Expert Systems for the Dissemination of Legal Knowledge among the General Population

Louis-Claude Paquin, Centre ATO•CI
Claude Thomasset, François Blanchard, GRID
Robert Dupuis, dept. math. info

Université du Québec à Montréal, C.P. 8888 succ. A
Montréal (Québec) CANADA H3C 3P8

Abstract. Our research which aims to build an expert system prototype, Loge-expert, to communicate legal knowledge on *Québec Housing Law* to non-lawyer users and to produce an ongoing evaluation of the effects of legal knowledge computerisation on the legal system and on the users of legal services. This contribution aims to briefly present remarks on three questions 1) the knowledge elicitation; 2) the communication of legal knowledge to the general public; and 3) the evaluation of the resulting expert system.

1. The general context

The legacy of a social democratic government is one of several public institutions created to soften the impact of a liberal economy on the people. Besides *Unemployment Insurance*, in Québec we have *Social Welfare* to provide a minimum income to those who are poor and unable to work, *Medicare* to give everybody adequate health care, *Work Standards* to avoid the exploitation of workers, *Consumer Protection*, among others. In terms of its case load, the *Rental Board* is the most important tribunal (70,000 to 90,000 housing actions are adjudicated each year).This is not surprising; according to the 1986 Canadian Census, 78% of the population of the City Montreal are tenants.

Behind all these institutions there are laws and regulations determining eligibility to the programs, solving conflicts, etc. To take advantage of the services, the general population needs information about its rights and redresses. The burden of the public deficit and repeated recessions have made the government decrease the budget of these institutions in such a way that their information services, mostly provided by phone have been severely cut. In search of alternative means of providing information, we have turned to the computer.

This is the general context of our research which aims to build an expert system on *Québec Housing Law* - Loge-expert[1] - and to produce an ongoing evaluation of the effects of legal knowledge computerisation on the legal system and on the users of legal services. This research, initiated in 1986, has been funded by the Canadian Donner Foundation, the Social Sciences and Humanities Research Council of Canada (SSHRCC), and the «Ministère de l'Enseignement supérieur et de la Science du Québec», 'Étalez votre science' program.

This contribution aims to briefly present three questions that arose during the construction of Loge-expert and the solutions we devised. The questions are: 1)

knowledge elicitation; 2) the communication of legal knowledge to laymen; and 3) the evaluation of the resulting expert system. Those questions seem to be inherent to the construction of expert systems within legal or administrative domains where the goal is to disseminate information to the general public.

2. Knowledge elicitation

The process of knowledge elicitation in building an expert system has been well documented for technical domains[2]. It is far less the case for managerial domains because most of the knowledge source lies on texts. Generally speaking, the management is to be accomplished from a written policy. If this policy is complex, it comes with documents giving explanations. When it is governmental management of law, the scope of the textual corpus is broad: law, regulations, decree, tribunal decisions, etc. In our case, the Québec Housing Law is to be read with the lights of the decisions, specially to figure out the proper acceptation of the terms of the law.

To deal with this textual material efficiently, we conceived a database to make the law, relevant regulations and most of the decisions available for information retrieval. That database intervened at different stages in the building of the expert system. At the input stage, it provides definition of key concepts, support for inference rules. At the output stage, pertinent decisions are used to give the user explanations and contextualizations.

None of the decisions were available in a computer-readable format and few were published. Under an agreement with the tribunal, we received for two years most of the decisions related to our field of interest. Some were translated into computer-readable format by means of an optical character recognition system and these decisions can be accessed in a "free text" mode way. Relevant contexts are retrieved by a Boolean request formulated with the words of the texts.

This technology substantially speeds up the early stages of the reading process and guarantees thoroughness, but offers no help for content interpretation. The key concepts cannot be extracted directly from the text surface without interpretation. Far beyond linguistic process, this cognitive process involves comparative analytic and classificatory operations of the retrieved contexts, the result of which is put into relation with the relevant legal conceptual hierarchy. The practical knowledge of this legal conceptual hierarchy is only acquired by proper training.

Furthermore, the necessary legal concepts and rules still need to be clarified. On the one hand, more legal knowledge is usually required than that specifically related to the question dealt with. For example, even if a residential lease is an exception to the principles of the general theory of obligations, we have to include some of the latter's concepts in order to modelize the rent payment question. On the other hand, since sometimes no definition is provided in the legal texts, one has to specify the meaning of some legal concepts in a very restrictive way. For example, the concept "landlord" is problematic in situations related to rented dwelling repossession; a specific meaning had to be derived from reviewing relevant contexts extracted from the tribunal decisions database.

