![HAL open science]

Facilitating Engineers Abilities to Solve Inventive Problems Using CBR and Semantic Similarity

Pei Zhang, Denis Cavallucci, Zhonghang Bai, Cecilia Zanni-Merk

HAL Id: hal-02279761
https://inria.hal.science/hal-02279761

Submitted on 5 Sep 2019

Facilitating Engineers abilities to solve inventive problems using CBR and Semantic similarity

Pei ZHANG[1,3], Denis CAVALLUCCI[3], Zhonghang BAI[4], Cécilia ZANNI-MERK[2]

[1] CSIP @ ICube (UMR-CNRS 7357), 67084 Strasbourg Cedex, FRANCE
[2] LITIS, Norm@stic (FR CNRS 3638), INSA Rouen Normandie, 76800, France
[3] INSA de Strasbourg, 67084 Strasbourg Cedex, France
[4]Hebei university of Technology, 300131, Tianjin, China

Abstract. Our industry currently undergoes a period of important changes. The era of computerization implies to companies to change not only through their organization, but also in automating as much as possible their internal processes. Our research focuses on the computerization of the problem-solution couple when facing inventive situations in R&D. The method used is based on Case-Based Reasoning (CBR) that has already been proven to be useful in routine design. On the other hand, CBR is hardly used in inventive situations because the latter require reasoning outside the circle of knowledge recorded in a database. Our proposal consists in coupling CBR with semantic similarity algorithms. The aim is to resolve a new problem based on its semantic similarity with the old problems. Then the old solution can be adapted to solve the new problem. We postulate that a multidisciplinary case base sufficiently populated of multidisciplinary problem-solution couples is likely to considerably improve the performance of R&D engineers to solve inventive problems. This being possible by bringing them alternative solutions based on the semantically similar problems, which are distant from their field of origin. In this way, we provide the possibility to enhance the inventiveness of solution. This type of reasoning, largely inspired by the TRIZ theory, is the subject of this paper. The methodology, the experiments and the conclusions that we develop here validate that this type of approach produces the claimed effects on designers although limited to the context where it has been conducted.

1   Introduction

Our industry is moving through a period of important changes. The era of computerization pushes companies to change not only through their organization, but also through their internal physical functioning. Among the research performed around the notion of industry 4.0, a large part is dedicated to the computerization of workshops, machines, controls and computer flow that optimizes and responds to customer demand by its physical functioning. However, there is a sector of the company, which is only little researched in industry 4.0, its R&D department. As the most important stage that leads

to an inventive product, a methodology adopted by R&D departments is the theory of inventive problem solving (TRIZ) [1].

TRIZ offers various components for solving different types of inventive problems, such as the 40 Inventive Principles, the 76 Inventive Standards. According to the researches respectively carried by [2–4], the Contradiction Matrix with its 40 Inventive Principles and the Su-Field analysis with its 76 Inventive Standards are among the most popular tools used by the TRIZ community.

However, the "intellectual" cost for an enterprise to use a tool such as TRIZ, which is based on the users' experience, is exceptionally high. In addition, in the era of Industry 4.0, where the world is engaged in sharing data and automating the manufacturing technologies, R&D departments seem to be behind schedule in the way they automate problem solving.

Therefore, a legitimate axis of research is to find ways to facilitate the inventive problem solving process. Researches can be mainly categorized into two groups. One group addresses the problem solving model, and finding ways to ease the use of the models, while another group addresses on the knowledge sources used by TRIZ and attempts to facilitate the use of TRIZ by knowledge modelling.

The first group of researchers regard problem solving as an analogical process with the objective of finding inventive solutions to their problems [5]. This process is composed of two essential activities, with the help of TRIZ tools. One is the analogy reasoning between the problem and the problem model, the other is the analogy reasoning between the problem model and the solution model. To cope with the former problem, the work of [6] provided a way to systematically map design parameters with the Generic Engineering Parameters in axiomatic design. Moreover, to ease the use of the Contradiction Matrix, the authors in [7] matching considered the aspects regarding to the human factors issues of the Generic Engineering Parameters. To cope with the latter problem, ASIT [8] grouped 32 Inventive Principles into 5 thinking tools; the analysis conducted in [9] classified the Inventive Principles into clustered principles assuming the fact that there might be losses as compared to the use of TRIZ in a classical way.

