the growth of energy consumption.

It is possible to define what linked CSFs will conflict by analysing their partial derivatives by a physical parameter  $P$ . Table 4.2 illustrates possible states.

TABLE 4.2.

Condition	Result
$\partial F_i^+ / \partial P > 0$ and $\partial F_j^+ / \partial P > 0$	no conflict
$\partial F_i^+ / \partial P > 0$ and $\partial F_j^+ / \partial P < 0$	conflict
$\partial F_i^+ / \partial P < 0$ and $\partial F_j^+ / \partial P > 0$	conflict
$\partial F_i^+ / \partial P < 0$ and $\partial F_j^+ / \partial P < 0$	no conflict
$\partial F_i^+ / \partial P > 0$ and $\partial F_j^- / \partial P > 0$	conflict
$\partial F_i^+ / \partial P > 0$ and $\partial F_j^- / \partial P < 0$	no conflict
$\partial F_i^+ / \partial P < 0$ and $\partial F_j^- / \partial P > 0$	no conflict
$\partial F_i^+ / \partial P < 0$ and $\partial F_j^- / \partial P < 0$	conflict
$\partial F_i^- / \partial P < 0$ and $\partial F_j^- / \partial P < 0$	no conflict
$\partial F_i^- / \partial P < 0$ and $\partial F_j^- / \partial P > 0$	conflict
$\partial F_i^- / \partial P > 0$ and $\partial F_j^- / \partial P < 0$	conflict
$\partial F_i^- / \partial P < 0$ and $\partial F_j^- / \partial P < 0$	no conflict

In the situation when one of the partial derivatives can have both negative and positive signs (within different time intervals, for instance), this situation is labelled *partial conflict*.

### 4.4.3 Problem solving

A potential conflict between two CSFs is expressed as the relation  $goal(P(X_1), P(X_2))$ , which can be interpreted in the following way: “a physical parameter  $P$  has to have two different values at the same time:  $X_1$  (to satisfy  $F_i$ ) and  $X_2$  (to satisfy  $F_j$ )”. This is similar to what Altshuller defines as a physical conflict (Section 3.6.1): the same object has to possess two contrary physical properties at the same moment in time or to have two different values of the same physical parameter.

A procedure for conflict elimination consists in the search for an object from a repository of predefined physical objects that can satisfy both conditions. For instance, a spiral is regarded as the object which under certain conditions satisfy two conflicting values of the parameter “length”. The spiral is long (if to look at the length of a line forming the spiral) and short (if to look at the spiral diameter) at the same time. Therefore, the spiral possesses so-called *dual properties* and is capable of resolving the conflict.

### 4.4.4 Discussion of Glasunov’s approach

A major advantage of Glasunov’s approach is that it defines a conflict in a formal way by analysing a mathematical model of a design product. However, the concept of a Complex System Feature still remains ambiguous. Furthermore, given a set of Complex System Features, no reasoning about a system’s behaviour is possible. As a consequence, a search for an object with dual physical properties is only determined by two required values of a physical variable and there is no guarantee that the use of the selected object in an existing system will not violate the overall system behaviour. A repository should store components with all possible dual physical properties that does not seem possible due to a relative nature of proposed definition of duality. It is clear, that a spring as the object with two different values of the same parameter might not be considered as a solution in every situation.

Another disadvantage is that the success of problem solving strongly depends on two factors: what CSFs are identified in every particular case and how many physical objects with dual properties are stored in the repository. Provided the repository is large, there are no means to verify the applicability of the solutions without the participation of a human designer. In this case, the use of a mathematical model of a design product seems to be useless during the synthesis phase since we can not calculate quantitative values of the solution.

As a conclusion, to build a computational model of innovative design based on the Glasunov’s approach is possible. However, its applicability will be restricted to those problems which can be solved by using predefined objects with dual properties.

## 4.5. Computer-Aided TRIZ and CAD/CAM software

Recently, a number of software packages supporting design problem solving with TRIZ have been developed. Among them are Invention Machine (Tsourikov [1992]), Edison (\*\*\*) [1992], and Innovation Workbench (Braham [1995]). Although all packages incorporate different approaches to represent TRIZ information and organize the problem solving process, they form a new category of computer-aided design tools which support a conceptual phase of engineering design.

While traditional CAD/CAM systems focus on processing and computing geometrical and material aspects of specific designs (Amirouche [1994]), TRIZ-based software packages structure access to previous inventive experience stored in the form of inventive principles and indexed physical principles. According to a given problem formulation, TRIZ-based packages propose information on what the generic behaviour of a design solution should be rather than what form and geometry the solution should have.

In summary, TRIZ-based packages organize mapping between the function and behaviour of a concept which is still to be found, whereas CAD/CAM systems map functional and geometrical specifications directly onto already known design solutions stored in the database.

In addition, CAD/CAM systems propose specific, “ready-to-manufacture” descriptions of solutions that makes such systems relatively easy to learn and use. TRIZ-based packages are well-organized interactive systems which only help with finding general recommendations on how to solve problems, or at best, indicating what physical principles to use. A designer should be able to interpret this information and translate it into a feasible solution. No sufficient computer aid has been available so far to support this step. This causes certain difficulties in the use of the software by many designers since the gap between general recommendation and specific solution can be very large. It is our belief that to be accepted by a wide audience, TRIZ-based software has to bridge this gap and be able to generate solutions in terms of specific problems instead of displaying general recommendations.

## 4.6. Discussion

Until now, it was unclear how to structure design and scientific knowledge in order to develop a knowledge-based system which will support reasoning about innovative and inventive designs. As follows from Altshuller’s studies of inventive activities, to make inventions, the outstanding inventor uses his/her own experience which is usually based on 5-7 empirically found inventive principles over the long time of practice. The inventor’s experience thus comprises knowledge of different engineering and scientific areas that makes it possible to establish high-order analogies when solving new problems. However, too much of this knowledge remains implicit so no systematic use of it is possible.

To automate conceptual engineering design by modeling all available design and scientific knowledge and developing a comprehensive knowledge base capable of storing all this knowledge does not seem to be possible at the moment. It would require an incomprehensible amount of work to be done. In addition, there is no way available to establish links and relations between design solutions drawn from different engineering disciplines at the level of specific details. Although some general approaches to building very large engineering knowledge bases are mentioned in the literature (Tyugu [1995], Ishi et al. [1995]), their feasibility has not been proven yet. Apart from that, even if the creation of such a knowledge base becomes possible, no strategy is available to reason with this knowledge to obtain inventive solutions.

TRIZ tackles this problem by revealing high-order similarities between inventive solutions in different domains and making generic patterns behind these similarities explicit. However, the high degree of generalization makes a process of finding specific solutions difficult if not impossible.

Known TRIZ-based software packages do not incorporate AI elements due to a lack of computational models of innovative design. It is the authors' opinion that it would be premature to directly incorporate TRIZ knowledge available in TRIZ literature into a knowledge-based system unless we have clear understanding of how to support TRIZ with formal methods and represent TRIZ knowledge.

To summarize the discussion in the chapter, we can distinguish the following major shortcomings of TRIZ with respect to its possible computer implementation in the form of a knowledge-based system:

- TRIZ techniques are meant for personal use. They propose no formal methods for problem solving. Instead, empirical rules for restricting the search space of solutions are introduced.
- The wealth of knowledge available in TRIZ is necessary for solving a large variety of inventive problems but access to the needed specific knowledge might be troublesome.
- TRIZ does not operate with formal scientific categories thus making it impossible to apply quantitative constraints at the phase of problem formulation. However, it is often the case.
- TRIZ definitions of physical concepts such as substances and fields are ambiguous and can not be adequately interpreted.
- Using a recommendation proposed by TRIZ for solving a specific problem requires an extensive knowledge of different engineering domains and is not currently supported by TRIZ. As a consequence, the user of TRIZ is supposed to possess a high degree of expertise in engineering design.

- TRIZ does not include techniques for evaluation and verification of the obtained design concepts.
- It takes a very long time to master the necessary skills for working with TRIZ even with the available computer aid.

We believe that TRIZ knowledge should be restructured and redefined. As the basis for restructuring we have chosen a product-centered approach to design, i.e. strategic knowledge should be defined in terms of product knowledge.

To develop a TRIZ-based computational model of innovative design that could be incorporated into a knowledge-based system we defined three major milestones of INDES project:

1. To develop a formal framework for modeling and representing product knowledge for conceptual design in a uniform way.
2. To redefine major TRIZ concepts in terms of a product modelling framework.
3. To establish relations between TRIZ problem solving techniques and model strategic knowledge for generating new design concepts.

Throughout the remainder of this work, we will present the results of our study.

## *Chapter 5. A Knowledge-Centered Model of Systematic Innovative Design*

*“The important thing for you is not how much you know, but the  
quality of what you know.”*

*Desiderius Erasmus (1469-1536)*

### **5.1. Introduction**

The project INDES (stands for “Invention Designer”) was initiated in 1993 within the framework of a more general STEVIN project at the Knowledge-Based Group of the University of Twente. The STEVIN project has been aimed at studying what design methodologies are currently available and have been proven successful, and what AI modelling approaches and techniques are best suited to model various phases of engineering design. The results of the study are going to be utilized for developing knowledge-based tools supporting various phases of engineering design process. As a part of STEVIN, INDES focuses on modelling innovative phases of engineering design (Mars et al. [1993]).

Among the prerequisites which had the impact on making decision to launch INDES were the availability of the Theory of Inventive Problem Solving (TRIZ) discussed in the previous chapters and YMIR, an earlier development within STEVIN. YMIR is a domain-independent

ontology which introduced a number of knowledge concepts for formal modelling and representing generic design knowledge (Alberts [1993a, 1993b]).

Since both TRIZ and YMIR utilize a domain-independent approach to design and aim at similar goals - structuring knowledge for engineering design, we decided to study if it would be possible to model TRIZ knowledge in terms of YMIR concepts.

This chapter presents results obtained after this study. First, we describe the goals of INDES and a problem scope. Second, we introduce a knowledge-centered model of innovative design based on TRIZ. This model will be used as a generic model from which a number of more specialized design models can be drawn from.

The next part of the chapter discusses different categories of knowledge for innovative engineering design. We show how an ontological approach to knowledge conceptualization can be used to identify TRIZ object and strategic knowledge. In the next chapters, these knowledge concepts will be used to build a computational model of design.

At the end of the chapter, we present a knowledge-intensive model of innovative engineering design.

## 5.2. INDES Project overview

In this section, we present INDES goals, problem area and scope.

### 5.2.1 INDES Goals

The primary goal of INDES is to develop a formal framework which could be used for further development of knowledge-based systems supporting innovative engineering design. In addition, a prototype of a knowledge-based system (KBS) capable of generating new design concepts was planned to build.

Following a general methodology of KBS development presented in Hayes-Roth et al. [1983], INDES was divided into the following phases:

1. Analysis of available theoretical methods for innovative design. Identification of the problem area and scope, requirements and knowledge sources.
2. Building a conceptual model of innovative design on the basis of the problem area and scope defined.
3. Identification of key knowledge concepts and relations, task strategies and constraints.
4. Developing a technique for formal modelling of knowledge and mapping key concepts and relations onto formal representations.

5. Testing the feasibility of the framework through an industrial case study.
6. Developing a research prototype of a knowledge-based system for innovative design.

It was decided to restrict INDES to semi-automated reasoning about innovative solutions. At the moment it seems to be impossible to develop a KBS supporting full automated reasoning due to a lack of AI methods supporting reasoning about commonsense knowledge. However as seen from Altshuller's theory, innovative design relies heavily on this type of knowledge. Instead, INDES aims at bringing a more structured approach into developing knowledge-intensive systems in the area of innovative design support systems.

### 5.2.2 Problem area and scope

In one of the early works on developing KBS for design, Brown & Chandrasekaran [1985] divided the innovative design solutions into two categories: *inventions* (such as creation of the wing) and *innovations* (such as creation of a wing with changeable geometry). Although similar to Altshuller's definitions for inventions and innovations, they do not address the prescribed strategy for design. They just state what types of design solutions can be obtained with respect to already existing space of design solutions.

Using these definitions within the context of Altshuller's explanation of a systematic nature behind the process of generating innovative solutions we defined the scope of INDES as *study and development of a knowledge-based support for conceptual engineering design*.

Therefore, INDES should provide a general domain-independent methodology for organizing a process of conceptual design. To do this, we formulated a number of more specific questions which defined milestones of INDES:

- What design method(s) and knowledge sources are available for innovative design?
- What general model of a design process can be derived from the design method?
- What knowledge sources are needed to support the design model?
- How knowledge about design products can be modelled to enable reasoning with it?
- What techniques for knowledge formalisation and representation are available and best suited to formalise knowledge for conceptual design?
- What computational models of design can be built on the basis of defined knowledge sources and reasoning strategies?

Further in the work, we make an attempt to answer these questions.



### 5.2.3 INDES definition of conceptual design

To define a problem area more clearly we introduce definitions of *conceptual design* and *design concept* in the context of INDES.

**Definition 5.2.**

**Conceptual design** is a process of mapping of new societal demands and requirements onto a description of a physically realisable artefact which is presented in terms of a *design concept*.

**Definition 5.3.**

**Design Concept** is a description of a physical structure of a designed artefact at the conceptual level which lacks specific details, but consistent and informative enough to provide correct instantiation of the concept into the description of a feasible design product.

### 5.2.4 INDES limitations

TRIZ is limited to finding innovative and inventive solutions to problems in areas of technology which deal with material and energy transforming systems. As a consequence, an INDES-based design model will not be suitable for solving problems in areas of information or, for instance, business systems.

As mentioned in the previous chapter, as a problem-solving tool TRIZ does not include a technique for comparison and evaluation of obtained solutions. We have therefore decided to limit ourselves to building a model of a design process which does not include means for solution evaluation. On the other hand, given the systematic nature of the search for most appropriate knowledge using TRIZ, the number of solution concepts resulting from the search is relatively low. This makes it easy to perform an analysis of alternatives. We believe that a procedure for concept evaluation should be independent of the problem-solving methodology and might thus be easily adapted to any problem-solving methodology without changing the methodology itself. As an example, we can refer to Nordlund [1994] and Killander et al. [1995] to name a few works studying how TRIZ can be extended with an evaluation technique on the basis of Design Axioms (Suh [1990]).

## 5.3. Knowledge-centered model of innovative design

The key factor for developing effective design models that could be further incorporated as KBS is the availability of theoretical design methodology. As mentioned before, studies with

AI in Design are divided between two groups of models: process-oriented and product-oriented. The first group concentrates on studying cognitive qualities of a designer or a design team and modeling them in the form of directives. Since such studies result in inevitably fuzzy and broad theoretical frameworks (see, for instance, Wilson et al. [1995]), specific models of design may not be derived from them. Knowledge-based systems implementing a process-oriented model of design can therefore be regarded as *user-centered*.

The advantage of the product-oriented approach is that it results in a designer-independent model of design which focuses on product qualities rather than on process organization. A philosophy behind this product-oriented approach thus coincides with Altshuller's approach to innovative design which considers psychological inertia to be a major obstacle to finding innovative solutions. Numerous TRIZ applications have proven Altshuller's thesis that properly structured knowledge wealth is a necessary attribute of successful inventive design. Using TRIZ *any* engineer no matter what creative abilities he/she possesses is able to solve inventive problems in a relatively short time provided access to the TRIZ knowledge repository. As a conclusion, we believe that KBS for early design phases should be *knowledge-centered* rather than user-centered.

In the previous chapter, we presented an overview of the set of recommended practises to come up with innovative solutions to the problem given based on the Theory of Inventive Problem Solving. We also described the Algorithm of Inventive Problem Solving (ARIZ) which integrates different TRIZ problem-solving techniques. Despite its undeniable practical usefulness in providing a designer with methodological support, ARIZ can hardly be regarded as an appropriate model of design process ready to be transferred onto a computer. A major drawback of ARIZ is that it mixes various problem representations and operates with different models of problems and technical systems. Finally, it is unclear how to translate between one or another representation performing ARIZ steps.

To build a consistent model of TRIZ-based design that could be implemented in KBS, we propose to separate between two categories of innovative design tasks different in the types of initial design specifications and knowledge sources able to deal with each problem representation:

1. **Innovative Redesign.** The distinguishing feature of this type of design is that a problem is formulated on the basis of the desired improvement to an existing design product. It might be either as a contradiction between two parameters or a substance-field model which contains an undesired interaction. In the first case, principles for contradiction elimination should be used to produce a recommendation on how to solve the problem. Alternatively, the Inventive Standards are to be used to modify the substance-field model.

2. **Innovative design from physical principles.** A problem is represented as a function a new design product has to deliver. No existing (or known) design product or component is able to fulfil the function properly. In this case, the Pointers to natural phenomena are then to be used to identify the function with one of the predefined functions and map the function onto a generic design concept.

Figure 5.1 depicts a general model of TRIZ-based innovative design. Partly, this model is supported by the INVENTION MACHINE system (Tsourikov [1992]). However, INVENTION MACHINE does not perform reasoning in terms of user-defined problems. To enable such reasoning, TRIZ-based KBS must be able to deal with *specific task knowledge*, select an appropriate knowledge source and propose a solution in terms of the problems given.

Since we defined knowledge as a key point in our approach, in the next sections, types of knowledge we use to model and represent knowledge of TRIZ will be discussed.

## 5.4. Knowledge categories for innovative design

In this section, we discuss a modern AI approach to distinguish various categories of knowledge. We show how knowledge for innovative engineering design can be divided into four categories: deep, shallow, object and strategic. In turn, we distinguish two subcategories for object knowledge: primitive and complex.

### 5.4.1 Deep and shallow knowledge

The early success of knowledge-based systems refers to the mid-seventies when expert systems based on knowledge expressed in the form of rules were designed and applied to solve practical problems in various domains. The majority of such systems served diagnostic purposes. Their distinguishing feature was the use of shallow knowledge, that is, knowledge acquired by an expert without understanding the underlying reasons. As a consequence, a knowledge-based system built on shallow knowledge was unable to give an explanation of its reasoning nor validate produced results.

In the 1980s, a major paradigm shift was observed in AI when Lenat and Feigenbaum [1987] introduced the **knowledge principle**<sup>1</sup> which dramatically changed the view on knowledge-based systems development. According to the principle, an AI program performs well if it knows a great deal about the task and domain where it operates.

According to the knowledge principle, the broader the domain is, the more complex and extensive a knowledge base for this domain has to be. On the other hand, if the knowledge base only incorporates shallow knowledge, it will inherit all the disadvantages of shallow knowledge no matter what size the knowledge base is. The only known way to tackle this sit-

---

1. Also known as “knowledge is power”

uation is to introduce *deep* knowledge into the AI program. In contrast to shallow knowledge, deep knowledge is a fundamental building block of understanding. A good example of deep knowledge is a fundamental knowledge of the laws of nature. The role of deep knowledge is twofold: first, to provide casual links between shallow knowledge to enable explanation and second, to decrease the degree of inconsistency in the knowledge base.

Basically, in any domain both types of knowledge can be distinguished which are relative to each other. In TRIZ, both shallow and deep knowledge can be distinguished too. For instance, the principles for engineering contradiction elimination can be regarded as shallow knowledge with regard to the trends of technical systems evolution. Suppose the rule for contradiction elimination is defined as:

**IF the length of a movable object is to be increased  
AND the weight of the object is not allowed to increase  
THEN the object must be fragmented.**

Then, despite this rule of thumb being applicable to most inventive problems containing the same contradiction, it still does not explain why the object must be fragmented. In some cases, it is unclear how to produce such fragmentation despite the fact that the word “fragmentation” might be interpreted in many ways. Moreover, the fragmentation is not always possible, especially when the system has reached its final phase of evolution and tends to be replaced with another system.

At the same time, the trend of transition to the microlevel explains why fragmentation is proposed to resolve this particular type of contradiction. Therefore, the trend of transition to the microlevel can be regarded as a part of the deep knowledge level in TRIZ. As a consequence, all TRIZ trends of evolution are deep knowledge with respect to the specific rules of contradiction elimination.

Most KBS, especially those developed for commercial purposes, do not require a deep level of knowledge analysis. As a consequence, they are built on codifying relatively shallow knowledge. However, the more knowledge-intensive a system is to be built the greater the depth of analysis is necessary to provide an effective reasoning about domain and task knowledge. As seen from the previous chapter, TRIZ incorporates knowledge of various degree of deepness. The use of more deep knowledge results in solutions with a higher degree of feasibility. Besides, deep knowledge must be presented in a KBS to provide a check against violation of basic physical laws when inferring a new solution.

We must note that in INDES, no design model based on explicit use of the trends of evolution is proposed. Instead, we use Altshuller’s principles for inventive design which implicitly incorporate the evolution trends with respect to specific tasks.

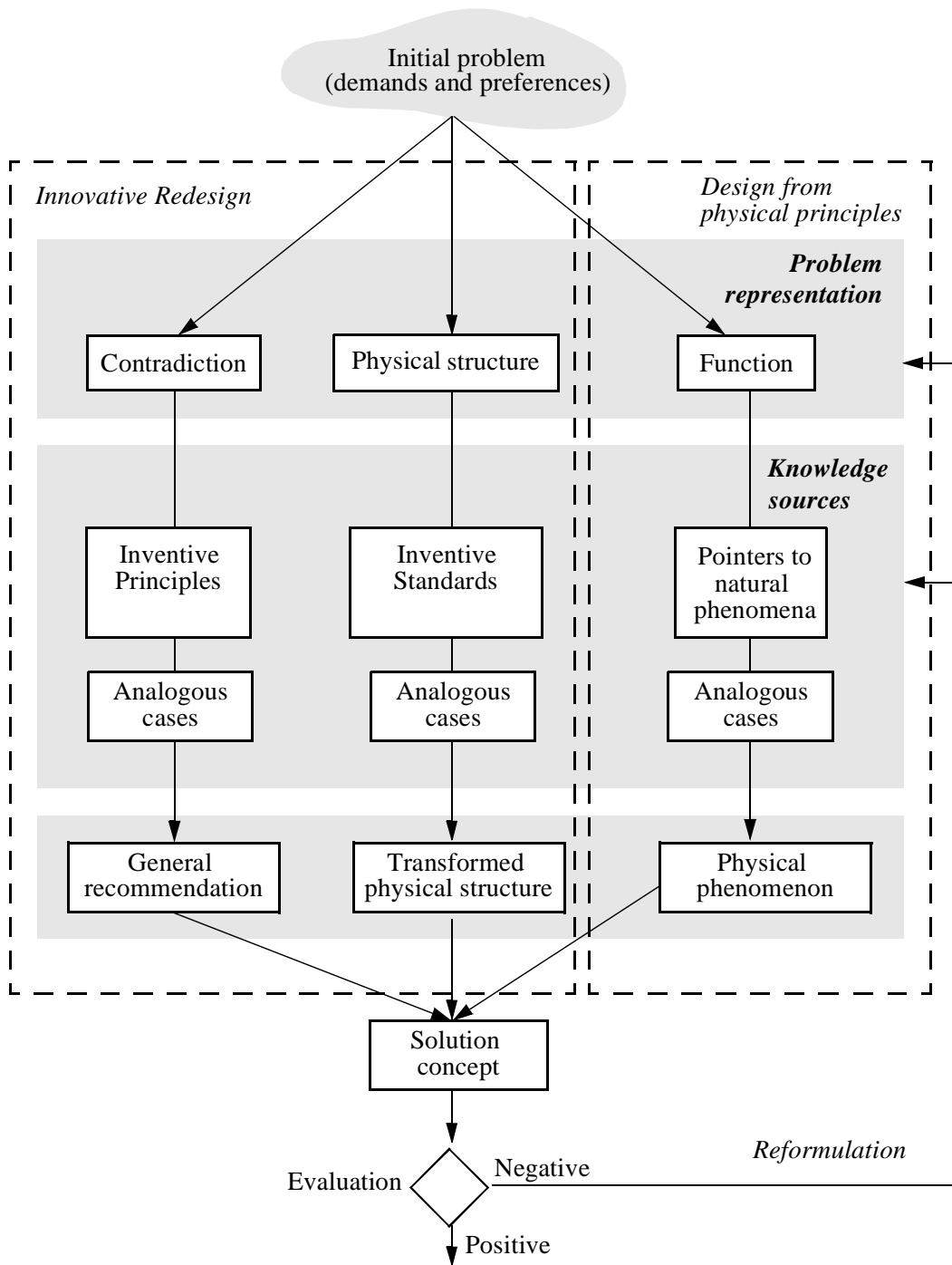


Figure 5.1: Knowledge-centered model of TRIZ-based innovative design

### 5.4.2 Object and strategic knowledge

Apart from the division into deep and shallow knowledge, a more specific distinction can be made between knowledge presenting a solution to a certain problem and knowledge describing a method for obtaining solutions. Since a design solution is a description of an artefact resulting from the design process conducted in accord with a certain method, this type of knowledge is called *object knowledge*. It is a knowledge structure which presents information on material and immaterial components the artefact consists of as well as relations between these components. Depending on the type of problem and the required level of details, design object knowledge might include constraints-related information and a description of operating conditions.

