Integrations of Rule-Based and Case-Based Reasoning

Jim Prentzas, Ioannis Hatzilygeroudis
Dept. of Computer Engin. & Informatics, University of Patras
26500 Patras, Hellas (Greece)
&
Research Academic Computer Technology Institute
Patras, Hellas (Greece)

ABSTRACT

Rule-based and case-based reasoning are two popular approaches used for problem solving in intelligent systems. They are natural alternatives in knowledge representation. Rules usually represent general knowledge, whereas cases encompass knowledge accumulated from specific (specialized) situations. Each approach has advantages and disadvantages. Due to their interchangeable nature, rules and cases can be easily integrated and thus produce effective knowledge representation formalisms, surpassing the disadvantages of each component method working individually. In this paper, we present a survey of approaches integrating rule-based reasoning with case-based reasoning. We describe a categorization scheme for the integrated formalisms and briefly present representative approaches of each category.

Keywords: Hybrid intelligent systems, rule-based reasoning, case-based reasoning, rule-based and case-based integrations.

1. INTRODUCTION

The integration of (two or more) different knowledge representation methods is a very active research area in Artificial Intelligence [24]. The aim is to create hybrid formalisms benefiting from each of their components. It is generally believed that complex problems can be easier solved with hybrid systems. The effectiveness of the various hybrid approaches has been demonstrated in a number of application areas [4], [5], [21], [22], [27], [30]. In most of the hybrid approaches, two knowledge representation methods are being integrated. This is due to the fact that the integration of three or more knowledge representation methods is more complicated.

One of the most popular types of integration involves the combination of rule-based with case-based reasoning approaches [2], [12]. The efforts to combine symbolic rules and cases have yielded advanced knowledge representation formalisms. The effectiveness of those approaches stems from the fact that rules and cases are alternatives in representing application domains and solving problems. Rules represent general knowledge of the domain, whereas cases specific knowledge. Rule-based systems solve problems from scratch, while case-based systems use pre-stored situations to deal with similar new instances. Therefore, the integration of both approaches turns out to be natural and useful.

In this paper, we present a categorization of the approaches integrating rule-based with case-based reasoning and describe representative approaches for each of the categories. Till now, to the authors' best knowledge, a survey of such hybrid approaches including the most recent advances is missing. A serious effort is presented in [14], but it is not quite recent.

This paper is organized as follows. Section 2 briefly describes the main characteristics of rules and cases and presents the three primary categories of the hybrid approaches. Each of the subsequent three sections (Sections 3, 4 and 5) describes approaches belonging to each of the primary categories. Finally, Section 6 concludes.

2. INTEGRATION CATEGORIES

Symbolic rules constitute a popular knowledge representation scheme used in the development of expert systems. Rules represent general knowledge of the domain. They exhibit a number of attractive features such as naturalness, modularity and ease of explanation. One of their major drawbacks is the difficulty in acquiring them. The traditional process of eliciting rules through the interaction with the expert may turn out to be a bottleneck causing delays in the system's overall development. Furthermore, the acquired rules may be imperfect and not covering the full complexities of the domain. Rule induction methods deal with these disadvantages but may still be unable to recognize exceptions in small, low frequency sections of the domain. Finally, rules may contain terms that are not well defined (e.g. in legal domains).

Case-based reasoning offers some advantages compared to symbolic rules and other knowledge representation formalisms [1], [15], [18]. Cases represent

specific knowledge of the domain. Cases are natural and usually easy to obtain. New cases can be inserted into a knowledge base without making changes to the preexisting knowledge. Incremental learning comes natural to case-based reasoning. The more cases are available the better the domain knowledge will be represented. Therefore, the accuracy of a case-based system can be enhanced throughout its operation, as new cases become available. A negative aspect of cases compared to symbolic rules is that they do not provide concise representations of the incorporated knowledge. Furthermore, the time-performance of the retrieval operations is not always the desirable.

Approaches integrating rule-based and case-based reasoning have given interesting and effective knowledge representation schemes [2], [12]. The goal of these efforts is to derive hybrid representations that augment the positive aspects of the integrated formalisms and simultaneously minimize their negative aspects. In [14] the approaches integrating rule-based and case-based reasoning are distinguished into two basic categories: efficiency-improving and accuracy-improving methods. The former concern integration methods in which rules and cases are dependent, meaning that one representation scheme was derived from the other (i.e., rules derived from cases or vice versa), and the efficiency of the integrated scheme exceeds the efficiency that could have been achieved with rules or cases alone. The latter involves approaches in which the two representation schemes are independent and their integration results in improved accuracy compared to each representation scheme working individually.

