A Dialog system for controlling question/answer dialogues

José Gabriel Lopes
Universidade Nova de Lisboa
email:gpl@di.fct.unl.pt

Paulo Quaresma
Universidade Nova de Lisboa
email:pq@di.fct.unl.pt

Irene Pimenta Rodrigues
Universidade de Évora
email: ipr@dmat.ce.uevora.pt

Abstract

In this paper we present a dialogue system that is able to maintain a conversation with an user who poses questions about the existence of Supreme Court texts with some particular characteristics.

The system we propose, is able to handle ambiguities in the user questions, clarifications made by the user and the situations where the user starts parallel discourses by introducing incompatible characteristics in its questions.

1.Introduction

We have build a set of knowledge bases where each knowledge base has the information conveyed by a Supreme Court Text. Each knowledge base is built through the pragmatic interpretation of a text, so a text is represented by a set of text temporal structures, one for each possible interpretation.

The dialog system controls the dialog with a user that is interrogating the system. The goal of the user is to obtain the set of Supreme Court Texts that have a set of characteristics phrased by the user in his question. An example of a user question is:

Q1- How many Supreme Court Texts there are where an unemployed murdered his wife?

A1- 1831

In order to give an answer the system should check which are the knowledge bases that entail (the pragmatic interpretation of) the sentence: ”an employed murdered his wife”, and count them.

We assume that the user starts by presenting some characteristics of the set of texts that he wants to see and, depending on the number of texts, he may want to add more characteristics. In the previous example the user could ask:

Q2: And where he killed her with poison.

A2: 321

With question Q2 the user intention is to continue the characteristics presented in Q1. This means that Q2 should be interpreted as:

How many Supreme Court Texts there are where:

An unemployed murdered his wife. He killed her with poison.

In order to control this kind of dialogs our dialog system represents and reasons the intentions of the user for the interpretation context of each question. This feature allows the system to handle with elegance and with the right cognitive attitudes the following dialog phenomena:

- Ambiguity in the interpretation of the characteristics phrased in each question. As we shall present later, this can happen very often and the system must me able to commit itself with one interpretation in order to supply an answer and it must phrase the main differences of the possible interpretations in its answer. Note that in the criteria for choosing one interpretation the system may take into account the easiest difference to phrase, not all differences can be phrased.
- Parallel sub-dialogs, this phenomena happens when the user phrases some characteristics that are incompatible with others that were phrased in a previous question.
- Clarification, an intervention of the user that:
    - **intends** to inform the system that the interpretation that the system chosen is not the one that the user intended.
    - and phrases the differences between the system interpretation and the intended interpretation.

The interpretation context of a user question is represented with a set of temporal text structures. For this particular domain (characteristics of supreme court texts supplied by an user in his questions) the temporal text structure is well suitable because:

- it allows us to solve some discourse such as:
    - temporal anaphora
    - pronominal and nominal anaphora
    - detecting inconsistent discourse (namely when there are incompatible characteristics conveyed by the user questions)

- to represent ambiguity, this is done by representing each interpretation by a different temporal structure.
- to obtain the differences between interpretations in order to the system be able to phrase them in its answers when it needs to made a commitment.
- to check if an interpretation entails the clarification made by the user. When the user makes a clarification, in the clarification there is a phrase that is not entailed by the interpretation that the system assumed to be the one intended by the user. So the system must check if any of the alternative interpretations entails the user phrase. If none of the alternatives entails it then the system must choose one interpretation that is consistent with the phrase in the user question an it must interpret the phrase in that context in order to obtain the user intended interpretation.

2.Overview of the system

We are building a system that is able to give sets of supreme court texts that describe situations that are specified by a user. The system interacts with the user through a natural language dialog that is controlled by our dialog system.

2.1.The Database with the Supreme Court Texts

The Supreme Court Texts are processed one by one in order to obtain a knowledge base for each of the texts.