Another type of knowledge expresses a strategy of how design objects are reasoned about. It comprises the principles codified in the form of both generic and production rules which are used to infer a target solution from the task-specific and predefined object knowledge given. This type of knowledge qualifies as *strategic knowledge*.

There are no unique criteria defining what object or strategic knowledge should consist of. It depends on specific features of a domain where objects and problems are identified. In the case of engineering design, there might be several levels of design object knowledge which are dependent on what level of abstraction the object's components and relations between the components are presented.

A division between strategic and object knowledge makes it possible to separate a knowledge base storing design objects from the problem-solving structure. Besides, it allows the logical design to be transformed into a physical design with fewer changes and modifications.

## 5.5. Object knowledge for innovative design

In this section, we present several categories of object knowledge concepts for engineering design. Object knowledge describes a design object as a whole or any of its parts independently of how the object or the part was designed. The goal of this section is to discuss why we decided to distinguish several categories of object knowledge and to present the categories in general. A more detailed description of object knowledge concepts will be given in Chapter 6.

### 5.5.1 Basic principles for defining and structuring object knowledge

Knowledge about all design products created during the history of mankind would form the universe of discourse (UoD) of artificial systems. In other words, the UoD includes all

known artificially created things we can reason about: objects, materials, shapes as well as various types of relations between entities constituting the artefacts.

As discussed earlier, it would not seem wise to start building a knowledge base (or a database) that would store all known specific design solutions. The task would be considerably simplified if we could store only certain aspects of the design solutions that would be necessary to generate new design solutions. However, before deciding what aspects of design solutions the knowledge base should store, we must define: a) what the purpose of such a knowledge base will be and b) how to present, structure, and reason about knowledge stored in it.

An example of a badly designed database of designs (limited to inventive solutions) is a patent collection. Due to a vague classification and a lack of uniform methods for information structuring and representation it provides a user with little help in the design process. The main purpose of the patent collection is to help with finding analogous solutions and verifications of the novelty of solutions proposed but not with designing new artefacts.

On the other hand, there is no need to present design objects by describing every material and immaterial detail. Various objects can be grouped into larger groups following certain criteria, such as the function performed, similarity of shapes, kinematics, etc. Besides, every object can be regarded from a multitude of views focusing on specific aspects of the object's organization and functioning. In turn, each view can define what aspects are essential and necessary and what aspects are excessive. Besides, a view can define a type of model which describes the system: physical, mathematical, functional, logical, etc.

Another well known way to group solutions is to collect them into taxonomies via a domain feature. For instance, it is possible to distinguish between technological domains for cars, boats, medical tools, engines, etc. In the taxonomy based on the domain separation principle, various objects such as internal combustion engines and steam engines will belong to the same domain despite the fact that they incorporate different physical principles. Such taxonomy will be valuable for routine design when no new solution is needed and the task is to find an appropriate previous solution. However as follows from Altshuller's principle of cross-domain search for previous solutions, the domain-specific taxonomy appears to be virtually useless for innovative design. The reason is that only shallow knowledge is stored in the domain taxonomy.

The degree of knowledge deepness might be increased by grouping known solutions according to physical principles behind the solutions. From this point of view, we can present object knowledge for engineering design as a object-oriented hierarchy consisting of several knowledge layers. Each layer is to present knowledge with a different degree of details and generalization: the upper layer is formed by fundamental facts and laws of nature and the lower layer contains specific objects obtained by instantiating fundamental knowledge into specific solutions to meet particular needs and demands. Such classification makes it possible to organize the instantiation of deep knowledge into shallow knowledge by adding specific information on each lower level and inheriting the properties of the whole class of objects.

In turn, object knowledge can be divided into subcategories of *primitive* objects and *complex* objects. Complex objects can be assembled from primitive objects according to prescribed rules. However, the same object might be considered as a primitive component or a complex system where a number of primitive components can be networked into a system which will be able to deliver a given functionality.

### 5.5.2 Primitive object knowledge

Pahl [1984] distinguishes three types of knowledge structures involved in evolutionary engineering design: *physical effect* which maps a *function* into a physical law independent of the design solution, *physical principle* which is a combination of the physical effect and a function and *solution principle* which indicates what form features are needed to implement the physical principle (Figure 5.2).

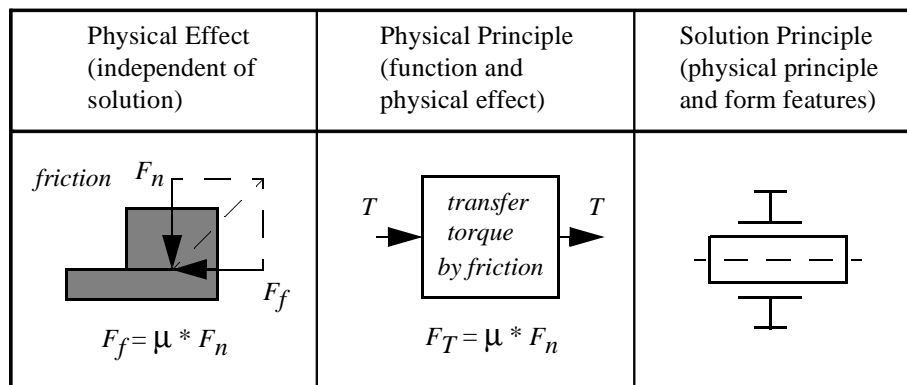


Figure 5.2: Translation between physical effect, physical principle and solution principle (adapted from Pahl [1984])

The German School for Systematic Design proposes to use this scheme as a basic means for organizing innovative design by mapping a required function needed into physical or solution principles stored in the Design Catalogues. A translation between levels is performed by adding more specific information at each level.

However in practise, obtaining new design solutions by only adding form features to physical principle might be difficult to perform due to a big difference between deepness of representations for physical and solution principles. The problem is that a physical principle comprises general physical relations which are difficult to interpret within the context of a particular design problem. On the other hand, a solution principle is too specific since it contains a predefined configuration and assembly of components thus limiting a range of possible interpretations and instantiations of a physical principle. The same physical principle might be instantiated into a multitude of design solutions that might have different form and



material features but perform the same physical function. No references are proposed how to perform such instantiation either in Design Catalogues or in TRIZ. In other words, there is a gap between physical and design views on the same object knowledge.

We believe that the problem can be overcome by introducing an intermediate knowledge structure between the physical and solution principles - design concept. In addition to information on its physical structure, the design concept captures information on what technical requirements can be met by the physical principle and what *generic* form and material features are needed. A design concept is represented as a view of how specific aspects of the overall behaviour of the physical principle can be related to engineering needs.

Summarizing, we believe that four layers of primitive object knowledge separated according to the degree of conceptualization should be distinguished for engineering design:

- fundamental physical laws,
- generic physical principles,
- generic design concepts,
- specific design descriptions.

#### **Fundamental physical laws.**

Knowledge of this layer expresses facts about the laws of nature in terms of relations between two or more physical variables and describing under what conditions the relations are valid. For instance, in classical physics, Newton's second law  $F=ma$  establishes a relation between the variables of force and acceleration via the constant value of mass. A distinguishing feature of this type of knowledge is that it is independent of the design context. The necessity of separating this layer from others is obvious: this is a layer of fundamental knowledge, and if we are interested in obtaining feasible solutions we must be able to check every new solution proposed by some procedure of design against its physical validity. On the other hand, if specific knowledge is obtained by instantiation of prior valid high-level knowledge, the former automatically inherit the feasibility of the upper level.

#### **Generic physical principles.**

A generic physical principle is an instance of a more general physical law regarded within the context of what behaviour a system of physical components based on the law possesses. The generic physical principle is more specific than the fundamental knowledge since it represents a system of physical components interacting under specific conditions. As a consequence, the behaviour of such a system might involve several relations based on several physical laws which are needed to provide the overall behaviour of the system. In addition to information presenting a behaviour of physical objects under specific conditions, the generic physical principle includes a description of a physical function. In turn, the physical function is a relation between two or more physical variables.

### Generic design concepts

A generic design concept is an interpretation of a physical principle according to what engineering needs can be satisfied by the use of the physical principle. The design concept provides specialization of the physical principle by adding a purpose that constrains a range of possible applications of the design concept.

If presented properly, this level can provide a link between a knowledge-intensive system for conceptual design and CAD systems for routine design. This makes it possible to automatically instantiate selected physical principles into specific design descriptions in accord with given design specifications.

### Specific design descriptions

This level stores all known instances of design concepts into detailed design descriptions. A space of solutions forming this layer consists of all known design descriptions. In addition to the information on the system's behaviour, specific design descriptions must contain all information needed to manufacture products and maintain their life-cycle.

TABLE 5.1. Layers of design object knowledge

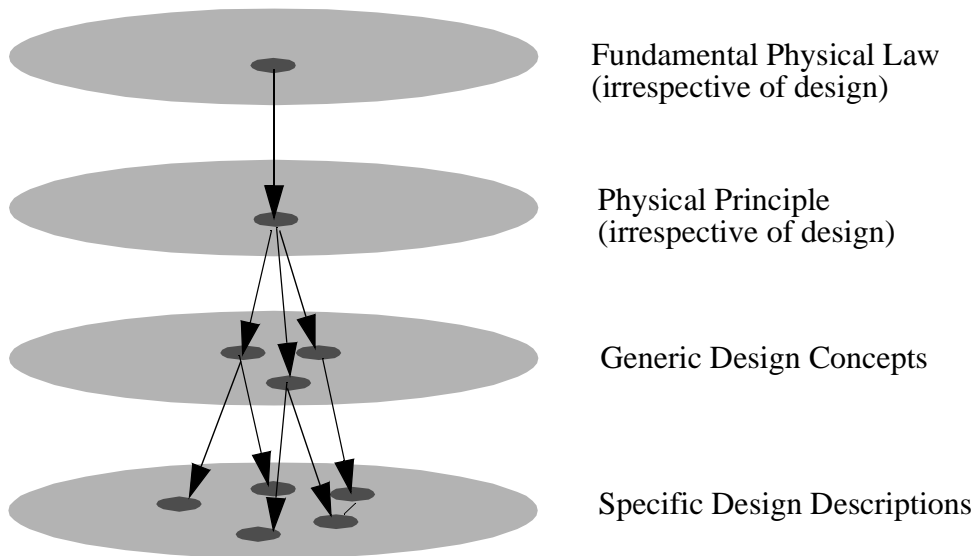
Layer	Contents	Example
Fundamental physical laws	<ol style="list-style-type: none"> <li>1. Qualitative description of a physical law.</li> <li>2. Specific mathematical expression establishing relations between quantities involved into the physical law.</li> </ol>	The equation of thermal expansion of materials
Generic physical principles	<ol style="list-style-type: none"> <li>1. Specification of physical components a physical phenomenon consists of;</li> <li>2. Specification of essential relations between components which realize the main physical function of the phenomenon.</li> <li>3. Specification of general physical constraints within which the phenomenon occurs.</li> </ol>	Mechanical displacement is a function that results from the expansion of solid bodies subjected to heating.
Generic design concepts	<ol style="list-style-type: none"> <li>1. Description of a generic system consisting of all the required physical components to form the phenomenon.</li> <li>2. Specification of relations constituting the overall behaviour of the system.</li> <li>3. Specification of engineering purpose.</li> </ol>	A metal body and a source of heat provide the function "mechanical displacement" with the purpose "precise displacement".

**TABLE 5.1.** Layers of design object knowledge

Specific design description	<ol style="list-style-type: none"> <li>1. Specification of structure and configuration of a specific design solution</li> <li>2. Specification of materials.</li> <li>3. Specification of shape.</li> <li>4. Specification of operational conditions</li> <li>5. Specification of constraints.</li> </ol>	An element of a component for precise positioning of a table in a microscope: a metal rod with a description of its geometry, material, and temperature intervals.
-----------------------------	---	--

It can be that every lower layer in this hierarchy adds new components and specific features thus making them different in the degrees of detalization and specialization (Table 5.1).

As clear, amount of knowledge and data that forms the layer of specific design knowledge is enormous. If we attempt to codify and store it, it would be unclear how to represent this knowledge to enable a cross-domain search. However, it might be important if we want, for instance, to establish analogical reasoning with specific knowledge. Even if it would be able to develop a proper knowledge representation method to store knowledge of this layer in a uniform way, the cost of the development of such a knowledge base would be unacceptable high - only patent collections count over 5 millions patents.



*Figure 5.3: Growth of the degree of specialization during translation of physical laws into specific design descriptions*

However if we could generate new conceptual design solutions without accessing this layer at the phase of innovative design, the task of automating the innovative design might be sig-

nificantly simplified. Hence, to organize reasoning about existing design concepts in order to be able to generate new solutions one has to codify all information at the layer of conceptual design knowledge. This layer stores significantly less data.

It is important that objects stored at the layers of generic knowledge must not be separated by domain features. We believe that such separation creates barriers to the use of previously used physical principles within a new context. As follows from a TRIZ interpretation of innovative design, this layer is most important for organizing the search for innovative solutions through high-order knowledge transfer over domains.

Apart from a better understanding of how knowledge for design is to be organized, the division between knowledge layers proposed provides an explanation of discovery, invention and innovation at a knowledge level:

1. A solution satisfying the given requirements is known and available at the level of generic design concepts. It can be selected and adapted to meet the requirements by a routine design procedure or synthesized from existing design objects. This situation results in innovative solutions.
2. No design object which is a solution to the problem is available. However, it can be generated by instantiating existing fundamental knowledge or physical principles into a new design object. Such type of design results in inventions.
3. No physical knowledge is available to meet the given requirements. This situation indicates that requirements should be reformulated otherwise a new scientific discovery at the level of fundamental physical knowledge is required to solve the problem.

### 5.5.3 Complex object knowledge

It is a rare case when a specific design description only utilizes a single physical effect. The majority of new specific design descriptions as well as generic design concepts consist of combinations of several objects. In many cases, to meet a new requirement is not possible by directly mapping the function given onto a single physical principle. To perform well, the system must have a source of energy and deliver some auxiliary functions to provide the required performance of system's main function.

A criterion which can be defined for dividing between primitive and complex objects is *complexity*. Similarly to knowledge deepness, complexity is a subjective factor and it may only be regarded with respect to a predefined level of system decomposition. For instance, a physical phenomenon of thermal expansion might be regarded as a primitive object whereas for other purposes, the same phenomenon can be decomposed into more primitive objects such as phenomenon of heat transfer, phenomenon of molecular movement and so on.

In engineering design, a concept of object complexity plays an important role in organizing bottom-up and top-down models of design process. Any design product can be represented

as a set of interrelated primitive objects. Vice versa, primitive objects can be arranged and related in such a way that a resultant composition of objects is a new design product.

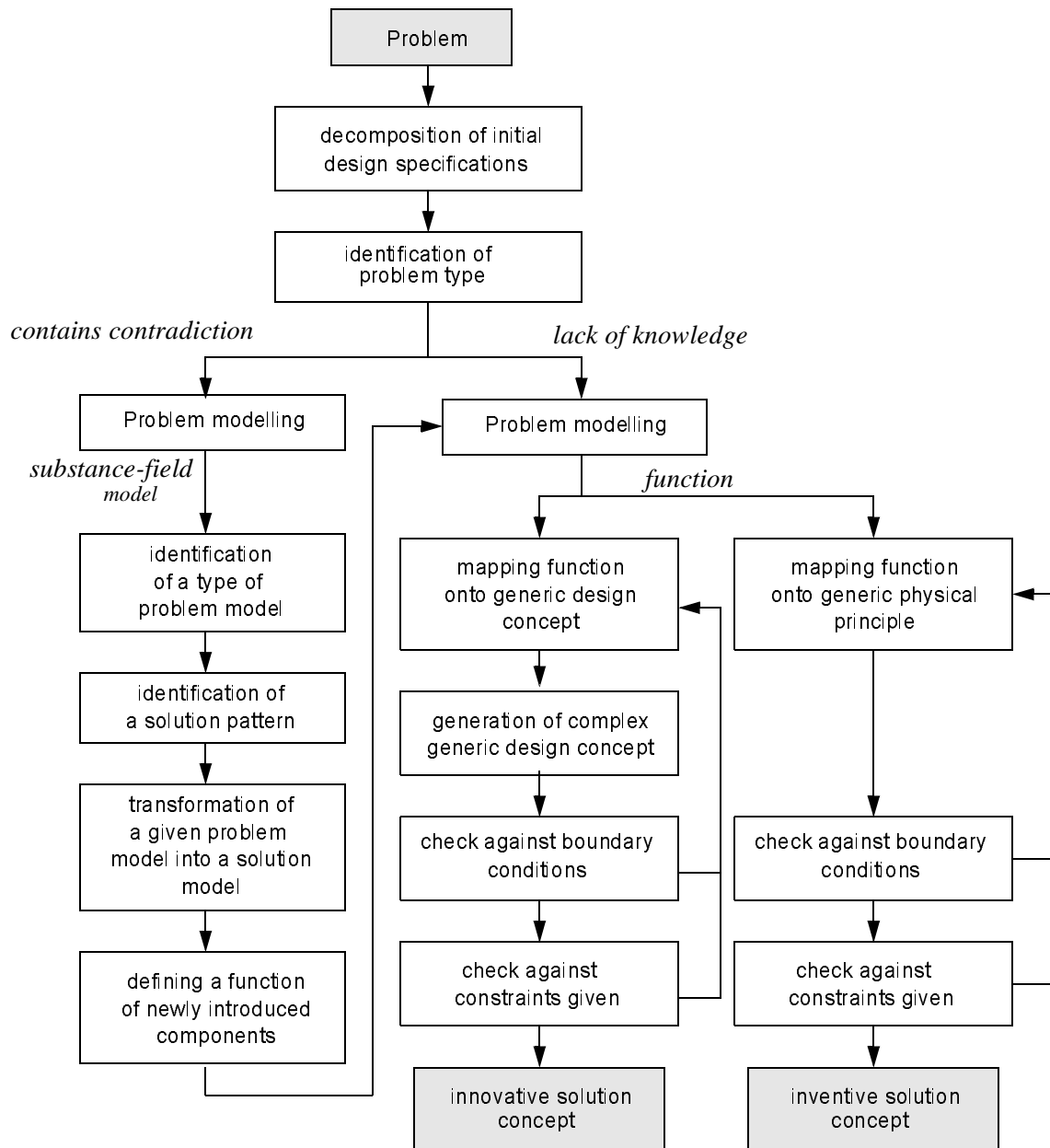


Figure 5.4: A knowledge-intensive model of innovative design

Summarizing, complex knowledge can be: i) decomposed into primitive objects to perform system analysis ii) synthesized from primitive objects to perform system synthesis.

#### **5.5.4 Object knowledge base**

Both primitive and complex object knowledge might be represented in KBS independently of the inference mechanism and strategic knowledge. Therefore, a knowledge base using classes, objects and relations between them might be sharable between various applications, such as diagnosis, generation of explanations, etc.

One of the INDES goals is to provide a theoretical framework for building a knowledge base of object knowledge for conceptual design. We argue that such a knowledge base should consist of at least two knowledge layers: a layer of generic physical principles and the layer of generic design concepts to obtain innovative and inventive design solutions respectively. A general modelling framework, knowledge representation methods and knowledge base contents will be presented in detail in chapter 6.

A similar approach to structuring design knowledge is used by several research projects worldwide which are trying to build knowledge bases which would be able to store physical features of design objects. For instance, a KBS based on qualitative process theory and developed at the University of Tokyo will contain 10.000 chunks of knowledge (Tomiyaama et al. [1994]). Similar efforts are being undertaken at Stanford University within the project "How things work" for building a knowledge base of physical devices used in engineering design and diagnosis (Iwasaki et al. [1993]).

### **5.6. Strategic knowledge for conceptual design**

In general design theory (Yoshikawa [1981], Tomiyama et al. [1987]), a design is considered as mapping from the functional space where design specifications are described in terms of functions onto the attribute space where design solutions are presented in terms of exact attributes. In innovative design, this mapping may not be realized directly since no design solution in terms of exact attributes is available before the mapping. Innovative design involves an evolutionary process: functional space is first mapped onto conceptual design space and only then conceptual attributes can be instantiated into specific attributes.

The role of strategic knowledge is to organize a reasoning process with object knowledge and task knowledge. A strategy can be derived from a general model of the problem solving process. In turn, the strategy defines what problem-solving strategy a KBS should be able to respond with to a particular problem.

Based on a study of TRIZ, we recognize the following groups of strategic knowledge for conceptual design:

1. Meta-level rules defining what particular reasoning strategy is to be used.
2. Principles for handling ill-defined initial design specifications available at the phase of inventive design and decomposing them into specifications recognisable by problem-solving.
3. Principles for mapping obtained specifications onto generic design concepts or generic physical principles.
4. Principles for the synthesis of alternative design concepts.
5. Principles for the evaluation and selection of the best design concept.

### 5.6.1 An integrated model of innovative design

Before we present an overview of strategic knowledge for innovative design, we would like to present a model of a design process we developed within INDES. The INDES design model is derived from the knowledge-centered model of innovative design, described in section 5.4.

Two types of designs are possible: design as synthesis and design as problem solving. Despite successful application within industries, this model incorporates a significant disadvantage if we want to build a KBS upon it: after working with Inventive Principles or Inventive Standards, we do not obtain a solution in terms of exact physical components to be introduced into the existing design or how the design has to be physically changed. Only abstract models of solutions are available.

Therefore we believe, that a solution concept obtained after design process must be represented as a design object in terms providing reference to what physical components are to be used no matter what problem solving strategy has been selected.

Another problem is caused by knowledge inconsistency. As follows from the analysis we conducted on TRIZ problem solving techniques, it would not be possible to model and represent engineering contradictions in a formal way if the formulations available in TRIZ were used. The principles for the elimination of engineering contradictions are very general and they are not intended to operate in terms of specific problems. On the other hand, the comparison of Inventive Principles and Inventive Standards revealed that the latter represent contradictions implicitly. Also solutions resulting from Inventive Standards are of higher quality since Inventive Standards are able to eliminate both physical and engineering contradictions. As a conclusion, Inventive Standards are able to cover the whole class of innovative problems which could be solved by Inventive Principles and there is no need to model both Inventive Principles and Inventive Standards.

To eliminate the above mentioned disadvantages, we have decided to combine both models of design into an integrated model of knowledge-intensive design shown in Figure 5.4. A choice of strategy for problem solving is defined after decomposing the initial design specifications. If this decomposition results in the necessity to change the existing design product

then a solution model can be obtained by applying Inventive Standards. Then, a function of newly introduced or modified components has to be defined to identify specific substances and fields. However, if the designer lacks knowledge of how to deliver the function, a direct mapping of the function onto existing generic object knowledge has to be performed.

### 5.6.2 Meta-level strategy

The meta-level strategy determines what type of inference will be relevant to solve a particular problem. As follows from previous discussion, after a decomposition of IDS, the design can proceed in two ways: transformation of the existing system or direct mapping of the function given onto generic object knowledge.

Summarizing, meta-level strategy rules are used to:

- evaluate what type of problem is specified;
- determine what type of problem model results from the decomposition of IDS;
- determine what problem-solving strategy and what knowledge sources are to be used to organize an inference with respect to a particular problem.

### 5.6.3 Initial design specifications

One of the disadvantages of a product-oriented approach to design process is that the quality of the obtained products depends solely on the validity and completeness of the input specifications. For this reason, getting the “right” specifications becomes a difficult task. It is common in different areas of engineering design that specifications of a product to be designed are expressed in the form of functions the product has to deliver, values of parameters of the functions and various constraints restricting possible design variants. As mentioned in chapter 2, exact functional specifications might not be available at the phase of innovative design. However, no design process is possible without defining the functionality of a future product.

To obtain a problem model expressed in the form of functional specifications, we believe that a technique for obtaining functional specifications from general problem formulations has to be developed. In INDES, we distinguish between several levels of initial design specifications (IDS), where the upper level contains commonsense expressions reflecting our wishes and desires, and the lower level contains functional specifications which can provide an adequate reference to some generic physical principle or design concept stored in a knowledge base.

Summarizing, the principles for IDS decomposition aim at coming up with two possible problem models:

- a function identifying a physical action with respect to some material component or physical field;



- a substance-field model of a physical system where undesired interaction between components implicitly indicates a contradiction.

#### 5.6.4 Problem modelling and problem solving

This type of knowledge defines a procedure of mapping between the specifications obtained after IDS decomposition and generic physical principles or design concept(s) available in the object knowledge base.

Problem modeling rules are defined on the basis of Inventive Standards. They are designated to specify a procedure for building correct and consistent models of design products that have to be improved but might not be since any attempt to make a change leads to a contradiction. A collection of Inventive Standards proposes specific patterns of solutions expressed in the form of black-box components. Therefore, a principal requirement is that a problem model must be formulated in terms adequate to the solution patterns.

Object knowledge presenting a design product that should be changed (design prototype, in other words) is defined as *task knowledge* in INDES. Task knowledge represents a description of a design prototype. The role of task knowledge is to provide a KBS with core information on a problem to be solved. A way of modelling and representing task knowledge is similar to the way of representing generic design concepts. The suitability of these representations makes it possible to reason about task knowledge in terms of task-independent object knowledge stored in a knowledge base of generic physical principles and design concepts.