The knowledge elicitation challenge is first to isolate the key concepts and the reasoning among the linguistic and textual systems. Once the concepts and the reasoning have been isolated, an out of context reconstruction into frames and

inference rules. Unfortunately, data found in the textual material, isn't structured in neat little zones and fields marked with unambiguous delimiters. The structure is a textual one, in words, phrases, sentences, etc.

Questions arise. What is specific to a text ? to a legal text ? How textual analysis assisted by a computer could help knowledge elicitation from textual sources ?

What is specific to a text?

A text is both inflexible and unsettled: the text is an entanglement of systems that could take several forms. We only find in it what have been written: words and punctuation. A minimal definition of a text could be the following: an ordered set of segments written in a natural language recorded on a material (paper or magnetic). A more substantial definition of text should include and present as such the set of linguistic systems, namely: the voicing system (phonological); the outside world referential system (lexical); the internal structure and forms system (morphological) and the organisation and relationship of words within groups, phrases, clauses and sentence system.

The higher levels of the linguistic model are less defined. For instance, the semantical level is thought as a kind of calculus on lexical properties of the words and/or their morpho-syntactic position in the segment. It is important to take note that the gradual complexity is for didactic presentation. Under real life conditions we have to deal with the unavoidable entanglement of the systems. For example, it is virtually impossible to automatically select between two or more potentially contradictory surface categorisation without an throughout description of the deep structure of the text.

After more than 30 years of research in the filed of automatic natural language processors, J. Sowa, an IBM System Research team member stated that:

> "... the successes of language processors on small domains and their failure on unrestricted domains result from the fundamental nature of language. In particular, a large grammar and dictionary are not sufficient to scale up a small system to an unrestricted natural language processor"[3]

Problems arise when those systemic descriptions intersect higher or at other levels of description we called textual. Among the descriptions of proper textual systems, there is the figure of speech, the network of argumentation, the communicational environment, etc. Furthermore, not only a minimal knowledge of the particular universe of the text is needed, but also the reader has to be familiar with the social conventions according witch the text had emerged. This dimension of the texts make it to "decode" beyond its linguistic structures.

A text is both inflexible and unsettled. Its life is unpredictable: it could be destroy; it could be duplicated; it could be quoted in another text with or without indication. This evanescent trajectory is called interdiscursivity and should be taken into account. There is several form of text depending on its aim: report, study, directive, free-text answers to opinion pool, interview's retranscription, etc. In the legal domain, the texts have characteristics of their own.

What is specific to a legal text?

We should be aware of some characteristics of the legal texts, characteristics that have been already pointed out by different authors. For instance, both theoreticians (computer scientists or logicians that tackled the field of law as a mean to test their research program) and practicing lawyers that had tried to develop a “machine-like” understanding of legal texts have come across several difficulties. For those dealing with statutes, it became apparent that: “the structure of the text of the statute is no longer irrelevant but dictates the nature of the formalisation[4]” and that the “definition of a concept” may have been distributed “ across a number of fragments”. Others[5] still, have tried to describe the interactive play of cognitive agents and the legal corpora involved in the legal reasoning process.

The reading as an act of knowledge elicitation

We have seen that the knowledge doesn't lay at the text surface. The only way to overcome the lack of certainty concerning the meaning of a text and still use the computer is to include the reader into the model of the meaning construction process. This intuition is confirmed both by the psychological and sociological latest theory stating that an univocal meaning is not set down in the text but constructed by the reader through his cognitive structures and his socialisation. But reading alone prove to be an inconsistent knowledge elicitation device at the low level of recognition; we need the computer regularity in complex pattern matching situation to improve the reader making of better judgmental choice. Gathering information is what human perform worst, make inferences; take decision on uncertain facts is what human perform best. In this perspective, the computer accumulates every context needed by the expert-reader to construct a meaning by correlating information gathered throughout the corpus. This model is attractive because it includes the implicit possibility of a partial integration of the reading "expertise" (with the knowledge engineering inquiry techniques ?).

How texual analysis assisted by a computer could help knowledge elicitation?