A disadvantage of this categorization scheme is the fact that it may be restrictive to future hybrid methods. A more general categorization scheme may be more proper to classify the hybrid approaches. Most of the recent papers describing methods integrating rules and cases imply or mention a general approach to classifying other hybrid methods. In this paper, we distinguish three basic categories of the approaches integrating rule-based with case-based reasoning. The categorization is based on the importance of each of the two component schemes in the inference process.

The first category includes approaches in which the rule-based component prevails in the inference process, whereas the case-based component plays a complementary role. The approaches belonging to this category usually focus on the rule-based component and invoke the case-based component only when rules are unable to deal with specialized situations. We call the approaches belonging to this category, *rule-dominant approaches*. Examples are [13]-[14], [30]-[31], [19]-[20], [16]. The second category consists of approaches in which the role of the integrated components is balanced, which means that none of the integrated components plays a supportive role. We call these approaches, *balanced approaches*. Example such approaches are [25]-[26], [7]-[10], [4], [5]-[6], [3], [11]. The third category consists of approaches in which the case-based component plays a more important role and the rule-based component is less significant. In this paradigm, the rules play a supportive role to case-based reasoning, useful for instance when the case library contains a limited number of cases. The approaches belonging to this last category will be referred to as *case-dominant approaches*. Examples are [17], [27], [22]-[23].

Roughly speaking, the inference pattern of the rule-dominant or case-dominant approaches is somewhat predefined as far as the invocation order of the integrated components is concerned. Rule-dominant approaches usually invoke first the rule-based component and call the case-based component when there seem to be difficulties in reaching conclusions. Case-dominant approaches, on the contrary, focus on the case-based component and invoke the rule-based component only when assistance is needed in order to find a solution or to improve the solution proposed by the case-based component. Balanced approaches, on the other hand, follow a 'mixed' paradigm, where the invocation order of the integrated components is not preset and usually during inference one component dynamically calls the other and vice versa.

Legal reasoning seems to be the most popular application field for the approaches integrating rules and cases. Such examples are GREBE [7]-[10], CABARET [25]-[26], IKBALS II [30]-[32]. This is because legal reasoning concerns rules containing terms, which are not well defined and need an integration with cases to reach (or enhance) a conclusion. These terms are called open-textured terms. Other application fields are medicine [17], [4], [5], [6], surname pronunciation [13]-[14], part-of-speech tagging [19]-[20], music [27], design of nutrition menus [22]-[23]. In the following, some of the most popular integrated approaches are briefly described.

3. RULE-DOMINANT APPROACHES

ANAPRON is one of the classical integration methods. It involves integration of independent rules and cases in an innovative way. The idea is to use cases to enhance a set of symbolic rules, which are only approximately correct. Cases are used only to amend the conclusions reached by rules. The existence of domains represented by incomplete or partially correct rules is quite often. On the one hand, cases are used as exceptions to symbolic rules, filling their gaps in representing domain knowledge. On the other hand, symbolic rules perform indexing to cases, facilitating their retrieval. So, this approach results in accuracy improvement of the rule-based component and efficiency improvement of the case-based component.

Inference focuses mainly on the symbolic rules. Firing a rule is suspended when a sufficient number of its conditions are satisfied. Its exception cases are checked and if a similar case is found with adequate 'weight', the rule is not allowed to fire and the conclusion proposed by the case is considered valid instead. Cases supporting rules are also employed in order to assess whether the similarity between the new case and an exception case is compelling.

ANAPRON constitutes a general architecture for integrating rule-based with case-based reasoning consisting of a number of modules [14]. This approach has been used in an application field defining the pronunciation of American surnames. This can be a difficult task due to the diverse national origins of the surnames. Experimental results have demonstrated the effectiveness of the integration since ANAPRON approximates the performance of commercial systems in the domain.