After selection of an appropriate pattern, a solution is presented in terms of “empty” components introduced into the model of a design product. A specific-task instantiation of “empty” components into conceptual design solutions is provided by finding what material components and energy flows (physical fields) that constitute physical principles are capable of delivering a function specified for a “blank” component. The applicability of a selected physical principle is then verified by checking against boundary conditions between existing components of a design prototype and the components to be introduced.

#### 5.6.5 Design concept generation

Provided that a problem model is expressed in terms of exact functions, two situations are possible after mapping functional specifications onto object knowledge:

1. The object knowledge completely meets functional specifications. Depending on what level object knowledge has been found, physical or design, a specific design solution might be evolved from a set of abstract descriptions into a set of more concrete descriptions.
2. The object knowledge partly meets the functional specifications. In this case, a complex design concept has been generated from predefined primitives - generic physical principles or generic design concepts. The rules for new design concept synthesis

define a procedure for how relevant primitives are selected and combined into complex structures and what optimization methods exist to come up with an optimal concept.

We must note, however, that evolving the specific design description from generic design knowledge is not part of INDES. Therefore, the output of the design process performed with the INDES-based KBS results in the description of a primitive generic knowledge in the first case. The output in the second case is a complex generic design concept.

## 5.7. Summary

In this chapter, we presented a knowledge-centered model of innovative engineering design that emerged from studying Altshuller's approach and the TRIZ knowledge repository from a knowledge-based point of view. Based on this model, we distinguished and gave a brief overview of object and strategical knowledge that might be incorporated into a knowledge based system supporting innovative design.

We explained a difference between concepts of discovery, invention and innovation at a knowledge level since it is important for better understanding of how to structure knowledge for innovative design.

To make a model of innovative design more consistent, it is proposed to build an integrated model of innovative design allowing for both design as synthesis and design as problem solving to be supported by the same knowledge sources.

In the next chapters we will present both object and strategical knowledge in detail. We will also explain a framework for modelling and representing object knowledge for design in an uniform way based on an ontological approach.



## *Chapter 6. INDES: An ontological approach to modeling object knowledge.*

### **6.1. Introduction**

In supporting knowledge-intensive human activities, such as engineering design, AI systems capable of reasoning with a large diversity of knowledge provide better results compared to systems based on a narrow, domain-specific method. Such systems usually incorporate various different knowledge representation methods allowing for modeling knowledge with different degrees of abstraction and completeness. This requires communication between various knowledge sources. As a consequence, the development of such systems using traditional AI approaches, for instance, rule-based technology together with relevant techniques for knowledge acquisition, would be very complex, costly and time-consuming.

Over a few recent years, the ontological approach has played an increasingly important role with tackling the problem of describing knowledge domains. In general, the ontological approach belongs to a branch of philosophy concerned with what really exists as opposed to what appears to exist but does not. An ontology is the set of real objects or events the theory consists of described by referring to them without reference to an observer.

However, in AI, an ontology has slightly different meaning. The ontology is used to introduce a set of fixed concepts thus enabling factual information about domain to be modeled in uniform way. The ontology is separated from a knowledge representation method. Such separation makes it possible to make knowledge concepts presented within the same ontology to be sharable between different application frameworks. In AI, to create a domain ontology

means to perform knowledge categorization and conceptualization and to establish explicit relations between domain knowledge concepts.

As we pointed in Chapter 2, ontologies are especially useful for representing knowledge components of very large domains or metaknowledge that can be generalized over several specific domains. With respect to engineering design, massive research efforts are being undertaken at the University of Tokyo, where a collection of ontologies has been developed to create a Knowledge-Intensive Engineering Framework (Tomiya et al. [1994], Ishi et al. [1995]). Similarly, in the Knowledge System Laboratory of Stanford University, a collection of ontologies has been developed for design knowledge representation, sharing and reuse (Gruber et al. [1992], Gruber & Olsen [1994]).

Development of the ontology is a necessary step for organizing innovative design based on sharable physical knowledge. While in the previous chapter we discussed a role of object knowledge for organizing knowledge-centered innovative engineering design, in the present chapter the INDES ontology aimed at modeling object knowledge for conceptual engineering design is presented. The role of INDES is twofold: first it introduces a fixed set of concepts allowing for uniform modeling of existing design products, and second, the same set of concepts is used to design new product concepts.

TRIZ Pointer to physical effects (PPE) presented in Chapter 3 has proven to be effective in supporting innovative engineering activities. Knowledge stored in PPE can be viewed as a useful collection of factual information. On the other hand, the way in which design knowledge is organized and represented is a critical factor in the computer-supported design process. We therefore defined the task of INDES as the provision of a framework for restructuring, reorganizing and modeling knowledge stored in PPE. To achieve this, we use the ontological approach.

The chapter is organized in the following way: first, we present a critique of TRIZ approach to modeling physical knowledge; second, basic INDES concepts for modeling generic physical knowledge are discussed; and in the end, we illustrate our approach by several examples of modeled physical phenomena and generic design concepts.

## **6.2. INDES modeling framework**

### **6.2.1 Physical knowledge and System Theory**

The vast majority of designers refer to three types of information sources when designing new product concepts: physical handbooks, engineering handbooks and previous design documentation (or patent descriptions). While physical handbooks and encyclopaedia single out generic properties of interacting physical entities and present them in terms of quantitatively related variables and constants, engineering handbooks describe specific relationships between technical parameters of ready-to-use solutions. In addition, previous design docu-

mentation presents every aspect of a specific solution. Such organization of previous knowledge makes finding the right conceptual description of design product or concept within a large mass of information difficult if not impossible.

Therefore, structured collections of generic building blocks, presented in the form of conceptual design primitives could be of great help to the designers. The task of creating such collections addresses the task of *conceptual modeling* since accurate representation of *every* aspect of mechanical, electrical and other types of systems in a uniform way does not seem to be possible.

For this reason, our attention was drawn to the question: how can kinematic, electronic and other types of conceptual models widely used in engineering be presented in such way so that various domain-specific principles could coexist in the same diagram regardless of the domains used? What degree of conceptualization is necessary to provide sharability of the models and their compatibility to synthesise new design concepts from such models? The answer to this question requires a modeling method that will allow every physical effect and phenomenon to be presented in a uniform way.

Modeling is a key concept in problem solving. A large part of the difficulty of a computer-aided problem solving process is due to the difficulty of selecting an adequate modeling method and building a proper problem model. With respect to engineering design, one known framework which allows for such modeling are Bond Graphs. However, we believe that to present information on energy transformations is not sufficient. Although any transformation of matter can be expressed through a change of energy, we argue that both types of transformations should be presented explicitly. If a functional requirement is formulated as “to vary electrical resistance”, it may not be directly mapped onto an energy transformation principle since the requirement defines an operation with strictly material property. Hence, the first motivation for developing INDES was therefore ability to model information on material transformations.

The second motivation behind developing INDES was that when targeting at domain-independent computer-aided innovative design, a proper modeling framework should not be dependent on a particular problem-solving method. For this reason, models should explicitly incorporate the generic features of physical systems which are important for any type of problem-solving process in conceptual engineering design. Such independence can be provided by using the ontological approach to modeling knowledge.

Among a number of existing theoretical backgrounds for developing ontologies for conceptual design, our opinion has been that System Theory and network modeling is best suited for our task (Shearer et al. [1967], Martin & Allen [1969]). A major advantage of System Theory is that it regards entities of the physical world as systems and proposes adequate tools for establishing mathematically accurate relations between the physical entities.

System Theory also provides tools for the correct modeling of systems which belong to different domains. Apart from mathematical accuracy, representing different physical effects and phenomena in terms of systems gives the possibility to instantiate high-order system

models into more specific solution models, and to check against the feasibility of resulted models at very early stages of design. Therefore, INDES is intended for modeling of all known physical phenomena and effects in terms of systems by using common ontological concepts.

### 6.2.2 Macroscopic and microscopic observations

A complexity of modeling is related to, at least, two factors:

1. *Degree of detalization* which defines how deep modeling is performed.
2. *Assumptions* which define what information can be omitted during modeling since it has no influence on the use of the resulting model.

There are at least two relative levels of detalization at which the modeling of a physical phenomenon might be performed: microscopic and macroscopic. Take, for instance, the effect of Joule heat: generation of heat as a result of passing an electrical current through a conductive material. In this case, modeling at the microscopic level implies establishing relationships between energy characteristics of migrating electrons and the general increase of internal energy of the material. To model the effect under macroscopic observations means that first we have to make a decision on what information is crucial to achieve the externally observed result. In terms of System Theory, we have to establish the relation between the change of electrical energy and thermal energy generated, and to specify what parameters are involved in the relation. Information on internal processes can therefore be of no interest to the designer and is thus hidden inside a black box.

The latter way of modeling gives no clear understanding what internal physical processes cause the effect. Instead, it indicates what is needed to obtain the required result and how a change of energy and material properties can be used practically. In that case, the microlevel physical processes can be modeled as a black box. More detailed explanation of how modeling is performed on the basis of the effect of Joule heat will be given in section 5.6.1.

A black-box approach to modeling may not be identified with absolutely precise modeling. It does provide, however, information enough for organizing a reasoning process with modeled knowledge which is needed to generate new solution concepts. We can argue that such idealized models are the best way to represent information on various domains in a uniform way. Therefore, a choice of the level of detalization defines the degree of the ideality of the model.

### 6.2.3 Requirements for a model of physical phenomena

Given System Theory and network modeling as theoretical backgrounds, we formulated the following requirements for a conceptual model of physical effects and phenomena:

1. A physical effect has to be modeled as an abstract, simplified concept of a system providing context-independent semantics.

2. The model has to be generic with respect to possible specific solutions that might be obtained on the basis of the model.
3. The model has to explicitly incorporate all the necessary information which establishes relations between inputs and outputs of the system.
4. The model must be mathematically correct with respect to the chosen modeling level and not violate physical laws.
5. The model has to include constraint information to enable calculating quantitative values of input data when the output data is given and output data when the input data is given respectively.

#### **6.2.4 YMIR: a sharable ontology for modeling design knowledge**

As a basis for modeling generic physical knowledge for innovative design according to the requirements formulated above, we use YMIR -- a sharable ontology for modeling design knowledge in a uniform way. YMIR was developed in the Knowledge-Based Systems Group of the University of Twente (Alberts [1993]). YMIR defines a taxonomy of concepts for the formal description of design knowledge in different domains. The concepts in YMIR for the elements from which to synthesize technical system descriptions are labelled “generic system models” (GSM). These generalized concepts have been defined in terms of network models in System Theory. Generic system models explicitly incorporate the relation between features of an engineering system such as behaviour and form.

YMIR distinguishes multiple levels of abstraction with corresponding sets of generic system models. The result allows for the gradual refinement of the design description at levels ranging from the original problem specification to the final artifact description. The resulting collection of knowledge structure concepts makes it possible to systematically organize the knowledge to support its reuse. The advantage of YMIR is that the modeling framework is applicable to all domains in which technical systems can be described as system-theoretical network models.

Whereas Bond Graphs are limited to modeling energy aspects of systems, YMIR enables modeling form properties of systems. Using material and geometrical aspects of designs to model their behaviours makes it possible to incorporate necessary information on material properties of designs. This claim is important since, in many situations, the initial key requirements are formulated in terms of parameters of material components. As a consequence, the requirements can not be directly expressed in terms of energy transformations without knowing what effect is able to meet these requirements.



## 6.3. INDES Concepts

### 6.3.1 Physical Components

The INDES ontology defines two sets of concepts: concepts for modeling physical components and concepts for modeling the relations between the physical components. Among the concepts presenting the physical components in INDES are: material component, energy flow, generic physical principle and generic design concept. By introducing these concept, INDES covers the whole class of physical entities and generic systems that exist or can be obtained from more primitive components.

A generic design concept consists of a number of more primitive components such as material components and energy flows.

#### *Definition 6.1.*

**Material component** is an idealized physical object possessing mass, properties of materials and occupying space. We distinguish between lumped and distributed components (such as solid, gas or liquid).

A material component is an object modeled under the assumption that it may not be decomposed into more primitive components.

#### *Definition 6.2.*

**Energy flow** is an idealized flow of energy either spatially oriented or distributed in space.

In general, the energy flow might be represented through physical parameters both in vector and scalar forms. In INDES, we use scalar form. Similarly to the material component, the energy flow is an idealized component which may not be decomposed into more primitive components.

The material components and energy flows can be composed into a generic physical principle which is a central concept of INDES.

#### *Definition 6.3.*

**Generic Physical Principle** is a spatial arrangement of material components and energy flows that results in continuous or discrete change of parameters identified with an energy flow or material component for the time given.

Analogously, generic physical principles can be composed into a more complex structure labelled the *Generic Design Concept*. The generic design concept is a model of a final design product or some part that interacts with the environment or other systems and fulfils a practical purpose.

**Definition 6.4.**

**Generic Design Concept** is a physically valid network of generic physical principles arranged in a predetermined way and which achieves a purpose specified by design specifications.

Formally, a generic design concept is a tuple  $DC = \langle S_F, E_F, P \rangle$ , where  $S_F$  is a set of physical phenomena occurring within a system,  $E_F$  is a set of relations between phenomena, and  $P$  is the purpose(s) of the design product.

A detailed description of modeling both generic physical principles and generic design concepts will be presented further in this chapter after definition of concepts for establishing relations between the physical components.

### 6.3.2 Relations: behaviour and function

Physical effects result from specific spatial and temporal arrangements of material components and energy flows. When arranged in a specific manner and provided that there is a time interval, the physical components interact with each other. Such an interaction results in a change of certain parameters identified with the objects. An ontological concept for presenting such changes over the time by specifying the relations between the objects is *behaviour*. A system's behaviour serves as a means to:

1. Establish the relations between system parameters involved into description of the physical components comprising physical principles.
2. Provide a mapping of functional specifications onto a physical system that is capable of meeting the specifications.

A behaviour specifies *how* a system works by specifying relationships between the components of the system. These relationships specify how a system's components interact with each other according to an external change. As components of the system may possess their own behaviour, the overall system's behaviour might be decomposed into more specific behaviours.

Before giving definitions of behaviour and function, we must analyse how these two concepts are interrelated. This analysis is important since no common agreement has been achieved in AI in Design on what is dominant: function or behaviour.

A study of the literature shows that discussions arise about the relative places of function and behaviour because of two different interpretations of the concept of function. The first point of view is widely accepted in the design community which defines a function as a relationship of input and output material, energy or information and was introduced by Rodenacker [1971]. On the basis of his definition, Koller [1976], Roth [1982] and Krumhauer [1974] introduced sets of so-called "elementary functions" that might be used for modeling of any type of physical system in terms of energy-transforming components.

Within such approach, a behaviour of a system is defined as a combination of such elementary functions. However, there is still no common agreement between designers as to which set of elementary functions to use (van den Kroonenberg & Siers [1996]).

A totally different approach defines function as an expression of the designer's intent(s) regardless of physical or informational semantics. Value Engineering, for instance, defines a function in the verbal form "to do something" (Miles [1972], Litvin & Guerassimov [1991b], Sasajuma et al. [1995], Kowalick [1996]). Under this view, function is represented as a verbal expression of what an object, which is a part of a system, is supposed to do with respect to other system components. Therefore, such function as "to prevent an aircraft from shaking" is correct in Value Engineering.

As a conclusion, without prior commitment on a context in which the word "function" is understood, its use might be misleading. From our point of view, the ambiguity of definitions for function and behaviour observed within the different AI and Design schools is an attempt to use the same word to describe both physical processes occurring in the system and their interpretation in the form of a designer's intents. However, the designer's intents, and, as a consequence, functions formulated as the intents might have nothing to do with the physical context.

To eliminate this ambiguity, we argue that it is important to distinguish between two views under which any artificial system might be viewed: *physical-oriented* and *solution-oriented*. These two views define different concepts which can be used to model both information about physical processes occurring in the system and the expression of the designer's intents. Under a physical-oriented view, function specifies how two physical entities are related, that is, an unambiguous, reproducible relationship between input and output. For instance, it might be a relation between heat and electrical resistance as well as between temperature difference and a mechanical displacement with an exact specification of all variables participating in the relations. Therefore, our notion of a physical function is similar to the one used in classical mathematics, but in the context of specific physical parameters.

A solution-oriented view interprets such relations from the particular point of view of a designer. Usually such functions are expressed in a qualitative or verbal form. In that case, function can be regarded as an abstract interpretation of the relation between two physical parameters regardless what physical parameters participate in the relation. For instance, if a function is expressed as "to cool an object", it may refer to a multitude of physical phenomena that can be used to change the object's temperature.

To avoid confusion between the various interpretations of the term "function", we proposed to call a solution-oriented view of a physical function a *technical function*. The technical function therefore is a task-oriented, qualitative interpretation of the relation between two physical parameters under a certain context expressing a designer's intent. Observing the system from a physical point of view, we can define behaviour and function as two physical properties of the system which are independent of the engineering context.

For instance, the physical function of heat transfer  $T = f(H)$  does not specify itself if the temperature grows or drops unless the sign of dynamic change of heat flow is taken into consideration. Therefore, the technical function “to cool the object” is a particular interpretation of the function of heat transfer and is an instance of a more general physical function.

Since we agreed that we should regard physical function as a specific occurrence of an energy transformation process within a system, we will define the physical function as a mathematical relation between input and output energy flows.

**Definition 6.5.**

**Physical function** specifies a particular relation between input and output of a system presented in the form of energy flows.

Formally, physical function can be defined as a tuple  $F = \langle P_1, P_2 \rangle$ , where  $P_1$ , and  $P_2$  are two related physical parameters describing energy flow, and  $P_1 = f(P_2)$ .

Note that physical function<sup>1</sup> itself only indicates a type of relation between parameters without specifying constraints on parameters participating in the relation. For this reason, the same function can produce multiple behaviours. In Alberts [1993], two types of functions are introduced: generic and specific. A generic function constrains the values of parameters to a limited set of intervals, whereas a specific function constrain the values of the parameters to one particular interval. When modeling existing design prototypes with INDES concepts, we will use the concept of specific function.

A specific function can be further decomposed into *subfunctions*. Subfunctions are needed to clarify the representation of specific aspects of the physical process behind the physical phenomenon. Decomposition of the function into subfunctions is important to analyse problems to reveal a specific role of individual components. On the other hand, explicit reference to subfunctions provides mapping not only between functional specifications presented in terms of energy transformations but in terms of material transformations as well. There are two types of subfunctions in INDES:

1. A subfunction which defines how a material parameter depends on a change of energy flow for the time given:  $P_M = f(P_{E\ in})$ .
2. A subfunction which defines how a certain parameter of energy flow depends on a material parameter for the time given:  $P_{E\ out} = f(P_M)$ .

To give a definition of behaviour we will assume that technical systems are specific instances of generic physical systems considered in a particular engineering context. As described in previous chapter, such instantiated systems can be modeled as a set of physical objects interacting with each other through energy and informational flows among the objects constitut-

---

1. Further in the text, we will refer to the physical function as “function” and the term “technical function” will be used explicitly

ing the system. At two particular moments in time, a system can have different values for its parameters or its components can be in different physical states according to the specifications given. Then, from a physical point of view, an overall behaviour can be represented as a set of related functions between physical components combined in such a way that the desired effect is produced.

**Definition 6.6.**

**Behaviour** of a system is a set of relations between parameters presenting input and output energy flows established through the use of material parameters.

In other words, the overall behaviour of a system comprises functions of individual components. YMIR uses two types of variables to describe a behaviour of a system:

1. *System variables* which describe energy flows in a system. In turn, two types of system variables are distinguished (Table 6.1):
  - 1.1. *Implicit* variables which describe the energy flow through the system (e.g., electrical current and force flow).
  - 1.2. *Explicit* variables, which describe the potential differences across the system. (e.g., voltage and displacement respectively).
2. *Form-related variables* which describe principal material and geometrical properties of a design. These serve as parameters in the equations relating the system variables (e.g., length, modulus of elasticity).

In addition, the derived forms of both types of variables can be used as arguments of functions. For instance, power and kinetic energy are derived system variables and area and volume are derived form-related variables.

**TABLE 6.1.** Classification of system variables (adapted from Shearer et al. [1967]).

System	Implicit variable	Integrated implicit variable	Explicit variable	Integrated explicit variable
Mechanical translational	Force	Translational momentum	Velocity difference	Displacement difference
Electrical	Current	Charge	Voltage difference	Flux linkage
Fluid	Fluid flow	Volume	Pressure difference	Pressure momentum
Thermal	Heat flow	Heat energy	Temperature difference	Not used in general

### 6.3.3 Structure.

To separate between two or more GPPs, their borders are denoted with the help of *terminals*, that is, ports for input and output energy flows. Therefore, a structure of the model is defined by a list of input and output ports.

Since we have chosen a black-box modeling approach, we will define two types of primitive components: two-terminal GPP and multiterminal GPPs. An example of a two-terminal component is an electrical resistor when it is used to drop a voltage difference. According to the effect of Joule heat, the same electrical resistor produces heat when an electrical current passes through it. Viewed from the position of both electrical and thermal domains the same physical object can be represented as a GPP converting electrical energy into thermal energy. In this case, the model of the effect of Joule heat consists of four terminals. Another example of a four-terminal GPP is the piezoeffect: generation of a difference of electrical potential as a result of mechanical displacement.

#### *Definition 6.7.*

**Structure** of the generic physical principle is a list of input ports where energy flow comes in and output ports where the energy flow comes out.

### 6.3.4 Constraints

Taking into consideration that each parameter which belongs to a real system must be defined within a certain interval, or to have different discrete values during different states, all quantitative constraints on parameters can be described by the space of values. Given any particular moment in time, actual values of the parameters can be defined. Therefore, the following notions are necessary to present numeric information constraining the system's behaviour:

- 1 *Generic space of values of system parameters:*  $S = \{P^S_1, \dots, P^S_n\}$ ; where  $P^S_i$  is a range of allowed values of energy or material parameter  $P_i$ . Two types of intervals are possible: discrete and continuous intervals.
- 1 *Specific space of values system parameters at a given time:*  $A = \{P^A_1, \dots, P^A_n\}$ ; where  $P^A_i$  is an actual value of energy or material parameter  $P_i$  at a selected moment of time.

These types of constraints represent information on a system through numeric intervals or discrete values of parameters under which the system's behaviour corresponds to system purposes and meets all requirements. To represent dynamic aspects of a system we also need to take into account dynamic changes of variables specifying functions. If a static constraint is expressed through interval values of parameters, then dynamic constraints specifies how a

variable describing the system parameter depends on time. The dynamic constraint is a qualitative expression of a function in terms of qualitative reasoning.

The best way to represent dynamic constraints is to use the mathematical notion of a partial derivative, as it is proposed in Qualitative Process Theory. We distinguish between four possible types of dynamic constraints which present the required dynamic change of a system parameter. These constraints can be formulated both for energy flows and material variables (Table 6.2).

TABLE 6.2. Dynamic change of a system parameter.

Situation	Interpretation
$\partial P / \partial t \leq 0$	Increasing the value of the parameter.
$\partial P / \partial t \geq 0$	Decreasing the value of the parameter.
$\partial P / \partial t \neq 0$	Constant value of the parameter.
$\partial P / \partial t = 0$	Variable value of the parameter.

## 6.4. Generic Physical Principle (GPP)

### 6.4.1 Model of generic physical principle

The basic idea behind the GPP is that the behaviour of a physical phenomenon can be expressed as a set of related physical functions, each of which, in turn, can be instantiated into a multitude of technical functions within specific contexts. Therefore, a GPP can be regarded as a high-level model of generic design knowledge which can be instantiated into a design description by adding constraints and domain-specific information.

In YMIR, the generic system model is used as a basic concept to represent generic design knowledge. The same framework can be used to model fundamental physical knowledge from the engineering point of view. A physical principle is represented as a system that has input and output ports and its behaviour comprises different physical functions and subfunctions. A GPP in INDES is more generic than a GSM in YMIR which represents design knowledge because the GPP has no influence on design constraints and other specific limitations.

Any part of a system performing some specific function can be modeled as a set of fundamental physical phenomena occurring in the design. This makes it possible to apply the same physical model to represent different groups of physical phenomena. From a knowledge-based point of view, any physical phenomenon might be represented as a tuple:

$$\Phi = \langle E, M \rangle$$

where  $E$  is a set of energy parameters and  $M$  is a set of material parameters. As a consequence, GPP is a model of a system based on a specific physical phenomenon.

A GPP model has three parts:

- *Behaviour*: a set of relations between input and output energy flows.
- *Form*: list of material and geometrical parameters which determine the behaviour of a GPP.
- *Structure*: lists of ports for input and output energy flows.

We distinguish between two categories of GPP:

- *Homogeneous GPP*: input and output flows are of the same type of energy (Figure 6.1). Here,  $C_m$  is a particular physical property of a material which provides the change of the energy flow  $E$  and the potential difference  $\Delta A$ .
- *Heterogeneous GPP*: input and output energy flows are of different types of energy. The elements converting one type of energy into another are called *transducers* in System Theory (Figure 6.2). Here,  $C_m$  is a particular physical property of a material which provides the transformation of an input energy flow of one type  $A_{in}$  into an output energy flow of another type  $B_{out}$ .  $\Delta A$  and  $\Delta B$  are explicit variables that describe the potential differences across the system.

