Training of the Learner in Criminal Law by Case-Based Reasoning

Simon Bélanger, Marc-André Thibodeau, and Esma Aïmeur

University of Montreal
Department of Computer Science and Operational Research
C.P. 6128, succ. Centre-ville
Montreal, Canada H3C 3J7
Tel: (514) 343-6794; Fax: (514) 343-5834
[belanger, thibodea, aimeur]@iro.montreal.ca

Abstract. FORSETI is a system that uses Case-Based Reasoning applied to Criminal law. It has two functional modes: the **expert** mode and the **tutorial** mode. The first mode makes use of its knowledge in order to resolve new cases, i.e. to determine a sentence (imprisonment, delay before request for parole, etc.). Canadian jurisprudence of similar cases to the one presented constitutes the basis for FORSETI's judgement. An expert may also consult the case base in order to corroborate the result obtained and supply the base with new cases. The second mode is an educational tool for professionals in Criminal law. In virtue of this mode, two types of exercises are possible. The first one focuses on developing the user's judgement in determining sentences and the second one on improving his jurisprudence analysis. FORSETI uses an experimental method for adaptation that we call *planar interpolation*. The goal of this method is to improve the level of consistency in the knowledge base and produce significant results in adaptation of new cases.

Keywords: Training, Case-Based Reasoning, Tutorial systems, Criminal Law, Sentence.

1. Introduction

In the course of his everyday life, a human being is often confronted with problems that he can solve by using his memory by looking into his past to find situations that are similar to the one to which he is confronted. These situations all contain, at least in part, knowledge required to the resolution of a new problem. This illustrates the basic idea of a relatively new field of interest in artificial intelligence research called ***Case-Based Reasoning*** (or CBR) [4]. Its goal is to resolve problems in a particular field of

I. Imam et al. (Eds.): IEA/AIE-99, LNAI 1611, pp. 398-408, 1999.

activity. The two basic problems in CBR are the indexing and the definition of an appropriate heuristic for estimating the similarity between two or more cases. To this end, it requires a base of cases from the past linked to their solutions. When it is confronted with a new problem, it searches in this base for the cases that contain the most similarities, then attempts to adapt their solutions in order to create one which will satisfy the aspects of the new situation.

It is possible to apply this Case-Based Reasoning to an array of different fields. Law (JUDGE) [2], medicine (CASEY) [5], architecture (ARCHIE) [7], design (NIRMANI) [8] and military planning (Battle Planner) [3] are only a few examples. Their development has been growing rapidly, as is shown by the systems described by [11].

FORSETI[1], the system which we have conceived, is specialized in the field of Criminal law. Its first objective is to determine sentences for crimes on the basis of cases similar in jurisprudence. Its second, is to help lawyers further their analysis of criminal cases in a continuing development of Criminal law decision-making. FORSETI uses Case-Based Reasoning to these purposes.

Our work has consisted in conceiving and developing the FORSETI system and to establish a case base linked to jurisprudence. First, this document presents the basic concepts of the theory on Case-Based Reasoning, such as representation of cases, indexation and adaptation of new cases presented to the system. This last method which is adaptated presents a new resolution approach called planar interpolation. Secondly, it is an incursion into the field of Criminal law. Finally, we will discuss with more emphasis the tutorial function of FORSETI.

2. The Elaboration of FORSETI

This section presents the basic principles which have guided the elaboration of FORSETI and explains why Case-Based Reasoning has been chosen for its realization. This segment ends with a brief presentation of Criminal law in Canada.

2.1 Use of CBR for Training

Jean Piaget, an epistemologist, was the one to unite the two great philosophies of innatism and empiricism in laying the foundation of the theory of Constructivism [10]. Today, the theory of Constructivism is the most credible, well-rounded and accepted theory in the field of education. It integrates the new concept of construction of knowledge. For Piaget, it is by his interaction with his environment that a human being elaborates the mental structures which permit him to assimilate the reality that surrounds him and enable him to reflect on this reality. Therefore, forming an individual implies putting him in contact with a problematic situation, then bringing him to integrate this situation in his memory (assimilation) and to modify his internal cognitive structures (accommodation).

