Expert Systems With Applications 58 (2016) 143–154

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Flexible knowledge representation and new similarity measure: Application on case based reasoning for waste treatment

Philippe Chazara*, Stéphane Negny, Ludovic Montastruc

Universite de Toulouse, Laboratoire de Genie Chimique UMR CNRS/INPT/UPS 5503, BP 34038, 4 allee Emile Monso, Toulouse 31030 Cedex 4, France

ARTICLE INFO

Article history:
Received 24 October 2015
Revised 8 February 2016
Accepted 8 March 2016
Available online 1 April 2016

Keywords:
Case-based reasoning
Similarity measures
Ontology
Dynamic cases
Flexible case representation
Waste management

ABSTRACT

In Case Based Reasoning the representation of a case and the similarity measures are two difficult steps in the conception of a system. Often, these steps are developed to resolve one kind of problem. However, in some of them such as recovery treatment processes generation, it is necessary for the system to be able to modify and adapt the representation of a case and the similarity measures with respect of the context and also the kind of solutions proposed. In this paper, authors introduce a new method to represent cases with a flexibility based on a structure in a connectionist model. This flexibility is needed due to the complexity of cases, the number of possible options and to ensure the durability of the system. In a second main contribution, authors introduce a method for the selection of source cases using abstraction, conceptualisation and inference mechanisms. Finally, authors test their system in a CBR developed on SWI-Prolog with different problems. The CBR is applied to find new recovery processes and try to estimate the new upgraded product generated.

1. Introduction

The problem of waste and in particular the problem of waste management has increased sharply during the last decades, producing three kinds of effects. First, the problem of waste treatment is becoming more and more important due to the quantity produced with the increase of human population size and consumption. Second, the prices of some raw materials are growing sharply due to the phenomenon of depletion. It becomes more and more difficult to find new sources and their exploitation costs enhance. Third, the treatment of waste can have a strategic dimension. Actually, it can reduce the raw material dependency for some countries, it can develop new industries and create new jobs. But currently, waste is considered as a pollution source for environment and as a costly burden for companies because of the loss of material and the waste treatment. Consequently, it is necessary to propose new recovery processes and new ways to manage waste. However, some elements induce limitations. First, contrary to a new product, a waste has not essence by definition. Therefore, the first question is to find one or more essences for it. The second question is how to transform a waste into new valuable products. To solve these questions, authors propose to use an artificial intelligence system, and more particularly case based reasoning (CBR). CBR is relevant for this kind of problems because it allows solving problems without a clearly defined knowledge of the process needed for the resolution. The reasoning can rely on a vast number of cases, with their precise description of previous solved problems and their associated solutions (Cordier, Mascret, Mille, 2009). Secondly, in the domain of waste treatment, cases may contain different information: valorisation processes and essences for the new created objects. In the literature, case based reasoning systems are used in different waste treatment problems and in processes research. For example, López-Arévalo, Bañares Alcántara, Aldea, Rodríguez-Martínez, and Jiménez (2007) describe a tool based on CBR for the generation of process alternatives. Yang and Chen (2011) propose a classical CBR retrieve method used for Eco-innovation Kuo (2010) gives an example of CBR used to determine a recyclable index of some components. Liu and Yu (2009) use CBR for problems linked to environmental topic. Zeid, M. Gupta, and Bardasz (1997) propose a model dedicated to disassembling problems.

As detailed in Section 2, CBR method is decomposed in different steps: *Retrieval, Adaptation, Memorisation or Learning* as explained by Aamodt and Plaza (1994) and Napoli, Lieber, and Curien (1996), similarity measure is one key cornerstone of a CBR system and of

* Corresponding author. Tel.: +330534323663.
E-mail addresses: philippe.chazara@ensiacet.fr (P. Chazara), stephane.negny@ensiacet.fr (S. Negny), ludovic.montastruc@ensiacet.fr (L. Montastruc).

http://dx.doi.org/10.1016/j.eswa.2016.03.014
0957-4174/

the *Retrieval* part in particular. This measure allows finding close and relevant cases to solve the new problem. Therefore, with our goal to reuse the knowledge related to recovery methods for new waste valorisation, it is important to propose new approaches for this step respecting the constraints imposed by this category of problems.

On this topic, authors tackle several problems related to the similarity question. The first one is how to represent a case and more particularly for the domain of application, how to represent a waste. There are many kinds of waste and they need different representations. Moreover, the domains of waste and waste treatment have an important dynamic. Indeed, these domains change quickly i.e. the composition of waste, or the waste treatment processes evolve over time. To take into account these points, it is necessary to develop a flexible case representation to ensure a precise description of problem, knowledge reuse, CBR system efficiency and durability. Another consequence of theses points is that the problems of waste cannot be considered as routine problems. However, CBR systems are developed to resolve only routine problems i.e. problems which are very similar. Consequently, a system used for these kinds of problems need to go beyond this limitation by the introduction of flexibility. Another point is how to take into account that there are different possibilities of valorisation for a same waste. For example, in the case of used tyres, they can be burnt to produce energy, reused as tyres, transformed by crunching into material for different kinds of new object, transformed by fermentation to produce syngaz. For each solution, the same description parameters are not selected: for some solutions is the chemical composition; for other ones is the form or the functionality, for other ones mechanical properties. Therefore, as showed by Lieber (2002), problems and their solutions depend on their use. As a consequence, authors think that problem representation and similarity measure depend on the solution or the kind of solution targeted.