The knowledge base represents one text as a set of text temporal structures. Each one represents one of the possible interpretations of the text. As we present in the next section, given a text temporal structure it is possible to check if another text temporal is entailed or consistent with the first one.

Let Stts be a set where each element is a text temporal structure that represents one of the possible interpretations of the situation described by the user.

Let kbtexti be the knowledge base that has all the text temporal structures for the Supreme Court Text number i.

Let DbTexts be the Data base that has all the knowledge bases with the texts.

- Given a text temporal structure ttsj in Stts, the set of numbers of Supreme Court Texts that describe that interpretation of situation j is:

  {k| kbtextk in DbTexts and kbtextk entails ttsj}

- A situation described by the user is ambiguous if there exists ttsi, ttsk in Stts such that {k| kbtextk in DbTexts and kbtextk entails ttsi} <> {l| kbtextl in DbTexts and kbtextl entails ttsj}.
- A given kbtexti entails ttsj if there exists a ttsx in kbtexti such that ttsx entails ttsi..
- So to check if a Supreme Court Text describes a situation specified by a user:

1. We build the set of text temporal structure that represent the specified the situation (see section IV)
2. And check if any of the text temporal structures that are in the knowledge base of the supreme court text entails all the temporal structure in 1. To conclude that the text describes the situation it is enough to find one structure that entails all the structures in 1.

2.2. The Interface with the users

The user interact through a dialog that is able to maintain the conversation context in order to be able to recognize the user intentions.

By now we consider as possible user intentions:

- to be informed on the number of Supreme Court texts that describe a situation (How many ....)
- to be informed on the numbers (names) of Supreme Court texts that describe a situation (Which ....)
- to clarify the system about a previous intention that was not correctly recognized by the system (No ....)
- to inform the interpretation context of the question. (the interpretation context is not marked by a clue word)

And we consider the following speech acts:

- request

- inform

The dialog system recognizes the speech act conveyed by the user question, the recognition of the speech act gives rise to a user intention that system will try to satisfy.

The recognition of the user speech act and intention requires the recognition of the user intended context for the interpretation of the question.

Some of the nice features of our dialog system are:

- That it is able to revise user intentions when they are inconsistent with the system believes or when the system believes that the user has not conscience of the consequences of is intentions. For instance when the user asks for the names of the texts and there are 1834 of them the system will first tell him the number and only if the user insists it will tell their names.
- That it only deals with the ambiguity of the user questions when it is strictly necessary to do it. If the description of a situation may be understood as ambiguous, there is more then one text temporal structure to represent it, the dialog system considers the description ambiguous only if the set of Supreme Court texts that satisfy the structures is different
- That in case of ambiguity it has the possibility of defining a criteria to choose one interpretation based on different cognitive aspects such as:
    - the best answer to the question
    - The facility of expressing (phrase) the choice made by the system
- That it is able to deal with user questions that are contradictory with the previous dialog context. In this case, the system is able to obtain the previous non-contradictory dialog context. Moreover, in the next questions the system mantains a set of possible dialog contexts continuations and chooses the best one (using some preference rule).

2.3.Example of a user session

The following dialog is a simulation of a possible dialog in our system

Q1- How many Supreme Court texts there are where a man murdered the woman that he lived with.

A1- 143

Q2- And where at the time of the murder he was drunk.

A2- 89

Q3- They were married.

A3- 34

Q4- They were divorced.

A4- 24

Q5- He killed her with a knife

A5- 9. They were divorced.

Q6- No. Where they were married.

A6- 17

3.The Text Temporal Structure

It is our goal to process texts in order to extract the information conveyed by the sentences of these texts. In order to obtain the meaning of a text sentence, we must relate the entities of each sentence with those introduced by the previous sentences aiming to solve discourse phenomena such as: definite nominal anaphora, pronominal anaphora, verbal and nominal ellipsis, sentence tense and aspect interpretation, etc. The process that enables a computer to relate the entities of each text sentence with those introduced by previous text sentences is referred to in the literature as *pragmatic interpretation of discourse entities* (Per91,Gs86,Hob90). This is carried out by taking into account world knowledge, the situation described by the text already processed and the salience of the entities introduced by the sentences in the discourse (We88,Gs86).