[1]In Scandinavian mythology, Forseti is the God of Justice

This conception of learning is exactly the one adopted by Case-Based Reasoning. It aims at permitting the system learning by the construction of knowledge base. This construction is realized by presenting new situations to the system to which it must react conveniently. If let doesn't do so, then the system must adapt its internal structures. It finally integrates this situation and the newly acquired knowledge in its database in order to be able to use them in the future, when other situations occur.
For example, the intelligent tutorial systems HYPO [1] and SARA [9] use Case-Based Reasoning for training. The first is used for the training of students while the other is applied to the teaching of mathematics.

2.2 Principles behind CBR

Representation of cases: A case is, conceptually, a piece of knowledge representing a certain experience. Typically, a case should contain:

- The state of the problem in its original context
- The solution to the problem or the new situation of the case when the problem has been resolved.

Cases can be represented in a variety of forms including frames, objects, predicates, semantic networks and rules.
Indexation: Each case posseses different attributes. Some are used to index the case in order to limit and accelerate research of similar cases, whereas others are used to characterize the cases. *Indexation based on difference* and *indexation based on explanation* are two examples of methods that are commonly used.
Search of similar cases: To extract similar cases from the database, two general methods exist: *algorithm of the nearest neighbor* and *method by induction*.
Adaptation: It is a central step of Case-Based Reasoning. It consists of finding a solution for a new case by modifying the solution of a similar case. The manner in which the adaptation is realized can become a difficult problem. It is in fact one of the major stakes of CBR for which there has been enormous research. David Leake is one of the researchers brought attention to the study of adaptation. Substitution and transformation of solutions are methods used for adaptation [6].

2.3 Criminal Law

Criminal law aims at the repression of certain conducts, law enforcement and the protection of public wealth. It is characterized by a consequence; a sentence which can be a fine, an interdiction or an ordinance.
FORSETI is applied to Criminal law, more precisely to crimes against the person under the Canadian Criminal code. All cases used for the FORSETI system have been extracted from the Quebec Penal law jurisprudence index which provides abstracts of decisions in Quebec Penal law. The justice system must conserve a coherent link between present and past sentencing. Therefore it is of most importance to analyze the judicial system in order to identify the most influent criteria that lead to the judgement of a case. These are the seven principles that must guide the courts in the imposition of a sentence in Canada:

- Effective gravity of an infraction with regard to the law,
- Effective gravity of the infraction,

- Subjective gravity of the infraction, criminal record, age and reputation of the accused.
- Frequency of the same category of crime,
- Circumstances extenuating and increase,
- Rehabilitation of the accused,
- Salutary or exemplary effect of the imposed appeal.

3. The FORSETI System

Java, an Object-oriented programming language, was used in the realization of the FORSETI system. Object-oriented programming permits a better structure and helps maintain systems which use a high complexity level. Furthermore, Java brings the advantage of portability to FORSETI, permitting it to be executed on any given platform.
The description of a case is constituted by thirteen descriptive attributes, two indexation attributes (the type of crime and the plea) and four solution-oriented attributes (the sentence, the delay, the extra fine and the interdiction). The attributes of age, sentence and delay are the only numeric attributes, the others being hierarchic (for example, the precision «spouse» concerning the attribute of victim can only exist if the precision «known» is already attributed to the victim). Furthermore, certain attributes can be of different levels of importance.

3.1 Calculating Similarity

The algorithm used to calculate the degree of similarity between two cases is a variation of the *closest neighbor algorithm*. It consists of comparing the value of the descriptive attributes of two cases while taking note of the different levels of importance given to each attribute. These levels of importance have been determined with the collaboration of a law expert and reflect the importance of the attribute in the decision of the court (see section 2.3).

3.2 Adaptation

The adaptation of new cases consists of determining the value of the four solution-oriented attributes of a new case by examining similar cases in the database.