A general approach similar to ANAPRON is described in [28]-[29]. Once again, the knowledge base contains rules representing standard situations and cases representing exceptions or non-standard situations. The inference process first checks the rules and then the cases if the rule-based approach does not give any solution. The contents of the knowledge base are produced from an initial case base whose cases are split into two types: standard cases and exception cases. Standard cases are used to induce the rules of the knowledge base by employing a standard rule induction method. Splitting the initial case base is performed using heuristic approaches.

The advantages of the approach as demonstrated by various experiments are the acceptable accuracy of the inference process, the good explanatory ability and the convenient knowledge acquisition process. The explanatory ability is due to the high level of comprehensibility of the rules. This is based on the fact that the induced rules from the standard cases are closer to expert rules than the rules produced from the whole dataset of cases (standard and exceptional). However, as has been shown in [29] the splitting policy of the initial case base plays an important role in the accuracy and comprehensibility levels of the approach.

IKBALS II is an approach applied to a legal application domain [30]-[32]. It extends a previous object-oriented rule-based system (i.e. IKBALS) with the use of cases. Cases are independent from rules. More specifically, the case library contains exceptional cases for which the rules have proven inadequate. Thus the system's reasoning is primarily rule-based and cases are used when rules run out or prove insufficient in drawing conclusions. Rules correspond to knowledge extracted from statutes. IKBALS II also contains structures to facilitate the interleaving of rules and cases, the retrieval of cases and the performance of analysis tasks (e.g. what if situations).

The approach described in [19]-[20] integrates first-order rules with cases in order to improve the overall accuracy. It is used for part-of-speech-tagging (i.e. morpho-syntactic disambiguation). Rules are induced from examples in an iterative procedure. The quality of the induced rules degrades from iteration to iteration. When the quality of induced rules falls below a threshold, case-based approach is used with the examples not covered by the induced rules. A learning process is employed for the cases as well. The learning process produces sets of explanations for each case corresponding to different views. The explanations are used during case-based reasoning to analyze the similarity between a new case and a preexisting case. In this way, the case-based reasoning process of the approach is differentiated from standard case-based reasoning.

The system in [16] combines cases with neurules. Neurules are a type of hybrid rules integrating symbolic rules with neurocomputing. In this way, an effective scheme combining three types of knowledge representation formalisms (i.e. symbolic rules, neural networks and cases) is created. The integration scheme is similar to ANAPRON. However, the use of neurules instead of symbolic rules enhances the inference performance, because neurule-based inference is more efficient than simple rule-based reasoning, acting as the indexing component of case-based reasoning. The integration of neurules with cases improves their accuracy, when the symbolic rules, acting as source knowledge to the neurule base, do not cover the full complexities of the domain.

4. BALANCED APPROACHES

GREBE [7]-[10] is a representative of this second category. It is an approach described in Branting's PhD Thesis. This scheme attempts to achieve a coherent interleave of rules and cases in order to effectively handle legal reasoning. At any level of the inference process it can invoke the rules or the cases of the system.

CABARET [24]-[25] is an approach dealing with legal reasoning as well and can also be classified into the first category consisting of rule-dominant approaches. The architecture consists of two co-reasoners, the rule-based component and the case-based component having an equal status. There is also a controller, which observes the operation of the whole system and each co-reasoner separately and decides how they will proceed in the reasoning process as a whole and individually. The controller assigns tasks to each co-reasoner.

The approach described in [5]-[6] proposes a multimodal reasoning framework for the close

integration of the different knowledge base entities. The rules and cases are described using a special representation language. The condition and action part of the rules have the same representation with the problem situation and solution part of cases respectively. In this way, during inference, the knowledge base can be searched in parallel for applicable rules and cases. Pattern matching and case-based retrieval is performed in parallel and the conflict set may simultaneously contain rules as well as cases. Conflict resolution chooses the most applicable entities of the conflict set according to the similarity with the new case and the type of the entities. Therefore, a reasoning cycle tightly integrates the different knowledge base entities. This approach has been applied to a medical domain and more specifically to post-transplant patient care. A Web-based system has been developed for this purpose.

