A Collaborative Legal Information Retrieval System Using Dynamic Logic Programming

Paulo Quaresma and Irene Pimenta Rodrigues
Departamento de Matemática, Universidade de Évora
Largo dos Colegiais, 7000 Évora, Portugal
{pq|ipr}@dmat.uevora.pt

Abstract

We propose a framework for a collaborative legal information retrieval system based on dynamic logic programming. In order to be collaborative our system keeps the context of the interaction and tries to infer the user intentions. Each event is represented by logic programming facts which are used to dynamically update the previous user model. User intentions are inferred from this new model and are the basis of the interaction between the system and the legal texts knowledge base. As legal texts knowledge base we are using the documents from the Portuguese Attorney General (Procuradoria Geral da República). In this paper we will show some examples of the obtained collaborative behaviour.

1 Introduction

In this paper we propose a framework for a collaborative legal information retrieval system based on dynamic logic programming [1]. In our framework the user model is represented by a logic program and each event is described by logic programming facts which are used to dynamically update the previous user model. The result of this update action is a new logic program which models the interaction status. Collaboration is achieved through the inference of user intentions using rules describing the legal domain. The main idea is to help the user to access the legal text knowledge base. For instance if the user asks the system to be informed about documents with a specific property P and that property can be a consequence of other properties A and B, the system will answer the user question but it will add information about documents where properties A and B hold.

As legal texts knowledge base we are using documents from the Portuguese Attorney General (Procuradoria Geral da República) and we are working with examples using the legislation that defines when a person has a right for a pension for exceptional services. Our information retrieval system is based on SINO, a search engine for legal text databases [2].

After a user event, the following actions are done:

1. Update the user model (dynamic logic programming module);
2. Infer the user intentions (attitudes LP module);
3. Infer other possible collaborative queries (legal knowledge LP module);
4. Interact with the IR system (SINO);
5. Answer the user.

In the next section we will describe the IR framework (SINO). In section 3 the user attitudes representation module is described. In section 4 we will describe the dynamic logic program module and in section 5 we will show how legal knowledge can be represented using logic programs. Finally, in section 6 an example of a collaborative IR session is presented.

2 Information Retrieval System

The information retrieval system is based on SINO [2] from the AustLII Institute. We have changed SINO in order to be adapted to the Portuguese Language. Namely, the new system uses the Portuguese lexicon (more than 900,000 words) in order to handle morphological errors and to obtain the base queried word. For instance, if the user asks to be informed about documents where a specific word appear, the systems also searches for documents containing derived words (plurals for nouns, verbal forms for verbs, ...).

3 User attitudes

In order to be collaborative our system needs to model user attitudes (intentions and beliefs). This task is achieved through the use of well-founded semantics of extended logic programs with explicit negation, WFSX. In this framework the system mental state is represented by an extended logic program that can be decomposed in several modules (see [4, 3] for a complete description of these modules).

In this paper we will only consider the *inform* speech act. The rule which describes the effect of an inform speech act from the point of view of the receptor (assuming cooperative agents) is:

$$bel(A, bel(B, P)) \leftarrow inform(B, A, P).$$

4 Dynamic Logic Programming

As it was pointed out in the previous section, after each event it is necessary to update the logic program which models the system mental state with the new logic program facts. This update process is done using dynamic logic program updates which allows us to update a logic program P with another logic program U ([1]). Using this approach a new event in the interaction process forces an update of the logic program creating a new state of knowledge (a new mental state).

In our framework the system initial mental state is defined by the logic program which includes the rules for the speech acts and for the collaborative behavior (see previous section). The action rules have to be changed in order to deal with time (update states):

$$bel(A, bel(B, P)) \leftarrow inform(B, A, P)/before.$$

$P/before$ means that the predicate P should be verified in the previous update state. So, these action rules mean that the effects of an action are valid on the states after the action.

5 Legal Knowledge Representation

In order to more collaborative our system needs the capability to model legal knowledge. We are using the extended logic programming framework described in the previous sections to represent some aspects of this knowledge. For the moment we have chosen only some very specific domains, namely, the legislation that defines when a person has a right for a pension for exceptional services.

Our main goal is to handle situations like this one:

- The user asks for documents were O is valid;
- The system knows (using the legal knowledge) that O can be valid in two different situations:

$$O \leftarrow A$$
$$O \leftarrow B$$

- The system gives the answer for documents where O is valid but also for documents where A and B are valid:
  - X documents where O;Y documents where O and A; Z documents where O and B

Using this approach the system is trying to be collaborative dividing the answer in clusters and predicting possible future user questions.

This system behavior can be modeled through the use of abduction in the extended logic programming framework. The goal is to abduce the predicates $\triangle$ needed to support the observation (O):

$$P \cup \triangle \models_{WFSX} O$$

where P is the logic program and each predicate in $\triangle$ belongs to the abducible set of predicates.

6 Example

In this section we will show an example over the legislation that defines when a person has a right for a pension for exceptional services. However, due to its extension and complexity, we will make many simplifications over the legislation:

$$pension \leftarrow exceptional\_war\_action.$$
$$pension \leftarrow action\_benefits\_country.$$

These rules mean that a person has the right for a pension for exceptional services if has done an exceptional war actions or an action which benefits the country.

Analyzing these rules we obtain the set of abducibles (predicates that can not be derived from the rules):

$$\{exceptional\_war\_action, action\_benefits\_country\}$$

Suppose the user wants to be informed about pensions:

$$int(system, inform(system, user, X : documents(pension)))$$

The abduced models are:

$$\{exceptional\_war\_action\}, \{action\_benefits\_country\}$$

And the system will create the following SINO queries:

search pension;
search pension and exceptional_war_action;
search pension and action_benefits_country;

However, there are some problems with these SINO queries: for instance how should the concept "action benefits country" be searched? For the moment we use the juridical thesaurus and the synonymous dictionary to expand each word and we search for the set of words. This a naive approach and, as future work, we will need to analyze parts of the documents and to transform them into a logical representation (DRS [5]) which can be used by our logic programming framework.

References

[1] José Júlio Alferes, Jo ao Alexandre Leite, Luís Moniz Pereira, H. Przymusinska, and T. C. Przymusinski. Dynamic logic programming. In S. Shapiro A. Cohn, L. Schubert, editor, *Proceedings of KR'98 - Knowledge Representation*, pages 98–109, 1998.

[2] G. Greenleaf, A. Mowbray, and G. King. Law on the net via austlii - 14 m hypertext links can't be right? In *In Information Online and On Disk'97 Conference, Sydney*, 1997.

[3] Gabriel Lopes, Paulo Quaresma, and Irene Pimenta Rodrigues. Question/answer dialogues for interfacing a database with supreme court decisions. In *Proceedings of ICAIL'97 - International Conference on AI and Law - (poster)*, 1997.

[4] Paulo Quaresma and José Gabriel Lopes. Unified logic programming approach to the abduction of plans and intentions in information-seeking dialogues. *Journal of Logic Programming*, 54, 1995.

[5] Irene Pimenta Rodrigues and José Gabriel Lopes. Discourse temporal structure. In *Proceedings of the COLING'92*, 1992.