The pragmatic interpretation of sentence tense and aspect requires a discourse structure (We88, Hob90, RL92, RL93a, Las93c), a knowledge base with general knowledge about events, states and time intervals, and the representation of the events and states already described by the text (KR83, Pa84, Las92, RL92, RL93a, Kr93). One of the results of the pragmatic interpretation of tense and aspect of a sentence is the temporal anchoring of that sentence's main eventuality. A temporal anchoring of an eventuality is a

temporal relation between the eventuality's time interval and the time interval(s) of eventuality(ies) of previous text sentences (We88,MS88,Pas93,Eb92).

The pragmatic interpretation of a text sentence in the context provided by the interpretation of the previous sentences, namely for temporal anaphora resolution, requires a set of previously introduced referents which can be used as anaphoric referents (KR83, Pa84, Las91, Las92, RL92, RL93a, KR93}. In order to obtain the set of possible referents for the pragmatic interpretation of each text sentence's tense and aspect, we must build a discourse structure: the text temporal structure that we have presented in our earlier work (RL92, RL92b, RL93a, RL93l).

This structure provides the visible entities (the most salient ones) for the pragmatic interpretation process.

This structure is made up of discourse segments that reflect how the eventualities introduced by new sentences relate temporally to salient (visible) eventualities of the temporal discourse structure. Each text segment has features for kind, time, eventuality and subsegments. According to the kind of a segment, its subsegments must obey some temporal constraints, and the features time and eventuality are computed bottom-up from the features of its subsegments. Every sentence is represented by a segment, its pragmatic interpretation will involve inserting it into the text temporal structure. Since the segmentation is based on the temporal properties of the segments, there are some temporal relations that must be inferred in order to be able to insert a new segment into the text temporal structure.

Regarding the pragmatic interpretation process, we follow the proposal of Hobbs et (hobs90) and view pragmatic interpretation as abduction. This means that the interpretation of a text is the minimal explanation of why the text would be true. So, the sentence tense and aspect interpretation is the explanation of why the main sentence eventuality would be true in the context of the previously acquired temporal facts. As a sentence eventuality is always a state or an event, the explanation of the eventuality is always a set of assumptions that will allow the prediction or explanation of the eventuality from its interpretation context.

In (RL92,RL93l) a sentence interpretation process is driven by the attempt to justify the sentence's main eventuality. So the tense interpretation theory, given an intermediate sentence semantic representation, gives rise to a set of abducted predicates that explain the sentence from knowledge acquired with the previously interpreted text sentences. However, the sentence tense is interpreted in the sentence context, i.e. only the information contained in visible segments of the text temporal structure is taken into account.

Some of the solutions for the tense interpretation can be contradictory with the information conveyed by previous text sentences that are no longer in visible segments. To eliminate these solutions we have to check the consistency of the knowledge base with the information conveyed by the previous sentences updated with the new sentence representation.

To perform this consistency checking operation on the knowledge base we shall implement the knowledge base in a logic programming framework, a contradiction removal system (LAD94). This implementation also allows us to check if an expression is a logic consequence of the knowledge base. This operation is important during the pragmatic interpretation of sentences' tense, namely for testing the preconditions for the abduction of a contingency relation (cause, consequence, etc) between two eventualities.