In this paper, authors propose to explain their methods for representing knowledge and cases, and for selecting relevant cases. These methods try to take into account the solution and therefore to adapt the similarity measure in function to the important parameters according to a kind of solution. Moreover, these methods do not produce a metric value of distance or similarity measure but, it determines if a case is similar to the current problem or not, i.e. if the case can be used to generate an original solution for the problem. Contrary to Perner (2003), the method is not based on graphs, and it does not use threshold or other metric value, but it is based on logical deductions. In conclusion, the major contributions of this paper are the following:

- The introduction of a flexible representation for knowledge.
- A dynamic construction of cases, which allows going beyond the limitation of routine problems.
- A new method for similarity measure, without calculation and with a limited need of knowledge.

In the remainder of this paper, the Section 2 explains some elements about CBR systems and develops some ideas for the realisation of each step finding in the literature. In Section 3, the proposed flexible representation of a case is described and more specifically the management of the knowledge is explained. Then, the core of this method is introduced with the presentation of the main assumptions, and the retrieve part is described step by step in Section 4. The Section 5 highlights the method capabilities through a case study, where some tests have been realised to assess the proposed method. Section 6 issues opinions about the positive points and the limitations of the method, and underlines some difficulties met during its implementation. Finally, Section 7 draws conclusions and summarises the presented work, and proposes different perspectives to improve it.

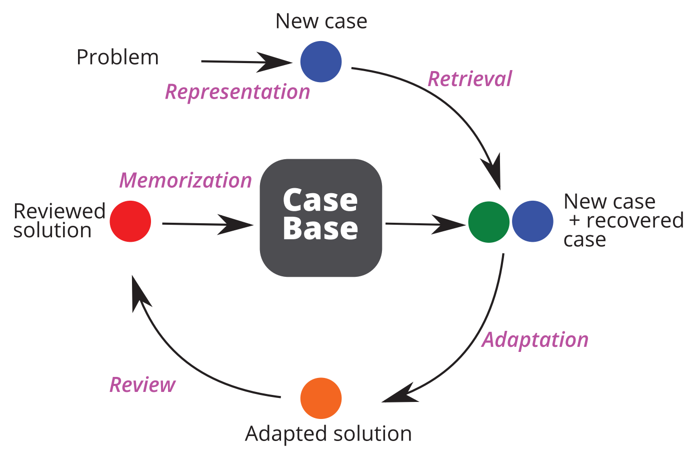

Fig. 1. Steps in classical CBR.

2. Case-based reasoning : different related steps

As explained in the introduction, a CBR system is based on different steps (each of them decomposes in sub processes not detail here) (Reyes, Negny, Robles, & Le Lann, 2015) (Fig. 1).

However, the realisation of one step impacts all the CBR's processes. The representation of the knowledge or cases impacts the sub-processes in the retrieval step, for example the similarity measure or the mapping phase. Therefore, it is necessary to represent knowledge by taking into account that retrieval step uses it, i.e. the definition of all the sub-processes depends on the choice of a kind of representation. Finally, the last sub-process of this retrieval step of the CBR is the selection of the relevant case in order to revise its solution to match to the target case requirements. One mechanism used is the analogy. Cornuéjols (1996) has studied the fundamental of this mechanism. He defined analogical reasoning as the way to find the expression which allows passing from a previous problem to its solution and to apply it to a new target case. Here too, the representation of cases is important.

In traditional CBR, the knowledge is often represented as a set of spaces. Napoli et al. (1996) explain that there is a space for the problem and another one for their solution. Mougouie and Bergmann (2002) define a query in CBR system as a point in these spaces. Therefore, each point of these spaces has to be represented with a common method. Kokinov (1994) explains that a cognitive mechanism is based on representation, memorisation. In CBR and in general for all artificial intelligent systems, representation is only a partial description of the reality. As a consequence, Mougouie and Bergmann (2002) explain that a query is only partially described. For Peschl, it is an interpretation of the world which allows the construction of a behaviour (Peschl & Riegler, 1999). Under this idea, Amailef and Lu (2013) link an ontology to a CBR system to facilitate the understanding of a situation and the retrieve step. This interpretation is very important in the resolution phase as Richard highlights because a modification of the interpretation can improve the efficiency of solving methods (Richard, 1979). Finally, representation can be symbolic, based on connexions (Kokinov, 1994), defined as vector features, or complex as semantic network (Branting & Aha, 1995). Whatever, the manner to represent knowledge, it is a reduction of the reality. But, the choice of the representation approach impacts the similarity measure step. For example, Branting and Aha (1995) and Garey and Johnson (2002) explain that the utilisation of semantic network for the representation of cases in CBR causes that the mapping step is