What could a computer do to simulate legal expert reading patterns? There are two reading patterns, according to types of legal documents a legal expert is confronted with. If the legal document is law or regulation, the reading pattern includes the analyse of contexts and co-texts of the written document in order to establish the specific meanings of words expressed in a legal language. If the legal document is a court or tribunal decision, we are confronted with a multi-dimensional analysis of that kind of legal discourse. The reading pattern has to cope with the argumentative nature of a decision, which implies that the judge has to make explicit the reasons which justify his decision. The legal expert usually finds all the references of the judgement motivation in the decision. He has not to search for them because the adversarial procedure facilitates the expression of the pro and contra arguments during the proceedings. The judge has to consider them and to make his own opinion in order to write his decision. The contexts and co-texts are included in the decision which give it its multi-dimensioned size.

For the knowledge elicitation process purpose, we need to identify legal concepts and common sense words which characterise specific situations; we need also to isolate the particular reasoning. Since our aim is not the exhaustive description of all the

textual systems, but the dissociation of the information from it's enunciation conditions, textual analysis with both its linguistic and discourse sides offers a suitable theoretical framework. The control of the analysis should not be let to the automata. By taking the exclusive point of view of the author's (legislator for instance) intentions, the analysis will be normative. To avoid this bias, the reader must have the control over the sequence and the modulation of textual manipulations and measurements.

The knowledge elicitation process covers the transformation of textual data into knowledge expressed into logical forms suitable for the expert system calculation. The communication of the knowledge from the expert system to the user will be covered in the next section.

3. Communication of the legal knowledge

The goal of making a legal expertise accessible to the general public adds two additional difficulties to the knowledge-based system building. First of all, the user should be motivated to undertake his search for legal expertise without the requirement of any technological knowledge. The general population is not interested in learning how to use a computer to get access to legal expertise. We want people to have access to legal expertise through devices like automated tellers at the bank. The problem is that much more wording is necessary for this type of consultation than to get money or to pay the electricity bill. Secondly, besides the job of legal knowledge modelisation and formalisation, it must also be translated into plain language.

To a certain extent, communication of legal knowledge to the uninitiated could be considered as a special field in itself. In the lawyer's interaction with his or her client, a double translation occurs. First, some facts are extracted from the "story" expressed in plain language by the client and translated into legal language (qualification). Afterwards, the legal reasoning is carried out and the results are translated back into plain language understandable by the layman. Here we present our thoughts on this communication process and the (partial) solutions we implemented in the Loge-expert prototype.

Questions arise. Up to what extent the legal reasoning process should be dependent to the communication process ? Should the user be aware of the legal expertise *per se*? How to tailored the communication process to the user ? What is the optimal technological implementation ?

Up to what extent the legal reasoning process should be dependent to the communication process ?

We made the working hypothesis of the modularity between the “legal nucleus” of the expert system and the “communicational layers” which convey plain language explanations. The aims of those two modules are totally different. The “legal nucleus” of the expert system aims at a simulation of skilled research and analysis strategies upon legal knowledge laying in tightly intermingled textual sources. The reasoning in terms of inference rules, the question asked of the user and the legal statement production are made within this "legal nucleus". The explanations written in plain language aim at making legal concepts understandable by non-lawyers as well as at encouraging learning of legal language in order to familiarise them with it.

The modularity principle enabled us to go on concomitantly with the formalisation of legal concepts in legal language in order to improve the Loge-expert knowledge base, as well as with the writing of explanations in plain language of the legal concepts. The building of the expert system could than be divided between two teams, each having its own speciality: knowledge engineering for the legal nucleus, linguistic and psycho-sociology of communication for the communicational layers. A specific solution could then be coined according to the problems particular to each of those components.

Should the user be aware of the legal expertise *per se* ?

The explanations written in plain language aim at making legal concepts understandable by non-lawyers as well as at encouraging learning of legal language in order to them familiarise them with it, should they consult a lawyer later on with regard to their conflictual situation. Therefore, we could not avoid the necessity of having them available to the user, but nor could we sacrifice the constraints of legal language by translating it in plain language. Non-lawyers should have access to alternative explanations according to their personal level of understanding of legal language without interfering with the legal reasoning or with the speed of the answers.

According to this principle, as many communicational layers as needed should be optimally tailored to targeted users. The tenant association user, for example, has a good background in the legal concepts which concern him or her and none at all for the relative to the opposite position. Nevertheless it might be difficult to determine which legal concepts are controlled by a specific class of users. The grouping of users on the basis of the similarity and divergence of the need expressed may be difficult.