3.1. Comments on the example

Q1- Gives rise to the following temporal structure:

Eventuality: e_1 Tense: perfect -**s_1**

x_1, x_2, e_1, e_2, t_1, t_2:

Man(x_1), Woman(x_2), $t_1 \subset t_2$
evt(e_1,murder(x_1,x_2), occurs(e_1,t1),
evt(e_2,liveswith(x_1,x_2), occurs(e_2,t_2)

Q2- Can not be interpreted alone, because of the use of the pronominals. The interpretation for this question e the following structure:

Eventuality: e_1 Tense: perfect -**s_{12}**

Kind: subset

Eventuality: e_1 Tense: perfect -**s_1**

x_1, x_2, e_1, e_2, t_1, t_2:

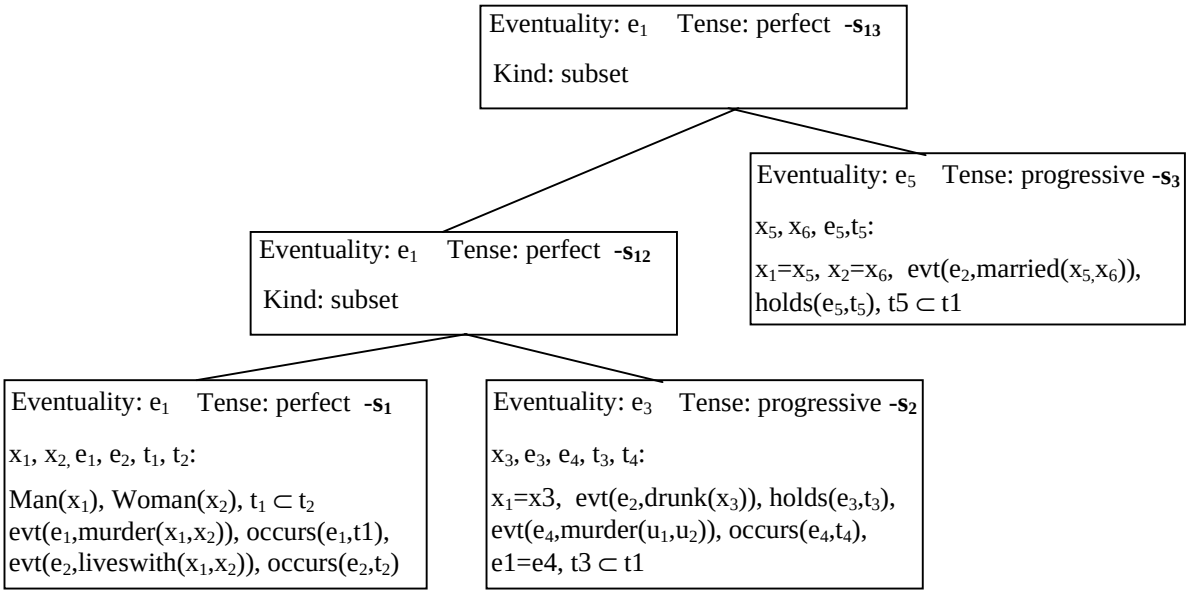
Man(x_1), Woman(x_2), $t_1 \subset t_2$
evt(e_1,murder(x_1,x_2)), occurs(e_1,t1),
evt(e_2,liveswith(x_1,x_2)), occurs(e_2,t_2)

Eventuality: e_3 Tense: progressive -**s_2**

x_3, e_3, e_4, t_3, t_4:

x_1=x3, evt(e_2,drunk(x_3), holds(e_3,t_3),
evt(e_4,murder(u_1,u_2)), occurs(e_4,t_4),
e1=e4, t3 $\subset$ t1

Q3- Again, only can be interpreted in the context of the characteristics phrased in the previous questions



Q4- Can not be interpreted in the context of structure s_{13} because it gives rise to a structure that is not consistent, married and divorced can not be true at the same time (the time of the murder). So the dialog system will define the structure s_{12} as the interpretation context for part of question Q4. We do not present it for lack of space, but new structure will be similar to s_{13} except that the term evt will have divorce instead of married.

Q5- will be inserted, by command of the dialog system in the structures s_{12}, s_{13}, s_{14}. When the dialog system asks for the differences between those structures the answer will be: the structure s_3, because it is not present in structures obtain from joining s_5 to s_{12} and s_{14}, the structure s_4 because it is not present in structures obtain from joining s_5 to s_{12} and s_{13}. So the system my commit itself to structure s_{135} or s_{145} because the differences are easy to be phrased. In the example the system chooses s_{145}.