The other direction of research addresses the aspect of taking advantage of the knowledge sources of TRIZ and knowledge modeling. The work of [10] proposed the TRIZ Technical System class ontology with its four sub-ontologies provided a framework that will enable the storage of knowledge found by other problem solving applications. The work of [11] connect the TRIZ knowledge sources with different abstract levels and designed rules to facilitate the problem solving process using ontology and its related rules.

Compared with the researches above, we propose an approach to collect and represent the problem-solution pairs from the problem solving know-how of experts (their professional life experience). In addition, we adopt Case-based reasoning to reuse available

solution at hand in order to support the problem solving process. This paper is organized as follows: section 2 discusses the importance of experience, section 3 details the proposed approach, section 4 validates the proposed approach by a case study and finally section 5 gives the conclusion and the future perspectives.

2   Is experience important?

With the aim of finding out how experience influences the problem solving, we did an experiment about how it can influence the time used for problem solving. We asked 28 students of the same grade in an engineering department of our university, and separate them into two groups with respectively 14 students: the experimental group and the control group.

For the experimental group, we designed a set of online forms with 10 simple cases and assigned them to 14 students. Each form contains the problem description and multiple-choice questions to guide students to find a solution model for each given problem. Once a student solves a problem and submits an online form, his/her problem solving time is recorded. For the control group, we ask the students to use the rule-based problem solving prototype [12] to solve the same 10 problems. The time used by each student for finding the solution model is automatically recorded by the prototype. The data we gathered concerning the time spent for problem solving is presented in Fig.1. The x-axis represents the average time used by the two groups using the assigned approach and the y-axis represents the ID of the problems.

As it can be observed in Fig.1, the average time used by the experimental group (the green bars) is decreasing as the students are solving more problems while the control group (the blue bars) remains stable. This result indicates that experience is a crucial factor influencing the time used of inventive problem solving, the more problems the students solve; the faster they tend to be. Furthermore, the students of the control group who uses the rule based approach; the time they use for solving the first problem until the last problem stays stable. This is because when using the rule-based approach, the problem solving is facilitated by if-then rules and the needed knowledge for each step, which makes the problem finding process easier. Therefore, the students in the control group use less time than the students in the experimental group for solving problem No.1. However, the problem with the rule-based approach is that every time the student solves a new problem, they have to go through the problem solving process from the beginning. In addition, this is why the problem solving time used by the control group stays balanced.

We can conclude that the collection and reuse of experience have the potential to improve problem solving, at least in terms of time. In addition, it is worth mentioning that in this experiment, we only ask the students to find the solution model in order to make the time used for problem solving comparable. In fact, the rule-based approach is not

able to provide the knowledge about the solution in its knowledge base. Therefore, the user with no experience would have trouble to interpret the solution model into a solution. Consequently, our research is focusing on finding a proper way to collect and reuse experience in the previous problem solving with the aim of providing users with the knowledge needed for solution finding.
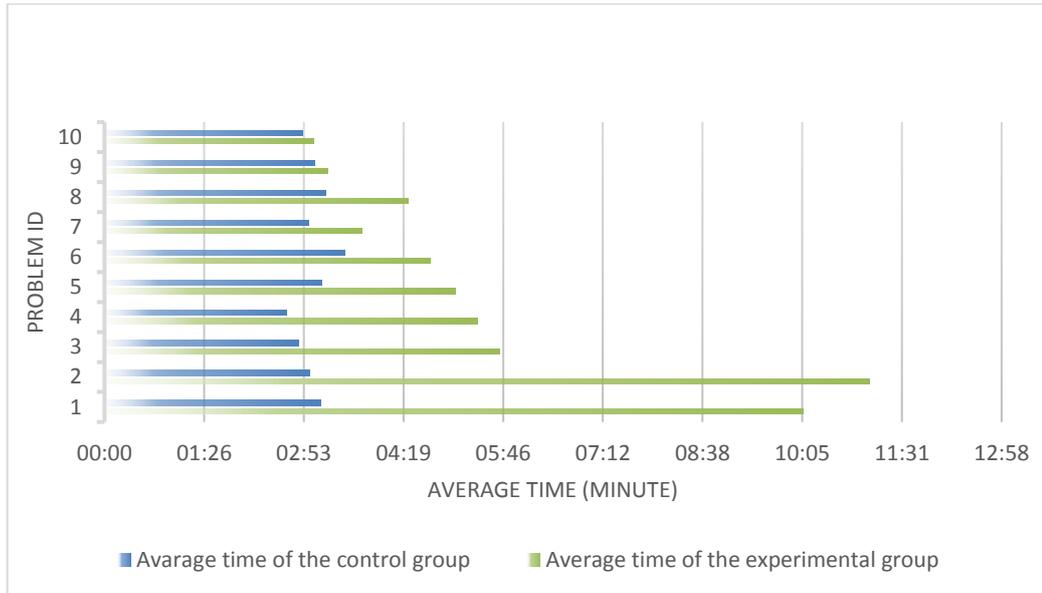