How to tailored the communication process to the user ?

In order to succeed in matching human being, legal knowledge and computer, the socio-cultural characteristics of the users have to be taken into account in the ergonomics of building the communicational layers. We know very little about the general public as users. We need to understand their behaviour when they are seeking information. We should pay attention to multiple human factors such as the user' cultural background, his values, his level of skill. Prerequisites for the use of those systems, such as training to run them or to understand the wording of a given answer, should be finely tuned to the expectations of the user, since this could be as important a factor in the system's success as the consistency of the legal nucleus.

The ergonomics of a given system should not be only a technological issue but also a linguistic one because even if the system is made easy to work, if the wording of its objects, actions and instructions are not easily understandable, it could not be considered user friendly. Thus the complexity and the diversity of the wording of the concepts, questions and statements involved should be reduced to a minimum. The vocabulary and the syntax should have been previously normalised. The wording should be designed with state of the art techniques developed on "plain language" and readability of legal language

What is the optimal technological implementation ?

Instead of following a deterministic path giving systematically the user plain language information about what is going on during an expert system session, we adopted an

"upon user request" approach. That means the communicational layer is virtually accessible at anytime upon request from the user. The technology of hypertext appeared suited for the implementation of the communicational layer. We designed a restricted hypertext device called HYPER_AID; an illustration of a screen is given after the end of this text. It is restricted because the given explanations for concepts, questions, and messages are divided into four fields:

1) a plain language translation and a succinct definition;
2) some references to complete the previous definition;
3) some decisions to give a contextualization; and
4) links with a short list of related concepts.

Activated by a hot key or by pointing to a specific icon, HYPER_AID displays a card presenting textual or graphical information in a specific window answering to such usual requests as: Why is the system asking me this question ? How does the system draw this conclusion ? What does this concept mean ? Mapping the legal concepts, questions and statements with their plain language elicitation is done simply by means of a table of strings.

But the HYPER_AID covers only one part of the needs: the explanation in plain language of questions asked by the system and the results both expressed in legal language. The printing of a brief report was added. This report summarises the legal consultation and the correlative explanation. It could be a useful tool for the user to remember the consultation, to compare with the results of other inquiries, to share the results with others, or to facilitate mediation.

4. Evaluation of the expert system

The building of a system expert in the Housing Law domain has for main purpose an extensive evaluation. We devised an evaluative process over a two-year period, with two cycles of trial, experimentation and modification of the prototype. A four-pronged approach was developed to evaluate:

- if the system is legally correct ;
- if the public can understand and use the system;
- what are the implications for the tenant, landlord and the legal advisor;
- which contexts are optimal for the diffusion of this type of system.

In this paper only the two first aspects are covered.

Functional feasibility

The objective is to evaluate the legal aspects of its dialogue, the reasoning of the program and the available explanations. The experts were chosen by snowball sampling: each expert designates one or two other experts given that in a small field, the experts know each other. We choose experts known for their work either with tenants or landlords. The experimentation protocol has three steps : first the expert uses the system, experiments with scenarios; then he studies the documentation; finally, in a second interview, the expert shall gives his/her evaluation of the system and of the documents. The data were collected by means of observation during the experiment with the system and of semi-directed interviews both after the first experiment, and after the study of the documentation.

Here are the first results of the evaluation. Only one legal issue has been discussed by an expert. The documentation is made of a schematic and a textual presentation of the rules. Both forms of presentation are necessary even if they appeared unusual for the legal experts. Furthermore there is a problem in analysing the logical structure of the production rules, particularly the links between their premises and their conclusions without a graphical representation of the virtual network.

Operational Feasibility

The three objectives of this evaluation made in co-operation with tenant associations were the following: to see if the user in a conflict situation can use and understand the system; to see if the use of the system brings a perceptible help; and to see if the general public understands the system and its use. The context of the experiment is the use of the system with the help of an advisor in conjunction with the normal consultation. The association's personnel makes sure that the tenant/landlord understands the situation. The user may know absolutely nothing about computers. The data were collected from the users by means of a questionnaire included in the program: socio-demographic data and an automatic log of the consultation. The data were collected from the advisors by interviews taken before and after the experimentation period.

The first results are that the program cannot be used directly by the tenants/landlords who normally seek advice from the Association. The program is too cumbersome to be used constantly by the advisor. They suggest the development of a direct consultation version, to be used by the advisor without the necessity to submit a complete scenario.