Planar interpolation The idea is to interpolate a numeric value by the equation of a plane in a tridimensional space. Finding such an equation implies the knowledge of the coordinates (x, y, z) of three non-aligned locations in a given space. These three locations will each correspond to a case chosen for the adaptation of the solution. The method for determining the coordinates of a location based on a case is relatively simple. The user that has just entered a new case to adapt chooses two attributes that he judges the most important, and that should most influence the decision of the judge in this particular case. It should be noted that it is possible to make another choice if the adaptation obtained is not satisfactory. These two attributes become x and y. It is

then necessary to calculate the value of these variables x and y, for each case, according to the value of these attributes. To this end, we have determined, with the assistance of a lawyer, a scale of «gravity» which indicates the relative gravity of different possible values for the four most important attributes: circumstances, motive, consequences on the victim and character of the victim. For example, an «accidental» circumstance will be much less important on the scale of gravity than the «planning of a crime».

Then, we must determine the z axis, the one relative to the attribute for which we have decided to interpolate the value. For example, we can interpolate the length of imprisonment. In this case, the number of months of imprisonment in similar cases will be the z coordinate of the corresponding points.

Example: If we take three locations (3,10,80), (8,8,90) and (10,5,80). The obtained equation is 50x + 70y - 11z = -30 and the corresponding interpolation plan is presented by Figure 1.

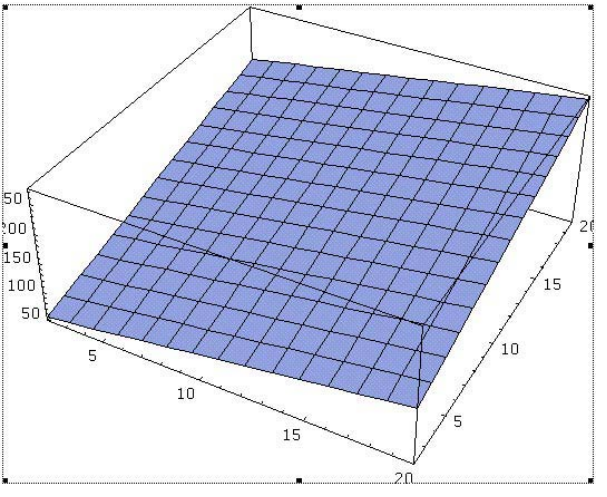

Fig. 1. Interpolation plan

The result obtained for the value of the attribute of the new case also considers the value of this attribute in the three most similar cases, instead of only one case, as in JUDGE. This solution cannot contradict those of the three cases used for interpolation, thus enhancing the coherence of the decisions taken.

It is also interesting to observe the potential of this method. Two important improvements could be made. First, the number of cases used for interpolation could be increased arbitrarily. Many interpolation methods that are more complex are known to use more locations. It is for instance the case of the *Collocation surfaces method* which consists of finding a polynomial whose degree varies according to the number of locations. This polynomial determines a surface which passes through all the locations and gives an interpolation whose realism increases with the number of locations used.

A surface obtained with this method for nine locations can look for example like Figure 2.

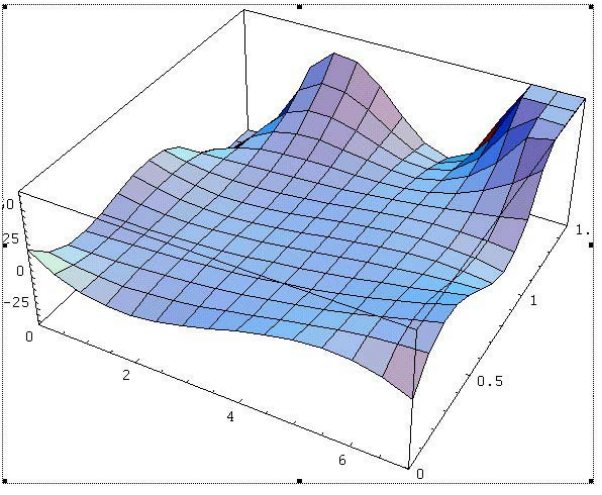

Fig. 2. Collocation surface calculated with nine points

It would be possible to use all similar cases in interpolation, maximizing at the same time the coherence of the derived solution with the one of similar cases. But the problem is as follows: implementing the Collocation surfaces method in a non-functional programming language such as Java is a complex task which has not yet been realized. It would certainly be an interesting extension to our research.