Q6 - This is a clarification, so the dialog system wants a structure that is compatible with the characteristics phrased by the user (“He killed her with a knife”). Structure s_{145} is the only one that entails the Q6.

4.The Dialog system

The dialog system tries to recognize the users intentions and it acts in order to satisfy them. For instance, in the user’s first question of the example in section 2.3, the dialog system recognizes a request speech act from the user to be informed of the number of Supreme Court texts with a specific characteristic. This

request speech act transmits a user intention and the system should act in order to satisfy it: it should obtain and inform the user about the knowledge bases that entail the situation described.

In order to represent and to reason about the users intentions the system needs to represent two different attitudes: believes and intentions. In fact the representation of its own believes (including the system believes about the users' believes and intentions) and its own intentions is the basis of the planning and acting process. This process has as input the system intentions and it executes the intended actions whenever is possible (there are no unsatisfied pre-requisites).

The dialog system is built over a logic programming framework that allows non-monotonic reasoning: (well-founded semantics of extended logic programs with explicit negation, WFSX, from the work of Pereira et al.) and it is represented by an extended logic program that can be decomposed in several modules (see (Qua95,Qua97) for a complete description of these modules):

- Description of the effects and the pre-conditions of the speech acts (inform and request) in terms of believes and intentions;
- Definition of behavior rules that define how the attitudes are related and how they are transferred between the users and the system (cooperatively);
- Temporal formalism (Event Calculus).

When the user poses a question the agents' model (logic program) is updated with the description of the event that occurred. This description captures the speech acts associated with the question and the set of text temporal structures that represent the possible interpretations of the situation described by the user. Using this information it is possible to calculate the users' intentions (using the speech acts definitions) and to interact with the database in order to obtain the information needed to answer the user (using the set of text temporal structures). A complete description of these processes is presented in Qua97.

In this section, we will show how the dialog process is done using the example presented in section 2.3:

- After the first question the system recognizes a request speech act:
  - request(u,s,inform(s,u,(how_many(X):stts1))); where u means user, s means system, and stts1 is the set of text temporal structures for question Q1. As this question was the first one there is no previous dialog context and stts1 is obtained only from Q1.
- The recognized speech act has the following effect:
  - bel(s,int(u,inform(s,u,how_many(X):stts1))); meaning that the system believes that the user wants to be informed of how many Supreme Courts there are where stts1;
- As the system is a cooperative one, it tries to satisfy the user and it asks the database the desired information: query(how_many(X):stts1) -- X is the number of knowledge bases that entail stts1;
- The system informs the user about the desired information:
  - inform(s,u,how_many(143):stts1).
- The second user's question (Q2) is similar with Q1 except from the fact that it is necessary to join the question with a previous dialog context (set of text temporal structures). This joining process is done by obtaining the previous set of text temporal structures (stts) that is compatible with the question. In Q2, this set is stts1 and Q2 is joined with stts1 obtaining stts2. The answer process is similar with Q1.
- The user's Q3 question is similar with Q2 and the joining process joins stts2 with Q3 obtaining stts3.
- After Q4 there is a new problem: Q4 can not be joined with stts3 because it is incompatible. In this case, Q4 should be joined with stts2 obtaining stts4. The answer process is identical to the previous ones.
- With Q5 the user poses a question that can be joined with three sets of temporal structures: stts4 (the previous one), stt3 (the other branch of the incompatible dialog), or stts2 (the root of the incompatible dialog). The system must choose between the sets and this process may take into account pragmatic and cognitive knowledge. In the exemple we have chosen the previous set (stts4). Note that the system must identify the chosen set and phrase the answer.
- Q6 is a clarification question. In this situation the system recognizes two speech acts:
  - inform(u,s,no).

- request(u,s,inform(s,u,how_many(X):stts6))

- In a clarification question the set of temporal structures (stts6) is obtained from one of the other previous compatible stts. In this example the user informs that the intended set is stts3. The answer process is identical to the previous ones.

