TRIZ Future 2012, Lisbon, Portugal

ARIZ85 and Patent-driven Knowledge Support

Niccolò Becattini[a*], Yuri Borgianni[b], Gaetano Cascini[a], Federico Rotini[b]

[a] *Politecnico di Milano, Dipartimento di Meccanica, Via G. La Masa 1, 20156 Milan, Italy*
[b] *Università di Firenze, Dipartimento di Meccanica e Tecnologie Industriali, Via S. Marta 3, 50139 Florence, Italy*

Abstract

The growing complexity of technical solutions, which encompass knowledge from different scientific fields, makes necessary, also for multi-disciplinary working teams, the consultation of information sources. Indeed, tacit knowledge is essential, but often not sufficient to achieve a proficient problem solving process. Besides, the most comprehensive tool of the TRIZ body of knowledge, i.e. ARIZ, requires, more or less explicitly, the retrieval of new knowledge in order to entirely exploit its potential to drive towards valuable solutions.
A multitude of contributions from the literature support various common tasks encountered when using TRIZ and requiring additional information; most of them hold the objective of speeding up the generation of inventive solutions thanks to the capabilities of text mining techniques. Nevertheless, no global study has been conducted to fully disclose the effective knowledge requirements of ARIZ. With respect to this deficiency, the present paper illustrates an analysis of the algorithm with the specific objective of identifying the different types of information needs that can be satisfied by patents. The results of the investigation lay bare the most significant gaps of the research in the field.
Further on, an initial proposal is advanced to structure the retrieval of relevant information from patent sources currently not supported by existing methodologies and software applications, so as to exploit the vast amount of technical knowledge contained in there. An illustrative experiment sheds light on the relevance of control parameters as input terms for the definition of search queries aimed at retrieving patents sharing the same physical contradiction of the problem to be treated.

Keywords: ARIZ85; patent; information retrieval; explicit knowledge

1. Introduction

Technical experts, whenever facing problems, usually rely on their tacit knowledge to interpret the physical phenomena underlying the technical system they are analyzing and for an intuitive definition of solution concepts. However, science and technology are exponentially developing with increasing levels of specialization and the exploration of new domains, so that it is more and more difficult to master all the needed subjects. As a consequence, tacit knowledge, even if leveraged by a multidisciplinary team of experts, could not be sufficient to tackle inventive problems. To this purpose, explicit knowledge represents a valuable support for both experts and non-experts as a means to acquire new knowledge and design innovative solutions. Among the available sources of explicit knowledge, patents show two main advantages. On the one hand, it has been estimated that patents contain a huge amount of information not available elsewhere (almost 80% of the whole content [1]); on the other hand, their codified structure facilitates the retrieval of relevant documents and the automatic extraction of information by means of text mining algorithms.

A growing number of scientific contributions aim at improving the performances of procedures for Information Retrieval and Extraction from different sources and some of them are directly addressing the support of the problem solving process. However, the obtained results are still far from the intended objectives. Within the TRIZ context, the main efforts in the field have been dedicated to the analysis of patents with different goals, as briefly summarized in [2]. Some of them propose to use patents in order to map the technical background for both clarifying system advancements and determining intellectual property strategies. Others focus on the search for contradictions by identifying already used inventive principles, as well as characteristics parameters. At last, some of them point to support forecasting techniques with the purpose of both determining technology transfer opportunities and the maturity level of existing technologies. Section 2 first introduces the role of information in different design strategies, then presents a state of the art of the most relevant patent-based applications aimed at supporting design tasks performed through TRIZ.

The survey of Section 2 supports the thought that all these efforts have a marginal impact on the problem solving process, since they can just provide elements of knowledge for supporting decision-making or technology transfer. Additionally, the results obtained so far are promising, but not yet completely reliable, also because of the prototypal stage of development of almost all the proposed contributions. However, the main limitation appears to reside in the lack of tools to retrieve and update the knowledge to be employed along the TRIZ problem solving process, for which patent information clearly represents a significant candidate source. According to [3], current Information Retrieval systems are not compliant with the exigency of generating knowledge bases capable to inspire TRIZ users. Besides, the scope of the few efforts, as in [4], dedicated to ease the selection of relevant patents within a TRIZ-based design activity, does not go beyond aiding users (markedly those with a good level of expertise in the field of Intellectual Property searches) and speeding up the iterative retrieval process.

In this framework, the overall purpose of the present work is to determine if and how the surveyed contributions can be referenced to the different stages of ARIZ85, highlighting opportunities of development for already available solutions and future directions of research. The results of this investigation are reported in Section 3, which highlights the major deficiencies in the perspective of assisting a problem solving process driven by ARIZ85 with explicit external knowledge (namely knowledge not yet acquired or fully mastered by the designer). Moreover, the authors suggest a preliminary set of criteria for driving patent retrieval during some steps of ARIZ85 as a means for supporting knowledge acquisition, measuring the outcomes of patent searches in terms of precision (Section 4).