Fig. 1. Average time respectively used by the two groups

3 Facilitating inventive problem solving using Case-based reasoning and semantic similarity

3.1 Case-based reasoning

Case-based reasoning (CBR) is a methodology that has been widely adopted for problem solving and learning. It can benefit from the specific knowledge of the previous experience rather than making associations between general problems and its general solutions. In this way, when a problem needs to be solved, it is able to retrieve similar cases in the past and solves the new problem by reusing it in the new problem [13]. Applying case-based reasoning consists in retrieving a similar past case, reusing the past case and retaining the new case.

3.2 Semantic similarity

If we want to find a similar past problem to a new inventive problem, the problem is that they are expressed in terms of natural language. Depending on the way the problem

is described, the words used can vary but the meanings are similar. For example, something that is big and small is similar to something that is large and tiny. This type of problems are semantically similar but do not share a single word. In order to solve this problem, we need some external knowledge that enables us to find the semantically similar problems.

3.3 Facilitate inventive problem solving using Case-based reasoning and semantic similarity

With the aim of using the specific knowledge from experience to facilitate the inventive problem solving, we propose an approach using case-based reasoning and semantic similarity.

Restricted by the length of this paper, we only introduce here the resolution of a contradiction. The CBR approach comprises four activities: Case representation; Case retrieval; Case reuse and Case retain.

Case representation
First, we need to represent the case in order to retrieve similar cases. The case representation is introduced in the previous work of [14].

Table 1. Features to describe a case.

	Input Features	Output Features
Problem features	Value	
	Negative value	
	Action Parameter	
	Evaluation Parameter to improve	
	Evaluation Parameter to degrade	
Solution features		Inventive Principle
		Concept Solution

Case retrieval
In order to retrieve similar cases, we need to find semantically similar cases. Therefore, we adopted WordNet[1] as the knowledge base for calculating the semantic similarity between the new problem and the old problem. We apply the semantic similarity algorithm based on the short-text similarity proposed in [11].
To calculate the similarity consists in two steps: calculating the similarity between the new problem and the old problems and then calculating the weighted similarity between the new problem and the old problems. The former is composed of five sub-steps, which are:

---

[1] https://wordnet.princeton.edu/

1. Pre-processing: It includes a sequence of actions. Firstly, split the shot text into single words (for example "smoothing the surface" becomes "smoothing", "the" "surface"). Next, remove the stems of the words (for example, the words obtained in the previous step becomes "smooth", "the" and "surface"). Finally, eliminate stop words (words like the, of, or etc., are eliminated). The output of this step is a set of terms.
2. Sense search: For each term, we look for its corresponding senses using WordNet (e.g. the word "smooth" has eight different senses).
3. Sense similarity: We adopt Lin's method [15] to calculate the semantic similarity between two terms. Using this method, the higher rate of sharing information, the more the two terms are similar.
4. Term similarity: The maximum value of the sense similarity value between two terms is the term similarity.
5. Semantic similarity: Based on the obtained term similarity, we can calculate the semantic similarity. Let's assume that a new problem, P_1, includes a sequence of words P_{11} , P_{12} ... P_{1n} and an old case P_2 , includes a sequence of terms P_{21}, P_{22}... P_{2m}. $s(P_{1i}, P_{2j})$ represents word similarity between P_{1i} and P_{2j} , $1 \leq i \leq n$, $1 \leq j \leq m$. We can build the matrix of similarity $M (P_1, P_2)$:

$$\begin{bmatrix} s(P_{11}, P_{21}) & s(P_{11}, P_{22}) & \cdots & s(P_{11}, P_{2m}) \\ \vdots & & \ddots & \vdots \\ s(P_{1n}, P_{21}) & s(P_{1n}, P_{22}) & \cdots & s(P_{1n}, P_{2m}) \end{bmatrix}$$

In general, we select the most similar terms in P_2 for each term in P_1 and then we calculate the average value as defined in Equation 1:

$$\frac{\sum_{i=1}^{n} \max_{1 \leq j \leq m} (s(P_{1i}, P_{2j}))}{n} \tag{1}$$

Then, we apply the tf*idf method [16], [17] to assign weight to the obtained similarity. Applying the tf*idf method, we can obtain the word weight ww_{1i}, that is the *ith* word in P_1 and the word weight ww_{2i}, that is the *ith* word in P_2. In addition, to calculate the weighted similarity, we should apply the word weight to all possible situations. One is the most similar word in P_2 for P_1. The other is the most similar words in P_1 for P_2. In both situations, $s(P_{1i}, P_{2j})$ represents the word similarity between new problem P_{1i} and the old problem $P_{2j} (1 \leq i \leq n, 1 \leq j \leq m)$.

$$s_{weighted}(P_1, P_2) = \frac{\sum_{i=1}^{n} ww_{1i} \max_{1 \leq j \leq m} (s(P_{1i}, P_{2j}))}{n} + \frac{\sum_{i=1}^{m} ww_{2i} \max_{1 \leq j \leq m} (s(P_{2i}, P_{1j}))}{m} \tag{2}$$

Case reuse and retention
When the similar cases are retrieved, the next step is to reuse the retrieved cases. If the user finds an identical case and reuse its solution without adaptation, the solved case will not be retained. However, if the user selects a similar case and reuse its solution with certain adaptations, the new case will be retained in the case base.

4   Case study

Clothes hangers are used in our daily life to hang clothes vertically. Their design should make them versatile enough to fit for hanging clothes with different sizes. Therefore, length of the arm should be long enough to support the cloth, but short enough to ease their entering into the clothes without efforts. The traditional design of the cloth hanger is often in the shape of a triangle that has this type of fixed structure design has disadvantages. In this section, we apply the proposed approach to solve this problem.

First, we describe the problem in terms of a conflicting situation as the follows: *The length of the clothes hanger has to be both short to satisfy the ease of hanging and long to satisfy the clothes stability.* Then, we apply the semantic similarity calculating to find the similar problems in meaning. The retrieving results are illustrated in Table.1. We present in Table.1 the first 10 old cases similar to the new one. The similarity values are given in the first column in Table.1.

We can observe that the retrieved similar cases are similar in meaning rather than using the same words for describing the problem. For example, *large* is similar to *big*; *freedom of movement* is similar to *ease of hanging* and *length* is similar to *distance, and so on*. Among the similar cases we retrieved, suppose we want to reuse the solution of case 102, which is *making the arrangement of the shopping cart s-shaped*. Inspired by the solution, we can design the solution for the hanger problem as it can be seen from Fig.2. The design of the s-shaped hanger, inspired by the S-shape of shopping carts case, is an interesting way of solving our new problem of cloth hanger. On the one hand, if the cloth has a tight neck, we can first stick the hanger vertically by the long edge, and then rotate it to the horizontal position so that the short edge can be used to hang the cloth easily without damaging it. On the other hand, the long edge supports the cloth sufficiently so that it will not produce wrinkles.

