Using Speech Acts in Logic-Based Rhetorical Structuring for Natural Language Generation in Human-Computer Dialogue

Vladimir Popescu[1,2], **Jean Caelen**[1], **Corneliu Burileanu**[2]
[1]Grenoble Institute of Technology, France
[2]University "Politehnica" of Bucharest, Romania
{vladimir.popescu, jean.caelen}@imag.fr

Abstract

Usually, human-computer dialogue systems rely on ad-hoc solutions for the component performing speech turn generation, in natural language. However, with the advent of novel dialogue planning techniques, integrating task-specific and general world knowledge in order to provide a more reliable and natural interaction with humans, more sophisticated language generation techniques become necessary. In this paper we present performance improvements of a module approximating a part of Segmented Discourse Representation Theory (SDRT) for language generation in dialogue, using first-order logic (FOL) enforced by a task-independent discourse ontology. These improvements concern reductions in computational costs and enhancements in rhetorical coherence for the discourse structures obtained and are obtained using speech-act related information for driving rhetorical relations computations.

1 Introduction

Early human-computer dialogue systems relied on handcrafted, usually template-based, language generation modules for producing machine's utterances (McTear, 2002). However, in the last decade or so, with the emergence of research results and ideas from the multi-agent systems community (Minker and Bennacef, 2001), dialogue systems became more sophisticated, aiming at better responses to user's requests, via a greater naturalness and pertinence of the speech turns produced, in relation to the context of the dialogue and to the users involved (Caelen and Xuereb, 2007), (McTear, 2002), (Minker and Bennacef, 2001). Hence, the natural language generation component itself started to evolve towards more contextualized and pragmatically situated language productions, involving consideration of rhetorical and actional aspects of language production; this way, two research trends became distinguishable: (i) a rhetoric-based approach, using formal accounts of discourse originally designed for language interpretation: thus, theories such as Rhetorical Structure Theory -RST (Mann and Thompson, 1988), Discourse Representation Theory - DRT (Kamp and Ryle, 1990) or, more recently, Segmented Discourse Representation Theory - SDRT (Asher and Lascarides, 2003) have been adapted and / or approximated in computational implementations of natural language generators, either rule-based (Danlos and El-Ghali, 2002), or corpus-based (Marcu, 1997), mostly in monologue situations; (ii) a speech-act based approach, combining traditional speech act theory (Searle, 1969), (Vanderveken, 1990-1991) with results provided by the Belief-Desire-Intention - BDI approach in modelling virtual agents (Bratman, 1987), (Poesio and Traum, 1997): thus, computational implementations oriented towards generation in dialogue have appeared, based on several taxonomies of speech acts (Traum and Hinkelman, 1992), derived from traditional theory; for instance, (Stent, 2001) is relevant to this research trend.

However, few research efforts aimed at combining insights from speech acts theory with formal accounts provided by rhetorical structuring theories, mostly for pragmatic interpretation purposes (Xuereb and Caelen, 2005). Therefore, work reported in this paper addresses the issue of improving the performances of a first-order logic-based rhetorical structuring component of a natural language generation module included in a task-oriented spoken dialogue system (Caelen and Xuereb, 2007). The general research goal of our team resides in the generic character of the processing stages involved, at all levels, in spoken human-computer dialogue. Hence, a general framework has been developed and instantiated in several illustrating applications, such as meeting room reservation (Xuereb and Caelen, 2005), or book reservation (Popescu *et al.*, 2007).

Regarding the natural language generation component in the generic framework mentioned above, a first-order logic-based rhetorical structuring component has been designed and presented elsewhere (Popescu *et al.*, 2007); this component is based on an adaptation and approximation of SDRT for speech turns generation in dialogue, using first-order logic doubled by a task-independent discourse ontology. In this paper, performance improvements of this rhetorical structuring component are shown; these reside in: (i) reductions in computational costs involved by determining the possible rhetorical structures connecting the utterance to be generated, to the dialogue history and (ii) improved selection capabilities so as, out of several valid discourse structures, most coherent (Asher and Lascarides, 2003) ones are retained. These performance improvements are achieved integrating the constraints induced by the speech acts characterizing each utterance to be linked to another utterance, via a rhetorical relation.