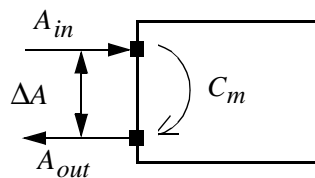


Figure 6.1: Graphical form of homogeneous GPP

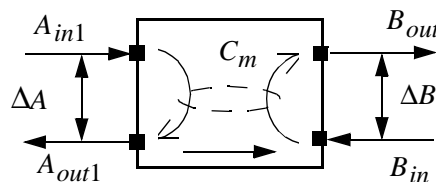


Figure 6.2: Graphical form of heterogeneous GPP



The modeling task consists of describing the behaviour of a physical system formed by material properties and energy flows which a GPP consists of with respect to its applicability in engineering. The behaviour of any GPP can be observed from two points of view which may be of a particular interest to a designer:

- The physical property  $C_m$  of the material components provides the transformation of the input energy flow  $E_{in}$  which results in the output flow  $E_{out}$ .
- The transformation of the input energy flow  $E_{in}$  into the output energy flow  $E_{out}$  always results in a change of some material property.

These two points of view can be used to decompose the overall behaviour of a GPP into various physical functions dealing with changing of energy as well as material parameters.

One of the important characteristics of a GPP is a type of relation between a change of system variables and time. Two types of GPPs can thus be distinguished: *continuous* systems and *discrete* systems. Continuous systems in which values of variables change continuously can be represented by differential equations. In turn, discrete systems in which variables change only at discrete moments can be represented by difference equations.

To provide physical realisability of a future design based on a GPP, we apply the law of energy conservation when describing the behaviour of the GPP. According to this law, losses of energy within a system must be conserved. Hence, the modeling task consists in establishing correct relationships between input and output energy flows according to the conditions following from the law of energy conservation.

The use of a GPP might therefore be organized through mapping between the GPP and a set of physical functions which can be produced by the GPP.

Summarizing, a GPP incorporates both high-level energy transformation knowledge as a capability of some physical phenomenon to produce physical functions and material aspects that make this transformation possible. The availability of material-related information in the GPP enables one to translate between a GPP and a more specific design description.

## 6.4.2 Types of Generic Physical Principles

As we pointed in Chapter 5, conceptual design might be performed in two ways: finding a mapping of the required function onto a physical phenomenon regardless of existing designs and producing a change of the existing design prototype. Since we defined the scope of INDES as modeling both types of design, the same modeling framework is to be used in both situations. As a consequence, the same ontological set of concepts for generic physical principles can be used to model the existing design prototypes.

TABLE 5.3. Categories of Generic Physical Principles

Component	Symbol	Overall function	Examples
Converter		Converts one type of energy A into another type B.	Thermal expansion, piezoeffect
Conductor		Conducts energy flow A in space	Water pipe, fibre optics
Variar		Energy flow A is controlled by energy flow B. Might be both linear and non-linear.	Variable electrical resistor
Resistor		Subclass of variar: decreases a system variable identified with energy flow A over time.	Electrical resistor, hydraulic filter
Amplifier		Subclass of variar: increases a value of a system variable identified with energy flow A by applying an external energy flow B	Electrical transistor, thermits
Source		Generates energy flow A	Any converter regarded as two-terminal component
Storage		Accumulates and stores energy A over the time	Electrical capacitor, water container
Switch		A discrete type of the conductor of energy flow A controlled by energy flow B	Curie point

All natural phenomena behind generic physical principles utilize the same general principle of energy transformation. We can, however, distinguish between several categories of natural laws, and generic physical principles respectively, which differ in a way we view them according to their functional destination. This division is important for modeling design prototypes.

A number of component categories is distinguished in System Theory (such as resistors, inductors, capacitors) as well as in Bond Graphs (such as changer, varier, connector, channeler, store and source). INDES classification of GPP (Table 6.3) introduces eight categories of components based on the analysis of TRIZ collections of physical effects and physical functions.

### 6.4.3 Generic Design Concept

A uniform way of modeling different GPPs and the law of energy conservation used to check against boundary conditions make it possible to connect several GPPs into a more complex structure that is labelled the *Generic Design Concept* (GDC). A GDC is a conceptual description of a new design consisting of the network of a GPP. The amount of GPPs that can be networked into a GDC is unlimited.

Two GPP are allowed to be connected when:

1. The output energy flow of the first GPP and the input energy flow are of the same type.
2. The values of the parameters of both flows are within the same generic intervals;
3. The boundary conditions are provided:
  - 3.1. For explicit variables:  $E_{2in} = -E_{1out}$
  - 3.2. For implicit variables:  $\Delta A_1 = \Delta A_2$

Example of a GDC consisting of one energy source and two converters is shown in Figure 6.3.

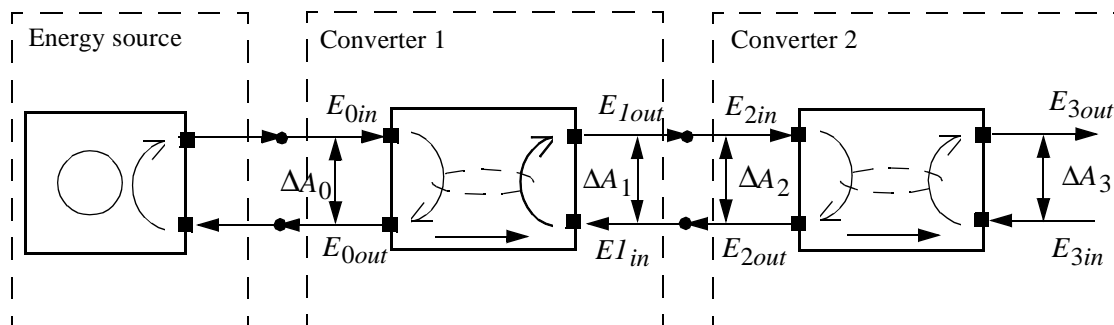


Figure 6.3: Graphical form of generic design concept

We must however note, that there is no clear distinction between a GPP and a GDC. The same physical phenomenon can be represented as a GPP whereas under another level of

abstraction the same effect can be regarded as a combination of more primitive effects and thus represented by a set of connected GPPs. We therefore have decided to limit ourselves to modeling those physical effects and phenomena which are available in physical encyclopaedia and can be represented as system-theoretical models.

The concept of a GDC in INDES is similar to the concept of a compound prototype in YMIR. The difference is in the way a GDC and compound prototypes are obtained: while YMIR proposes to model existing designs in terms of compound prototypes, INDES proposes to synthesise new GDCs each time when the new design specifications are available. Another difference is that GDC may consist of physical principles which belong to different domains, in contrast to YMIR.

## 6.5. Examples of modeled Physical Principles

In this section, we present the examples of modeled generic physical principles. Each physical phenomenon is modeled as a GPP at the appropriate level of abstraction. The choice of the level of abstraction depends on what energy and form-related information on a physical effect is crucial to translate between physical and engineering levels and to check the applicability of translated knowledge against functional requirements. To illustrate our approach to modeling physical effects in details, we will analyse the effects of Joule heat and Thermal Expansion.

### 6.5.1 Joule Heat

Joule heat arises in a conductive material when an electrical current passes through the material. The heat is generated as a consequence of increasing internal energy of the material due to collisions of migrating electrons which lose their energy. This effect is used in many technical systems, like electric stoves, toasters and hair dryers.

According to INDES, the behaviour of a generic system model is expressed as a set of relations between input and output system variables, where form-related variables serve as relating parameters. The physical behaviour of the effect of Joule heat can be defined by the ordered tuple:

$$R = \langle \langle U_{in}, I_{in} \rangle, \langle T_{out}, H_{out} \rangle \rangle$$

where  $U_{in}$  and  $I_{in}$  are input voltage and current,  $T_{out}$ ,  $H_{out}$  are the output temperature difference and the heat flow.

A graphical model of the effect of Joule heat is depicted in Figure 6.4. The line with the arrow indicates the direction of the energy flows. The line with arrow at the bottom of the box indicates the direction of the energy transformation.

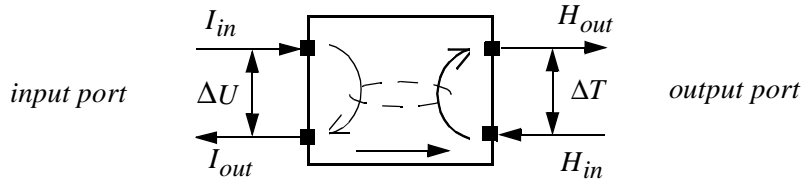


Figure 6.4: General model of the effect of Joule heat

Relations between the pairs of input or output system variables at each side of the GPP can easily be established in a correct way easily whereas to establish the overall relation between two different energy flows using variables of a single domain is not possible. The problem arises because we may not describe the same physical property of a material in terms of both domains explicitly.

For instance, in terms of an electrical domain, the relation between the voltage  $U$  and current  $I$  is established by Ohm law  $I=U/R$  whereas an electrical resistance  $R$  may not be used as a form-related variable to establish the relation between system variables characterizing the domain of heat and mass transfer. To establish this relation one must use appropriate form-related variables from the thermal domain, like thermal capacity or thermal resistance. To establish a relation between two different energy flows one should investigate how form-related variables of both domains are interrelated.

To solve this problem in an explicit way, a deeper level of modeling must be chosen because a precise physical behaviour specifying the capability of current to produce a heat flow may only be modeled at the molecular level. Therefore, the relation between the system variables of both energy flows can be established under macroscopic observations, whereas the relations between both flows can only be precisely modeled under a microscopic view.

A solution to this problem consists in explicit modeling of external effect caused by the combination of internal effects and hiding microscopic-level information within a black box. Therefore, the overall behaviour of the effect can be defined as a relation between the system variable at the left side of the GPP and the system variable at the right side. In the effect of Joule heat, such a system variable is voltage difference at the left side and the difference between values of the initial temperature of the material and the temperature of the material after some period of time on the right side. Thus the observed temperature difference is a time-dependent function of the internal process of heat and mass transfer occurring inside the material.

Since the effect of Joule heat belongs to the class of converters, we can explain why this way of modeling would be applicable to all types of the converters. An external energy flow amplifies some internal physical process inside the material which is too weak to be observed or does not provide the required physical function. In other words, the source of energy already exists in a system and may produce the required energy flow under the condition that this energy source is activated. It turn, the existence of the weak energy source is provided by some other external energy flow acting upon the material (like gravitational forces or thermal radiation from the environment). If this external flow has no influence on the system within the required range of parameters, it might be omitted while modeling.

To establish the relation between the generated internal heat flow which leads to a growing external temperature difference we propose to model the conductive material producing Joule heat as an energy source and to analyse the behaviour of the physical processes which occur inside this source. The material component can be modeled by three elements: thermal source, thermal resistor and a thermal capacitor which accumulates heat and thus increases the temperature of the material (Figure 6.5).

Two types of energy sources are known in System Theory: A-type sources, in which an across-variable is a function of time and T-type sources, in which a through-variable is a function of time. In the effect of Joule heat, the internal source of thermal energy is a T-type source since it generates the heat flow which, in turn, leads to the temperature growth. In this model, we can separate between the source of heat and the rest of the material components which acts as a thermal capacitor capable of accumulating heat energy and thus increasing the overall temperature of the object.

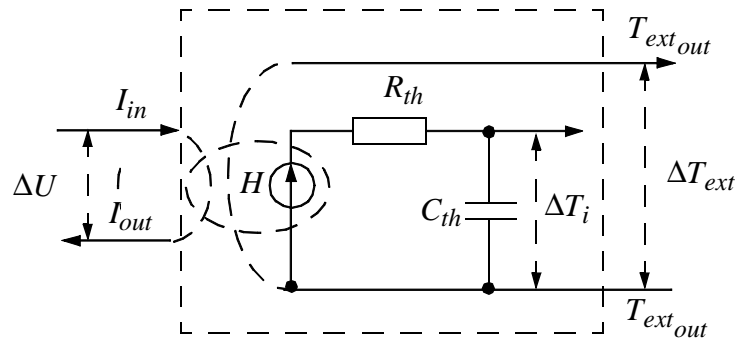


Figure 6.5: Internal processes in the effect of Joule heat

In this model a growing temperature difference  $\Delta T_{ext}$  is a function of time. In the diagram,  $R_{th}$  is the internal thermal resistance of the internal source, and  $C_{th}$  is the thermal capacitance of the material. According to the law of energy conservation, we assume that all electrical energy is transformed into a thermal energy generated by a dot source of heat, and the ther-

mal capacitance is created by the conductive material from which the object is made. The dashed box defines the processes omitted while modeling under macroscopic observations.

Nevertheless, information on the internal processes may have no practical value since we are not interested in comprehensive modeling of the phenomenon. Thus we have to know how to relate the observed difference of a temperature at the component and input voltage and not in the internal heat flow which causes this difference.

### 6.5.2 Behaviour

As known from System theory, thermodynamic power is defined as a product of entropy flow and temperature. However, entropy was introduced as an associated parameter with the flow of heat through a thermal resistance, since a thermal resistance dissipates no energy and the net heat flow is always zero. Therefore, the heat flow is itself the thermal power, and a relation between electrical-thermal transformation can be written as  $H=IU$ , where  $H$  is the heat flow through the system being generated. As a result of the heat flow through a given material, this material stores internal energy by virtue of temperature rise. For a real thermal system including a non-dot material component the relation between the temperature rise and amount of heat flowing through the component is:

$$\Delta T_{out} = \int_{t_1}^{t_2} \frac{H}{M_m C_m} dt, \quad (6.2)$$

where the system variable  $\Delta T_{out}$  is the temperature difference for the time interval from  $t_1$  to  $t_2$ , and form-related-variables are  $M_m$  and  $C_m$ , the mass and the specific heat of the resistive material respectively.

To define the external behaviour of the effect of Joule heat we relate both pairs of system variables as

$$\frac{dT}{dt} = f\left(\frac{dU}{dt}\right), \quad (6.3)$$

and substituting the heat flow variable with a relation for electrical power in terms of input voltage and electrical resistance, we obtain the equation for a particular behaviour of the effect of Joule heat:

$$\Delta T_{out} = \int_{t_1}^{t_2} \frac{U_{in}^2}{R_m \cdot M_m \cdot C_m} dt \quad (6.4)$$

Although the effect of Joule heat is non-reversible, this relation is correct unless the rise of temperature influences the relation between input system variables. At the same time, the model of the effect would not be complete if we did not take into account another physical effect which always accompanies the effect of Joule heat: the change in electrical resistivity

of a conductor as a consequence of heating the conductor. This influence can be regarded as feedback and must be included in the model of the system forming the effect as a part of its behaviour.

To model this part of the effect behaviour, the electrical resistance is first defined as a form-derived variable in terms of form variables:

$$R = \frac{L_m}{A_m} \rho_m, \quad (6.5)$$

where:  $\rho_m$  is electrical resistivity of the material,  $L_m$  is the length of the conductor and  $A_m$  is its cross-sectional area.

For temperature intervals of as much as a few hundred degrees, electrical resistivity is related to temperature by a linear expression of the form:

$$\rho = \rho_0(1 + \alpha_m \Delta T) \Big|_{\Delta T = T - T_0}, \quad (6.6)$$

where  $\alpha_m$  is the temperature coefficient of the resistivity, and  $\rho_0$  is the electrical resistivity at a temperature  $T_0$ .

Substituting  $R = \rho_0 L/A$ , we can define the behaviour of the system based on the Joule effect as:

$$\Phi = \Delta T_{out} = \int_{t_1}^{t_2} \frac{U_{in}^2}{\rho L_m A_m M_m C_m} dt, \quad (6.7)$$

with the feedback relation defined as:

$$\rho = \rho_0(1 + \alpha_m \Delta T_{out}) \quad (6.8)$$

### 6.5.3 Form

The form of the GPP of Joule Heat is a list of material and geometrical variables (both are known as form variables) involved in the relations presented above.



### 6.5.4 Overall model

In summary, an overall model of the effect of Joule heat includes three parts: form, structure specified by input and output ports, and behaviour as a set of relations between system variables. The overall model of the effect of Joule heat is presented in Tables 6.4, 6.5. and 6.6.

TABLE 6.4. Behaviour of GPP of Joule heat

Type of relation	Function	Mathematical expression
Input system variables	$\Phi_{in} = \langle I_{in}, U_{in} \rangle$	$I = \frac{A_m U}{\rho_m L_m}$
Output system variables	$\Phi_{out} = \langle H_{out}, T_{out} \rangle$	$\Delta T_{out} = \int_{t_1}^{t_2} \frac{H}{M_m C_m} dt$
Conversion relation	$\Phi_{tr} = \langle T_{out}, H_{out}, U_{in}, I_{in} \rangle$	$\Delta T_{out} = \int_{t_1}^{t_2} \frac{U_{in}^2}{\rho L_m A_m M_m C_m} dt$
Feedback relation	$\Phi_f = \langle I_{in}, T_{out} \rangle$	$I_{in} = U / \rho_0 (1 + \alpha_m \Delta T_{out})$

TABLE 6.5. Form of GPP based on Joule heat

Form parameter	Form variable
Mass of an object	$M_m$
Specific heat of material	$C_m$
Electric resistivity of material	$\rho$
Initial resistivity of material	$\rho_0$
Length of the object	$L$
Cross-sectional area of the object	$A_m$
Temperature coefficient of resistivity	$\alpha_m$

TABLE 6.6. Structure of GPP based on Joule heat

Input Ports	Output Ports
$I_{in}, U_{in}$	$H_{out}, T_{out}$

### 6.5.5 Thermal Expansion

The same approach of black-box modeling is used to model the effect of thermal expansion. Thermal expansion is a process where solid or liquid substances change volume as a consequence of undergoing changes in temperature. In engineering, the effect of thermal expansion is widely applied in various technical systems where a precise change of mechanical parameters is needed. For instance, it can be used to precisely control the displacement of a table in a microscope.

The black-box model of the effect of thermal expansion for changing the length of a material is depicted in Figure 6.6. In this model, an input heat flow  $H_{in}$  causes an elementary mechanical deformation in the crystal lattice of the material. Under macroscopic observations, these elementary deformations result in the change of length for the whole material.

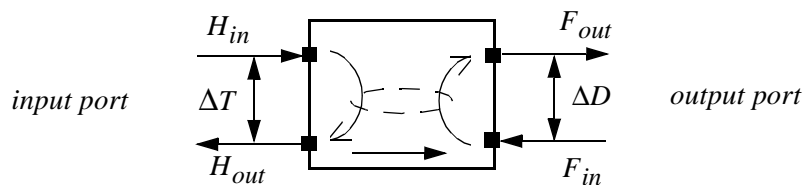


Figure 6.6: Black-box model of the effect of thermal expansion

In terms which relate thermal and mechanical domains, the behaviour of the effect can be expressed as:

$$R_{tm} = \langle \langle H_{in}, T_{in} \rangle, \langle D_{out}, F_{out} \rangle \rangle$$

where  $H_{in}$  and  $T_{in}$  are the input heat flow and temperature,  $D_{out}$  and  $F_{out}$  are the output displacement and mechanical force respectively.

The basic relation between linear size of a material and temperature is:

$$\Delta L = \int_{T_1}^{T_2} \alpha_m L_0 dT, \quad (6.9)$$

where  $\Delta L$  is the increase in length,  $L_0$  is the initial length of the material element,  $T$  is the temperature.  $T_1$  and  $T_2$  are the initial and final values of the temperature, and  $\alpha_m$  is the coefficient of thermal expansion for the material.

In terms of system theory, and following the law of energy conservation we can write the following relation for the behaviour of the effect of thermal expansion:

$$\frac{dD}{dT} = f\left(\frac{dT}{dt}, F\right), \quad (6.10)$$

or in exact form:

$$\Delta D = \int_{T_1}^{T_2} \alpha_m L_0 dT, \quad (6.11)$$

where  $\Delta D$  is the linear displacement of the element subjected to heating.

The overall model of the GPP presenting the effect of thermal expansion can be presented analogously to the GPP of Joule heat (section 6.6.4).

### 6.5.6 Example of Generic Design Concept

A generic design concept is synthesized on the basis of a causal chaining of several GPPs into a physical system whose overall behaviour achieves the function required on the basis of the energy resources given. Two GPPs may be connected if two conditions are satisfied:

The GDC includes at least two components: an energy source and a GPP that provides the required transformation of an energy or material parameter.

A topology for a Generic Design Concept can be synthesized which would be capable of performing the function “to control displacement” on the basis of the two effects described above: Joule heat and thermal expansion (Figure 6.7). The overall behaviour of the GDC is defined by the following relation:

$$R = \langle \langle U_{in}, I_{in} \rangle, \langle F_{out}, D_{out} \rangle \rangle$$

A mathematical expression comprising the overall GDC behaviour can be obtained by substituting the input variables in an equation (5.10) with equation for the output variables (5.6):

$$\Delta D = \alpha_m L_0 \int_{t_1}^{t_2} \frac{U_{in}^2}{R_m \cdot M_m \cdot C_m} dt \quad (6.12)$$

The form of the GDC includes form-related variables which belong to both GPP.

To control the change of the resulting mechanical displacement, a variable resistor can be introduced into the network. Therefore, an electrical resistance is a subfunction of mechanical displacement that controls, for instance, the length of the resistive material:  $R_m = f(D_{ext})$ . The energy flow  $D_{ext}$  is provided by an additional source of mechanical energy (figure 6.8).

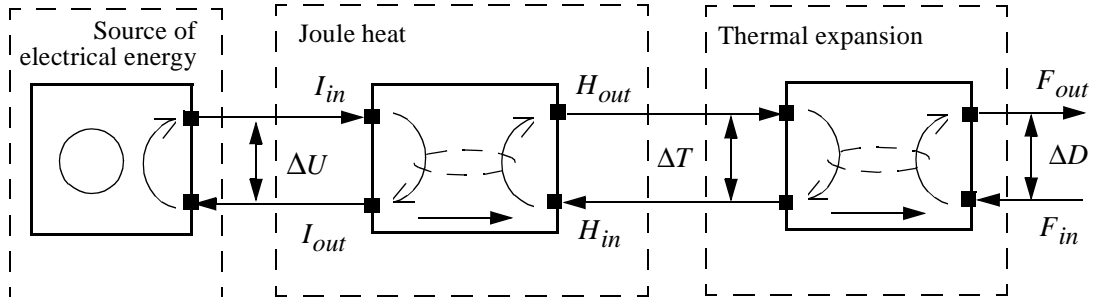


Figure 6.7: GDC of a system performing a mechanical displacement as a function of changing the voltage difference

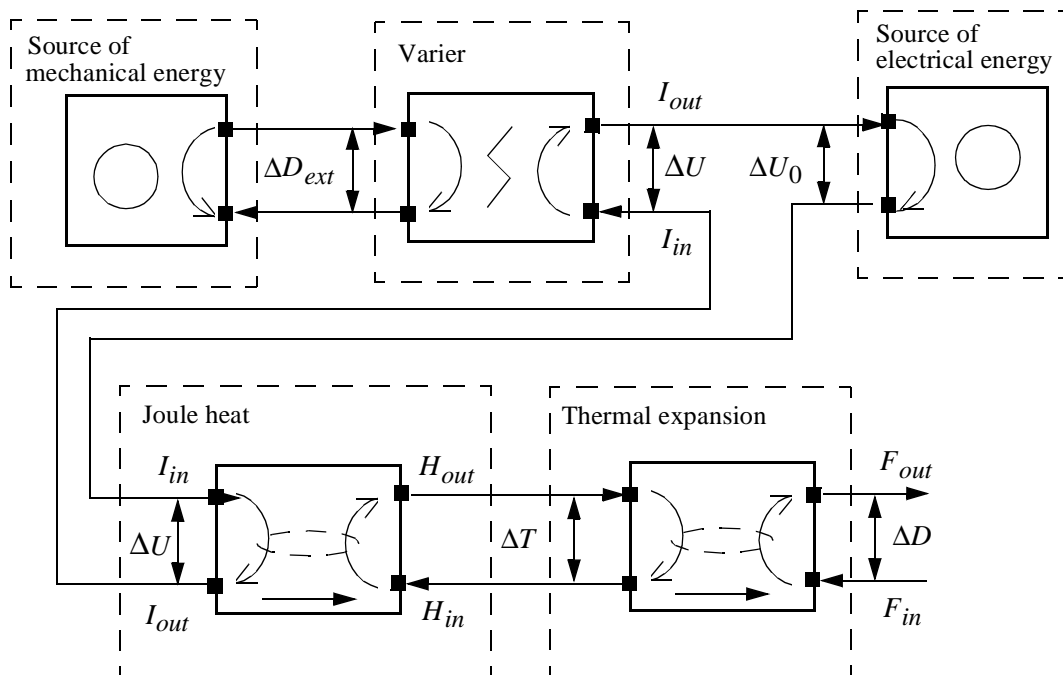


Figure 6.8: GDC of a system performing a mechanical displacement as a function of the change of the voltage difference via a variable electrical resistor (varier)

## 6.6. Summary and Discussion

We discussed the framework for modeling sharable physical knowledge in uniform way. To achieve uniformity in modeling we proposed to use an ontological approach for modeling domain knowledge. The INDES ontology contains two sets of concepts: concepts for modeling physical components and concepts for modeling relations between the physical components.