The second improvement that could be made would be to use more attributes. This would translate into increasing the dimension of the interpolation space. For example, we could, with four axes, interpolate a value in a space with four dimensions by using three attributes instead of two. More locations would obviously be necessary, as well as more similar cases for each equation. This can again be generalized at *n* dimensions. The implementation of the «spatial» interpolation which would result would be of a certain interest, since the use of all the attributes for the adaptation would certainly give even superior results. Nonetheless, this goes beyond the objective of this research.

Adjusting parameters When the attribute of the solution to be adapted possesses a numeric value, but that there are less than three similar cases found, then an adjustment method of parameters is launched. This method considers four factors we have already submitted. The degree of gravity is calculated for each of those four attributes by using the same scales as for planar interpolation. The gravity of each attribute is then compared to the corresponding attribute in the most similar case. The parameter is increased if the gravity is superior, decreased if the gravity is inferior.

3.3 Validation of the System

The quality of the results obtained with FORSETI has been measured qualitatively by tests made on a database of a hundred cases. The first test aimed at verifying the calculating of similarity. It consisted in calculating the similarity of each case with all the other cases of the database. This test revealed that a relatively small importance should be given to attributes with non-determined values so that the value of the calculated similarity is not biased.
The second test was used to evaluate the quality of the adaptation. We have adapted the solution of each case of the database as if it was a new case by using other cases of the base as reference. We then compared the solution obtained with the real solution, for each case. The principal information furnished by the test is an indication on the adjustment of attributes and mainly, on the importance in the choice of locations for the planar interpolation method. It is important to choose, as locations, attributes that for the adapted case have played a determining role in the judge's decision-making.
An evaluation of the system performance based on a comparison between the solutions given by FORSETI and the solutions given by the expert is underway.

4. FORSETI as Tutorial

FORSETI also aims at bringing the user that is not an expert to exercise his judgement and to help him in understanding jurisprudence. The lawyer, in the exercise of his work, must frequently search for similar cases to the one he is defending. It is of great importance for him to develop an ability to rapidly extract permanent information from jurisprudence. Furthermore, he must train to evaluate cases in order to determine which sentence would be acceptable for each of them. FORSETI possesses all the information necessary to help the training of a law professional. As such, in tutorial mode, this system offers two types of exercises to the user.

4.1 First Type of Exercise

FORSETI presents jurisprudence to the learner as it could be found in indexes. The learner is then invited to analyze texts taken from jurisprudence and to determine which attributes are most important, by filling in a form (Figure 3). Once filled, the form is assimilated by the system as a new case. This case is compared to the case corresponding to the text which has been presented. This comparison is made possible by calculating the similarity between the cases. If the similarity obtained is high, then the description created by the user is close to the one made by the expert and which is contained in the system. The user is finally brought to reflect on each difference encountered between the two cases. As such, the system questions the learner on the value that he has given to litigious attributes. Has he given this value simply because he has seen a certain term used in a case? Has he considered the context in which the term has been used in the text? Could he have regrouped a few terms by a more general term, for example «troubles-personality(aggressive) + troubles-personality(dangerous)» by «troubles-personality(criminal)? Or on the other hand, should he have given more previsions, for example by replacing the value «physical-

sequels» by «physical-sequels(major-injuries)» for the attribute of «consequence»? In these terms, FORSETI's goal is to bring the learner to reflect on his mistakes.
Furthermore, the system displays the percentage of efficiency of the learner so that he can be aware of his level of learning and of his progression. This percentage of efficiency is calculated simply: *S* being the maximal similarity level that can be obtained and *s* the level of similarity between the case and the description given by the user, then the percentage of efficiency is:

$$T = s/S * 100$$

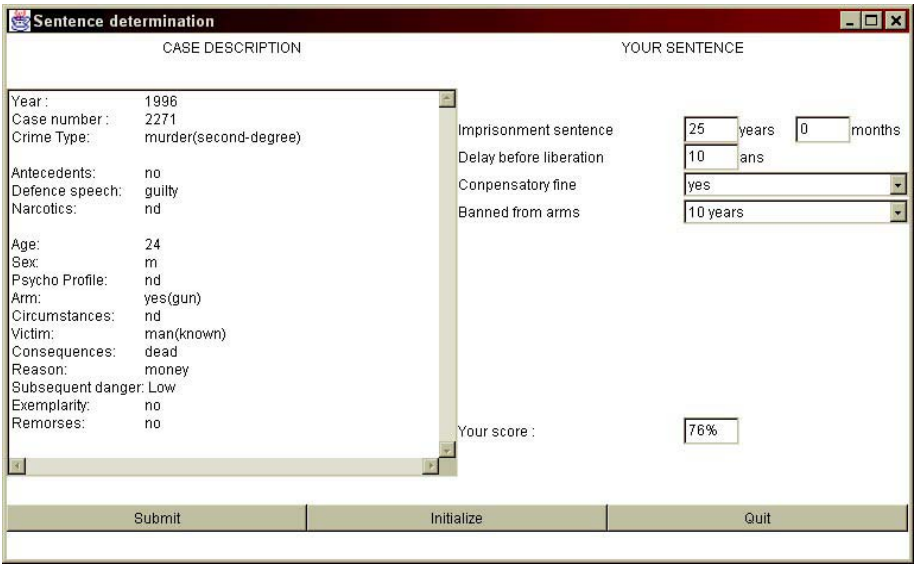

Fig. 3. First type of exercise

4.2 Second Type of Exercise

FORSETI presents cases chosen randomly in a base of unresolved exercises. Theses cases possess solutions, contained in a «solutions» folder, but are not part of the database from which FORSETI works. The learner must then, after analyzing the attributes of the case, determine a solution, i.e. enter a punishment that he finds appropriate (Figure 4). The system then tries itself to complete the exercise. It proceeds to the adaptation of the case, as if it was a new case, and also determines a solution. The three solutions are then presented to the learner: his own, FORSETI's, and the expert's (the real judgement that was given for the case, as described in the «solutions» folder). The success obtained by the system is an indication of the difficulty of the exercise and the efficiency of the user.
In fact, if FORSETI determines a solution which is far from the real solution, it is either that there are only a few similar cases in its database, or that the case is exceptional, having obtained a judgement that contradicts those of similar cases (in this case, the level of difficulty of the exercise is very high). The learner should be

encouraged if his solution is closer to the real solution than the one FORSETI has put forward. This brings us to determining a method for measuring the difference δ between two solutions. We calculate this difference by a function:

$$\delta = w_p \max(p_1, p_2)/ \min(p_1, p_2) + w_d \max(d_1,d_2)/ \min(d_1,d_2) + sw_s + iw_i$$

where w_k determines the relative importance of each attribute in a solution, according to the expert. p_k represent the values determined for the sentence. d_k are values determined for the delay before request for parole. *s* is the difference between the value for the estimated compensatory extra fine and the real value (*s* equals 0 if this value has been determined correctly, 1 if not) and *i* is the difference between the values of interdiction to possess firearms (i = number of levels of difference for the interdiction. For example, «life» and «5 years» have two levels of difference, so *i* equals 2).

As such, a value of δ is calculated for the difference between FORSETI's solution and the real solution, and another for the difference between the real solution and the learner's.

The learner's efficiency rate is in this case calculated as such:

$$T = \delta_F/(\delta_F + \delta_a) * 100$$

where δ_F is the error in FORSETI'S solution and δ_a the error in the learner's solution. A high value for δ_F (and so of the difficulty of the exercise) will bring the *T* rate closer to the maximum (100%) while a high value of δ_a will have the contrary effect. With this formula, an exact answer from the learner gives a rate of 100%. The *T* rate considers simultaneously the learner's answer and the level of difficulty of the exercise.

As in the first type of exercise, the learner is brought to reflect on his mistakes. This time, the tutorial gives him the opportunity to consult similar cases extracted from the database during adaptation. As such, the user can see if the judgement he has made corroborates the reality of jurisprudence. He is also questioned on the reasons for his false reasoning. Each learning session can be saved at any time by the user so that he can pursue it at another moment.