The paper is structured as follows: the next section provides a brief overview of the baseline (FOL-based) rhetorical structuring component, the third section provides motivational discussion regarding the usage of speech acts in rhetorical structure update, enforced by an algorithm in this respect; the fourth section presents a detailed example allowing comparisons between the baseline approach and the one integrating speech acts; the fifth section concludes the paper and gives pointers to further research.

2 Logic-Based Rhetorical Structuring Component

Our team has designed a rhetorical structuring component integrated in a natural language generation module of a task-oriented spoken dialogue system. In this context, seventeen rhetorical relations have been chosen, in the framework of SDRT, namely:

– first-order rhetorical relations - *Q-Elab*, *IQAP*, *P-Corr* and *P-Elab*, with informal semantics as in (Asher and Lascarides, 2003), that are strongly related to *temporal* aspects in dialogue, hence used in an approximate manner, specific to the type of dialogue concerned (i.e., conversations involving negotiations on time intervals of resource availability, as in Verbmobil corpus (Schlangen *et al.*, 2001);

– second-order rhetorical relations - $Background_q$, $Elab_q$, $Narration_q$, *QAP*, *ACK* and *NEI*, with informal semantics as in (Asher and Lascarides, 2003), that are less constrained by the temporal aspects of the dialogues concerned, hence used in a manner closer to that specified in vanilla SDRT;

– third-order rhetorical relations, specific to monologues and used to relate utterances within a speech turn, generated by one of the speakers (either the human or the machine) - *Alternation*, *Background*, *Consequence*, *Elaboration*, *Narration*, *Contrast* and *Parallel*, with semantics as in vanilla SDRT (Asher and Lascarides, 2003).

Each of these 17 rhetorical relations is expressed as a predicate in first-order logic; each such predicate is expressed in terms of other predicates instantiating actions, operations and relationships between entities. These entities are objects either in a task-independent discourse ontology, or in a task ontology, as described in (Popescu *et al.*, 2007); for example, the first-order rhetorical relation *Q-Elab* is expressed in terms of the binary predicates (in the discourse ontology) equals, Disjoint and of the unary predicates answer and SARG - Speech Act Related Goal (Asher and Lascarides, 2003); each of these predicates takes as arguments objects either in the discourse ontology (enounce, question, Δt), or in the task ontology (the entities expressing the semantics of the two utterances due to be related via *Q-Elab*). The predicates expressing the semantics of the rhetorical relations are linked through the usual connectors in first-order logic, namely $\wedge$ ("and"),

$\vee$ (“or”), $\neg$ or $\Rightarrow$ (implication); furthermore, each predicate in the discourse ontology is expressed in terms of several predicates in the same ontology and of objects in either of the two ontologies (task ontology or discourse ontology). For example, the predicate answer is expressed in terms of the predicates greater , equals , topic , MemberOf in the discourse ontology[1]. Further details and working examples in this respect are given in (Popescu *et al.*, 2007).

3 Speech Acts in Rhetorical Structure Computation

Previous studies of our team advocated for the correspondences that exist between pairs of speech acts (in the sense of Searle’s taxonomy (Searle, 1969), customized for human-computer dialogue (Caelen and Xuereb, 2007)) and mapping tables have been proposed, using a spoken dialogue corpus, acquired via the Wizard-of-Oz method in the context of a meeting room reservation task (Caelen and Xuereb, 2007), (Xuereb and Caelen, 2005), (Nguyen, 2005).

The taxonomy of speech act *types* proposed by our team (and described in detail in (Caelen and Xuereb, 2007)) supposes that human-computer dialogue is a coordination of actions (linguistic or not) according to some rules (in order to reach a present or future goal), with mutual knowledge and *know-how* construction. Hence, the interaction proceeds through an exchange of acts (our notion of acts generalizes Searle’s notion of speech acts but only speech acts will be considered in this paper); each act has two components: (i) a propositional content, expressing the semantics conveyed by the utterance produced (by the human speaker or by the machine), and (ii) an illocutionary act, which accompanies the propositional content and characterizes the latter in terms of language *use*. The illocutionary act is governed by specific axioms and rules (as specified in (Vanderveken, 1990-1991)), having preconditions and effects. Certain acts are *performed* in order to determine changes in the state of things - F^{A}: performing an action (denoted by “DO”), F^{F}: determining (a speaker) to perform an action (denoted by “MAKE-DO”); other acts are epistemic in nature, that is, they aim at determining changes in the discourse state or mental states of the speakers - F^{S}: informing a speaker about certain facts (denoted “MAKE-KNOW”), F^{FS}: asking (a speaker) about certain facts (denoted “MAKE-DO-KNOW”). Finally, there are two act types that are deontic in nature, i.e., they create obligations (necessities) or give choices (possibilities) - F^{D}: compel (a speaker) to do something (denoted “MAKE-MUST”), F^{P}: give a speaker choices of doing something (denoted “MAKE-CAN”).