The introduction of a concept of a Generic Physical Principle as a basic building block for conceptual engineering design makes it possible to develop a domain-independent, sharable collection of physical phenomena. The reuse of the phenomena might be organized according to the multitude of physical functions each GPP is capable of performing.

In some sense, INDES can be viewed as extension of YMIR since it uses the same basic modeling approach. However, there are three additional aspects which are unique features of INDES:

1. YMIR regards specific design solutions as a source for generalization and modeling whereas INDES uses collections of physical effects known in natural sciences.
2. The degree of conceptualization in INDES is higher than in YMIR: physical effects are more generic in nature than models of specific designs.
3. INDES allows for cross-domain modeling whereas YMIR is limited to modeling only domain knowledge.

Among the principal features of INDES are:

1. The basic knowledge structure is introduced for modeling sharable physical knowledge -- Generic Physical Principle. It incorporates both energy and material aspects of physical principles that are needed to make instantiations of physical knowledge into design concepts.
2. The concept of a generic physical principle makes it possible to relate both material and energy parameters of generic physical systems.
3. An ontological approach for modeling design knowledge is applied to model different types of physical principles in a uniform way and to evaluate future designs against physical realisability at the earliest stages.
4. To organize and structure a knowledge base of GPPs the available TRIZ collections of physical effects and phenomena can be used.
5. A uniform way of modeling and representing GPPs makes it possible to combine GPPs into more complex sharable structures labelled Generic Design Concepts.

In the rest of the book we will discuss how the INDES ontological concepts might be incorporated into appropriate models of innovative design.

## Chapter 7. Design Conflict

### 7.1. Introduction

In the previous chapter, we presented a basic concept for modelling object knowledge for conceptual engineering design - Generic Physical Principle (GPP). A GPP deals with a collection of tuples and assumes that specific values of variables involved in the GPP behaviour are supported but not identified with precise intervals. Such genericity ensures that the GPP can be instantiated into a multitude of different design product descriptions. Our assumption behind the proposed framework is that we are able to model any design product and its behaviour in terms of its generic physical principles.

The INDES-based innovative design is possible when a repository of GPPs is available. The overall conceptual design process consists of two steps: first, matching a function required to deliver with a function which is part of GPP(s) behaviour and second, synthesising a new design concept by networking the relevant GPPs.

This design procedure addresses the synthesis of a new design product based on the utilization of a new physical principle. However, this type of design does not support another important design activity: the problem of improving existing products. This type of design problems occurs more frequently in engineering than designing completely new products. Due to this we may not neglect the type of engineering design which we labelled “innovative redesign” (see Chapter 5). By innovative redesign we mean a design process which results in solutions that can be regarded as innovative modifications of existing products and which may not be obtained by routine redesign.

To organize a reasoning process about problems of innovative redesign, we introduce a number of INDES strategic knowledge concepts. These are used for modelling innovative redesign problems as well as providing rules for knowledge-based innovative redesign.

This chapter presents a key concept of innovative redesign - *design conflict*. This concept is similar to Altshuller's concept of the contradiction discussed in Chapter 3 and is used to formulate and model innovative redesign problems. However, the INDES notion of the design conflict is claimed to be more accurate than Altshuller's notion of contradiction. We have therefore decided to use another term to avoid confusion with the original TRIZ concept of contradiction.

In this chapter, we discuss how the design conflict can be used in modelling innovative design. First, we explain the role of contradiction in decomposing innovative design problems. Second, we present a new concept for modelling a design product: a Design Prototype Model. Next, we show how design conflict can be defined in terms of behaviour of the Design Prototype Model. Finally, we give a formal definition of a design conflict and show how the design conflict can be represented in terms of formal language.

## 7.2. Role of conflicts in problem decomposition

In general, the engineering design can be regarded as a process consisting of two phases: analysis and synthesis. Analysis involves obtaining and structuring information on a problem, as well as dissecting it and investigating the characteristics of individual elements and the interrelations between them. Analysis requires recognition, definition, structuring and ordering. Synthesis involves the processing information obtained by forming links and connecting elements to produce totally new effects, bringing them together in an orderly summary. During the synthesis phase, the problem solving process gradually moves from a qualitative concept to a formal quantitative model. Synthesis can be regarded as a process of searching and finding plus composition and combination.

If the problem is complex and extensive, it can be decomposed into a set of manageable sub-problems. The decomposition process is not difficult to perform if the semantics of the problem are clearly stated and both the space of involved variables and the solution space are well-defined. In this case, the decomposition process can easily be described in formal language. However, the decomposition is only possible when a criterium, according to which the decomposition is performed is known.

In the case of innovative problems, the situation appears to be more difficult. So far, no formal representation of a decomposition procedure for innovative problems was possible due to a lack of such criteria and an ill-defined solution space. To solve this problem, the TRIZ concept of a contradiction is a means to decompose innovative design problems. To define what the space of possible solutions is, we will assume that innovative solutions are combinations of generic physical principles obtained on the basis of an existing set of generic physical principles.

As known from TRIZ, the same problem can be translated into a multitude of different contradictions. A choice of a contradiction to resolve is however made intuitively. For this rea-

son, many iterations are required to choose the “right” contradiction and come up with a feasible solution. In conclusion, we argue that a contradiction should only be constructed on the basis of analysis of *model* of a design product but not the product itself.

### 7.3. Design Product Model

To check against possible design conflicts that will interfere with the desired improvement of a design product, the latter should be formally modelled. We propose to perform such modelling at a conceptual level in terms of Generic Physical Principles. The use of abstraction in modelling is useful for two reasons: first, it enables focusing on certain aspects of a problem, second, it limits the number of variables involved thus simplifying the process of mathematical modelling. Abstraction has long been recognized as a powerful AI technique for general problem solving as well as for solving non-trivial design problems using reasoning by analogy (Bhatta & Goel [1994], Wolverton & Hayes-Roth [1995]).

To model existing design products in a uniform way recognizable by computer we introduce a new INDES concept labelled a *Design Product Model (DPM)*. Like a Generic Design Concept (GDC, Chapter 6), the Design Product Model is a network of GPPs. There is no difference in modelling both the Generic Design Concept and the Design Prototype Model apart from identifying the generic values and intervals of variables in the GDC and specific values and intervals in the DPM. The labelling also indicates a way in which a model is obtained: GDC represents a solution, and DPM represents a problem.

#### *Definition 7.1.*

**Design Product Model** is a conceptual model of a design product presented as a network of Generic Physical Principles which constitute the existing design product or its part.

As a conceptual model of a real design product, the DPM preserves semantics of some part of a system which constitutes the product. Due to a black-box approach, detailed aspects of system dynamics are neglected.

Assuming that it is possible to build a knowledge base of GPPs which stores all known physical effects and phenomena, a new DPM can be composed of predefined building blocks by mapping between functions defined for the design product components and the available generic physical principles stored in the knowledge base. The advantage of translating a real design product onto a network of predefined generic physical principles is that the resulting DPM will incorporate consistent and correct information about the generic behaviour of each GPP the DPM consists of.



## 7.4. Theory of Design Conflict

One of the basic INDES assumptions is that a new requirement, no matter in what form it is formulated, can always be expressed in terms of a system's behaviour particularly in terms of functional specifications. The task of translation between various forms of requirement specifications is outside the scope of the book but we refer to the TRIZ-based Functional Analysis technique (Litvin & Guerassimov [1991]). This technique was tested on a large number of inventive problems and has proven successful for the exact identification of the functional context of a system's components.

Regarding innovative design as the task of mapping between functions and physical principles, two types of problems can be identified:

1. To change the value of an existing form or system variable (a particular case is to eliminate the variable, that is, to make its value equal to zero).
2. To Introduce a new variable with the certain value or the interval of the values.

In both cases, in order to meet the new requirement, a selected design product needs to be redesigned. A redesign process consists of performing a modification to the product's structure and behaviour so that the modified product will meet the requirements.

Our contribution to the definition of innovative redesign (first proposed in Goel [1989]) is that innovative design can be distinguished from routine redesign according to the presence of design conflict. We argue that it is possible to make a conclusion on which general problem solving method should be chosen by verifying if a problem contains a conflict. According to this, a procedure of routine redesign may be applied if no conflict is identified while innovative redesign is required when the problem involves a design conflict.

INDES regards design conflict as a concept which establishes a specific type of relation between two particular variables involved into the DPM's behaviour. This relation defines what variables will have conflicting values if the new requirement is met by the routine redesign of the product. The design conflict can be specified when two behaviours are compared:

1. The DPM's behaviour which corresponds the specifications but does not satisfy the designer.
2. A behaviour the DPM will possess after a routine redesign would have been performed to satisfy the designer's demand.

These two DPM states can be evaluated by the correctness of the overall system's behaviour. Two types of behaviour can be distinguished:

- *Correct system behaviour:* At any moment in time, the values of all system parameters belong to the overall space of specified parameters.

- *Incorrect system behaviour*: at a particular moment in time, the value of some system parameter do not belong to the interval of specified values.

If routine redesign results in correct system behaviour, and all constraints are satisfied, then no innovative redesign is required.

The conflict between two parameters is defined on the basis of analysing what parameter would cause an incorrect behaviour of a system as if routine redesign of the system was performed.

**Definition 7.2.**

**Design conflict** is a situation when the required change of the physical variable is not possible due to a violation of the overall system behaviour.

The design conflict usually occurs under the following circumstances:

1. A change in value of some physical parameter is required;
2. The required change might be achieved by changing other physical parameter(s);
3. A change of other physical parameter(s) is not allowed.

To represent a given problem as the design conflict, one should first identify what physical parameter is to be changed. This parameter should be identified with a system or form variable in a mathematical relation describing a part of the system's behaviour. The next step is to analyse if change to any of the other variables involved into the relation provides the required change of the parameter. If it does, but this is not allowed, such situation can be identified as the design conflict.

**Example 7.1.**

Suppose, we need to redesign a coffee maker to provide faster water heating. A Design Prototype Model of the coffee maker consisting of two GPPs is presented in Figure 7.1. The technical function of the heating element can be formulated as "to heat water", and it can be identified with the physical function of an energy-transforming element which generates Joule heat. The physical function can be expressed as  $\Delta T = I^2 R \Delta t$  (temperature difference  $\Delta T$  related to an electrical current  $I$  and electrical resistance  $R$  for the time  $\Delta t$ ).

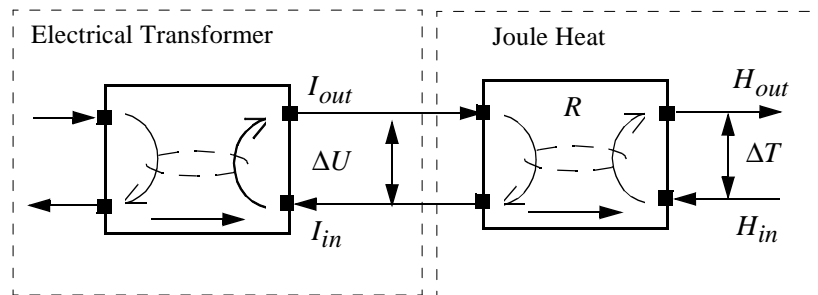


Figure 7.1: A DPM of a coffee maker

A new requirement is to reduce the time needed for heating ( $\Delta t$ ). Since it is only allowed to express the requirement in terms of system or form variables, the requirement can be reformulated so as to provide a higher value of temperature difference  $\Delta T$  for the fixed time  $\Delta t$ . Analysing the equation, it is possible to come up with two suggestions:

1. To increase the value of the electrical current  $I$ . However, since the electrical current is proportional to electrical voltage such a coffee maker will consume more energy. Under the condition that no more energy consumption is allowed, the conflict arises between the required value of the temperature difference and the existing value of the electrical current which it is not possible to increase.
2. To make the heating element of a conductive material which has a higher value of electrical resistance  $R$ . In this case, a conflict arises under the condition that the use of another material is forbidden.

In both situations, we can formulate the problem as a conflict: given a new requirement formulated in terms of a new value of a variable, the new value may not be obtained by directly changing the values of other variables related to the required one.  $\square$

This example shows that innovative redesign can be distinguished from routine redesign according to the following criteria: a direct change of values of variables involved into the system behaviour is not allowed according to the constraints given.

### 7.4.1 Conflict elimination by innovative redesign

According to the classification of inventive solutions presented in section 3.2, the innovative redesign results in the solutions of 3th and 4th levels if:

1. A design conflict which interferes with further system evolution is eliminated in full.

2. A solution resulting from the design conflict elimination utilizes a physical principle that has never been used in the system before.

A comparative study of the TRIZ problem solving techniques led us to the conclusion that all of them can be subdivided into three general methods for eliminating conflicts applicable to a conflict-based model of innovative design problem as discussed in the previous chapter:

- *Spatial elimination of conflicts.* This method states that potentially conflicting parameters should be separated in space thus eliminating the relation between conflicting parameters.
- *Temporal elimination of conflicts.* This method means that potentially conflicting actions should be performed within different time intervals.
- *Elimination of conflicts by changing physical contents.* To avoid the conflict, new physical components are introduced into the system or certain components which cause the conflict are replaced with those that do not cause the conflict.

All three methods can not be performed without a conceptual change to the physical contents of a design prototype. As a conclusion, in order to build a computational model of innovative redesign, we first need to formulate principles for changing the physical contents in terms of previously introduced INDES concepts.

In the next chapter, we will formulate a general methodology for innovative redesign utilizing a framework of model-based problem solving.

## 7.5. Implementation

To give an example of how a design conflict can be used in a practical way, we can define two rules of conflict analysis in terms of first-order logics.

First, we formally define notions of correct and incorrect system behaviour:

- *Correct system behaviour:* At any moment in time, the space  $A$  of actual values of all system parameters belong to the overall space of permitted (in accordance with design specifications) parameters  $S$ , that is,  $A \subseteq S$ .
- *Incorrect system behaviour:* at a particular moment in time, the value  $P_i^A$  of some system parameter  $P_i$  does not belong to the interval of specified values  $P_i^S$  that is,  $P_i^A \not\subseteq P_i^S$ .

If we express a space of interrelated parameters within a system as  $P^I$  ( a subset of  $P^S$ ), and establish a goal of problem solving as expanding a range of possible values for some parameter  $P_i$ , then a formal representation of a relation defining the conflicting value is:

$$\begin{aligned}
CONFLICT\_VALUE(P_i^A, P_i, P^E, t) \leftarrow \\
& P_i^A \in P^E \\
& \text{and } P^E \not\subset S \\
& \text{and } \exists P_j \rightarrow P_j^A \notin P^E \\
& \text{and } P_i, P_j \in P^I,
\end{aligned}$$

where  $P^E$  is the required extension of existing space of values  $P^S$ ;  $P_j$  is the variable functionally related to  $P_i$ .

We can define the relation *CONFLICT* between two variables  $P_i$  and  $P_j$  if there is a functional relation between them  $RELATED(P_i, P_j) \leftarrow P_i, P_j \in P^I$ :

$$\begin{aligned}
CONFLICT(P_i, P_j, t) \leftarrow \\
& RELATED(P_i, P_j) \\
& \text{and } P_i^A \in P^E \\
& \text{and } P_j^A \notin P^E \\
& \text{and } P_j^A \in S \\
& \text{and } P^E \not\subset S.
\end{aligned}$$

Although this relation only defines a conflict between two particular variables, similar relations can be defined for  $P_i$  and any other variable of the set  $P^I$  if the value of the latter causes incorrect behaviour of a system when changed.

A possible method for preventing a system from appearance of the conflict consists of finding any such parameter which belongs to  $P^I$  so that it would be allowed to alter its value without producing an undesired result. As a consequence, the value of  $P_i$  will also alter. However, when none of values of the related variables may be altered to provide the required value of  $P_i$  then a conceptual change of the existing physical structure of the system is necessary.

### 7.5.1 Benefits from the INDES definition of design conflict

Apart from a more precise definition of the design conflict, our approach solves one of the most difficult problems of TRIZ: what components to include in a problem model. As seen in

the example of problem solving with ARIZ-85B presented in section 3.9, the concept of an operation zone is used to determine the borders of a place where the conflict arises and enlist those system's components which are present in this place. However in ARIZ-85B, the operation zone is constructed in an ambiguous and intuitive way. Using INDES, it is always possible to identify what variables are involved in the conflict, and, therefore, what components should be included in the design prototype model. Thus, the approach to defining conflicts presented above provides a better localization of a conflict.

## 7.6. Summary

In this chapter, we introduced and discussed the key INDES concept for representing innovative design problems: the design conflict. A presence of the design conflict is a feature that distinguishes between routine redesign and innovative redesign tasks.

Due to a lack of formal approach to defining a contradiction, Altshuller's methodology for innovative design does not seem to be computable. To eliminate this disadvantage, we proposed a formal model of design conflict.

To define a design conflict in real situations, the designer has to build a model of a design product to be improved in terms of INDES concepts for object knowledge and then to investigate what variables will have conflicting values.

In the next chapter, we will discuss a model of innovative redesign based on the use of the concept of design conflict and introduce principles for innovative redesign.



## Chapter 8. *Innovative Redesign*

*"Problems cannot be solved by thinking within the framework within which the problems were created"*

*Albert Einstein*

### **8.1. Introduction**

In the previous chapter, we introduced and discussed two INDES concepts for modeling innovative design problems: a Design Prototype Model and a Design Conflict. We also have shown in which way they can be used to formulate innovative design problems.

In this chapter, we present INDES-based model of knowledge-intensive innovative redesign which organizes the use of physical knowledge for solving those inventive problems that contain design conflicts.

First, we discuss why it is important to eliminate design conflicts from the point of view of the technology evolution. Second, we discuss the role of physical knowledge in this process.

Next, we introduce two axioms of innovative redesign which we will use to build a formal basis for the modeling innovative redesign. A number of domain-independent *innovative redesign principles* drawn from the axioms are also proposed. These principles are designed to solve particular design problems that result in identical problem models.

Then, we present a structure of INDES-based innovative redesign.

In the end of the chapter, we show how INDES-based innovative redesign can be used in practise by a case study.



## 8.2. Role of physical knowledge for design conflict elimination

The first artificial tools aimed at work simplification were stones controlled by the physical force of a man's hand. A growth of demands and requirements has resulted in further complication of the working tools: new purposes were thought of, disadvantages had to be removed. The physical force of a hand was replaced with a force of a horse, then water flow, steam, nuclear energy, and so forth. This resulted in the creation of what have been known today as design products: artificially created physical systems supposed to deliver specific functions.

In every situation, however, designers are looking for ways to improve the existing products. Studies of patent collections indicated that most inventions result from improvement of already known products delivering the same function. In most cases, such the improvement is often achieved by complicating the existing products.

If we define the degree of complexity of a design product as a relative factor involving of components and interactions between the objects constituting the product, then the trend of "complexity growth" can be observed in engineering. It states that attempts to achieve a better performance of any product and satisfy the growing number of design constraints leads to the growth of the number of objects, and, as a consequence, interactions between the objects (Devoino [1993]).

In summary, every new invention provides improvement of performance and productivity, but at the expenses of growing complexity and costs. Such a phase of the technology evolution is known as a *phase of system complexity growth*<sup>1</sup>. The better overall performance is required, the larger number of material-energy transformations in the product is observed.

However, evolution of any product without replacing a physical principle the product is based upon will finally face a barrier put by the physical limits of the principle. To evolve further, the physical limits must be eliminated. This is only possible by replacing a physical principle the product is based upon with a new physical principle which will fulfil the same function with a better performance.

### **Example 8.1.**

Systems for sound recording have evolved along expanding the frequency range. Each time, a new physical principle was used to expand the range: 1) a mechanical principle (wax roll) 2) a magnetic principle (metal tape), 3) an optical principle (digital recording - audio CD or metal tape).

The next step of evolution known as *the complexity reduction*<sup>2</sup> phase sets off when a new physical principle is proposed to overcome a barrier established by the nature of previous

---

1. The phase of system complexity growth in INDES can be identified with the trend of an increase the number of substance-field interactions in TRIZ (Chapter 3). However, as shown further in the text, we needed to change the label to make the whole picture of technology evolution more consistent.

physical principle. Such a conceptual change can be identified with resolving a fundamental design conflict which is not possible to eliminate by available domain-specific methods. Some essential component of a system (for instance, a mechanical cutter) is replaced with a new component utilizing a different physical principle (for instance, a laser beam).

However, during the evolution, a new product tends to develop its subsystems and find more and more applications. As a result, another phase of the complexity growth starts. New demands and requirements make the product develop the number of interactions again. (Figure 8.1).

From the INDES point of view, this trend can be formulated in terms of material-energy interactions. Therefore, evolution of any artificial system delivering a certain function passes two phases:

- **complexity growth:** any system tends to increase a number of material-energy interactions during its evolution;
- **complexity reduction:** a number of interacting components needed to fulfil a function is replaced with a fewer number of interactions due to the utilization of a new physical principle.

Such interpretation of the system evolution forms a theoretical foundation for our model of innovative redesign.

On the basis of our interpretation of technology evolution process, we can give more precise definition of innovative redesign:

**Definition 8.1.**

**Innovative Redesign** is a specific type of engineering design which involves elimination of design conflicts by a conceptual change to a physical structure of an existing design product.

Respectively, two types of innovative redesign can be distinguished:

1. Innovative Redesign by complicating an existing system: a physical principle behind its function is preserved whereas a number of material-energy interactions grows.
2. Innovative Redesign by utilizing a new physical principle behind the main function.

---

2. The trend of complexity reduction is similar to TRIZ trend of transition to microlevel, although such comparison is rather rough since TRIZ does not specify how to provide such transition in terms of interacting material objects and energy flows.

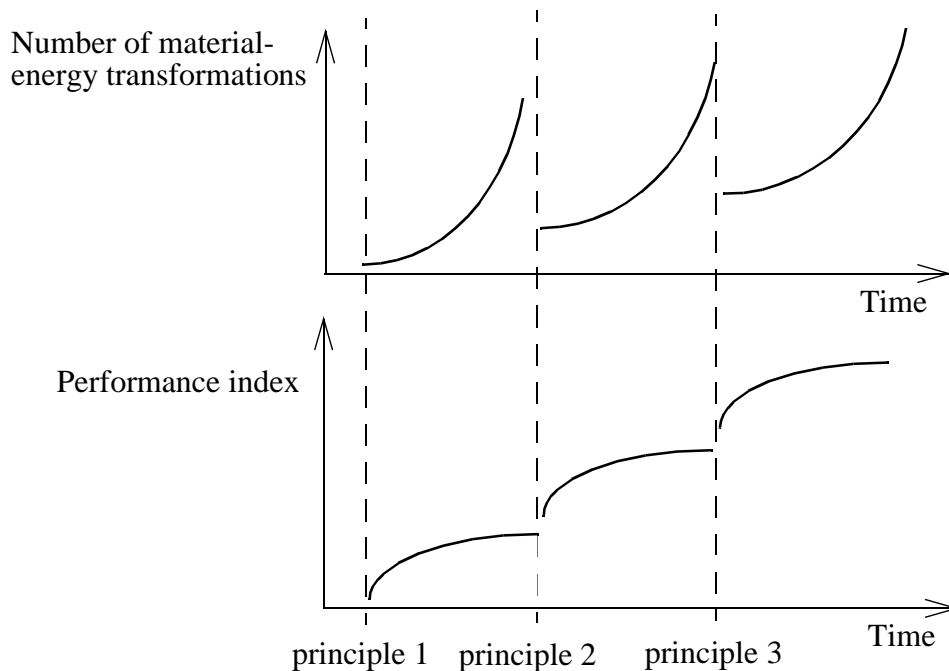


Figure 8.1: Correlation between system's performance growth and phases of complexity growth and complexity reduction.

### 8.3. Innovative Redesign Axioms

Studying the Inventive Standards (section 3.7), we made the conclusion that there were two general methods for changing the physical structure of a design prototype: achieving a change in a substance by introducing a field acting upon the substance and, respectively, achieving a change in a field by introducing a new substance into the system.

Therefore, to eliminate a potential design conflict means to change a part of a given design prototype in such a way that no conflict arises and the overall behaviour of a system is performed in accordance with design specifications.

However, as we discussed earlier, TRIZ definitions of a substance and a field are ambiguous. As a consequence, Inventive Standards can hardly be presented in a formal way. To eliminate this disadvantage, a formal approach to redefining Inventive Standards is needed.

Nevertheless, the two methods for changing the physical contents of a design prototype are very important to understand the nature of innovative redesign. Any design conflict (if it can be modeled as a pair of two related physical parameters) can be eliminated using one of the

two methods. Therefore, we decided to incorporate both methods into our model of innovative redesign. Since we aim at building a formal framework of innovative redesign, we used these two general methods as a basis to formulate two axioms of innovative redesign:

1. **Innovative Redesign Axiom 1.** The value of a form variable representing some material parameter of a system can be altered by introducing a new energy flow into the system that is capable of altering the value of the form variable.
2. **Innovative Redesign Axiom 2.** The value of a system variable identified with some energy parameter of a system can be altered by introducing a new material object into the system that is capable of altering the value of the system variable.