Eventually, the general conclusions of the paper and the illustration of the future activities are entrusted to Section 5.

2. Information in design: an overview of TRIZ-based contributions for patent exploitation

Design is an activity that requires a huge amount of cognitive resources: from the description of the design space to the decision making process aimed at excluding alternatives in order to converge towards a unique and effective solution [5]. The choices among different options are always driven by the need of satisfying specific requirements. Whenever two or more requirements appear as non mutually compatible, a new problem appears in the design space and the next design move requires a designer's creative leap to solve the problem. In order to identify the most efficient approaches for solving problems in design, Kruger and Cross [6] studied the design processes carried out by different designers: on this bases they distinguish four different cognitive strategies:

1. **Problem driven design**: main focus on problem definition using information and knowledge that is strictly needed to solve the problem, with the purpose of promptly generating a solution;
2. **Solution driven design**: little time spent on problem definition and main efforts put in the definition of solution concepts gathering information needed to further develop a solution;
3. **Information driven design**: generation of solution concepts on the basis of information gathered from external sources;
4. **Knowledge driven design**: generation of solution concepts on the basis of prior, structured, personal knowledge, with a minimal amount of information gathered from external sources.

Regardless of the followed strategies, it is clear that gathering information from external sources represent a key issue, even for the last approach. Indeed, as mentioned in the Introduction, individuals' memory cannot manage the fast changes occurring in science and technique [7], since the size of available information is exponentially growing. This, in turn, means that Information Retrieval may jeopardize the efficiency of the problem solving process, both because it is not always clear what kind of information could be useful for solving problems [8] and also because this activity results really time consuming [9].

As presented in [10], the authors believe that the most efficient strategy in design should exploit individual knowledge at the maximum extent, before gradually retrieving the needed information from external sources. This approach results beneficial also in an ARIZ-like problem solving process, avoiding costly and inefficient efforts for generating solutions through trial-and-error stages.

Since patent databases represent one of the richest sources of information and, as recalled in the Introduction, their codified structure eases the retrieval of documents and subsequent information extraction, Section 2.1 reviews available contributions in the field of TRIZ-based patent management. The survey includes several contributions, mostly operating by means of text-mining resources, oriented towards the extraction of relevant information from patents and its representation in a valuable and intuitive form for designers adopting TRIZ. Additionally, it illustrates proposals aimed at facilitating the creative problem solving process by exploiting the contents of the inventions with a more marked focus on the usability of the principles and the functions illustrated in the patents.

2.1. TRIZ-based systems exploiting patent content

With reference to TRIZ-based tools exploiting text-mining techniques, Liang and Tan [11] propose a strategy to classify patent documents according to the automatic identification of related inventive principles. The purpose of the research is to support TRIZ users in retrieving, also in different technical domains, knowledge useful to solve specific problems. In a subsequent study, the scholars [12] propose an algorithm to ease the extraction of valuable information from previously clustered patents that put into practice the inventive principles deemed to be beneficial for solving a specific contradiction. With a similar purpose, the work performed by Tong et al. [13], subsequently expanded in [14], is aimed at evaluating the capabilities of different classifiers in clustering a set of representative patents according to the inventive principle they exploit. Eventually, they present an original categorization approach that claims to outperform established classification modes [15].

Cascini and Russo [16] expand the capabilities of a prior software application, addressed at generating functional schemas of the inventions together with indications of hierarchical relationships between entities by semantically parsing patent contents. The paper explains how to extract the contradictions that are faced and solved in the patent by automatically analyzing the claims of the document. The achieved structured information represents an index about the complexity of the patent and the pertaining degree of inventiveness in the perspective of identifying the possible evolution patterns of the studied technical systems. The objective of assessing the evolution trajectories is strengthened by the automated building and confrontation of thesauri concerning different technical fields [17]; this kind of algorithm also allows to build ontologies with entities and related relationships and map key problems of the investigated field.

Cavallucci et al. [2] propose to improve entity/property-based ontologies with graphs capable to speed up the analysis of initial situation, with a clear representation of problems and available partial solutions.

With reference to forecasting purposes, Li et al. [18] illustrate a method to cluster patent information according to the metrics distinguishing between different levels of invention, as originally proposed by Altshuller. The retrieval of relevant patents is carried out by upgrading the search interface provided the United Stated Patents and Trademark Office (USPTO). Alternatively, Yoon and Kim [19, 20] demonstrate the capabilities of Natural Language Processing techniques to monitor the TRIZ trends followed by the inventive solutions contained in patents, thus allowing to build radar plots showing the potential evolution of designed products. The same methodological objective is achieved in [21], whose approach is addressed at retrieving the adjectives referred to a product within patent literature and subsequently linking the emerged attributes to the phases of particular trends.

In a recent paper, Prickett and Aparicio [3] try to collect the contributions which can facilitate the building of an ontology interfacing with engineering design tasks carried out by means of TRIZ approach. The methodology they propose clusters patents according to a set of classes pertaining the recalled ontology, which denote the maturity of the technical system, the employed resources, the performed function, the exploited inventive principles. However, the same authors remark how the current level of the research falls far from a target ontology fulfilling the actual exigencies of designers using TRIZ.