Each utterance is characterized, from an illocutionary point of view, by one speech act type, from those specified above; each pair of utterances is thus characterized by a pair of speech acts and, from a rhetorical point of view, by a set of rhetorical relations connecting the utterances[2].

The point we make here is that the set of rhetorical relations connecting a pair of utterances is conditioned not only by the semantics of the utterances (expressed as logic forms, as described in (Popescu *et al.*, 2007)), but also by the speech acts characterising these utterances from an illocutionary point of view; an extensive corpus study[3] in this respect is provided in (Xuereb and Caelen, 2005), hence, only an illustrative example will be given here; for detailed tables summarizing the relationships between pairs of speech acts and rhetorical relations the reader should refer to (Xuereb and Caelen, 2005).

Let us assume that we have two speakers, the human subject (denoted by U) and the machine (denoted by M), and that U tries to reserve a book in a library. Hence, U asks the librarian (i.e., the machine) where he can find the book “X”; the machine can give several answers, characterised by different speech acts and connected via different rhetorical relations to the utterance of U[4].

The situation is illustrated in Figure 1 (to the left, we show the speech act characterising the utterance, to the middle, the utterance itself, and to the right, the rhetorical relation connecting that utterance to

[1]The objects that these predicates take as arguments are, most of them, in the task ontology, since there are in fact variables expressing the semantics of the utterances involved.

[2]Usually, this set of rhetorical relations is reduced to one element, as it will be shown further in the paper.

[3]The corpus used in this study consisted of about 1500 speech turns in dialogues recorded in a Wizard-of-Oz manner, for a meeting room reservation task.

[4]The example provided here (in Figure 1) represents a translation in English from a French-spoken dialogue fragment.

$\mathrm{F}_{U}^{\mathrm{FS}}$:	Where can I find book "X"?	
Possible answers of M		
$\mathrm{F}_{M}^{\mathrm{S}}$:	It is at the end of this corridor	*QAP*
	Just next to the exit door, to the left	*IQAP*
	The plan of the book shelves is down the entrance hall	*P-Elab*
$\mathrm{F}_{M}^{\mathrm{FS}}$:	Is it for a scientific report you have to write?	*Elab$_q$*
	Is it a scientific book that you are looking for?	*Background$_q$*
	This book is to be found in the university library	*P-Corr*
$\mathrm{F}_{M}^{\mathrm{F}}$:	You go at the end of the corridor	*QAP*
	You go in another building, the "B" building	*P-Corr*
	You go look on the shelf that is just next to the DVD shelves	*IQAP*
$\mathrm{F}_{M}^{\mathrm{D}}$:	First, you have to validate your card, sir	*P-Elab*
$\mathrm{F}_{M}^{\mathrm{P}}$:	You can take either the hardcover edition, or the DVD edition of this book	*P-Elab*

Figure 1: Speech acts and rhetorical relations: illustrative example.

the utterance come from the speaker U)[5].

Using corpus-drawn examples of the type presented in Figure 1, our team has shown that for each pair of speech acts in dialogue, only some (usually, two or three) rhetorical relations (out of all the 17considered) are *authorized* to connect the utterances involved (Xuereb and Caelen, 2005).

These results are used for refining the set of candidate rhetorical relations in (segmented) discourse structure - SDRS update, according to the algorithm shown in Figure 2, which represents an improved version of the algorithm presented in (Popescu *et al.*, 2007), so as to take into account speech acts in rhetorical structure update. However, unlike the formal presentation given in (Popescu *et al.*, 2007), here a rather informal statement of the algorithm will be put forward; moreover, steps added in the present version of the algorithm are shown in boldface.

An example of rhetorical relation computation between two utterances has been presented in detail in (Popescu *et al.*, 2007); here, a rough estimation of computational cost reductions will be shown; then, two rhetorical structures, updated from the same starting SDRS, using and not using speech acts, will be presented.