As follows from the physical validity of these two axioms, innovative redesign is applicable to any design problem that involves material-energy transformations and can be formulated as a design conflict.

## 8.4. Innovative Redesign Principles

In this section, we introduce the concept of *Innovative Redesign Principle*. We also distinguish two particular methods for innovative redesign: Expansion Method and Replacement Method. Then, we present several specific Innovative Redesign Principles drawn from the Innovative Redesign Axioms.

### 8.4.1 Definition of the Innovative Redesign Principle

According to the two phases of “complexity growth-reduction” described above, we distinguish between two particular methods of innovative redesign:

1. **Expansion Method.** To eliminate a design conflict is possible by *expanding* the space of physical variable(s) used in a Design Prototype Model (DPM). A new variable(s) can be introduced into the DPM by adding new physical principles to the existing ones the DPM is based upon. The newly introduced variables should provide a necessary change in the required value of the variable and prevent the DPM’s behaviour from being violated.
2. **Replacement Method.** To eliminate a design conflict, a physical principle which is responsible for causing a design conflict can be replaced with another one physical principle. In this situation, a space of system variables might remain the same, but a part of the space of form variables is replaced with another parameters.

Both methods are possible to use through the *Innovative Redesign Principle* (IRP) which is the essential concept of the INDES ontology. An IRP belongs to the category of strategic knowledge since it defines how to process object knowledge.

An IRP is a rule which indicates how a physical structure of a design prototype should to be transformed to eliminate a design conflict. It consists of two parts. The right part specifies types of problem and Design Prototype Model; the left part specifies a Generic Design Concept which is an expanded (or transformed) Design Prototype Model.

**Definition 8.2.**

**Innovative Redesign Principle** is a problem solving rule which specifies how a physical structure of a design prototype model has to be transformed on the basis of the innovative redesign axioms.

A primary idea behind IRPs which belong to the category of the Expansion Method is that a new system or form variable is introduced into a DPM in order to establish a new functional relation between the newly introduced variable and a variable the value of which should be altered. A necessary condition is that this new relation should not cause new conflicts. As a result of applying the Expansion Method, the Generic Design Concept is generated. It consists of the same Design Prototype Model of a product and a new GPP connected to the DPM. It is assumed that the behaviour of the obtained generic concept will satisfy given functional specifications without causing conflicts between newly introduced variables and other variables in the DPM.

According to the Expansion Method, a new physical relation is established by introducing a new component(s) into a DPM. It can be a material component as well as an energy flow.<sup>1</sup> Each IRP proposes a pattern of transformation of a design prototype model into a generic design concept. Similarly an Inventive Standard, IRP does not define exactly what material components and energy flows are to be introduced. In any IRP, both DPM and GDC consist of blank components.

In the case of Replacement Method, the GDC is formed by a blank GPP which replaces either certain physical principles the DPM's part is based upon or the whole DPM.

As follows from the TRIZ trend of ideality growth (section 3.5.1), the use of the Replacement Method is always preferable. In most situations, the Replacement Method helps to create a new product that will provide a better performance of the required function. In reality, however, this does not always result in economically feasible products. In such cases, Expansion Method should be used to improve the existing product by identifying and eliminating design conflicts.

The INDES collection of Innovative Redesign Principles was defined on the basis of a current set of Inventive Standards which deal with substance-field transformations. Those Inventive Standards which contain recommendations regardless of specific product models, were omitted.

---

1. In most cases, a new energy flow can not be introduced without specifying a material object generating or carrying the flow. However, in some cases, the material object is not needed since it may not be regarded as a part of the design product (for instance when using gravitational forces of Earth).

The collection of IRP consists of rules structured respectively a type of a problem and design constraints which define permission on introduction of additional energy flows and material components.

In general, an IRP is a problem solving method which defines what modification of a physical structure of a design product modeled as a DPM should be made in terms of energy-material relations. Given a DPM and a description of a design conflict formulated with respect to the DPM behaviour, the IRP defines two product models: a source Design Product Model which contains the design conflict and a target GDC which does not contain the conflict.

In the next sections, we will give several examples of Innovative Redesign Principles based on the Expansion Method written in form of production rules.

### 8.4.2 IRP applicable to non-conflicting situations

In certain situations, engineering design might be not concerned with conflict situations. This happens when a source DPM is incomplete and consists of a single material component or an energy flow. Usually such problems address to the problems of new design synthesis. For instance, it might be a problem of processing a material when no design of a processing machine is yet available as well as a problem of preventing something from heating. Therefore, no conflict might be identified in such situations due to missing data. Such the problems appear to be difficult to solve due to a lack of knowledge of what physical principle to use to deliver the required function. In such situations, application of an IRP based on the direct use of Innovative Redesign Axioms should be considered first.

To build a target GDC for changing the value of form variable, IRP-1 can be used (Figure 8.2):

IREP-1:

IF it is required to alter the value of the form variable  $M_1$   
and no conflict can be identified

THEN the energy flow  $E_x$  should be applied to the material component  
so that the required value of the form variable  $M_1$  is obtained.

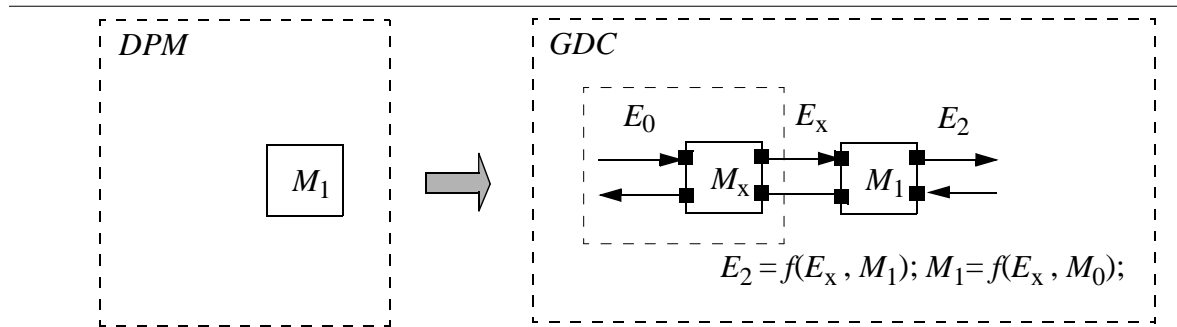


Figure 8.2: IREP-1

Respectively, IRP-2 is used to alter the value of a system variable (Figure 8.3).

**IRP-2:**

IF it is required to alter the value of the system variable  $E_1$   
and no conflict can be identified

THEN the energy flow should be connected to the material component  $M_0$   
which is capable of transforming the energy flow thus achieving  
the required value of the system variable  $E_1$ .

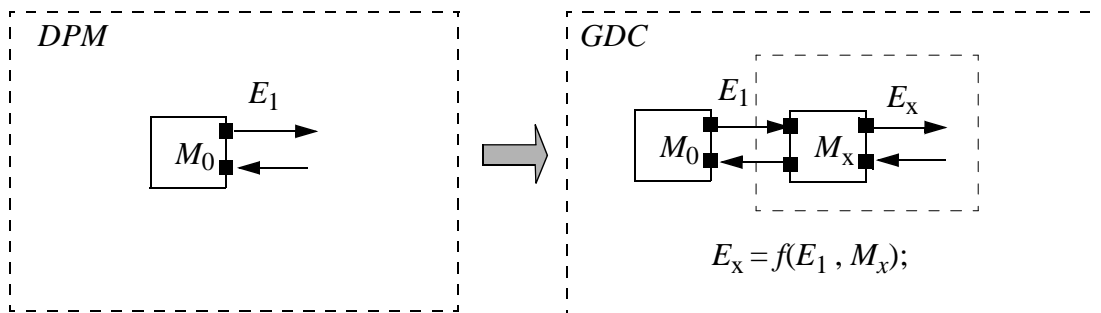
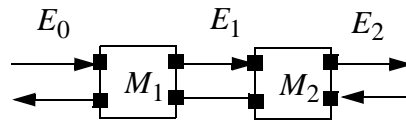


Figure 8.3: IREP-2

### 8.4.3 IRP for conflict elimination

Another part of the IRP collection is applicable to those problems that can be modeled in terms of a design conflict. A method for the design conflict elimination depends on what constraints are defined with respect to permitted modifications of a DPM. A diagram of a DPM that contains a design conflict  $CONFLICT(E_1, E_2)$  is shown in Fig. 8.4.



$$E_2 = f(E_1, M_2); CONFLICT(E_1, E_2)$$

Figure 8.4: A DPM with conflict

The rule IRP-1-1 is applicable in situations when there is the need to alter the value of a system variable, but the values of neither form nor system variables related to it are not allowed to alter. Available constraints do not limit introduction of other components between the two (Figure 8.5).

#### **IRP-1-1:**

IF it is required to alter the value of the system variable  $E_2$   
 and this can not be done by changing neither  $M_2$  nor  $E_1$   
 and there is a conflict relation  $CONFLICT(E_1, E_2)$ ,  
 THEN the material component  $M_x$  should be introduced to provide  
 the required value of the system variable  $E_1$ .



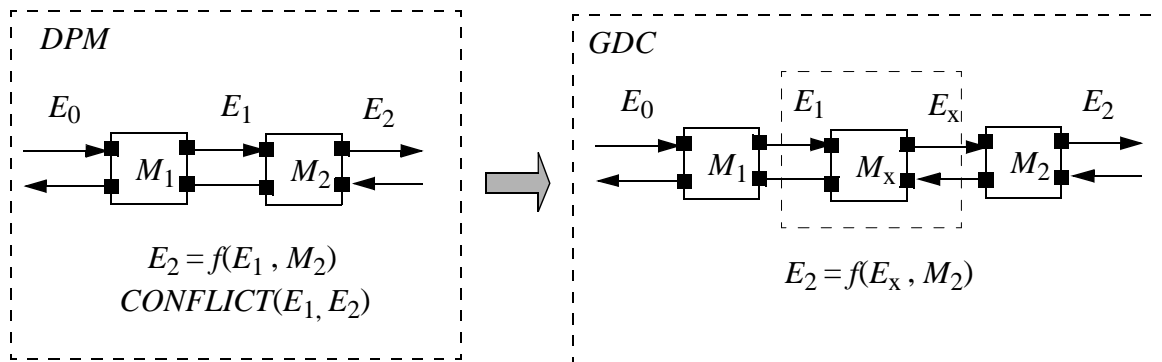


Figure 8.5: IREP-1-1

A particular constraint might be applied to a problem: a specification of how the value of a system variable should be altered. In the case of a decrease, a new resistive component should be introduced into a system (IRP-1-2, Figure 8.6). If an increase is required, a component which amplifies the current value of the system variable related to the needed one is introduced (IRP-1-3, Figure 8.7). Both IRP should be used in situations when no limitations on introduction of new components between the two are specified.

**IRP-1-2: (A particular case of IRP-1)**

IF it is required to decrease the value of the system variable  $E_2$   
and this can not be done by changing neither  $M_2$  nor  $E_1$   
and there is the conflict relation  $CONFLICT(E_1, E_2)$ ,

THEN the material component  $M_x$  (resistor) should be introduced  
to provide the required decrease of the value of the system  
variable  $E_1$ .

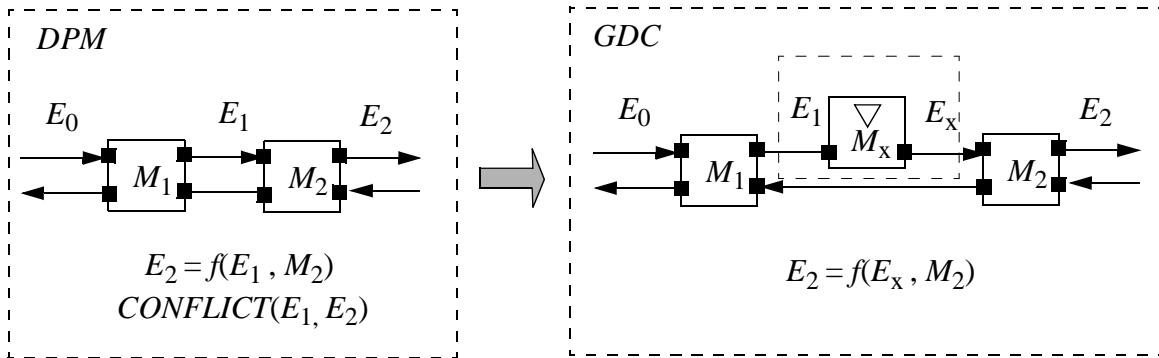


Figure 8.6: IRP-1-2

**IRP-1-3: (A particular case of IREP-1)**

If it is required to increase the value of the system variable  $E_2$  and this can not be done by changing neither  $M_2$  nor  $E_1$  and there is the conflict relation  $CONFLICT(E_1, E_2)$ ,

THEN the additional material component  $M_x$  (amplifier) should be introduced to provide the required increase of the system variable  $E_1$ .

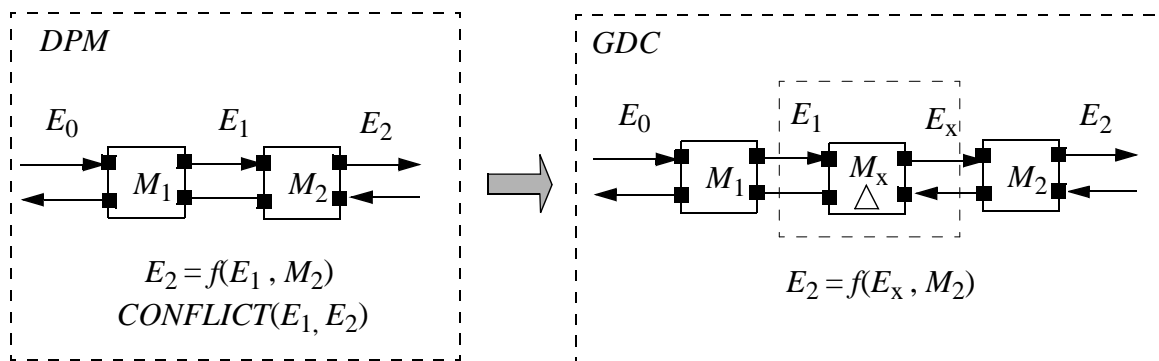


Figure 8.7: IRP-1-3

In situations, when no introduction of intermediate components is allowed, IRP-1-4 (Figure 7.8) and IRP-1-5 (Figure 8.9) should be used.

**IRP-1-4:**

IF it is required to decrease or increase the value of the system variable  $E_2$  and no component is allowed to be introduced between  $E_2$  and  $M_1$  and there is the conflict relation  $CONFLICT(E_1, E_2)$

THEN the new energy flow should be applied to the component  $M_1$  and effect of such combination should be provided by such physical phenomenon that the value of the system variable  $E_2$  changes.

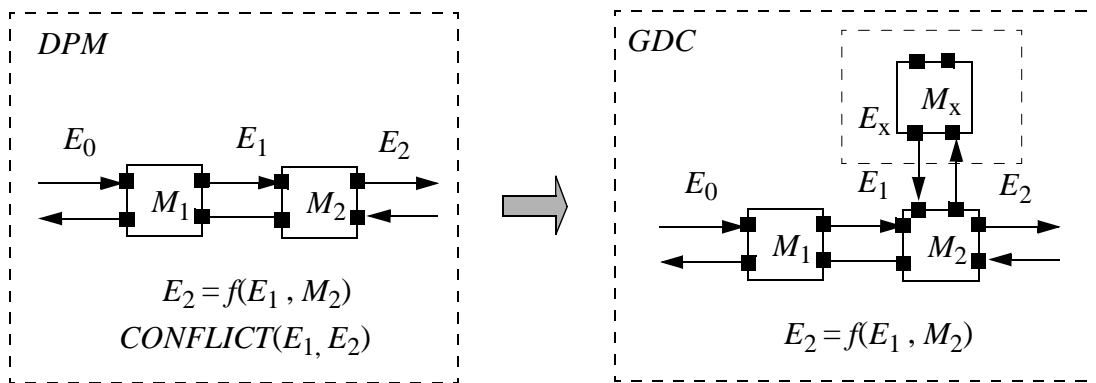


Figure 8.8: IRP-1-4

**IRP-1-5:**

IF it is required to decrease or increase the value of the system variable  $E_2$  and there is the a conflict relation  $CONFLICT(E_1, E_2)$  and none of previous cases is allowed

THEN connect the component  $M_1$  to the new component  $M_x$  producing the energy flow to change the value of  $M_1$ . This will results in change to the value of  $E_1$  and, as a consequence,  $E_2$ .

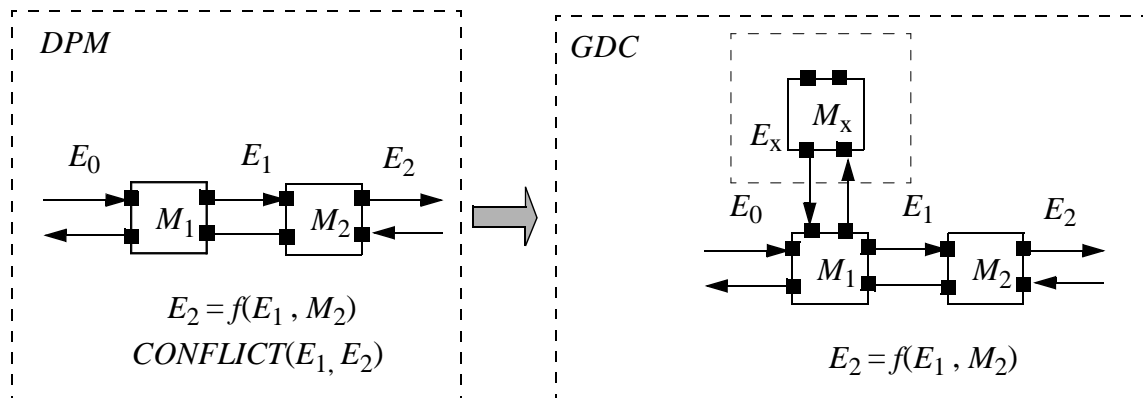


Figure 8.9: IRP-1-5

#### 8.4.4 Mapping between function and GPP

To solve a problem with IRP means to identify what GPP can be networked with existing design prototype model. This can be done through a mapping between a function defined for GPP.

The mapping involves finding a correspondence between a given description of the function and a physical structure which would be capable of delivering the function. IRP proposes how to introduce a new component, however does not specify what component should be introduced. The task is:

1. To identify a function that has to be fulfilled by a component.
2. To identify constraints.
3. To map the function identified onto a GPP(s) of available GPP knowledge base.

A procedure for mapping is outside the scope of the thesis. However, we should note that it can be solved

#### 8.4.5 Applicability of Innovative Redesign Principles

One can give numerous examples illustrating how the principles of innovative redesign presented above are widely used in engineering design. However, they were used intuitively by designers since no explicit expression of the principles has ever existed apart from Inventive Standards. As a consequence, no formal definition of such principles has been proposed so far.

A lack of a general modeling framework that would enable a designer to represent various types of design products in uniform way made it nearly impossible to formulate a reasoning strategy for a computational, knowledge-based innovative design.

Although the use of some pure AI methods (such as qualitative or case-based reasoning, for instance) can help with modeling and solving problems within well-defined solution space, we would still be unable to reason about innovative problems without appropriate problem-solving method. For this reason, the efficiency of pure AI methods which are not based on design principles would be very low. By defining axioms and drawing a number of design principles from them, we thereby introduced a strategy for solving innovative design problems. Therefore, using INDES strategic knowledge it is possible to build a knowledge-based system which will incorporate basic principles of innovative design.

## **8.5. General model of innovative redesign**

### **8.5.1 Model of innovative redesign**

Before performing the process of innovative redesign, a number of decisions has to be made by a designer. Among them are assumptions that limit the borders of a system to be modeled and assumptions that define the degrees of conceptualization and generalization of a given design prototype. During selecting components that has to be included into DPM, a special attention has to be paid to those objects and energy flows from the outer environment that interact or may interact with the system and thus might cause the design conflicts.

After making these decisions, the overall process of innovative redesign consists of the following steps:

#### **1. Analysis Part:**

- 1.1. Building a model of part of a selected design product in terms of DPM by identifying objects of a design prototype with appropriate GPPs from a GPP knowledge base;
- 1.2. Decomposition of the overall functions performed by GPPs into subfunctions;
- 1.3. Identification of what variable(s) should alter its value to meet the requirement given represented as a new value of the variable;
- 1.4. Identification of a pair of conflicting variables through analysis of the DPM;
- 1.5. Selection of a principle for innovative redesign according to the limitations on the new components use defined for a problem;

- 1.6. Identification of a blank GDC according to the selected innovative redesign principle;
- 1.7. Identification of a function that should be provided by a GPP.

## **2. Synthesis Part:**

- 2.1. Mapping the function identified for a GPP that can eliminate the conflict onto the available GPP(s);
- 2.2. Checking against the boundary conditions and integrity constraints to filtering out those specific GPP(s) that may not be incorporated into the DPM.
- 2.3. Building a GDC by instantiating the values of variables of introduced GPP(s).

As seen, our model of innovative redesign does not propose a means for the evaluation of obtained solutions. Assuming that the design process might result in a number of alternative GDCs, each of which meets the given requirement and does not contain conflicting values of parameters, their applicability should be verified against more specific limitations manually (for reasons see section 8.7). Our task was to ensure that the obtained solutions presented in form of GDCs are physically realisable.

## **8.6. Case study: Application of IRP**

To explain how the principles for structure extension can be used in practise, let us give an example.

A part of a system for outdoor water supply consists of two material components: a water flow and a steel pipe. During warm seasons, the value of water pressure is not enough to damage to the walls.

A problem occurs during cold seasons when a phase transition of water into ice takes place. If the pipe is not thermoprotected then water freezes and the value of pressure of expanding ice reaches a maximal allowed value of a mechanical stress limit for material of which the pipe is made thus causing growth of internal mechanical stresses inside the pipe wall. Elastic deformation of the wall turns out to be irreversible lattice expansion which results in slip of whole lattice areas occurring along the direction of maximum shear stresses.

According to a general procedure of innovative redesign, the system consisting of water flow and the pipe can be modeled as two interacting physical phenomena: phenomena of phase transition and phenomena of material deformation. As a conclusion, the problem model can be represented as a DPM consisting of two connected GPPs. The phenomena of phase transi-

tion results in transformation of initial energy into pressure which, in turn, exerts a force that produces the undesired deformation of the pipe.

After analysis of the situation, a conclusion can be made that the particular physical function which is responsible for causing incorrect behaviour is  $\Delta D = f(M_p, \Delta D_i, t)$ , where  $\Delta D$  relative lattice displacement caused by growing internal forces of mechanical stress inside the pipe material and  $M_p$  is a material parameter defining relation between input and output energy flows, that is, mechanical strength of the pipe. A DPM of part of the system is shown in Figure 8.10.

A conflict applicable to this situation can be formulated by introducing additional constraints: to change the ice pressure  $P_{out}$  or to replace a pipe material with another one which would have a higher value of mechanical stress limit  $M_p$  is not allowed.

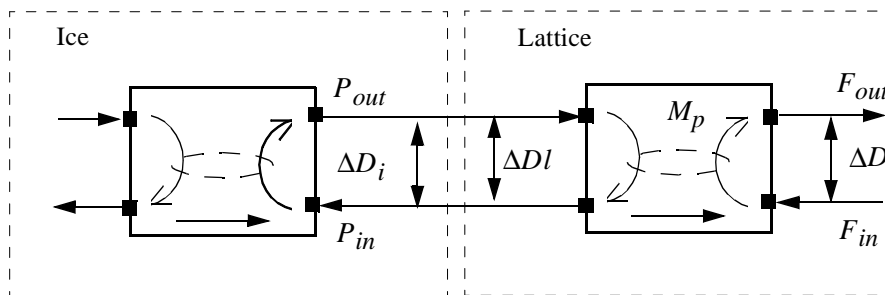


Figure 8.10: A problem of outdoor water supply

In terms of qualitative dynamics, the goal function which should be provided after problem solving can be expressed as  $\partial D/\partial t = 0$  while in the current situation is  $\partial D/\partial t > 0$ . This means that  $\Delta D_i$  should be equal to zero (or close to zero) and then it is an ideal situation which means that no lattice displacement inside the pipe material occurs. Within the borders of the DPM given, this is not possible.

According to IREP-1-2, a new physical variable should be introduced which will provide a condition  $\partial D_l/\partial t \ll \partial D_x/\partial t$ . In the given DPM,  $\Delta D_l = \Delta D_i$ . Such new variable  $\Delta D_x$  must provide a boundary conditions  $\Delta D_l + \Delta D_x = \Delta D_i$  at a given moment in time.

According to Innovative Redesign Axiom 1, once it is needed to alter the value of a system variable, one has to introduce a new material component which will provide necessary energy transformation thus altering the required value. Using a recommendation given in IREP-1-2, we can synthesise a blank GDC which includes the needed relation (Figure 8.11).