The approach described in [4] integrates rules and cases in an innovative way. The approach has been applied to a medical domain and more specifically to diabetic patient management. The rule base of the system contains different classes of rules. The innovative aspect is the ability to dynamically adapt rules belonging to specific classes in order to improve handling of a new situation. Refinement of the rules is performed with the use of cases and involves certain parameters of the rules, which are too general to deal with the specific situation. General structured knowledge is used to retrieve and adapt cases from the case base. If no case is found to be applicable for a specific situation (meaning there is a gap in case-based knowledge), inference uses only the rules and the case base is updated using the produced outcome. Therefore, the rule-based and the case-based components assist each other during inference. The integration makes the system more effective in detecting the patient's problems and making proper prescriptions reducing the time required to resolve the patient's problems.

5. CASE-DOMINANT APPROACHES

GYMEL [27] is a system for harmonizing melodies. Rules are invoked when the cases cannot produce a solution. The approach is useful in application domains for which it is difficult to acquire an adequate set of cases and the case-based reasoning component needs to be backed up by a rule-based component expressing general knowledge. In such an approach, the invocation frequency of the rule-based component will be high at the early stages of the system's operation. Subsequently however, it will decrease, as new cases will be incorporated into the case base.

CAMPER [22]-[23] is a nutritional menu planner built by combining the best features of independent case-based reasoning and rule-based reasoning menu planners, CAMP and PRISM respectively. Nutritional menu planning is a difficult task because there are many numeric constraints some of which conflict with others, menus can be evaluated only if they are entirely constructed and common sense must employed for combinations of foods that go or do not go together. CAMP and PRISM were evaluated and compared in order to locate their deficiencies and strengths. This analysis guided the construction of CAMPER. By and large, the case-based reasoning component constructs menus that are acceptable since they satisfy multiple nutrition constraints. However, the rule-based component can enhance the proposed menu by its creativity in considering new possibilities and performing 'what if' analysis. Therefore, in contrast to GYMEL, case-based reasoning always produces an output that is subsequently improved by the invocation of rules. Moreover, as in GYMEL, a significant reason for the usefulness of the integration is the difficulty in the acquisition of cases.

6. CONCLUSIONS

Rule-based and case-based reasoning are alternative ways of expressing knowledge. Approaches integrating the two formalisms have become popular in the last years. The hybrid approaches have managed to solve successfully problems in application domains where rules and cases are available and each representation formalism needs the assistance and/or completion of the other to work effectively. This trend is very likely to carry on the following years.

REFERENCES

[1] A. Aamodt and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches", *Artificial Intelligence Communications*, 7(1), pp. 39-59, 1994.

[2] D. Aha and J.J. Daniels (eds.), "Case-Based Reasoning Integrations: Papers from the 1998 AAAI Workshop". Technical Report WS-98-15, AAAI Press, 1998.

[3] A. An and N. Cercone, "Multimodal Reasoning with Rule Induction and Case-Based Reasoning" in [12], pp. 118-123.

[4] R. Bellazzi, S. Montani, L. Portinale and A. Riva, "Integrating Rule-Based and Case-Based Decision Making in Diabetic Patient Management" in *Proc. 1999 International Conference on Case-Based Reasoning*, pp. 386-400.

[5] I. Bichindaritz and K.M. Sullivan, "Reasoning from Knowledge Supported by More or Less Evidence in a Computerized Decision Support System for Bone-Marrow Post-Transplant Care", in [12], pp. 85-90.

[6] I. Bichindaritz, E. Kansu, and K.M. Sullivan, "Case-Based Reasoning in CARE-PARTNER: Gathering Evidence for Evidence-Based Medical Practice", in

Proc. 1998 European Conference on Case-Based Reasoning, pp. 334-345.

[7] L.K. Branting, "Reasoning with Portions of Precedents" in *Proc. 1991 International Conference on Artificial Intelligence and Law*.

[8] L.K. Branting, "Rules and Precedents as Complementary Warrants", in *Proc. 1991 National Conference on Artificial Intelligence*, pp. 3-9.

[9] L.K. Branting, "Building Explanations from Rules and Structured Cases", *International Journal of Man-Machine Studies*, 34(6), pp. 797-837, 1991.

[10] L.K. Branting, *Reasoning with Rules and Precedents*, Dordrecht: Kluwer Academic Publishers, 1999.

[11] N. Cercone, A. An, C. Chan, "Rule-Induction and Case-Based Reasoning: Hybrid Architectures Appear Advantageous", *IEEE Transactions on Knowledge and Data Engineering*, 11(1), pp. 164-174, 1999.