As for the reductions in computational costs, a rough estimate can be computed as follows: let us suppose that the SDRS to be updated already contains N utterances, that the total number of possible rhetorical relations between utterances is $R = 17$,

for each utterance α to be added to the dialogue SDRS:

1. read its corresponding logic form $K(\alpha)$, through a query to the *dialogue controller* (Caelen and Xuereb, 2007);
2. for each utterance β already in the dialogue SDRS:
   (a) read its corresponding logic form $K(\beta)$;
   (b) **read the pair $(\gamma_\alpha, \gamma_\beta)$ of speech acts for this utterance and the utterance at step 1.;**
   (c) **retrieve the set P of rhetorical relations *authorized* by the pair of speech acts read at 2.(a);**
   (d) for each rhetorical relation ρ in set P:
      i. read the semantics Σ_ρ of rhetorical relation ρ;
      ii. compute the truth value γ of the proposition $\Sigma_\rho(K(\alpha), K(\beta))$;
      iii. if $\gamma = \mathsf{FALSE}$, then go to step 2.(c);
      else add ρ to the set of rhetorical relations in the SDRS and α to the set of utterances in the SDRS and go to 2.(c).

Figure 2: Dialogue SDRS update algorithm.

and that the average number of rhetorical relations authorized by a certain pair of speech acts is M (usually, 3, according to our studies (Xuereb and Caelen, 2005)[6]). Furthermore, assuming that the time needed to read or retrieve logic formulas or speech acts is a negligible constant (since these elements are computed by the dialogue controller, independent of the language generation component (Nguyen, 2005)), the computational cost of updating the SDRS with one utterance is $N \times R$ proofs, since each of the R rhetorical relations needs to be

[5]The act F^{A} is not considered in this paper, because it is not essentially linguistic, supposing that the emitter performs some action.

[6]In fact, in (Xuereb and Caelen, 2005) it has been shown that, for certain pairs of speech acts, 1 to 5 rhetorical relations may be authorized, and the actual average number is 2.24 relations per pair of speech acts, hence less than 3.

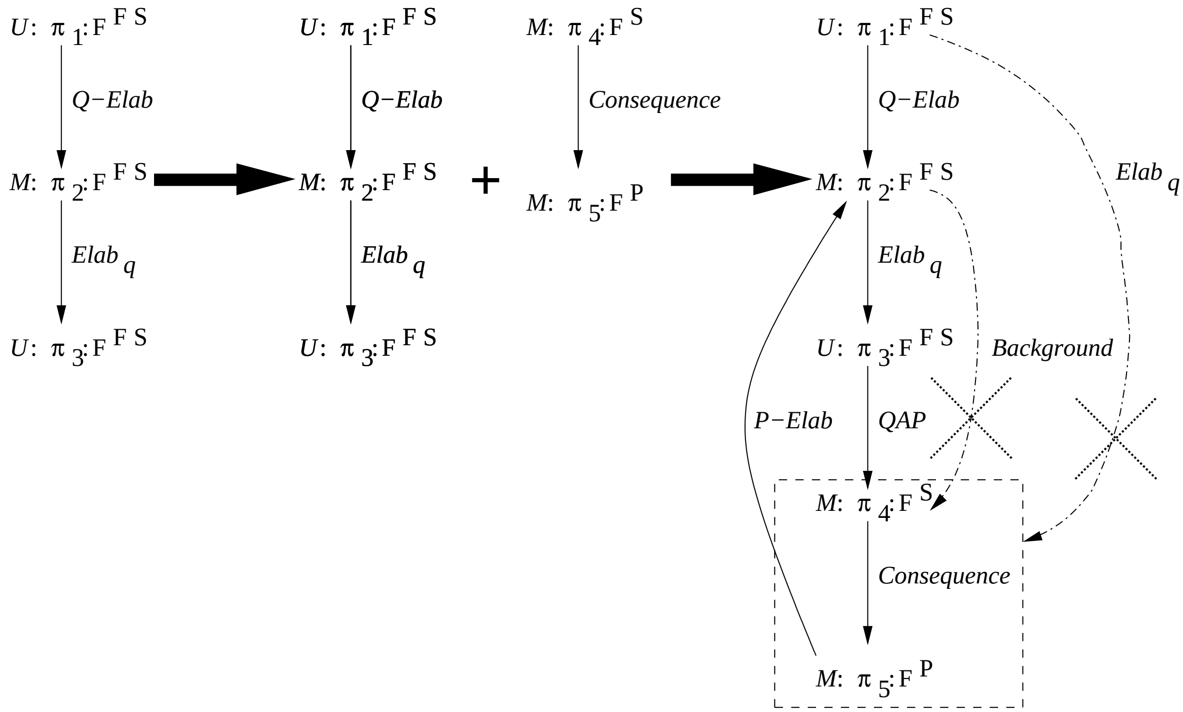

Figure 3: Discourse structure update process.