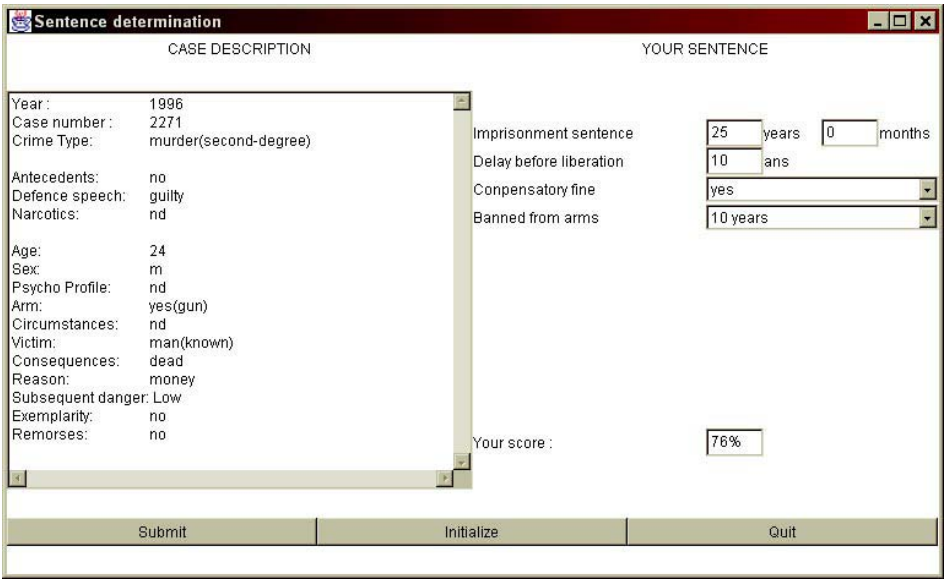

Fig. 4. Second type of exercise

5 Conclusion

FORSETI is a system that uses Case-Based Reasoning capable of determining sentences in cases of Criminal law. The cases presented in the system are those from 1984 to 1997. Its second function is the development, by exercises, of the ability of law professionals to analyze jurisprudence.

The system possesses characteristics which could be the object of future improvements, in particular with respect to the automation of the determination of the importance of the attributes. Nonetheless, FORSETI innovates by using an experimental method of adaptation for the parameters of new cases. This method helps the system conserve a high level of consistency and coherence in decision-making with prior cases. It could be generalized to be used for other fields of law.

6 References

1 Aleven, V., and Ashley, K.D. 1992. Automate Generation of Examples for a Tutorial in Case-Based Argumentation. In Proceedings of the Second International Conference on Intelligent Tutoring Systems. Berlin. Springer-Verlag.

2 Bain, W. 1989. Inside Case-Based Reasoning. Northvale, ed. C.K. Riesbeck and R.C. Schank.

3 Goodman, M. 1989. CBR in Battle Planning**.** In Proceedings : Workshop on Case-Based Reasoning (DARPA). pp.576-584, San Mateo, California.: Morgan Kaufmann.

4 Kolodner, J. 1993. Case-Based Reasoning. SanMateo. : Morgan Kaufmann Publishers.

5 Koton, P. 1989. Using Experience in Learning and Problem Solving. Ph.D. diss., Dept. of Computer Science, MIT, Boston.
6 Leake, D.B. 1996. Case-Based Reasoning, Experiences, Lessons & Futures Directions. Manlo Park. : American Association for Artificial Intelligence.
7 Pearce, M., Ashok, K.G., Kolodner, J., Zimring, C., and Billington, R. 1992. A Case Study in Architectural Design., IEEE Expert.
8 Perera, R.S., and Watson, I. 1995. A Case-Based Design Approach for the Integration of Design and Estimating, In Progress in Case-Based Reasoning. Berlin. : Springer-Verlag.
9 Shiri, M., Aïmeur, E., and Frasson, C. 1998. Student Modelling by Case-Based Reasoning, ITS 98 Conference, 4th International Conference on Intelligent Tutoring Systems. pp.394-403, San Antonio, Texas, 1998.
10 Wadsworth B.J. 1989. Piaget's Theory of Cognitive and Affective Developmen. New York. : Longman.
11 Watson, I. 1997. Applying Case-Based Reasoning : Techniques for Enterprise Systems. San Francisco, California. : Morgan Kaufmann Publishers, Inc.