Table 2. Retrieval results

Case	Similarity	Action Parameter	EP to degrade	EP to improve	Value	Negative Value
Case10	0.816	volume	ease of handling	hitting efficiency	large	small
Case2	0.752	size	vision	protection from rain	big	small
Case7	0.752	size	freedom of movement	comfort for cyclist	wide	narrow
Case9	0.739	length	precision	efficiency	long	short
Case102	0.728	distance	clearance of driveway	ease of storage	long	short
Case30	0.578	connection	cost	need of gas cooker	existing	missing
Case21	0.528	electrostatic	painting adhesiveness	less droplet rebound	existing	missing
Case1	0.509	volume	protection	weight	big	small
Case8	0.508	volume	less plastic consumption	storage volume	big	small

In this way, the application of the proposed approach facilitate problem solving by simply adapt the solution of the similar case that is retrieved.

Fig. 2. The shopping cart problem

Fig. 3. The designed solution of the clothes hanger problem

5   Conclusion and future perspectives

In this paper, we explored an approach to combine case-based reasoning and semantic similarity to facilitate inventive problem solving. Compared with the rule-based approach [18], the proposed approach provides the users with the specific knowledge from past successful cases. In this way, the user can solve the new problem by only adapting the old solution of the old problem. Therefore, the proposed approach can ease the problem solving process by reusing experience. Moreover, since the case base stores cases from different domains, the proposed approach can benefit from finding similar cases from different domains and reuse its solution.

As it has been discussed in Section 2, a large case base has the potential to increase the efficiency of the problem solving process. However, we are limited to a small case base and the future work will be dedicated to enlarge the case base with more cases.

References

[1] G. S. Altshuller, L. Shulyak, and S. Rodman, *The Innovation Algorithm: TRIZ, Systematic Innovation and Technical Creativity*. Technical Innovation Center, Inc., 1999.

[2] D. Cavallucci, "World Wide status of TRIZ perceptions and uses a survey of results," *Rep. TRIZ Future*, 2009.

[3] I. M. Ilevbare, D. Probert, and R. Phaal, "A review of TRIZ, and its benefits and challenges in practice," *Technovation*, vol. 33, no. 2, pp. 30–37, 2013.

[4] C. Spreafico and D. Russo, "TRIZ industrial case studies: a critical survey," *Procedia CIRP*, vol. 39, pp. 51–56, 2016.

[5] "Analogies Are the Way of Breakthrough Innovation," *The Triz Journal*. .

[6] J. R. Duflou and W. Dewulf, "On the complementarity of TRIZ and axiomatic design: from decoupling objective to contradiction identification," *Procedia Eng.*, vol. 9, pp. 633–639, Jan. 2011.

[7] D. A. Coelho, "Matching TRIZ engineering parameters to human factors issues in manufacturing," *Wseas Trans. Bus. Econ.*, vol. 6, no. 11, pp. 547–556, 2009.

[8] T. Toshio, "How people interact with objects using TRIZ and ASIT," *TRIZ J. Httpwww Triz-J. Comarc. Pdf*, 2003.

[9] H. Cong and L. H. Tong, "Grouping of TRIZ Inventive Principles to facilitate automatic patent classification," *Expert Syst. Appl.*, vol. 34, no. 1, pp. 788–795, 2008.

[10] P. Prickett and I. Aparicio, "The development of a modified TRIZ Technical System ontology," *Comput. Ind.*, vol. 63, no. 3, pp. 252–264, Apr. 2012.

[11] W. Yan, C. Zanni-Merk, D. Cavallucci, and P. Collet, "An ontology-based approach for inventive problem solving," *Eng. Appl. Artif. Intell.*, vol. 27, pp. 175–190, Jan. 2014.

[12] W. Yan, H. Liu, C. Zanni-Merk, and D. Cavallucci, "IngeniousTRIZ: An automatic ontology-based system for solving inventive problems," *Knowl.-Based Syst.*, vol. 75, pp. 52–65, Feb. 2015.

[13] A. Aamodt and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," *AI Commun.*, vol. 7, no. 1, pp. 39–59, 1994.