checked for each of the N utterances in the dialogue SDRS. We suppose that the time needed to prove a rhetorical relation between two utterances is a constant, T[7], thus the computational cost, when speech acts are not used, could be evaluated at $N \times R \times T$. When speech acts are used, the computational cost can be evaluated at $N \times M \times T$, hence a reduction of R/M is achieved. For the average values of R and M, the computational cost is reduced around 6 times when using speech acts.

4 Discourse Structure Update Example

In order to illustrate the augmentation of the pertinence for an updated SDRS, we consider the dialogue below[8] (here, U and M denote the human user and the computer respectively, and π_i denotes the label of the i-th speech turn):

U : π_1: Where can I find some book about "F"?
M : π_2: You want a book on "F" written by whom?
U : π_3: What's available?

[7]Strictly speaking, this is not true, since T depends on the logic formulas expressing the utterances and the semantics of the rhetorical relations.

[8]Translation in English of a French fragment of dialogue is considered.

From this point on, the machine is supposed to answer that books by authors "A" and "B" are available on the subject "F" and to give the user the opportunity to choose between these two authors; this drives M to produce two utterances, the first one as an act of informing the user (hence, a F^S) and the second one as an act of giving a choice to the user (hence, a F^P); for these two utterances, only logic forms (described according to the detailed discussion provided in (Popescu *et al.*, 2007)) are available (from the dialogue controller (Caelen and Xuereb, 2007)); however, for the ease of comprehension, two possible corresponding linguistic forms are given, in italics, below:

M : π_4: *We have books by authors "A" and "B".*
M : π_5: *Which one you like?*

Furthermore, the machine builds a sub-SDRS out of these two utterances, labeled π_4 and π_5 respectively and adjoins this sub-structure to the dialogue SDRS, formed with the utterances π_1 to π_3. This process is illustrated, in stages, in Figure 3, where we have represented, separated by colons, the speaker, the label of the utterance and the speech act performed by the utterance; the rhetorical relations between utterances are marked by directed arrows (as in vanilla SDRT (Asher and Lascarides, 2003)), their names being given to the right. In da-dotted lines are marked the rhetorical relations computed as valid by the logic-based SDRS update module, but not authorized by the pair of speech acts (this fact is also marked by the diagonal crosses intersecting these rhetorical relations). Thus, when the machine links π_4 and π_5 through a rhetorical relation, only *Consequence*, which is authorized by the pair of speech acts F^{S} and F^{P} in a monologue context, is found (via the logic-based SDRS update mechanism (Popescu *et al.*, 2007)) to hold between these utterances, hence the sub-SDRS containing utterances π_4 and π_5 is built. Then, this sub-SDRS is connected to the dialogue SDRS containing utterances π_1 to π_3 via several rhetorical relations that are found to hold by the logic-based SDRS update mechanism:

– *QAP* is found between utterances π_3 and π_4,

– *Background* is found between utterances π_4 and π_2,

– *P-Elab* is found between utterances π_2 and π_5,

– *Elab$_q$* is found between utterance π_1 and the pair of utterances (π_4, π_5).

From these relations, $Background(\pi_4, \pi_2)$ and $Elab_q(\pi_1, (\pi_4, \pi_5))$ are not authorized by the pairs of speech acts performed by the utterances concerned, which corresponds to our intuitions and to the informal semantics of the rhetorical relations, as provided in (Asher and Lascarides, 2003).