The synthesised GDC is a solution concept where a new material component  $M_x$  should provide sufficient resistance to mechanical pressure  $P_{out}$  which is a part of energy flow. the resulting output flow should not exceed the value leading to exceeding the mechanical strength limit.

The next step is to instantiate the partly blank GDC into a solution which satisfies major material and energy constraints. The procedure of instantiation is proposed in (Alberts [1993]) It consists in matching the function that should be provided with a function which can be achieved by the GPP(s) stored in the GPP knowledge base.

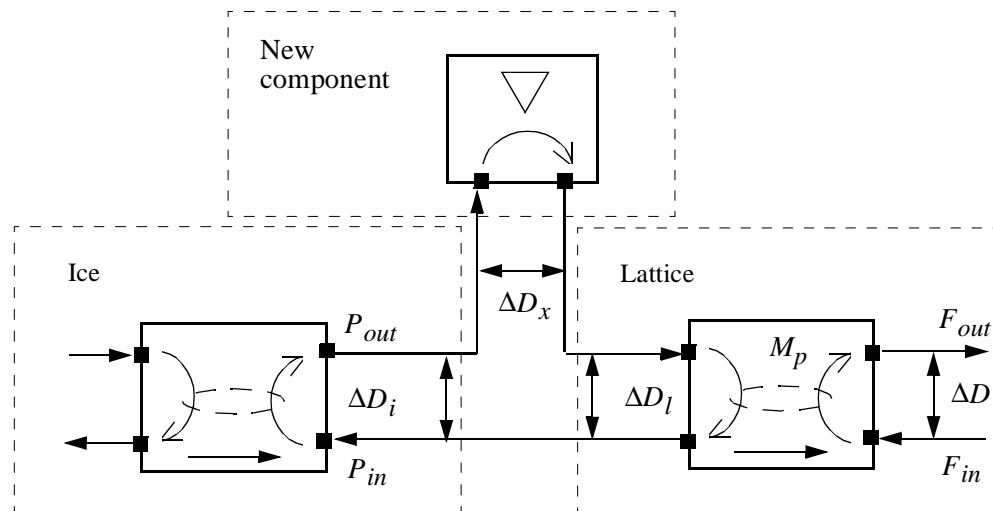


Figure 8.11: Solving the problem of outdoor water supply with innovative redesign

For instance, one of possible solutions would be to use a GPP based on the Hooke law which defines a relationship between stress and strain through the modulus of elasticity. It can be used as a resisting element to dynamically resist a mechanical force caused by growing ice pressure. In practise, one of the solutions based on the use of the Hooke law is to cover the inner part of the pipe with a of compound elastic rubber which would compress under a growing pressure of the ice.

Applying the other IRP and assuming that there are no other limitations on introduction of additional components, we could also obtain alternative design solutions.



## 4.7. Innovative redesign and ICAID

One of the objectives of AI research is to find such adequate methods to model and represent knowledge about real world that an AI system incorporating these models and knowledge would be able to automatically reason about new problems. Such reasoning process results in finding and evaluating solutions to the problems. With respect to innovative design, the Intelligent Computer-Aided Innovative Design (ICAID) system should be able to analyse problems defined within multiple constraint space and generate physically feasible solutions presented in design terms.

Apart from that, the ICAID system should be capable of multi-criteria evaluation of proposed solutions. As we already mentioned earlier, to achieve this is nearly impossible at the current level of AI since we are still unable to relate all specific and commonsense knowledge the designer reasons with about design products. This problem addresses to a large difficulty to foresee and explicitly represent every particular relation occurring within the system constituting the product as well as between the system and its outer environment.

Apart from this, many of designer's decisions are influenced by such situation-dependent criteria as manufacturability, ergonomics, recyclability, costs and so forth. All information concerning a specific design and its outer environment might thus not be available unless a solution concept has been proposed. For this reason, a model of the product may be regarded as consistent only under certain assumptions made by a designer. If, however, a solution obtained on the basis of previously made assumptions is invalid, these assumptions should be retracted from the product description. In AI, such type of problems dealing with difficulty to identify and update human's beliefs addresses to the so-called *Frame Problem* (Crockett [1993]) which is still to be solved. It was not planned within the framework of INDES research, and we therefore was left it out of the scope of this thesis.

Nevertheless, we believe that using INDES as a formal approach to structuring both TRIZ and physical knowledge, those parts of ICAID system that would support generation and evaluation of new design concepts by reasoning from physical knowledge are possible to develop. INDES proposes a formal background for conducting a design product analysis, synthesising new design concepts and evaluating the physical feasibility of the concepts generated. Those parts of ICAID which would support reasoning processes about decomposition of problems as well as multicriteria solution evaluation are left out to the responsibility of a human designer.

## 4.8. Summary

In this chapter, we proposed a framework for modeling innovative redesign based on design conflict elimination. To eliminate conflicts, the principles for innovative redesign were introduced. We only presented a part of all possible principles for innovative redesign.

The principles were drawn from two axioms for innovative redesign, which, in turn, were formulated on the basis of general properties of interacting physical entities and the trend of “complexity growth - complexity reduction”.

Based on the introduced concepts, a general model of innovative redesign was proposed. The model establishes consistent and formal methodology which relates INDES object and strategic knowledge. Finally, we discussed a possibility to incorporate the INDES model of innovative redesign into a system for Intelligent Computer-Aided Innovative Design.



## Chapter 9. Case Study

### 9.1. Introduction

This chapter presents a case study where the proposed framework is used to solve a real design problem in combination with function analysis technique for finding and formulating problems. Cordis N.V. (a Johnson & Johnson Company, Roden, The Netherlands) generously agreed to participate in the experiment and to provide all the necessary information. The company designs and manufactures a stent delivery system - a small metal tube inserted into an artery to treat a Coronary Artery Disease. This design is known as Palmaz-Schatz Stent.

The goal of the experiment was not to find a novel, feasible and patentable solution to the problem which is unknown yet but to verify if the approach proposed in the book has a practical value.

Specific technical details were omitted to make the description more structured and clear.

### 9.2. Problem Description

A coronary stent is a small, slotted, stainless steel tube mounted on a balloon catheter (Figure 9.1). It is introduced into the artery just after balloon angioplasty and is positioned at the site of the obstruction.

When the balloon is inflated with the liquid, the stent expands and is pressed against the inner walls of the artery. After the balloon is deflated and removed, the stent remains in place, keeping the artery open.

The stent is a permanent implant that remains in the artery. It helps hold the artery open, improves blood flow, and relieves symptoms of Coronary Artery Disease.

The list of technical parameters and design features of the stent is presented in Table 10.1

TABLE 9.1. Technical features of the stent

Design	Balloon expandible
Configuration	Slotted tube
Material	Stainless steel
Strut diameter (mm/inch)	0.08/0.003
Length (mm)	15
Metal surface area (%)	12
Maximum guidewire (in.)	0.014
Minimum guide internal diameter (mm)	0.084
Expanded diameter (mm)	3.0-5.0
Deployment pressure (atm.)	6-8

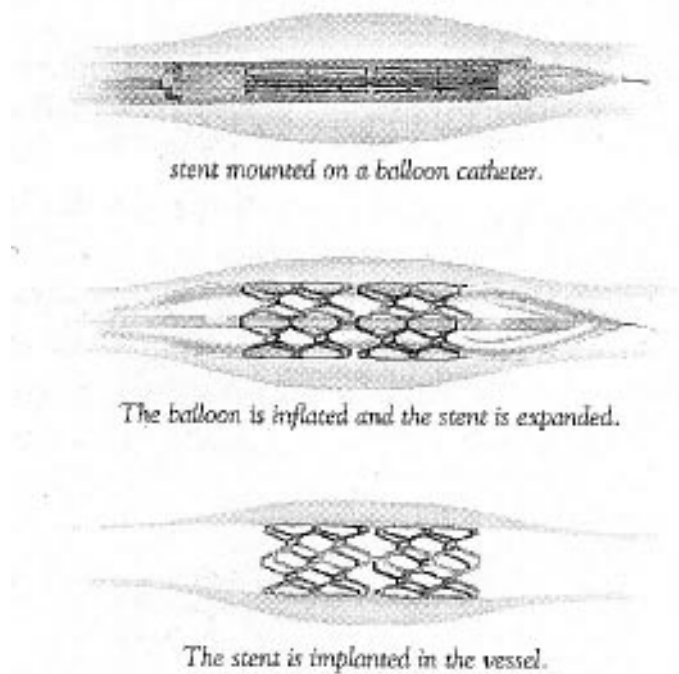


Figure 9.1: Palmaz-Schatz Stent

### 9.3. Problem model

Before formulating a problem, a function analysis of a stent delivery system was conducted accordingly the rules formulated in (Litvin and Guerassimov, 1989). Apart from the components which belong to the technical system “stent” and the stent delivery system, those components of the outer system which interact with the stent delivery system during the operation were included into the model. Among these components were the blood vessel (artery) and the blood stream. A part of the function model of the system which includes principal components is presented in Figure 9.2.

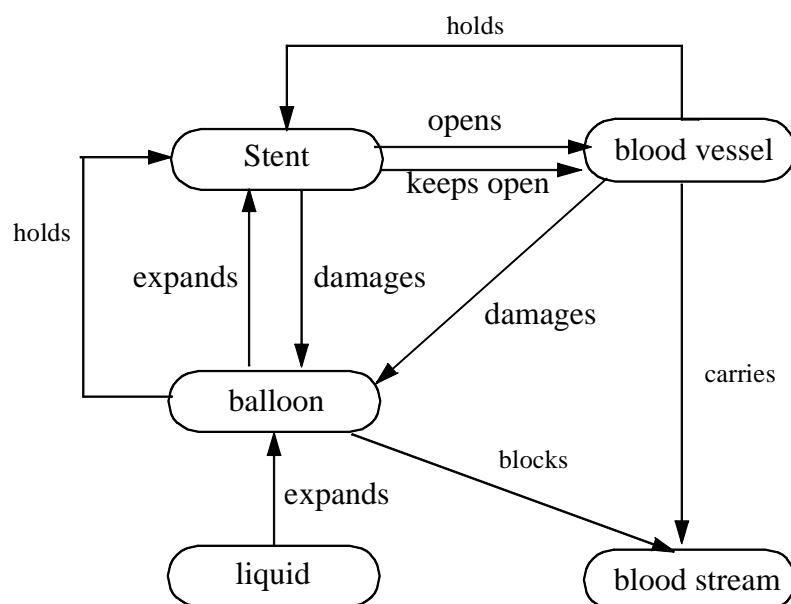


Figure 9.2: Function Diagram of the stent delivery system.

The next step was to analyse if the model contains the undesired, harmful or insufficient interactions. To do this, a degree of performance each function was matched against these criteria. The results are presented in the table 9.2

TABLE 9.2: Analysis of the functions

No.	Component	Function	Component	Performance
1	stent	opens	blood vessel	adequate
2	stent	keeps open	blood vessel	adequate
3	stent	damages to	balloon	harmful
5	blood vessel	holds	stent	adequate
6	blood vessel	carries	blood stream	adequate
7	blood vessel	damages to	balloon	harmful
8	balloon	holds	stent	insufficient
9	balloon	blocks	blood stream	harmful
10	balloon	expands	stent	adequate
11	liquid	expands	balloon	adequate

Analysis of the table indicates that four problems can be formulated for the stent delivery system - positions 3,7,8,9.

For further work, we formulated the following tasks:

1. To solve a particular problem. The problem of damaging to the balloon by the stent during the opening (position 3) was chosen.
2. To see if any other physical principles can be proposed to design a new stent opening system.

## 9.4. A Design Prototype Model

The function model presented in the previous section was translated into the INDES-model of a design prototype presented in Figure 9.3. The DPM of the stent delivery system includes four components: the liquid, the balloon, the stent and the blood vessel as well as corresponding energy flows between the components.

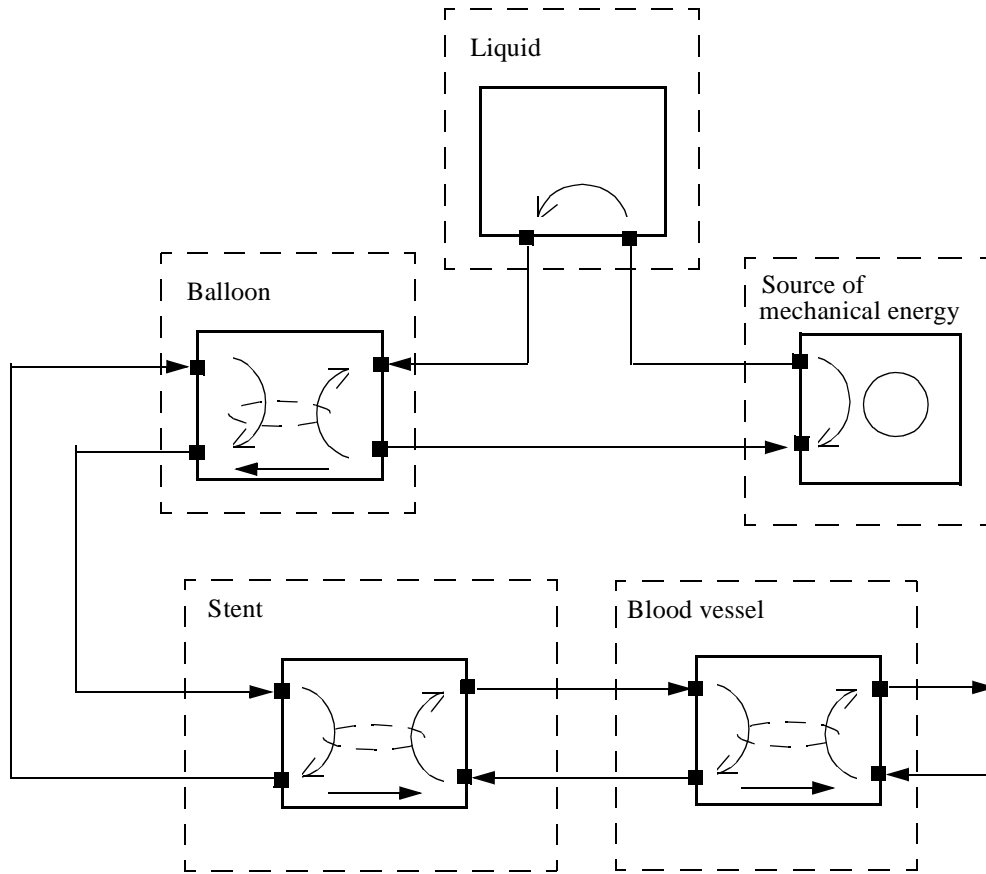


Figure 9.3: Design Product Model of a part of the stent delivery system.

## 9.5. Solving a specific problem

In section 9.3, the task of preventing a balloon from being damaged by the stent during the opening process. Since the balloon is a thin plastic shell, there is a possibility of breaking it during inflation since the metal mesh of the stent can intrude the balloon.

As pointed in Chapter 7, the first step is to define a conflict on the basis of a Design Product Model. To establish a conflict relation, we select a pair of components involved into the interaction - the stent and the balloon and represent them as a separate DPM (Figure 9.3). A goal is to find a solution concept in terms of a Generic Design Concept. To do this, we have to use the two-step procedure described in Chapter 7:



1. Defining the conflict relation by suggesting a straightforward method of solving the problem through changing the values of variables.
2. Applying the Inventive Rules to redesign the existent model and avoid the potential conflict.

To perform this procedure, one has to make a decision which physical parameters fall into the conflict. According to the Section 7.2, the design conflict occurs under the following circumstances:

1. A change in value of some physical parameter is required;
2. The required change might be achieved by changing other physical parameter(s);
3. A change of other physical parameter(s) is not allowed.

The decision was made that the value of force  $F_{I_{in}}$  (Figure 9.4) has to be low enough to prevent the balloon from being damaged. However, force  $F_{I_{in}}$  is produced by the expansion of the stent. In this case, no required result is achieved since the mechanical displacement  $\Delta D2$  is not enough to open the blood vessel. To decrease the effect of expansion it would be possible to decrease the value of the force  $F_{in}$  generated by the pressure of the liquid.

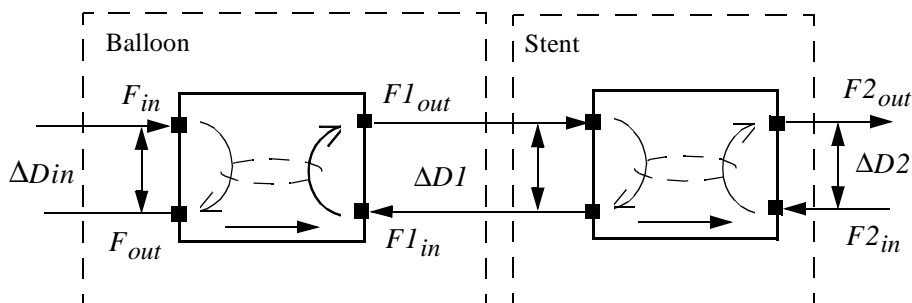


Figure 9.4: A part of the DPM causing the problem

Therefore, the conflict relation can be defined between the values of variables  $F_{2_{out}}$  and  $F_{I_{in}}$ : the force applied to the stent should be small enough to not damage the balloon and force applied to the walls of the blood vessel has to be large enough to open the blood vessel.

One of the problem condition defined at the phase of documenting the problem is that it is not allowed to introduce any additional components between the balloon and the stent. Therefore, it can be recommended to use the Innovative Redesign Principle IRP-1-4 (see Chapter 8).

**IRP-1-4:**

IF it is required to decrease or increase the value of the system variable  $E_2$   
and no component is allowed to be introduced between  $E_2$  and  $M_1$   
and there is the conflict relation  $CONFLICT(E_1, E_2)$

THEN the new energy flow should be applied to the component  $M_1$   
and effect of such combination should be provided by such physical  
phenomenon that the value of the system variable  $E_2$  changes.

After substituting the generic variables in this rule, we obtain the specific formulation of the rule:

**IRP-1-4:**

IF it is required to increase the value of the system variable  $F2_{out}$   
and no component is allowed to be introduced between  $F2_{out}$  and *the stent*  
and there is the conflict relation  $CONFLICT(F2_{out}, F1_{in})$

THEN the new energy flow should be applied to the component *stent*  
and effect of such combination should be provided by such physical  
phenomenon that the value of the system variable  $F2_{out}$  increases.

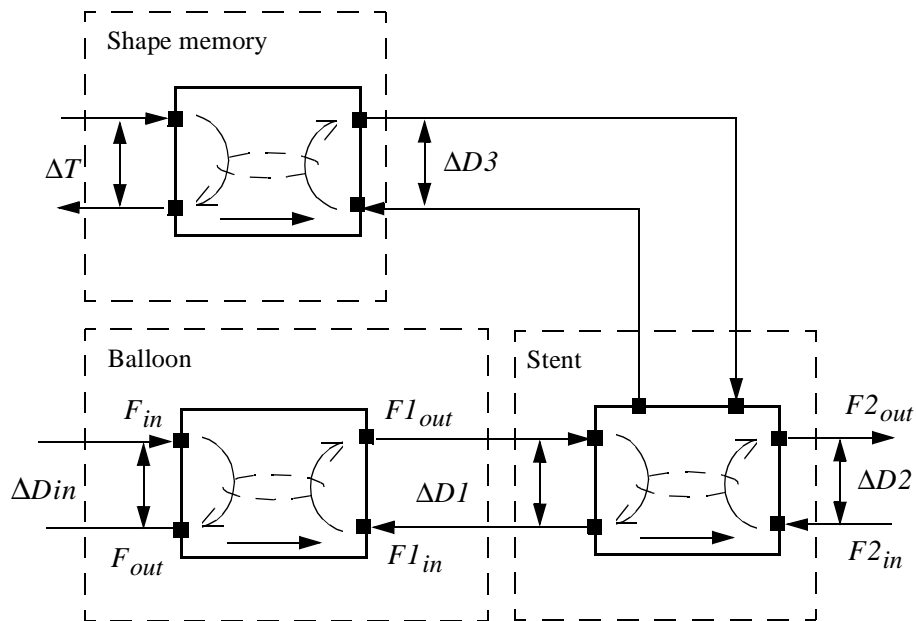


Figure 9.5: Generic Design Concept on the basis of using the memory shape effect

A role of a newly introduced component is to produce additional force towards the inner walls of the blood vessel provided that the value of the liquid pressure remains low enough.

The next step is to instantiate the unknown component by finding a relevant physical effect from the repository of effects (Appendix B). Effects that can deliver the required function have to be matched against the required intervals of values of the pressure needed to open the blood vessel.

One of the retrieved effects is the effect of memory shape - that is, changing of the shape of an object by changing the temperature of the object. Interpretation of this solution is that the links in the stent can be made of an alloy which has the memory shape property, so they can open the stent under the certain value of the temperature. In this case, the liquid can serve as a medium to heat the stent to the required temperature and its pressure might be kept lower.

The Generic Design Concept which results from the use of this principle will look as shown in Figure 9.5

## 9.6. New design concept generation

The second task formulated in section 9.3 was to check if there any other physical principles applicable to avoid the system which includes the balloon and liquid since those are the parts that cause most problems (Figure 9.6).

As seen, the function of the stent is to open the blood vessels walls by providing a mechanical displacement. A database of physical effects was scanned to find those effects that convert any other type of energy into mechanical energy which would directly result in the mechanical displacement. To limit the search, effects which have an input parameter as a pressure were filtered out.

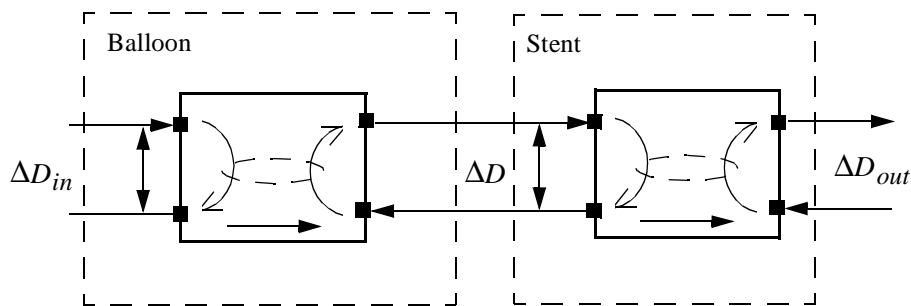


Figure 9.6: A DPM which causes most problems

A number off effects which are capable of delivering the required function were selected. The mechanical displacement can be achieved by the effects presented in Table 9.3:

TABLE 9.3: Physical effects for a new stent design

Effect	Input	Output
Bi-metal effect	temperature	displacement
Shape-memory effect	temperature	displacement
Inverse piezoelectric effect	electrical current	displacement
Thermal expansion	temperature	displacement
Magnetostriction	magnetic flux	displacement

A check against intervals of the required displacement also filtered out the effects of magnetostriction and inverse piezoelectric effects - the displacements produced by these effects are too small to provide the required displacement of the blood vessel wall ranging from 1 to 5

mm. Nevertheless, the effects of bi-metal and shape memory effects can be used to design a new system for opening the stent.

## 9.7. Conclusions

As shown above, the use of the proposed framework can be used both for problem solving and synthesis of new design concepts. Discussions with the research engineers from Cordis N.V. indicated that the concepts generated are worth to consider, however, it is unclear what will be the costs of realizing the solutions, however this question resides outside the scope of INDES research.

In addition, we found information that the idea based on the shape memory effect was proposed recently as a concept for future stent designs, but we could not find a reference to the use of bi-metals .

However, the major result from the experiment was shortening the cycle of the conceptual design phase by using the INDES framework. Although not implemented yet as a computer-based system, even applied in a “manual” way, the approach appeared to be very effective to model problems and generate solution concepts within a short time. The whole process took three days - from documenting the problem to generating a number of solution concepts.

## Chapter 10. Summary and Conclusions

### 10.1. Summary

A central question we attempted to answer in the book was: *“Is it possible to organize an automated knowledge-based support for the early phases of engineering design?”*

In order to obtain the answer we conducted a research which resulted in the development of INDES - an integrated framework for uniform modelling of domain and strategic knowledge for innovative design. As it had become clear soon after initiating the research, it was not possible to create INDES by only using techniques and methods of Artificial Intelligence and Knowledge-based modelling. We faced a situation when we had to construct a new design theory since no consistent theoretical foundation for innovative engineering design had ever existed.

To construct such the theory, we used TRIZ, a Russian Theory of Inventive Problem Solving. To make the reader of the book become familiar with TRIZ, we provided an introduction to the Altshuller’s approach (Chapter 3). However, TRIZ resulted from empirical studies and lacks a formal background. While being useful as a tool for “manual” work, TRIZ technique are too abstract and general to be represented in formal way. Since no formal background have ever been proposed to model TRIZ apart from several unsuccessful attempts, the goal of INDES was to restructure TRIZ knowledge and to develop a sharable domain-independent ontology for modeling knowledge for innovative design.

As a result, we defined several key INDES concepts on the basis of TRIZ. Among them are the concept of a conflict for finding and formulating problems, substance-field modelling, the approach to representing physical effects.