Choi et al. [22] experienced an application to extract patent information for the purpose of easing the use of Function-Oriented Search (FOS), hence to retrieve technologies which can be applied to overcome industrial problems. On the other hand, Montecchi et al. [23] try to update the FOS by allowing a complete search at different detail levels, with the aim of showing opportunities for technology transfer. By specifically focusing on bionics, Walter et al. [24] elucidate the functions described in a sample of 147 American patents through a semantic engine swivelling on the Subject-Action-Object structures and subsequently group the treated documents according to similarities in the above triads. The main aim of the research is to urge designers to use principles from bionics in order to support the employment of TRIZ and generate innovative solutions.

Eventually, the methodology and software application developed in [25] support the ideation of solutions by analogies, as mapped in a large sample of patents through text mining tools. The proposal aims at enhancing the effectiveness of design techniques, like TRIZ, that rely on an abstraction process of the encountered problem and subsequently need to contextualize the general ideas that have been produced.

In brief, according to the review, the literature is populated of patent-based tools, algorithms and software applications focused on supporting the users about particular issues encountered during the design process conducted by means of TRIZ heuristics. Along with the fact that most contributions have been published in the very last years, they are scarcely linked among them, so that no articulated proposal has been advanced to cover the spectrum of the whole problem solving task. In this context, an analysis of the “big picture” is required to shed light on the major deficiencies about the not exploited potential contribution of knowledge from patents in favour of TRIZ users.

3. Patent information supporting problem solving process through ARIZ

The importance of enriching the problem space with new information and knowledge along the problem solving process has been already highlighted at the beginning of Section 2. Moreover, the need of relying on technical information becomes crucial in non-routine design problems, especially in those situations where opportunities for generating novel ideas are hidden to the mind of the designer. After more than 25 years of application, it has been demonstrated that in such cases ARIZ85 fruitfully supports the cognitive processes of designers for solving problems. However, ARIZ is a method whose efficacy largely depends on the problem solver's capability of mastering it and on the owned technical knowledge. Whenever the solver's tacit knowledge does not sufficiently cover the knowledge relevant within the design domain, it is necessary to follow an information-driven approach as a means for knowledge acquisition. Therefore, the efficacy of ARIZ85 could be further boosted if the problem solvers may proficiently use available information, so as to carry out the activities suggested by the algorithm. For such a purpose, Section 3.1 will focus on the kind of information that is explicitly requested along the ARIZ85 problem solving process, taking into account both steps of the so called ARIZ85-A Part 0 and the whole ARIZ85-C algorithm. This detailed analysis allows the authors to point out what kind of information could be searched in a patent database and its potential contribution along the ARIZ process, as a support for carrying out critical steps. According to such an analysis, Section 3.2 aims at mapping if and how existing algorithms and systems for managing patent content according to TRIZ knowledge base can support the execution of ARIZ steps.

3.1. ARIZ85 and information needs

Both Part 0 of ARIZ85-A and the whole algorithm of ARIZ85-C are characterized by explicit requests of information to be retrieved within the TRIZ body of knowledge (e.g. 76 Standard Solutions) or outside the boundaries of the theory, e.g. priority searches in patent databases. Table 1 summarizes these explicit requests, with a preliminary classification of their content (right column). In details, the System of 76 Standard Solutions needs to be examined three times in the steps of ARIZ85-C, according to the updated formulation of problem model and, additionally, one time during ARIZ85-A Part 0 even before than building a detailed model of the problem to be solved. The exploitation of TRIZ instruments for solving problems (and obviously of the information there included) is strongly requested all along Part 5: besides the 76 Standard Solutions, solutions should be devised by using elements of knowledge outside the problem space and within the TRIZ body of knowledge, such as the Principles for resolving Physical Contradictions (PhC) and the Pointer to Physical Effects and Phenomena. Two requests of external information from patents occur during Step 0.8 and 7.3 with the purpose of sharpening the problem definition (0.8) and evaluating the eligibility of the generated solution concept by means of priority searches in Patent Databases (7.3).

According to the authors' vision, the efficacy of the problem solving process driven by ARIZ85 could be further increased if the execution of some of its steps gets supported by the introduction of information obtained from patents content. In details, patents may provide three different types of support along such a problem solving process:

1. "**Domain Identification**" (DI) – Information for supporting the identification of the characteristics of technical systems (entities, properties and relationships between them);
2. "**Solution Oriented**" (SO) – Information for supporting the generation of solution concepts (by exploiting analogies among solutions, by effects commonly used in different fields of technique, etc.);

3. "**Boundary Conditions**" (BC) – Information for supporting the analysis of the context in which the problem appears and the solution has to be introduced, also with reference to its potential adoption in different areas of technology (information from Technology Maturity Assessment, presence of analogous problem, etc.).