The detailed manner of computing the rhetorical relations between pairs of utterances is presented in (Popescu *et al.*, 2007); however, for the completeness of the paper, we present below the way in which the rhetorical relations *Background(π_4, π_2)* and *P-Elab(π_2, π_5)* are computed. For this, the logic forms of the three utterances involved are specified below (by $K(\pi_i)$ one denotes the logic form expressing the semantics of utterance π_i):

– $\pi_2 \mapsto K(\pi_2) ::= \exists X, Y, T, V : \mathsf{agent}(X) \wedge \mathsf{equals}(X, U) \wedge \mathsf{equals}(X, \neg\mathsf{emitter}(\pi_2)) \wedge \mathsf{object}(Y) \wedge \mathsf{equals}(Y, \mathsf{book}) \wedge \mathsf{feature}(Y, T) \wedge \mathsf{author}(T) \wedge \mathsf{equals}(T, \text{'?'}) \wedge \mathsf{feature}(Y, V) \wedge \mathsf{field}(V) \wedge \mathsf{equals}(V, \text{'F'}) \wedge \mathsf{want}(X, Y) \wedge \mathsf{equals}(\Delta t_{\pi_2}, t)$,

– $\pi_4 \mapsto K(\pi_4) ::= \exists X, Y, Z, T, V, W : \mathsf{agent}(X) \wedge \mathsf{equals}(X, M) \wedge \mathsf{equals}(\mathsf{X}, \mathsf{emitter}(\pi_4)) \wedge \mathsf{object}(Y) \wedge \mathsf{equals}(Y, \mathsf{book}) \wedge \mathsf{object}(Z) \wedge \mathsf{equals}(Z, \mathsf{book}) \wedge \mathsf{feature}(Y, T) \wedge \mathsf{author}(T) \wedge \mathsf{equals}(T, \text{'A'}) \wedge \mathsf{feature}(Z, V) \wedge \mathsf{author}(V) \wedge \mathsf{equals}(V, \text{'B'}) \wedge \mathsf{ExhaustiveDecomposition}(Y, Z, W) \wedge \mathsf{have}(X, W) \wedge \mathsf{equals}(\Delta t_{\pi_4}, t_\forall)$,

– $\pi_5 \mapsto K(\pi_5) ::= \exists X, Y, Z : \mathsf{agent}(X) \wedge \mathsf{equals}(X, U) \wedge \mathsf{equals}(X, \neg\mathsf{emitter}(\pi_5)) \wedge \mathsf{object}(Y) \wedge \mathsf{equals}(Y, \mathsf{book}) \wedge \mathsf{feature}(Y, Z) \wedge \mathsf{author}(Z) \wedge \mathsf{equals}(Z, \text{'?'}) \wedge \mathsf{want}(X, Y) \wedge \mathsf{equals}(\Delta t_{\pi_5}, t)$.

The two rhetorical relations concerned have the semantics specified below:

– $Background(\pi_4, \pi_2) ::= \mathsf{equals}(\pi_4, \mathsf{enounce}) \wedge \mathsf{entails}(\pi_4, \pi_2)$,

– $P - Elab(\pi_2, \pi_5) ::= \mathsf{good\_time}(\Delta t_{\pi_5}) \wedge \neg\mathsf{Disjoint}(\Delta t_{\pi_5}, SARG(\pi_2)) \vee \mathsf{bad\_time}(\Delta t_{\pi_5}) \wedge \neg\mathsf{equals}(SARG(\pi_2) \setminus \Delta t_{\pi_5}, \emptyset) \wedge \mathsf{ClassOf}(K(\pi_5), \mathsf{Plan}) \wedge (\mathsf{Plan} \Rightarrow SARG(\pi_2))$.

In these logic formulas, several predicates (defined in the discourse ontology or in the task ontology), functions and objects are used:

(i) unary predicates or functions: agent (predicate defined in the discourse ontology) - the agent performing the action reported in an utterance; object (predicate defined in the discourse ontology) - the object of the action reported in the utterance; author (predicate defined in the task ontology) - the author of an item in a library; field (predicate defined in the task ontology) - the field of a document in a library; emitter (function defined in the discourse ontology, using objects in the task ontology) - retrieves the emitter of the utterance given as argument; $SARG$ (function defined in the discourse ontology, using objects defined in the task ontology) - retrieves the speech act-related goal of the utterance given as argument; good_time (predicate defined in the discourse ontology) - specified that the time interval given as argument is "good" with respect to the SARG of the speaker whom the utterance is intended for; Δt_{π_i} (predicate defined in the discourse ontology) - specifies the time interval integrated to the SARG of the utterance labels π_i;