As it was shown the dialog system has the capability to represent the questions speech acts and the dialog contexts. Using these information it is possible to infer the users' attitudes and, as a consequence, the system's attitudes. The inferred attitudes are the basis of the interaction with the database and the answer process.

5.Bibliography

- [ADP95] Jos{\'e} J{\'u}lio Alferes, Carlos Dam{\'a}sio, and Lu{\'\i}s Moniz Pereira. A logic programming system for nonmonotonic reasoning. Journal of Automated Reasoning, 14:93--147, 1995.
- [AKPT91] James Allen, Henry Kautz, Richard Pelavin, and Josh Tenenberg. Reasoning about Plans. Morgan Kaufman Publishers, Inc., 1991.
- [AP96] Jos{\'e} J{\'u}lio Alferes and Lu{\'\i}s Moniz Pereira. Reasoning with Logic Programming, volume 1111 of Lecture Notes in Artificial Intelligence.Springer, 1996.
- [Bra90] Michael Bratman. What is Intention?, in Intentions in Communication. MIT, 1990.
- [CL90a] P. Cohen and H. Levesque. Intention is choice with commitment. Artificial Intelligence, 42(3), 1990.
- [CL90b] Philip Cohen and Hector Levesque. Persistence, Intention, and Commitment, in Intentions in Communication, pages 33--70. MIT, 1990.
- [Car85] Sandra Carberry. Pragmatic Modeling in Information System Interfaces. PhD thesis, University of Delaware, 1985.
- [Car88] Sandra Carberry. Modelling the user's plans and goals. Computational Linguistics, 14(3):23--37, 1988.
- [Ebe92] Kurt Eberle. On representing the temporal structure of a natural language text. In Proceedings of the COLING'92, pages 288-294, 1992.
- [Esh88] Kave Eshghi. Abductive planning with event calculus. In Proceedings of the International Conference on Logic Programming, 1988.
- [GS86] Barbara Grosz and Candice Sidner. Attention, intention, and the structure of discourse. Computational Linguistics, 12(3):175--204, 1986.
- [HM87] S. Hanks and D. McDermott. Nonmonotonic logic and temporal projection. Artificial Intelligence, 33, 1987.
- [Hin88] Erhard Hinrichs. Tense,quantifiers, and contexts. Computational Linguistics, 14(2):3--14, June 1988.
- [Hob90] Jerry Hobbs, Mark Stickel, Douglas Appelt, and Paul Martin. Interpretation as abduction. Technical Report 499, SRI International, December 1990.
- [KKT92] A.C. Kakas, R.A. Kowalski, and F.Toni. Abductive Logic Programming. Journal of Logic and Computation, 2(6):719--770, 1992.
- [KR83] Hans Kamp and Christian Rohrer. Tense in texts. In C. Bauerle, R. Schwarze and A. von Stechow, editors, Use and Interpretation of Language. de Gruyter, 1983.