Table 1. Analysis of ARIZ steps where the support of information is explicitly requested

ARIZ85 Parts	Number of Steps of ARIZ85 Part	Number of steps explicitly requiring external information	ARIZ85 steps explicitly requiring external information [*Step X.x – Information type*]
0 – Analysis of the Initial Situation	9	2	Step 0.7 – System of 76 Standard Solutions Step 0.8 – Patent searches for solutions to similar problems
1 – Analysing the problem	7	1	Step 1.7 – System of 76 Standard Solutions
2 – Analysing the problem model	3	0	-
3 – Defining IFR and PhC	6	1	Step 3.6 – System of 76 Standard Solutions
4 – Mobilizing and utilizing of SFRs	7	0	
5 - Applying the Knowledge Base	4	3	Step 5.1 – System of 76 Standard Solutions Step 5.3 – Principles for resolving Physical Contradictions Step 5.4 – Pointer to Physical Effects and Phenomena
6 – Changing or substituting the problem	4	0	-
7 – Analysing the method of resolving PhC	4	1	Step 7.3 – Patent search for priorities
8 – Applying the obtained solution	3	0	-
9 – Analysing the problem solving process	2	0	-

Table 2 summarizes where these contributions from patents can be proficiently used according to the sequence of ARIZ Parts in addition to the explicitly requested external information: the second last column enumerates how many steps could be supported by information retrieved and extracted from patents; the last one specifies which kind of contents is relevant with reference to the abovementioned numbered list.

Table 2. Authors' vision about the potential contribution of information from patent content along the problem solving process guided by ARIZ85.

ARIZ85 Parts	Number of Steps of ARIZ85 Part	Number of steps where patent content may support ARIZ85	Potential Explicit Knowledge support from Patent Sources
0 – Analysis of the Initial Situation	9	3	"Domain Identification" "Boundary Conditions" "Solution Oriented"
1 – Analyzing the problem	7	1	"Solution Oriented"
2 – Analyzing the problem model	3	1	"Domain Identification"
3 – Defining IFR and Physical Contradiction	6	1	"Domain Identification"
4 – Mobilizing and utilizing of SFRs	7	6	"Solution Oriented"

5 - Applying the Knowledge Base	4	4	"Solution Oriented"
6 – Changing or substituting the problem	4	1	"Solution Oriented"
7 – Analyzing the method of resolving the Physical Contradiction	4	2	"Boundary Condition" "Domain identification"
8 – Applying the obtained solution	3	1	"Boundary Condition"
9 – Analyzing the problem solving process	2	0	-

The overall number of steps where external information plays a role in supporting the problem solving process increases from 7 to 20, differently distributed according to the specific objectives of each ARIZ Part. Details about relevant steps considered in Table 2 are presented in the hereafter.

Part 0: Analysis of the Initial Situation

For what concerns Part 0, three steps could be supported by means of retrieved patent information. Step 0.2 requires to define new problems on the basis of the overall goal of the solution, as formulated during Step 0.1, at different hierarchical levels (both super-system and sub-systems). The purpose is to identify by-pass problems whose solution can be more easily generated or is even already available. In order to carry out this kind of task, it is therefore necessary to determine which elements compose the system, their mutual interactions and the environment in which the expected solution is going to work (DI). Besides, patents could not really support Step 0.7 as explicitly requested by ARIZ. Indeed, up to this step, the problem model has not been defined yet and also the search for solutions within the 76 Standards Solutions appears to be poorly systematic if not simply made through intuition. The only possible support to generate solutions so far is a goal-based search for patents addressing the same by-pass problems (SO), an investigation that can be carried out as soon as Step 0.2 is concluded. During Step 0.3, the user has to choose if it makes more sense to solve the original one or the by-pass problems, according to objective factors such as available opportunities for the evolution of the system. This decision making task can be therefore supported by means of information about the maturity of the technology under investigation (BC). Sets of relevant patents can serve for this purpose and different methods are already available to determine the residual evolutionary potential of given classes of products. Part 0.8 explicitly asks to sharpen the problem model through patent information. According to the prescription of this step, this problem reformulation should take into account information about "close", "similar" or "opposite" problems, even if the interpretation of these terms is fully susceptible of individual perspective. In order to clearly determine what could be considered "close/similar/opposite", the authors propose to support this kind of investigation with a requirements-based patent search (SO), according to the definition of requirements carried out during the Step 0.6.

Part 1: Problem Analysis

The authors consider that Step 1.7 is the only one where external information could give a valuable support to Part 1, since all previous steps rely on aspects already examined during Part 0. The possibility to apply the 76 Standard Solutions can be now focused on the two requirements (namely Evaluation Parameters – EP – according to the OTSM-TRIZ notation) selected for the model of contradiction. Thus, the search could be centred on patents where the solution satisfies the same pair of EPs or on patents revealing the applicability of 76 Standard Solutions.