(ii) binary predicates or functions: feature (predicate defined in the task ontology) - the variable specified as the second argument is a feature of the variable specified as the first argument; want (predicate defined in the task ontology) - the variable specified as the second argument is wanted by the agent specified as the first argument; have (predicate defined in the task ontology) - the variable specified as the second argument is possessed by the agent specified as the first argument; equals (predicate defined in the discourse ontology) - specified whether the arguments given are identical or not; Disjoint - specified whether the arguments given (as sets) are disjoint or not; entails (predicate defined in the discourse ontology, using objects in the task ontology) - specifies if the semantics of the utterance given as the second argument is implied by the semantics of the utterance given as the first argument, the time of the first utterance being unspecified;

(iii) ternary predicates: ExhaustiveDecomposition (predicate defined in the discourse ontology) - states that the entity specified by the third argument is the (set) union of the entities specified by the first two arguments;

(iv) objects: book (object defined in the task ontology) - specifies the *type* of an object referred to in an utterance; 'A', 'B', 'F' (objects defined in the task ontology) - specify the *instances* of several objects referred to in an utterance; t, $t_\forall$ (objects defined in the discourse ontology) - specify moments of time, stating present and indefinite time, respectively; $\emptyset$ (object defined in both the discourse and the task ontologies) - specifies the empty set or object; '?' (object defined in the task ontology) - specifies an unknown value for another object, when used as the second argument of the predicate equals.

In order to prove, for instance, that the relation *Background(π_4, π_2)* holds, the following steps are performed:

1. variable substitution, so that $K(\pi_4)$ and $K(\pi_2)$ on the one hand, and $K(\pi_2)$ and $K(\pi_5)$ on the other hand, do not have the same variable names; hence, in $K(\pi_2)$ the following substitutions are performed: X'/X, Y'/Y, Z'/Z, $T'/T, V'/V$;

2. explicit formal statement of the semantics of the rhetorical relation concerned: $Background(\pi_4, \pi_2)$ ::= equals(π_4, enounce) $\wedge$ entails(π_4, π_2) $\mapsto$ equals(π_4, enounce) $\wedge$ $(K(\pi_4) \Rightarrow K(\pi_2))$ $\wedge$ equals($\Delta t_{\pi_4}, t_\forall$);

3. verification of each "atomic" proposition in the semantics of the rhetorical relation (that is, of each proposition separated by the "$\wedge$" connector): for $Background(\pi_4, \pi_2)$:

– equals(π_4, enounce): this statement is true, since π_4 does not contain non initialised variables (that is, that are equal to '?');

– equals($\Delta t_{\pi_4}, t_\forall$): this statement is also true, from the semantics of the utterance π_4 (viz. the "atomic" propositions in $K(\pi_4)$);

– $K(\pi_4) \Rightarrow K(\pi_2)$: in order to prove this proposition, the following steps are performed:

(a) elimination of the existential quantifiers in $K(\pi_4)$ and $K(\pi_2)$:
 – $K(\pi_4)$ ::= equals(M, emitter(π_4)) $\wedge$ (author('A') $\wedge$ feature(book,'A') $\wedge$ have(M, book)) $\wedge$ (author('B') $\wedge$ feature(book,'B') $\wedge$ have(M, book)) $\wedge$ equals($\Delta t_{\pi_4}, t_\forall$),
 – $K(\pi_2)$ ::= equals($U, \neg$emitter(π_2)) $\wedge$ author('?') $\wedge$ field('F') $\wedge$ feature(book,'?') $\wedge$ feature(book, 'F') $\wedge$ want(U, book) $\wedge$ equals($\Delta t_{\pi_2}, t$);

(b) proof of $K(\pi_4) \Rightarrow K(\pi_2)$: $K(\pi_4) \Rightarrow K(\pi_2) \Leftrightarrow \neg K(\pi_2) \vee K(\pi_4)$, which is true if and only if $K(\pi_2) \wedge \neg K(\pi_4)$ is false; the falsity of the latter logic proposition results from its explicit statement (using the expression computed above) and from the fact that the set of values for the objects in the task and discourse ontologies is *closed* (Popescu *et al.*, 2007);

The rhetorical relation $P - Elab(\pi_2, \pi_5)$ is found to hold in a similar way; however, the truth values for these utterances are computed in a left-right manner, hence in the semantics of *P-Elab* only one part of the logic formula can be considered, namely: $P - Elab(\pi_2, \pi_5)$ ::= good_time(Δt_{π_5}) $\wedge$ $\neg$Disjoint($\Delta t_{\pi_5}, SARG(\pi_2)$) $\wedge$ ClassOf($K(\pi_5)$, Plan) $\wedge$ (Plan $\Rightarrow SARG(\pi_2)$), where the object Plan in the discourse ontology unites portions of the semantics of utterance π_5 such that it implies the SARG of π_2, want(U, book) $\wedge$ feature(book, 'F') $\wedge$ feature(book, '?') $\wedge$ field('F') $\wedge$ author('?').