- [KR93] Hans Kamp and Uwe Reyle. From Discourse to Logic: An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory. Dordrecht: D. Reidel., 1993.
- [KS86] Robert Kowalski and Marek Sergot. A logic-based calculus of events. New Generation Computing, 4:67--95, 1986.
- [LA87] D. Litman and J. Allen. A plan recognition model for subdialogues in conversations. Cognitive Science, (11):163--200, 1987.
- [LA91] Alex Lascarides and Nicholas Asher. Discourse relations and defeasible knowledge. In Proceedings of the 29th Annual Meeting of ACL, pages 55--62, 1991.
- [LA92] Alex Lascarides, Nicholas Asher, and Jon Oberlander. Inferring discourse relations in context. In Proceedings of the 30th Annual Meeting of ACL, pages 55--62, 1992.
- [LAD93] L.M. Pereira, C. Dam sio, and J.J. Alferes. Diagnosis and debugging as contradiction removal. In LuisMoniz Pereira and Anil Nerode, editors, Proceedings of the Second International Workshop on Logic Programming and Non-monotonic Reasoning, pages 316--330. The MIT Press, 1993.
- [LO93] Alex Lascarides and Jon Oberlander. Temporal connectives in a discourse context. In Proceedings of the 6th European chapter of ACL, EACL'93, 1993.
- [MS88] Mark Moens and Mark Steedman. Temporal ontology and temporal reference. Computational Linguistics, 14(2):15--28, June 1988.
- [Mis91] Lode Missiaen. Localized Abductive Planning with the Event Calculus. PhD thesis, Univ. Leuven, 1991.
- [PR93] J.Pinto and R.Reiter. Temporal reasoning in logic programming: A case for the situation calculus. In D.S. Warren, editor, Proceedings of the 10th ICLP. MIT Press, 1993.
- [Pa93] Megumi Kameyama, Rebecca Passonneau, and Massimo Poesio. Temporal centering. In Proceedings of the 31st Annual Meeting of the ACL, 1993.
- [Par84] Barbara H. Partee. Nominal and temporal anaphora. Linguistics and Phylosophy, 7:243--286, 1984.
- [Per90] Raymond Perrault. An Application of Default Logic to Speech Act Theory, in Intentions in Communication, chapter 9, pages 161--186. MIT, 1990.
- [Per91] F.Pereira and M.Pollack. Incremental interpretation. Artificial Intelligence, 50:40--82, 1991.
- [Pol86] Martha E. Pollack. Inferring Domain Plans in Question-Answering. PhD thesis, Dep. of Computer and Information Science, University of Pennsylvania, 1986.
- [Pol90] Martha Pollack. Plans as Complex Mental Attitudes, in Intentions in Communication, chapter 5, pages 77--104. MIT, 1990.
- [QL92] P. Quaresma and J. G. Lopes. A two-headed architecture for intelligent multimedia man-machine interaction. In B. de Boulay and V. Sgurev (eds). Artificial Intelligence V - methodology, systems, applications. North Holland, 1992.
- [QL95] Paulo Quaresma and Jos{\'e} Gabriel Lopes. Unified logic programming approach to the abduction of plans and intentions in information-seeking dialogues. Journal of Logic Programming, (54), 1995.
- [RL92a] IrenePimenta Rodrigues and Jos, GabrielPereira Lopes. Discourse temporal structure. In Proceedings of the COLING'92, 1992.
- [RL92b] IrenePimenta Rodrigues and Jos, GabrielPereira Lopes. A framework for text interpretation. In B.duBouley and V.Sgurev, editors, Artificial Intelligence V --- methodology, systems and applications, pages 181--190. North Holland, 1992.
- [RL93a] IrenePimenta Rodrigues and Jos,Gabriel Lopes. Building the text temporal structure. In Progress in Artificial Intelligence: 6th Portuguese Conference on AI. Springer-Verlag, 1993.
- [RL93b] IrenePimenta Rodrigues and Jos, Gabriel Lopes. Text processing in a logic programming framework: The interpretation of tense and aspect. In Proceedings of the fourth International Workshop on Natural Language Understanding and Logic Programming, NLULP4, 1993.
- [Rod94] IreneP. Rodrigues. Processamento de Texto: Interpreta\cc\ao Temporal. PhD thesis, Universidade Nova de Lisboa, submetida <E0> FCT/UNL em Novembro de 1994.

- [SC88] F.Song and R.Cohen. The interpretation of temporal relations in narrative. In Proceedings of the National Conference on Artificial Intelligence, AAAI 88, St. Paul, MN. Morgan Kaufmann Publishers Inc., 1988.
- [Son91] F.Song. A Processing Model for Temporal Analysis and its Aplication to Plan Recognition. PhD thesis, University of Waterloo, Waterloo, Ontario, Canada, 1991.
- [Web88] BoonieLynn Webber. Tense as discourse anaphor. Computational Linguistics, 14(2):61--73, June 1988.