Part 2: Problem Modelling

Poor or no support at all could be given to the first two steps of Part 2, where the individual knowledge is crucial to determine operative space and time. However, the definition of Su-Field Resources (SFRs) requires the identification of elements involved in the conflict as well as the ones in its surroundings, paying also attention to their properties. Patents can strongly support this task and a certain degree of automation can be expected thanks to their intrinsic ordered structure: components are numbered and text mining algorithms may build ontologies by semantic analysis aimed at determining component features as well as relationships of inclusions and interactions (DI).

Part 3: Ideal Final Result and Physical Contradiction Identification

Except for what concerns the generation of solutions through the 76 Standard Solutions prescribed at Step 3.6 (where patent information support is similar to what already claimed for step 1.7), Part 3 almost entirely relies on knowledge and information already emerged in previous steps. However, it is still possible to support the choice of the best candidate SFR for reformulating IFR-1 during Step 3.2. Indeed, this decision requires a comprehensive definition of available resources, including properties to be leveraged in order to satisfy the conflicting requirements. In this case, more than for Part 2, a patent-derived ontology including cause and effect relationships between its entities (DI) can support this kind of choice.

Part 4: Use of Substance-Field Resources

Part 4 offers the best chances for supporting the execution of ARIZ steps with information retrieved by patents. Step 4.1 is aimed at removing psychological inertia by means of Smart Little People (SLP): the attention of the solver is released from the specific structure of the system that, in turn, gets substituted by teams of tiny “individuals” showing a behaviour capable to overcome the conflict. Obviously patents get filed without mentioning SLP role in the inventive process. Even so, it is possible to help the generation of solution by patent searches aimed at retrieving documents where the already defined SFR shows the same behaviour expected by SLP (SO). On the other hand, Steps from 4.3 to 4.7 explicitly suggest the modification of SFRs as a means for driving the user towards the definition of solution concepts.

Part 5: Knowledge Base Application

Whenever the attempt of generating solutions in last steps of Part 4 is unfruitful, Step 5.1 prescribes to apply the 76 Standard Solutions with modified SFRs. In all these circumstances it is possible to support the identification of relevant information from patent by searching examples of inventions where the chosen modified SFR gets proficiently applied to solve a problem (SO). This kind of searches should take into account changes in the set of SFRs as they update from Step 4.3 to 4.7.

While the opportunity for patent support during step 5.1 has been exhausted for the last steps of Part 4, Part 5 still has several chances to be assisted by external information. Step 5.2 relies on previous successful experiences using ARIZ, with a specific reference to analogous problems sharing the same Physical Contradiction. Therefore, patent searches aimed at finding this kind of analogies in inventions where the opposite values of the SFR, chosen at Step 3.2, (SO) become paramount. At last, Steps 5.3 and 5.4 are actually supported by TRIZ instruments for solving problems. However, while for Step 5.3 it seems hard to retrieve information about separation principles in patents that have not previously examined along Step 5.2, the exploration of the Pointer to Effects at Step 5.4 can be also supported by means of patent searches. For instance, the identification of relevant Physical, Chemical and Geometrical effects proficiently applied in existing inventions eases the selection of a controllable working principle for the desired solution (SO).

Part 6: Check and Reformulation of the Problem

Part 6 (Steps 6.2-6.4) aims at changing the problem model after unsuccessful analyses, according to what has been previously carried out. Therefore, it seems that no valuable support can be provided by patent sources for these steps. However, when the designer generates a promising solution concept, Step 6.1 asks to translate it in a technical solution, specifying a preliminary structure that implements the identified working principle. Thus, a patent search aimed at retrieving examples of system embodiments working with the same behaviour can be usefully carried out (SO).

Part 7: Solution Assessment

The first two Steps of Part 7 check the eligibility of solution concept from the perspective of requirement satisfaction, smart use of SFRs and the actual overcoming of the Physical Contradiction. On the other hand, Step 7.3 and 7.4 aim at checking the eligibility of solution both in terms of its patentability and of the potential new issues that may emerge while embodying the solution concept. As suggested by the algorithm itself, Step 7.3 requires a classical priority search in patent databases (BC). The execution of Step 7.4, indeed, strictly depends on the designer's individual knowledge and on the capability to forecast the emergence of new problems before facing them. From this perspective, ontologies of problems in a given field of technique (e.g. in a given IPC class) could represent a fruitful support for carrying out this Step (DI).

Part 8: Solution Implementation

The overall purpose of Part 8 is to consider the implications of the adoption of the technical solution. This kind of analysis, even if carried out in just three steps, encompasses a wide range of different aspects such as the estimation of super-system changes and the identification of opportunities for further application of the solution in different contexts. According to this comprehensive perspective, it is hard to exactly determine the specific support role patents may play. However, it is possible to devise a preliminary set of information that eases such an investigation. Technology transfer opportunities, for instance, can be more easily determined

- on the basis of patent searches aimed at retrieving documents concerning analogous problems partially solved in different fields of technique (BC);
- according to time-correlations in the evolution of different technical domains (BC);
- by means of automatically generated indexes characterizing the maturity of given niches of techniques as retrieved by IPC classes (BC).