[14] P. Zhang, A. Essaid, C. Zanni-Merk, and D. Cavallucci, "Case-based Reasoning for Knowledge Capitalization in Inventive Design Using Latent Semantic Analysis," *Procedia Comput. Sci.*, vol. 112, pp. 323–332, 2017.

[15] D. Lin, "Extracting collocations from text corpora," in *First workshop on computational terminology*, 1998, pp. 57–63.

[16] G. Salton and M. E. Lesk, "Computer evaluation of indexing and text processing," *J. ACM JACM*, vol. 15, no. 1, pp. 8–36, 1968.

[17] K. Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval," *J. Doc.*, vol. 28, no. 1, pp. 11–21, 1972.

10

[18] W. Yan, C. Zanni-Merk, F. Rousselot, and D. Cavallucci, "A Method for Facilitating Inventive Design Based on Semantic Similarity and Case-Based Reasoning," *Procedia Eng.*, vol. 131, pp. 194–203, 2015.

Response to reviewers
Reviewer#1
Overall evaluation:
3: (strong accept)
Very interesting approach and a good demonstration.

Our answer:
Thanks for the comments.

Reviewer#2
Overall evaluation:
3: (strong accept)
Good paper. Addresses an important issue, proposes a solution and demonstrates it. Well structured.
There are some minor language mistakes. A review is suggested. Some mistakes were marked in the pdf file.

Our answer:
1. Page 2, 4th paragraph, line 4:
==While== another group addresses on the knowledge sources used by TRIZ and attempts to facilitate the use of TRIZ by knowledge modelling.

It has been changed to:
One group addresses the problem solving model, and finding ways to ease the use of the models, while another group addresses on the knowledge sources used by TRIZ and attempts to facilitate the use of TRIZ by knowledge modelling.

2. Page 2, 5th paragraph, line 12:
To cope with the latter problem, ASIT [8] grouped 32 Inventive Principles into 5 thinking tools; the analysis conducted in [9] classified the Inventive Principles into clustered principles assuming the fact that there might be losses as compared to classical TRIZ ==use.==

It has been changed to:
To cope with the latter problem, ASIT [8] grouped 32 Inventive Principles into 5 thinking tools; the analysis conducted in [9] classified the Inventive Principles into clustered principles assuming the fact that there might be losses as compared to the use of TRIZ in a classical way.

3. Page 2, 5th paragraph, first sentence:
==Another== direction of research addresses the aspect of taking advantage of the knowledge sources of TRIZ and knowledge modeling.

It has been changed to:
The other direction of research addresses the aspect of taking advantage of the knowledge sources of TRIZ and knowledge modeling.

4. Page 3, 2nd paragraph, line 4:

Once the students solve a problem and ==submits== an online form, his problem solving time is recorded.

It has been changed to:

Once a student solves a problem and submits an online form, his/her problem solving time is recorded.

5. Page 3, 3rd paragraph, line 7:

Furthermore, the students of the control group who uses the rule based approach; the time they use for solving the first problem until the last problem stays ==table==.

It has been changed to:

Furthermore, the students of the control group who uses the rule based approach; the time they use for solving the first problem until the last problem stays stable.

6. Page 4, section 3.1, line 4:

In this way, when a problem need to be solved, it is able to retrieve similar cases in the past and solves the new problem by reusing it in the new problem [13].

It has been changed to:

In this way, when a problem needs to be solved, it is able to retrieve similar cases in the past and solves the new problem by reusing it in the new problem [13].

7. Page 9, reference 1:

[1] G. S. Al′tshuller, L. Shulyak, and S. Rodman, The Innovation Algorithm: TRIZ, Systematic Innovation and Technical Creativity. Technical Innovation Center, Inc., 1999.

It has been changed to:

[1] G. S. Altshuller, L. Shulyak, and S. Rodman, The Innovation Algorithm: TRIZ, Sys-tematic Innovation and Technical Creativity. Technical Innovation Center, Inc., 1999.

Reviewer#3
Overall evaluation:
2: (accept)
The topic is relevant, the context of the study is well posed as well as the problematic however it could be interesting to detail a little more the experimentation during the oral presentation.

Our answer:

Thanks for the comments, we will address the experimentation details during the oral presentation.