The operational Feasibility evaluation directly with the general Public carries two main objectives: to collect as many comments as possible about the perception people have of the system and to identify the major areas in the system that may be improved, and how. The context of the experiment is one of general public places: social services centres, shopping malls and university students and professors, including some in law and in computer science. The data were collected by observation, by semi-directed interview after the experiment, by a socio-demographic questionnaire included in the program and an automatic log of the consultation.

The first results of the evaluation shows that it is difficult to convince the general public to participate in such an experiment. In one hand there are not as many repossession cases (the domain covered by the prototype) this year. On the other hand, many will not participate in an experiment with a computer. Most of those who stop spontaneously already use a computer at work or at home.

5. Conclusion

In conclusion, the attempt to formalise the legal knowledge raises some crucial issues which present legal and technological as well as social dimensions. On the legal side, what our main concern must be is to clearly understand and evaluate the impact of cumulative reductions of real conflict situations through the successive categorisation occurring during the standardisation process of the knowledge engineering. For instance, law offers open textured concepts that are subject to multiple and sometimes opposing interpretations (*good faith*, for example). On the technological side, a multi-

technological approach seems more productive in terms of a real-world scale legal information system, that is, a legal expert system in connection with hypertext and a textual database. Finally, we need to improve our knowledge of the general public as information systems users. We need to understand the public behaviour when seek information. We should pay attention to multiple human factors such as the user's profile, his or her cultural background, values and expectations. This information should be contextualised to its interaction with the legal and administrative instances and integrated into the system design.

Nevertheless, given the development of new media combining television and computer program execution, we are convinced that the road to solving problems of social justice lies in direction of systems like Loge-expert, providing the general public with information on the rights and services to which it is entitled.

Références

[1] We have published each stage of the construction of Loge-Expert in different papers:

THOMASSET, C. "Expert System and Legal Formalization: Evaluation of a Prototype in Québec Housing Law", Proceedings of the *International Conference on Intelligence and Society*, International Social Science Council, European Coordination Centre for Research and Documentation in Social Sciences, Vienna, 1988.

THOMASSET, C., " Expert System in Québec Housing Law: from HOME-Expert I to HOME-Expert II", Proceedings of the *International Conference on Law and Artificial Intelligence: Expert Systems in Law*, CIRFID, Bologna, May 3-5, 1989, 29p.

THOMASSET, C., BLANCHARD, F., PAQUIN, L.-C., "Legal knowledge elicitation: from textual databases to expert systems", Proceedings of *DEXA '90, Data Base and Expert Systems Applications*, Springler-Verlag, Wien New-York, 1990, pp. 167-173.

THOMASSET,C., PAQUIN, L.-C., "Expert Systems in Law and the Representation of Legal Knowledge: Can we Isolate it from the Why and the Who?", Proceedings of the *III International Congress: Logica, Informatica, Diritto: Expert Systems in Law*, Consiglio Nazionale delle Ricerche, Istituto per la documentazione giuridica, editore: Antonio A. Martino, Firenze, 1989, pp. 751-772.

THOMASSET, C., BLANCHARD, F., PAQUIN, L.-C. "Loge-expert : From a legal expert system to an information system for non-lawyers", *Proceedings of the Third Internation Conference on Artificial Intelligence and Law*, Oxford, UK, Baltimore, ACM Sigart, 1991, pp.254- 259.

DUPUIS, R. .THOMASSET, C., L.-C. PAQUIN, BLANCHARD, F. "Towards a Methodology for the Validation of Expert Systems: The Case of Loge-expert", *Proceedings of the Fourth Internation Conference on Artificial Intelligence and Law*, Amsterdam, UK, Baltimore, ACM Sigart, 1993, pp.66-71.

[2] Among others, a 1987 issue of the *International Journal of Man Machine Studies*, 26 J. Boose, B. Gaines eds.

[3] SOWA, J. F. "Multi-Domain Semantic Theory" draft given by the author at a Montreal conference dated November 28, 1988. See also: SOWA, J. F."There's More to Logic than the Predicate Calculus" draft given by the author at the same conference.

[4] Routen, Tom. "Hierchically Organised Formalisation". *The Second ICAIL*. Vancouver, BC: ACM, 1989. p. 242-250.

[5] Wahlgren, Peter. "Legal Reasoning _ A jurisprudential Description." *The Second ICAIL*. Vancouver, BC: ACM, 1989. p. 147-156.