Part 9: Reflective Stage

At last, Part 9 has the objective of analysing the problem solving process in order to enrich the knowledge base on which the algorithm itself is grounded. Then, its analysis is out of the scope of this paper.

3.2. Potential of systems and algorithms for supporting ARIZ Steps requiring external information

In order to point out further directions of development in this field, Table 3 shows which steps of ARIZ85 can be currently supported by available TRIZ-based contributions for patent management, according to the review presented in Section 2.1 and with reference to the needs of information from patent along the problem solving process driven by ARIZ as presented in Section 3.1.

Despite Table 3 shows that just few steps are not yet supported by TRIZ systems for patent management, most of the contributions matched with ARIZ85 Steps provide just partial support or even need to be further adapted for an effective use.

At last it is worth to mention that two of the most critical steps (e.g. Step 3.2 and 5.2) are still lacking any kind of support from patents. Section 4 shows a preliminary attempt to plug one of this gaps.

Table 3. Available TRIZ-based contribution showing some capabilities to support the execution of ARIZ85 Steps where patent information may come useful.

ARIZ85 Steps	TRIZ-based contributions potentially supporting ARIZ85 step	ARIZ85 Steps	TRIZ-based contributions potentially supporting ARIZ85 step
Step 0.2	[16] [17]	**Step 4.6**	[3] [4] [11,12] [13,14,15] [24]
Step 0.3	[3] [4] [14] [18][19] [20] [21] [23]	**Step 4.7**	[3] [4] [11,12] [13,14,15] [24]
Step 0.8	[16]	**Step 5.1**	[3] [4] [11,12] [13,14,15] [24]
Step 1.7	[3] [4] [11,12] [13,14,15] [24]	**Step 5.2**	-
Step 2.3	[2] [3] [16] [17]	**Step 5.3**	-
Step 3.2	-	**Step 5.4**	[22] [23] [24]
Step 4.1	[22] [23] [24] [25]	**Step 6.1**	-
Step 4.3	[3] [4] [11,12] [13,14,15] [24]	**Step 7.3**	[16] [17] [23]
Step 4.4	[3] [4] [11,12] [13,14,15] [24]	**Step 7.4**	[2]
Step 4.5	[3] [4] [11,12] [13,14,15] [24]	**Step 8.3**	[3] [4] [14] [18] [19] [20] [21] [22] [23]

4. Discussion and prior research directions

Table 3 illustrates several gaps of the existing contributions in the systematic support of the ARIZ process through the information contained in patent databases. The limitations involve several ARIZ steps, according to the purpose they should fulfil in the problem solving task.

As well as disparate approaches and text mining techniques have been used to retrieve and extract knowledge responding to different goals, the authors believe that no general purpose method can be fine-tuned to cover the aspects that have not been yet taken into consideration. In other words, each of the "blank cells" of the Table 3 has to be filled following a specific strategy.

Furthermore, in a broader perspective and according to the presented state of the art, the literature lacks tools supporting the individuation of relevant patents upstream their content analysis, the extrapolation of significant information, turning the extracted knowledge in a proper form to be rapidly interpreted by TRIZ users. It makes sense to also consider that people and organizations adopting TRIZ not necessarily master the Intellectual Property field and, more specifically, adequately manage the tools for patent retrieval. From this point of view, it could result useful to support the preliminary individuation of a set of patents from which to extract the required information. According to this objective, a procedure with a minimal human involvement would result a strongly desirable outcome. Thus, the most proper strategy would result in the automated building of a tailored patents sample upon the definition of the relevant factors regarding the technical problem under investigation and exploiting the research mask of free patent databases. Of course, the above factors (e.g. name of the system, industrial field, evaluation parameter, inventive principle), to be transformed in proper research keys to consult patent databases, will vary according to the specific objective of the ARIZ step and the consequent need of information. The following paragraph depicts a preliminary experiment aimed at individuating the most beneficial factors to be employed in a patent retrieval task concerning the identification of analogous problems sharing, in the best scenario, the same Physical Contradiction, which moreover represents one of the most disregarded issues within the previous art (see Table 3 for steps 5.2 and 5.3).

4.1. Extracting the terms for a beneficial patent retrieval within the search of akin contradictions

The experiment consisted in asking a group of 15 MS Engineering students, attending an University course about TRIZ and systematic innovation, to formulate queries in an assigned patent database (Esp@cenet from the European Patent Office) and to consequently evaluate the results in terms of their pertinence with a given contradiction. The conflict that was supposed to be faced, represented in Figure 1, regarded a common problem of washing machines drums and was explained as summarized hereafter:

- a small diameter of the holes is desirable, since a reduced circumference perimeter allows to limit the damage of the clothes;
- a large diameter of the holes is desirable, since a big circle area favours the draining of the water contained in the drum.