5 Conclusions and Further Work

In this paper we have presented several improvements concerning a first-order logic-based rhetorical structuring component for natural language generation in human-computer dialogue, approximating SDRT in a first-order logic interpretation. These performance improvements concern the computational costs involved by the update of discourse structures, as well as the pertinence of the updated discourse structures obtained. We have used constraints induced by the speech acts performed by utterances on the set of possible rhetorical relations connecting them, showing that, with a computational cost around six times lower than when speech acts are not used, segmented discourse representation structures in a greater agreement with human intuitions are obtained.

At present, this rhetorical structuring component is being implemented in ISO-PROLOG; more specifically, the addition of the constraints induced by speech acts on discourse structure update is in progress and a prototype is under development. In the near future, it would be interesting to compare the current SDRT-based rhetorical structuring component for natural language generation with other implementations of SDRT, adapted to natural language generation purposes, for instance the system reported in (Danlos and El-Ghali, 2002); however, this latter system, as most of the work concerning computational implementation of SDRT address either *monologue* generation (Danlos *et al.*, 2003), or dialogue and monologue *interpretation* (Schlangen *et al.*, 2001).

References

N. Asher and A. Lascarides. 2003. *Logics of Conversation*. Cambridge University Press.

M. Bratman. 1987. *Intention, Plans and Practical Reason*. CSLI Publications.

J. Caelen and A. Xuereb. 2007. *Interaction et Pragmatique - jeux de dialogue et de langage*. Editions Hermès-Lavoisier, Paris.

L. Danlos and D. El-Ghali. 2002. A Complete Integrated NLG System Using AI and NLU Tools. *Proceedings of COLING 2002*, Taiwan.

L. Danlos, B. Gaiffe and L. Roussarie. 2003. Document Structuring à la SDRT. *Proceedings of the International Workshop on Generation, ACL*, Toulouse.

H. Kamp and U. Ryle. 1990. *From Discourse to Logic: An Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Institute of Computational Linguistics, University of Stuttgart.

W. C. Mann and S. A. Thompson. 1988. Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text*. 8(3):243–281.

D. Marcu. 1997. *The Rhetorical Parsing, Summarization and Generation of Natural Language Texts*, Ph. D. Thesis, University of Toronto.

M. F. McTear. 2002. Spoken Language Technology: Enabling the Conversational User Interface. *ACM Computer Surveys*, 34(1).

W. Minker and S. Bennacef. 2001. *Parole et dialogue homme-machine*. Editions CNRS-Eyrolles, Paris.

H. Nguyen. 2005. *Dialogue homme-machine: modélisation de multi-session*. Thèse de doctorat, Université Joseph Fourier, Grenoble.

M. Poesio and D. Traum. 1997. Conversational Actions and Discourse Situations. *Computational Intelligence*, 13(3):309–347.

V. Popescu, J. Caelen and C. Burileanu. 2007. Logic-Based Rhetorical Structuring Component in Natural Language Generation for Human-Computer Dialogue. *Proceedings of TSD 2007*, Pilsen, Czech Republic (submitted).

D. Schlangen, A. Lascarides and A. Copestake. 2001. Resolving Underspecification Using Discourse Information. *Proceedings of Bi-Dialog 2001*, Springer.

J. R. Searle. 1969. *Speech Acts: An Essay in Philosophy of Language*. Cambridge University Press.

A. Stent. 2001. *Dialogue Systems as Conversational Partners: Applying Conversation Acts Theory to Natural Language Generation for Task-Oriented Mixed-Initiative Spoken Dialogue*. Ph. D. Thesis, University of Rochester, New York.

D. Traum and E. Hinkelman. 1992. Conversation Acts in Task-Oriented Spoken Dialogue. *Computational Intelligence*, 8(3).

D. Vanderveken. 1990-1991. *Meaning and Speech Acts*. Cambridge University Press.

A. Xuereb and J. Caelen. 2005. Actes de langage et relations rhétoriques en dialogue homme-machine. *Revue de l'université de Moncton, Canada*, 36(2):5–51.