[12] E. Freuder (ed.), "Multimodal Reasoning: Papers from the 1998 AAAI Spring Symposium". Technical Report SS-98-04, AAAI Press, 1998.

[13] A.R. Golding and P.S. Rosenbloom, "Improving Rule-Based Systems through Case-Based Reasoning", in *Proc. 1991 National Conference on Artificial Intelligence*, pp. 22-27.

[14] A.R. Golding and P.S. Rosenbloom, "Improving accuracy by combining rule-based and case-based reasoning", *Artificial Intelligence*, 87(1-2), pp. 215-254, 1996.

[15] J. Kolodner, *Case-Based Reasoning*, San Mateo, CA: Morgan Kaufmann Publishers, 1993.

[16] J. Prentzas and I. Hatzilygeroudis, "Integrating Hybrid Rule-Based with Case-Based Reasoning", in S. Craw and A. Preece (Eds), Advances in Case-Based Reasoning, Procs 2002 European Conference on Case-Based Reasoning, LNAI 2416, Springer-Verlag, pp. 336-349.

[17] P. Koton, "Reasoning about Evidence in Causal Explanations", in *Proc. 1988 National Conference on Artificial Intelligence*, pp. 256-263.

[18] D.B. Leake (ed.), *Case-Based Reasoning: Experiences, Lessons & Future Directions*, AAAI Press/MIT Press, 1996.

[19] A.A. Lopes and A. Jorge, "Combining rule-based and case-based learning for iterative part-of-speech tagging", in Proc. 2000 European Conference on Case-Based Reasoning, pp. 26-36.

[20] A.A. Lopes and A. Jorge, "Integrating Rules and Cases in Learning via Case Explanation and Paradigm Shift", in Proc. 2000 IBERAMIA-SBIA, pp. 33-42.

[21] J. MacIntyre, P. Smith, and T. Harris, "Industrial Experience: The use of Hybrid Systems in the Power Industry", in [23], pp. 57-74.

[22] C.R. Marling, G.J. Petot, and L.S. Sterling, "A CBR/RBR Hybrid for Designing Nutritional Menus", in [12], pp. 152-156.

[23] C.R. Marling, G.J. Petot, and L.S. Sterling, "Integrating Case-Based and Rule-Based Reasoning to Meet Multiple Design Constraints", *Computational Intelligence*, vol. 15, pp. 308-332, 1999.

[24] L.R. Medsker, *Hybrid Intelligent Systems*, Kluwer Academic Publishers, 1995 (second printing 1998).

[25] E.L. Rissland and D.B. Skalak, "Combining Case-Based and Rule-Based Reasoning: A Heuristic Approach", in *Proc. 1989 International Joint Conference on Artificial Intelligence*, pp. 524-530.

[26] E.L. Rissland and D.B. Skalak, "CABARET: Rule Interpretation in a Hybrid Architecture", *International Journal of Man-Machine Studies*, 34(6), pp. 839-887, 1991.

[27] J. Sabater, J.L. Arcos, and R. Lopez de Mantaras, "Using Rules to Support Case-Based Reasoning for Harmonizing Melodies", in [12], pp. 147-151.

[28] J. Surma and K. Vanhoof, "Integrating Rules and Cases for the Classification Task", in *Proc. 1995 International Conference on Case-Based Reasoning*, pp. 325-334.

[29] J. Surma and K. Vanhoof, "An Empirical Study on Combining Instance-Based and Rule-Based Classifiers", in [12], pp. 130-134.

[30] G. Vossos, J. Zeleznikow, Th. Dillon, and V. Vossos, "An Example of Integrating Legal Case-Based Reasoning with Object-Oriented Rule-Based Systems: IKBALS II", in *Proc. 1991 International Conference on Artificial Intelligence and Law*, pp. 31-41.

[31] J. Zeleznikow G. Vossos, and D. Hunter, "The IKBALS Project: Multi-Modal Reasoning in Legal Knowledge Based Systems", *Artificial Intelligence and Law*, 2(3), pp. 169-203, 1994.

[32] J. Zeleznikow and D. Hunter, *Building Intelligent Legal Information Systems: Knowledge Representation and Reasoning in Law*, Kluwer Computer/Law Series, vol. 13, 1994.