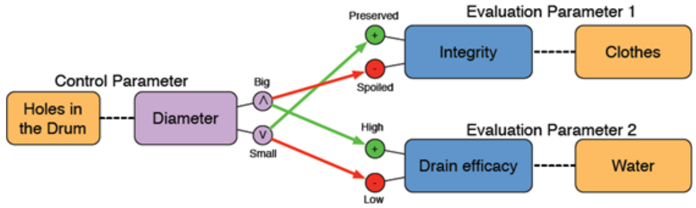

Fig. 1: representation of the contradiction according to which to search for analogies

The students were capable to devise about 100 different search queries, which were reported in a tailored table together with their comments about the obtained results (quantity of retrieved documents, number of relevant patents) and proper reformulation strategies to improve precision and recall. Among the alternatives, those showing a limited number of results (less than 300) and a not negligible percentage of pertinent patents (precision greater than 10%) were further analyzed, corresponding to 10 search attempts.

The terms, i.e. the keywords introduced to formulate the queries for individuating inventions with akin contradictions, were classified according to their "role" within the contradiction or the technical system (although it was not required to search for patents restricted to the assigned industrial context). The employed terms, with reference to the above mentioned 10 queries (see Table 4), dealt with:

- Control Parameter (e.g. size, diameter);
- Element of the Control Parameter (e.g. hole, drum, basket);
- Evaluation Parameters (e.g. draining, damage);
- Elements of Evaluation Parameters (e.g. water, fluid, garment, cloth);
- System, its Function and Behaviour (e.g. washing, spin, centrifugal);
- Characteristics of the Control Parameter aimed at determining a solution concept (e.g. multiple, varying);
- Values related to the Control Parameter (e.g. big, small, wide).

Some search strategies included also the indication of the International Patent Classification pertaining the household appliances as a criterion for formulation of the query, in a way varying from 1-digit to 6-digits fashion. Many students adopted the possibility to introduce synonyms as a driver for enlarging the patent set, obtaining satisfying improvements.

Table 4. Queries extracted from the texting activities with less than 200 results and precision greater or equal to 10%.

Quantity of resulting patents	Precision (approximate values)	Query
2	100%	Title or abstract: desirable effect according to EP2 (with synonyms) AND behaviour of the system Class: field of application related to the described artefact (4 digit) OR to collection of technologies to perform the main useful function (4 digit)
20	10%	Title: control parameter AND control parameter's element
290	15%	Title or abstract: control parameter AND control parameters' element AND further attribute concerning the control parameter aimed at obtaining a solution
10	30%	Title or abstract: behaviour of the control parameters' element Class: field of application related to the described artifact (6 digit)
43	10%	Title or abstract: name of the control parameters' element AND detail of the structure of the system AND behaviour of the system
36	15%	Title: name of the system Title or abstract: control parameter AND control parameters' element Class: field of application related to the described artefact (8 digit)
3	33%	Title or abstract: element of the control parameter (with synonyms) Title: control parameter AND first direction of the control parameter AND second direction of the control parameter
286	20%	Title or abstract: element of the EP2 (with synonyms) AND EP2 AND element of the EP1 (with synonyms) AND EP1
104	50%	Title or abstract: element of the EP2 (with synonyms) AND EP2 AND element of the EP1 (with synonyms) AND EP1 Class: field of application related to the described artefact (1 digit)
17	15%	Title or abstract: control parameter's element AND attribute related to the system AND behaviour of the control parameter's element

As an overall appraisal of the experience, the results can be assessed just at a qualitative level, because of a plurality of motivations. The outcomes can be evaluated just in terms of the precision index, since the determination of the global number of patents sharing the contradiction results a very difficult task, hence hindering to take into account also the recall term. Additionally, in order to look for a “best search strategy”, a full combination of the vast amount of terms resulting valuable for the research should be tested by introducing alternatively them in title, abstract or full text and by linking them with logical operators (i.e. AND or OR) when formulating the query. Such issue would require to organize an experiment involving thousands of alternative queries. On the other hand, it should be checked whether the results arising from a specific case study would match those of different examples regarding dissimilar technical fields.

Despite of the above limitations, the experiment provides sufficient evidence about the relevant role played by the terms pertaining the control parameter when attempting to retrieve patents characterized by a common contradiction.

5. Conclusions and future activities

The manuscript has highlighted the main missing capabilities of the existing proposals, reviewed in the state of the art Section, in terms of supporting the exploitation of ARIZ through the provision of explicit knowledge. The authors would be glad to discuss with other scholars and TRIZ experts the results of the presented investigation, with the objective of defining a shared picture about the knowledge requirements of the problem solving process with ARIZ and agreeing about the aspects to be supported with major urgency.