We also explained and demonstrated our approach to structure knowledge for innovative engineering design which was used to develop formal methods for modeling.

The second part of the monograph (Chapters 5-8) introduced and explained major INDES concepts and underlying theory of innovative engineering design.

The case study presented in Chapter 9 demonstrated a practical applicability of INDES for a synthesis of new design concepts and solving specific design problems. In addition, the case study helped to formulate a number of guidelines for further research.

## 10.2. Contribution of the research.

INDES defines sharable object and strategic knowledge concepts for innovative design and proposes a new model of conceptual design based on sharable physical knowledge. Below we outline a number of INDES results which can be regarded as our contribution to AI and Design research.

1. To model and represent basic information necessary for tasks of analysis and synthesis during conceptual design, INDES introduced a sharable knowledge concept labelled *Generic Physical Principle* defined on the basis of System Theory and network modelling. Generic Physical Principle is a black-box model of a physical phenomenon viewed as energy transformation process. Generic Physical Principle explicitly incorporates both parameters of energy flows and material aspects of components constituting any physical phenomenon.
2. INDES regards conceptual engineering design as a procedure of synthesis from the physical principles. Such view is based on the fact that two and more generic physical principles can be combined to a network. The resulting network is labelled a *Generic Design Concept* which is a compound INDES concept. The physical validity of a Generic Design Concept is provided by checking against boundary conditions following from the law of energy preservation.
3. All physical effects can be modelled as generic physical principles and organized to a knowledge base.
4. It is possible to organize a design from generic physical principles by decomposing given specifications and matching the required function and a predefined function which can be delivered by a particular generic physical principle.
5. In order to enable modelling of existing design products in terms of INDES, we proposed to model a selected design product as a *Design Prototype Model*. A Design Prototype Model is an instance of a Generic Design Concept obtained by constrain-

ing intervals of values of variables. Given a Design Prototype Model, one can reason about what conflicts may arise if the designer wishes to change a design product in innovative way.

6. We defined the key concept for problem finding and formulation as a *conflict*. Two types of behaviour of design products are distinguished: *correct behaviour*, when all values of system and form variables are within the required intervals; and *violated behaviour*, when a value of some system or form variable does not belong to the specified interval. Violated behaviour results from the designer's wish to change a parameter of a Design Prototype Model which is related to another parameter. Therefore, if no change of the value of the variable is possible without violating the overall behaviour, a conflict between two parameters takes place.
7. A process of redesigning a Design Product Model by conflict elimination is labelled "Innovative Redesign". INDES defines two axioms of Innovative Redesign based on today's state-of-the-art in physics research.
8. Based on the axiomatic approach, a number of *Innovative Redesign Principles* was proposed. The principles specify how a Design Prototype Model should be changed in such a way that no violation of behaviour will occur. We introduced two Basic Inventive Redesign Principles: *Extension Principle* which recommends to solve problems by adding new components to the existing system and *Replacement Principle*, which proposes to replace a physical principle behind a system component that does not meet new design requirements and specifications.
9. A general model of INDES-based design was proposed.

### 10.3. Conclusions

We believe that the INDES framework can be used to build the knowledge-intensive CAID system supporting the development of new product concepts. From our point of view, among the principal advantages of INDES are:

1. INDES object knowledge is based on objective physical laws.
2. INDES strategic knowledge is based on axiomatic approach.

Incorporating basic TRIZ principles, INDES eliminates inconsistency and informality of TRIZ and makes it possible to organize a reasoning process with formal models of design products.

Among the limitations of INDES are inability to reason about geometrical information and restriction to energy-transforming systems.



## 10.4. Further research

The ultimate goal is to develop a system for automated generation of new design solutions, we believe the following problems have to be solved:

1. A comprehensive knowledge base of physical effects has to be developed that will contain all physical effects represented in INDES terms.
2. A proposed collection of Innovative Redesign Principles (Chapter 8) has to be refined and added with more principles. This can be done by further analysing the TRIZ collection of Inventive Standards and selection of those rules which can be represented in terms of INDES strategic knowledge.
3. Since INDES misses an analytical part, a new framework for conceptual analysis of existing design products has to be developed to localize and formulate design problems that can be further solved with INDES. We believe that such framework has to be based on the function approach to the products analysis.
4. A new intermediary format for data representation has to be developed which will make it possible to automatically represent information about specific design components collected by a designer in terms of generic physical principles. A problem of identification of the specific design components with generic physical principles has also to be solved.
5. One of the remaining problems is how to interpret the solutions obtained by automated generation in specific design terms. In other words, it is necessary to understand how to instantiate proposed concepts into specific develop machines and machine parts. A cross-domain search for existing design solutions will not always help since a specific solution might not be available at all. Therefore, the research has to be focused on establishing links between INDES-based system and existing CAD systems.

## *Bibliography*

- V. Akman, P.J.V ten Hagen, & T. Tomiyama [1990], "A fundamental and theoretical framework for intelligent CAD Systems", *Computer-Aided Design*, 22(6), 368-376.
- V. Akman & P.J.W. ten Hagen [1989], "The power of Physical Representations", *AI Magazine*, Fall 1988, 49-65.
- L.K. Alberts, P.M. Wognum & N.J.I. Mars [1991], "Design as a science instead of art", University of Twente, Department of Computer Science, Enschede, The Netherlands.
- L.K. Alberts [1993a], "YMIR: a domain-independent ontology for the formal representation of engineering-design knowledge", in *Proceedings of IFIP Workshop on Formal Design Methods for CAD*, June 16-19, 1993, Tallinn, Estonia, 139-152.
- L.K. Alberts [1993b], YMIR: an Ontology for Engineering Design, Ph.D Thesis, University of Twente, Enschede, The Netherlands.
- G.S. Altshuller & R.V. Shapiro [1956], "About technology of creativity", *Questions of Psychology*, 6, 37-49
- G.S. Altshuller [1969], *Algorithm of Invention*, Moscovskiy Rabochy (in Russian), Moscow.
- H. Altshuller [1984], *Creativity as an Exact Science*, Gordon and Breach Scientific Publishers, New York.
- G.S. Altshuller [1986], *To Find an Idea: Introduction into the Theory of Inventive Problem Solving*, Nauka, Novosibirsk (in Russian).

- G.S. Altshuller [1988]. "A Brief Report on the Theory of Inventive Problem Solving", Typescript, Petrozavodsk, Russia.
- G.S. Althsuller, B.L. Zlotin, A.V. Sussman & V.I. Filatov [1989], *Search for new ideas: From Insight to Technology (Theory and Practise of Inventive Problem Solving)*, , Karte Moldaveniaske, Kishinev (in Russian).
- F. Amirouche [1994], *Computer Aided Design and Manufacturing*, Prentice Hall.
- K. Andersson [1993], "A Vocabulary for Conceptual Design - Part of a Design Grammar", in *Proceedings of IFIP Workshop on Formal Design Methods for CAD*, June 16-19, 1993, Tallinn, Estonia, 139-152.
- L. Apelskog Killander [1996], "Injection-Molding Inserts For Prototype Tools Made With Free Form Fabrication", Ph.D. Thesis, Department of Manufacturing Systems, Royal Institute of Technology, Stockholm.
- T. Arciszewski [1988], "ARIZ-77: A Method of Innovation Design", *Design Methods and Theories* 22(2).
- W. Beitz & K.-H. Kuttner (eds) [1994], *Dubbel Handbook of Mechanical Engineering*, Springer-Verlag, London.
- S. Bhatta & A. Goel [1994], "Innovation in analogical design: a model-based approach", in *Proceedings Artificial Intelligence in Design'94, August 15-18, Lausanne, Switzerland*, J.S. Gero and F. Sudweeks, eds., Kluwer Academic Publishers, 57-74.
- E. de Bono [1992], *Serious Creativity*, Harper Collins, New York.
- G. Boothroyd [1987], "Design For Assembly - The Key Design for Manufacture", *The International Journal of Manufacturing Technology*, 2, 3-11.
- P. Borst, J.M. Akkermans, A. Pos & J.L. Top [1995] "The PhysSys Ontology for Physical Systems", in *Working Papers of the Ninth International Workshop on Qualitative Reasoning*, 16-19 May 1995, Amsterdam, The Netherlands, 11-21.
- J. Braham [1995], "Inventive ideas grow with 'Triz.'", *Machine Design*. Vol. 67 No. 18. October 12.
- D. Brown & B. Chandrasekaran [1985], "Expert system for a class of mechanical design activity", in *Knowledge engineering in computer-aided design*, J.S. Gero, ed., North-Holland.

- K. Clibbon & E. Edmonds [1995], "Strategic knowledge in computational models of creative design", in *Preprints of the 3rd International. Roundtable Conference on Computational Models of Creative Design*, J.S. Gero, M.L. Maher & F. Sudweeks, eds., 3-7 December 1995, Heron Island, Queensland, Australia, 319-343.
- L. Crockett, [1994]. *The Turing Test and the Frame Problem: AI's Mistaken Understanding of Intelligence*. Ablex, Norwood, New Jersey.
- I.G. Devoino [1993], "Complication of Technical Systems", in *Proc. of ICED-93, International Conference on Engineering Design*, The Hague, The Netherlands.
- F. Dikker, G.B. Leemeijer, P.M. Wognum, & N.J.I. Mars [1992], "Towards computer-supported creative design based on Altshuller's theory of inventive problem solving", *Memoranda Informatica 92-49*, Enschede, Univeristy of Twente, The Netherlands.
- M. Dyer, M. Flowers & J. Hodges [1986], "EDISON: An engineering design invention system operating naively", in *Proceedings of the First International Conference on Application of Artificial Intelligence in Engineering Problems*, D. Sriram and R. Adey, eds., Springer Verlag, Southampton, UK, 327-341.
- K. Engelmaier [1910], *Theory of Creativity*, Severo-Zapad, St. Petersburg, Russia (in Russian).
- V.R. Fey, E.I. Rivin, I.M. Vertkin [1994], "Application of the Theory of Inventive Problem Solving to Design and Manufacturing Systems", *Annals of the CIRP* 43(1).
- K.D. Forbus [1984], "Qualitative Process Theory", *Artificial Intelligence*, 24, 85-168.
- J.S. Gero [1994], "Towards a model of exploration in computer-aided design", in *Formal Design Methods for CAD*, J.S. Gero & E. Tyugu, eds, Elsevier, 315-316.
- V.V. Glasunov [1993], *Synthesis of Novel Technical Devices*, Method, Moscow.
- G.L. Gleg [1960], *Design of Design*, Cambridge University Press, New York.
- A. Goel [1989], "Integration of Case-Based Reasoning and Model-Based Reasoning for Adaptive Design Problem Solving", PhD thesis, The Ohio State University, Ohio, USA.
- Yu. V. Gorin [1973], "A Pointer to Physical Effects", Typescript.
- T.R. Gruber [1993], "A translation approach to portable ontology specifications", *Knowledge Acquisition*, 5(2), 199-220.

- T.R. Gruber & G.R. Olsen [1994] "An ontology for Engineering Mathematics", in *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning*, Gustav Stresemann Institut, Bonn, Germany, Morgan Kaufmann.
- T. Gruber, J. Tenebaum & J. Weber [1992], "Toward a Knowledge Medium for Collaborative Product Development", in *Proceedings of the Second International Conference on Artificial Intelligence in Design*, J.S. Gero, ed., June 22-25 1992, Pittsburg, USA, Kluwer Academic Publishers.
- L. Harrisberger [1966], *Engineermanship: A Philosophy of Design*, Brooks/Cole, Belmont.
- B. Hayes-Roth, Waterman & D. Lenat [1983], "Building Expert Systems".
- P.H. Hill [1970], *The Science of Engineering Design*, Holt, Rinehart and Winston, New York.
- J.H. Holland [1975], *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*, University of Michigan Press, Ann Arbor.
- K. Holt.[1992], "Computer-aided creativity in engineering design", *Journal of Engineering Design*, 4(4), 371-376
- V. Hubka & W. Eder [1988], *Theory of technical systems*, Springer-Verlag, Berlin, Heidelberg.
- M. Ishi, T. Tomiyama & H. Youshikava [\*\*\*], "A synthetic reasoning method for conceptual design", in *Towards Worlds Class Manufacturing*, M. Wozny & G. Olling, eds., Elsevier Science, North-Holland.
- M. Ishi, T. Sekiya & T. Tomiyama [1995], "A Very Large-Scale Knowledge Base for the Knowledge Intensive engineering framework", in *Towards Very Large Knowledge Bases*, N.J.I Mars, ed., IOS Press, 123-131.
- Y. Iwasaki, R. Fikes, M. Vescovi, and B. Chandrasekaran. [1993], "How Things are Intended to Work: Capturing Functional Knowledge in Device Design.", in *Proceedings of the International Joint Conference on Artificial Intelligence*, Chambery, France.
- J.C. Jones [1981], *Design methods: seeds of human futures*, Wiley and Son, New York.
- D.C. Karnopp, D.L. Margolis & R.C. Rosenberg [1990], *System Dynamics: A Unified Approach*, John Wiley & Sons, New York.
- M.T. Keane [1988], *Analogical Problem Solving*, Ellis Horwood, Chichester.

- A. Killander & V. Sushkov [1995] "Conflict-Oriented Model of Creative Design", in *Proceedings of the 3rd Int. Roundtable Conference on Computational Models of Creative Design*, J.S. Gero, M.L. Maher & F. Sudweeks, eds, 3-7 December 1995, Heron Island, Queensland, Australia, 369-397.
- R. Koller [1976], *Konstruktionsmethode für den Maschinen-, Geräte- und Apparatebau*, Berlin, Springer.
- L.K. Kolodner [1991], "Improving human decision making through case-based decision making", *AI Magazine*, 10(1), 52-68.
- J. Kowalick [1996], "Tutorial: Use of Functional Analysis and Prunning, with TRIZ and ARIZ, To solve "Impossible-to-solve Problems", *The TRIZ Journal*, December Issue, The TRIZ Institute, available at <http://www.triz-journal.com>.
- T.J. Kowalski [1985]. *An AI Approach to VLSI Design*. Kluwer Academic Publishers, Boston.
- H. van den Kroonenberg & F. Siers [1996], *Methodisch Ontwerpen*, Stam Techniek, Educatieve Partners Nederland BV.
- W. Kruiskamp & D. Leenaerts [1995], "DARWIN: CMOS opamp synthesis by means of a genetic algorithm", In *Proceedings of 29th Design Automation Conference*, Association for Computing Machinery, New York, NY, USA, 433-438
- P. Krumhauer [1974], "Rechnerunterstützung für die Konzeptphase der Konstruktion", Ein Beitrag zur Entwicklung eines Programmsystems für die Lösungsfindung Konstruktiver Teilaufgaben, PhD thesis, Berlin.
- D. Lenat & M. Feigenbaum [1987], "On the thresholds of knowledge," *Artificial Intelligence*, 47, 185-250.
- H-J. Linde, K-H. Mohr & U. Neumann, [1994], "COIS - Contradiction Oriented Innovative Strategy", in *Proceedings of International Conference on Engineering Design ICED'94*, The Hague, The Netherlands.
- C.U. Linderstrom-Lang [1964], "Gas separation in the Ranque-Hilsch Vortex Tube"; *International Journal of Heat and Mass Transfer*, vol. 7 (1964), 1195-1206.
- S.S. Litvin & V.M. Guerassimov [1991], "Structured Principles For Contradiction Elimination", in *Materials on the Advanced Course in TIPS and VEA*, Minsk, Invention Machine Laboratory, Belarus (in Russian).
- S.S. Litvin & V.M. Guerassimov [1991b], "Methodical Recommendations on Conducting Value Engineering Analysis", in *Materials on the Advanced Course in TIPS and VEA*, Minsk, Invention Machine Laboratory, Minsk, Belarus (in Russian).

- B. Logan & T. Smithers [1993], "Creativity and design as exploration", in *Modelling Creativity and Knowledge-Based Creative Design*, J.S. Gero & M.L. Maher, eds., Lawrence Erlbaum, Hillsdale, New Jersey. \*\*\*
- M.L. Maher [1990], "Process models for design synthesis", *AI Magazine*, 1990, 49-58.
- M.L. Maher, S. Boulanger, J. Poon, & A. Gomez [1996], "Exploration and transformation in computational methods for creative design processes", in *Preprints of the 3rd Int. Roundtable Conference on Computational Models of Creative Design*, J.S. Gero, M.L. Maher & F. Sudweeks, eds, 3-7 December 1995, Heron Island, Queensland, Univeristy of Sydney, Australia, 233-265
- M.L. Maher, & D.M. Zhang [1993], "CADSYN: A case-based design process model", *AI EDAM*, 7(2), 97-110.
- J. Malmqvist [1993], "Computer-Aided conceptual design of energy-transforming technical systems, in *Proceedings of International. Conference on Engineering Design ICED-93*, The Hague, The Netherlands.
- J. Malmqvist, R. Axelsson & M. Johansson [1996], A Comparative Analysis of the Theory of Inventive Problem-Solving and the Systematic Approach of Pahl and Beitz. In *Proceedings of DTM'96*, Irvine, CA, USA.
- N.J.I. Mars, P.M. Wognum & V.V. Sushkov [1993], "AI and Creative Design", in *Artificial Intelligence & Creativity Papers from the 1993 Spring Symposium*, T. Dartnall & S. Kim, eds., Stanford, AAAI Press, 51-56 .
- H.Martin & D. Allen [1969], *Introduction To System Theory*, Merrill Publishing Company, Columbus, Ohio, USA.
- J.G. McGuire, D.R. Kuokka, J.C. Weber, & J.M. Tenenbaum [1993], "SHADE: Technology for knowledge-based collaborative engineering", *Journal of Concurrent Engineering: Applications and Research*, 1(2).
- L.D. Miles [1972], *Techniques of value and engineering*, McGraw-Hill, New York.
- D. Navinchandra, K. Sycara & S. Narasimhan [1991], "A transformational approach to case-based synthesis", *AI EDAM* 5(1), 31-45
- M. Nordlund [1994], "Applications of System Theories and AI Tools in Aircraft Design", in *Proceedings of the Conference Innovative Applications of Artificial Intelligence IAAI-94*, Florida, USA.
- J. Oullette [1996], "TRIZ-based tools promote innovation", *The Industrial Physicist*, 2(3), September, 1996.

- G. Pahl & W. Beitz [1984], *Engineering Design: A Systematic Approach*, Springer Verlag.
- N.T. Petrovich & V. M. Tsourikov [1986] *A Path to Invention*, Evrika, Molodaya Gvardia, Moscow (in Russian).
- L. Qian & J. S. Gero [1996], "An Approach to Design Exploration Using Analogy", in *Preprints of the 3rd Int. Roundtable Conference on Computational Models of Creative Design*, J.S. Gero, M.L. Maher & F. Sudweeks, eds, 3-7 December 1995, Heron Island, Queensland, Univeristy of Sydney, Australia, 3-36.
- W. Rodenacker [1971], *Methodishes Konstruieren*, Springer-Verlag, Berlin.
- K. Roth [1982], *Konstruieren mit Konstruktionskatalogen*, Berlin, Springer.
- Yu.P. Salamatov [1988], "Achievements At The Molecular Level", in *A Thread in The Labirynt*, A.B. Selutsky (ed.), Petrozavodsk, Karelia (in Russian).
- Yu.P. Salamatov [1991], "A System of Laws of Engineering Evolution", in *Chance for Adventure*, A.B. Selutsky (ed.), Petrozavodsk, Karelia, 5-174 (in Russian).
- M. Sasajima, Y. Kitamura, M. Ikeda, R. Mizoguci [1995], "FBRL: A function and Behavior Representation Language", Technical Report AI-TR-95-1, Univeristy of Tokyo.
- H. Schmekel [1992], "A system for conceptual design based on general and systematic principles of design", Doctoral Thesis, Department of Manufacturing Systems, The Royal Institute of Technology, Stockholm, Sweden.
- J.L. Shearer, A.T. Murphy & H.H Richardson [1969], *Introduction To System Dynamics*, Edison-Wesley.
- T. Smithers [1996], "On knowledge level theories of design process". In *Artificial Intelligence in Design'96*, J.S. Gero and F. Sudweeks (eds), Kluwer Academic Publishers.
- K. Sycara. & D. Navinchandra [1989], "Integrating Case-Based Reasoning and Qualitative Reasoning in Design", *AI in Design*, J. Gero, ed., Computational Mechanics, UK.
- N. Suh [1993], *The Principles of Design*, Oxford University Press.
- V. V. Sushkov [1991], "Analysis of the development of intelligent systems based on TRIZ". *TRIZ Journal*, 2,1, 35-41.
- V.V. Sushkov [1994]. Reuse of physical knowledge in creative engineering design. In: ECAI'94 Workshop Notes 'Models and techniques for reuse of designs', P.M. Wognum (ed), Amsterdam, 1994, pp. 121-126.
- V. V. Sushkov, N.J.I. Mars & P. M. Wognum [1995], "TIPS: Introduction to a Theory for Creative Design", *AI in Engineering* 9, 177-189.



- V. V. Sushkov, L.K. Alberts & N. Mars [1996], Innovative Engineering Design Based on Sharable Physical Knowledge”, in *Artificial Intelligence in Design '96*, J.S. Gero and F. Sudweeks, eds., Kluwer Academic Publishers, 723-742.
- V.V. Sushkov & N.Mars [1997]. INDES: An Ontology For Innovative Engineering Design, In: *Proceedings of International Conference on Engineering Design ICED-97*, Tampere, Finland, August 19-21, 1997, 215-218
- T. Taura, T. Koyama & T. Kawaguchi [1995], “Research on Natural Law Database”, in *Proceedings of Design Symposium*, Tokyo.
- T. Tomiyama [1994], “From general Design theory to knowledge-intensive engineering”, *AI EDAM*, 8(4), 319-333.
- T. Tomiyama & H. Yoshikawa [1987], “Extended General Design Theory”, in *Design Theory for CAD, Proceedings of the IFIP WG 5.2 Working Conference 1985*, North-Holland, Amsterdam, 95-130.
- T. Tomiyama, D.Xue & H. Yoshikawa [1994], “Developing an Intelligent CAD System”, in *Artificial Intelligence in Optimal Design and Manufacturing*, Z. Dong (ed), PTR Prentice Hall, Englewood, Cliffs, New Jersey, USA, 83-112.
- C. Tong & D. Sriram, eds [1992], *Artificial Intelligence in Engineering Design, Volume II: Models of innovative design, reasoning about physical systems, and reasoning about geometry*, Academic Press Inc.
- V.M. Tsourikov [1993], “Inventive Machine: Second Generation”, *AI & Society*, 7(1), Springer International, UK, 62-78.
- E. Tyugu [1995]. “Large Engineering Knowledge Bases”, *AI in Engineering* 8, 265-270.
- D. Verduyn & A. Wu [1995], “Integration of QFD, TRIZ, & Robust Design: Overview & "Mountain Bike" Case Study”, *Proceedings ASI Total Product Development Symposium*, November 1-3, 1995, Dearborn, MI, USA.
- I.L. Vikentiev & V.I. Efremov [1989], “A Curve That Resques”, in *Rules of The Game Without The Rules*, A.B. Selutsky, ed., Petrozavodsk, Karelia (in Russian).
- B.C. Williams [1990], “Interaction-Based Invention: Designing Novel Devices from First Principles”, in *Proceedings of the 2nd AAAI Workshop on Model Based Reasoning*, AAAI, Boston, MA, 168-175.
- D.R. Wilson [1980], *An Exploratory Study of Complexity in Axiomatic Design*, Ph.D. Thesis, Massachusetts Institute of Technology, Boston, MA, USA.

- B. Wilson, J. Teslow & R. Osman-Jouchoux, [1995]. The Impact of Constructivism (and Postmodernism) on ID Fundamentals. In *Instructional Design Fundamentals: A Review and Reconsideration*, B. Seels, ed., Educational Technology Publications, Englewood Cliffs New Jersey, , 137-157.
- L. Wills & J. Kolodner [1994], "Towards More Creative Case-Based Design Systems", in *Proceedings of AAAI-94*, August 1-4 1994, Seattle, Washington, USA, 50-55.
- M. Wolverton & B. Hayes-Roth [1995], "Finding analogues for innovative design", *Preprints of the 3rd Int. Roundtable Conference on Computational Models of Creative Design*, J.S. Gero, M.L. Maher & F. Sudweeks, eds, 3-7 December 1995, Heron Island, Queensland, University of Sydney, Australia, 59-84.
- H. Yoshikawa [1981], "General Design Theory and a CAD System", in *Man-Machine Communication in CAD/CAM, Proceedings of IFIP WG 5.2 Working Conference 1980*, North-Holland, Amsterdam, 241-253.
- H. Yoshikawa [1985], "Extended General Design Theory", in *Proceedings of IFIP W.G. 5.2: Design Theory for CAD*, North-Holland, Amsterdam.
- M. Zaripov [1988], *Energy-Informational Method of Scientific and Engineering Creativity*. VNIPI, Moscow (In Russian).
- F. Zwicky [1948], "The Morphological Method of Analysis and Construction", *Courant*, Anniversary Volume, Interscience Publishers, New York.