According to the authors' vision, a prior task to be carried out by the TRIZ community is the definition of suitable methods for retrieving patents from existing databases, resulting valuable for the purposes of the various ARIZ steps, e.g. the inventions sharing a contradiction, constituting the background of the given problem, etc. The proposed approach aims at automatically building search queries to explore patent databases on the basis of the terms characterizing the problem according to TRIZ formalism (e.g. evaluation parameters, operative space, etc). The goal is to speed-up the process of identifying relevant patents, providing a particular support for those individuals with a scarce experience within the field of Intellectual Property. Along with some evidences about the role of the control parameter in the retrieval of patents with analogous problems, an exploratory experiment carried out by 15 MS Engineering Students basically reveals the complexity and the extent of such an activity. Hence, among the planned future studies, the authors are intentioned to systematize the identification of "best research queries" by exploiting the computational capabilities of machines, so to quickly measure precision and recall of the alternative query formulations.

References

[1] Bregonje M. Patents: A unique source for scientific technical information in chemistry related industry?. *World Patent Information* 2005;**27**:309–15.

[2] Cavallucci D, Rousselot F. Zanni C. Using patents to populate an inventive design ontology. *Procedia Engineering* 2011;**9**:52-62.

[3] Prickett P, Aparicio I. The development of a modified TRIZ Technical System ontology. *Computers in Industry* 2012; **63**:252–64.

[4] Regazzoni D, Nani R. TRIZ-Based Patent Investigation by Evaluating Inventiveness. *Computer-Aided Innovation (CAI) - IFIP International Federation for Information Processing* 2008;**277**:248-57.

[5] Cross, N. *Engineering Design Methods: Strategies for Product Design*. Fourth edition. Chichester: John Wiley and Sons Ltd.; 2008.

[6] Kruger C, Cross N. Solution driven versus problem driven design: strategies and outcomes. *Design Studies* 2006;**27**:527–48.

[7] Bracewell R, Wallace K, Moss M, Knott D. Capturing design rationale. *Computer-Aided Design* 2009;**41**:173–86.

[8] Ahmed S. An approach to assist designers with their queries and designs. *Proceedings of the ASME IDETC* 2006,359-66.

[9] Aurisicchio M, Bracewell R, Wallace K. Characterising the information requests of engineering designers. *Proceedings of the 9th International Design Conference DESIGN* 2006;1057-64.

[10] Becattini N., Borgianni Y., Cascini G., Rotini F. Model and algorithm for computer-aided inventive problem analysis. *Computer-Aided Design* 2012;**44**: 961–86.

[11] Liang Y, Tan R. A Text-Mining-based Patent Analysis in Product Innovative Process. *Trends in Computer Aided Innovation* 2007;**250**:86-96.

[12] Liang Y, Tan R, Wang C, Li Z. A Computer-aided Classification of Patents Oriented to TRIZ. *Proceedings of the IEEE International Conference on Industrial Engineering and Engineering Management* 2009;2389-93.

[13] Tong LH, Cong H, Lixiang S. Automatic classification of patent documents for TRIZ users. *World Patent Information* 2006;**28**:6-13.

[14] Cong H, Tong LH. Grouping of TRIZ Inventive Principles to facilitate automatic patent classification. *Expert Systems with Applications* 2008;**34**:788–95.

[15] Cong H, Tong LH. Pattern-oriented associative rule-based patent classification. *Expert Systems with Applications* 2010;**37**:2395–404.

[16] Cascini G, Russo D. Computer-aided analysis of patents and search for TRIZ contradictions. *International Journal of Product Development* 2007;**4**:52–67.

[17] Cascini G, Zini M. Computer-Aided Comparison of Thesauri Extracted from Complementary Patent Classes as a Means to Identify Relevant Field Parameters. *Global Product Development: Proceedings of the 20th CIRP Design Conference* 2010;555-66.

[18] Li Z, Tate D, Lane C, Adams C. A framework for automatic TRIZ level of invention estimation of patents using natural language processing, knowledge-transfer and patent citation metrics. *Computer-Aided Design* 2012; in press.

[19] Yoon J, Kim K. An automated method for identifying TRIZ evolution trends from patents. *Expert Systems with Applications* 2011;**38**:15540–8.

[20] Yoon J, Kim K. TrendPerceptor: A property–function based technology intelligence system for identifying technology trends from patents. *Expert Systems with Applications* 2012;**39**:2927–38.

[21] Verhaegen P-A, D'hondt J, Vertommen J, Dewulf S, Duflou JR. Relating properties and functions from patents to TRIZ trends. *CIRP Journal of Manufacturing Science and Technology* 2009; **1**:126–30.

[22] Choi S, Kang D, Lim J, Kim K. A fact-oriented ontological approach to SAO-based function modeling of patents for implementing Function-based Technology Database. *Expert Systems with Applications* 2012;**39**:9129–40.

[23] Montecchi T., Russo D. FBOS: Function/Behaviour–Oriented Search, In: *Proceedings of the TRIZ Future Conference 2011*, Dublin, Ireland.

[24] Walter L, Isenmann R, Moehrle MG. Bionics in patents – semantic-based analysis for the exploitation of bionic principles in patents. *Procedia Engineering* 2011;**9**:620–32.

[25] Verhaegen P-A, D'hondt J, Vandevenne D, Dewulf S, Duflou JR. Identifying candidates for design-by-analogy. *Computers in Industry* 2011; **62**:446–59.