Model Generation for Discourse Representation Theory

Michael Kohlhase[1]

Abstract. We present a tableaux-based model generation calculus for DRT, which gives us an *incremental* approach to discourse processing in the presence of world knowledge. We show the usefulness of such as system for several discourse phenomena.

1 Introduction

Natural Language Processing (NLP) is one of the core areas of Artificial Intelligence. While some of the subtasks of NLP, such as speech recognition and syntactical analysis have made considerable progress towards practical applicability, the area of natural language semantics has so far mainly concentrated on representational issues and on the semantics construction process proper. The field of "dynamic semantics" studies logical systems that are well-suited for representing the meaning of discourses and in particular anaphoric binding. It has developed various "discourse logics", e.g. Discourse Representation Theory (DRT) [8] or Dynamic Predicate Logic (DPL) [6].

Semantic analysis, – inference on the basis of semantic information and world knowledge – still is largely uncharted territory in dynamic semantics. It is needed, among other things, for the reconstruction of linguistically unspecified parts of the discourse or for restricting ambiguities introduced by prior analysis processes, i.e., multiple syntactic readings or lexical ambiguities. Both need to take into account the pragmatics of the discourse and background knowledge about the specific situation.

For instance there are pragmatic principles that a discourse should be *consistent* and *informative*, that is, every contribution to the discourse should introduce new information which is not contradictory to previous parts. Thus a reading can be discarded if it breaks one of these (or similar) requirements. The consistency of an utterance U with respect to a prior discourse D can be checked by determining whether the logical forms $\mathbf{U}/\mathbf{D}$ of U/D together with an axiomatization $\mathbf{W}$ of the world knowledge are *satisfiable*, i.e. whether there is a model $\mathcal{M}$ for $\mathbf{U} \wedge \mathbf{D} \wedge \mathbf{W}$. Informativity can be modeled by *entailment* (U is informative for D iff $\mathbf{W} \wedge \mathbf{D} \not\models \mathbf{U}$, see [22, 2] for a linguistic account of the inference problems involved).

Thus, semantic analysis and automated reasoning in discourse logics are a crucial necessity in NLP from the application[2] point of view. The *dynamic* aspect is essential, since classical first-order logic (FOL) is incapable of expressing cross-sentential phenomena like anaphoric binding. Observe for instance that the conclusion of a logical consequence may be interpreted only with (anaphoric) reference to the premises, as in example (1) from [23]

$$\underbrace{\begin{array}{l}\textit{If a man owns a house, then he owns a garden.}\\ \textit{Suppose Socrates is a man who owns a house.}\end{array}}_{\models} \qquad (1)$$
He owns a garden.

In this paper, we will study these phenomena in the context of DRT. This is a dynamic variant of FOL, where so-called discourse referents take the place of bound variables to account for dynamic phenomena in natural language. In contrast to bound variables, the scope of referents is not governed by the subterm relation in formulae, but by the so-called accessibility relation induced by the dynamic connectives $\neg$, $seqmerge$, and $\Rrightarrow$.

Traditionally, DRT models discourse understanding as a process of *incrementally* building a Discourse Representation Structure (DRS; the boxes in 3) for a discourse. The DRS representing the meaning of the full discourse is computed by first adjoining the DRS for the new sentence using the sequential merge operator "$;;$" and then actually merging the DRSes by so-called τ-reduction. E.g. discourse (2) is represented and reduced in (3).

(2) *A man sleeps. He snores.*

(3) $$\begin{array}{|c|}\hline U\\\hline \mathsf{man}(U)\\ \mathsf{sleep}(U)\\\hline\end{array} \;;;\; \begin{array}{|c|}\hline V\\\hline \mathsf{male}(V)\\ \mathsf{snore}(V)\\\hline\end{array} \;\rightarrow_\tau\; \begin{array}{|c|}\hline U,V\\\hline \mathsf{man}(U)\\ \mathsf{sleep}(U)\\ \mathsf{snore}(V)\\\hline\end{array}$$

In a separate step, anaphoric bindings (in our example *He*) are reconstructed. In our example by adding the condition $U = V$ to the last DRS in (3) to the resulting DRS (see [8] for details).

This "dynamic" approach is certainly more natural than one based on FOL, where the first sentence would have been represented as $\exists X.\mathsf{man}(X) \wedge \mathsf{sleep}(X)$. Here the scope of the existential quantifiers is closed, preventing the resolution of the anaphor *He*. Unfortunately, current dynamic interpretation procedures are purely syntax-based and fail to take into account world knowledge.

Recently, the inference technique of *model generation* (initially developed for automated theorem proving [15]) has been very successfully applied as a semantic analysis technique. It is an *integrated*, *logic-based* approach to sentence understanding in the presence of world knowledge (see e.g. [4, 1, 12]). In this paper, we present a model generation calculus for DRT, that extends the results above to DRT (section 4). We demonstrate the usefulness of such as system on several discourse processing examples in section 5.

2 Discourse Representation Theory

We use a variant DRT^+ of DRT. As in classical DRT [8], we have two syntactic categories of objects in DRT: DRSes and conditions. A **DRS** is a pair $\delta\mathcal{X}.\mathbf{C}$[3] where $\mathcal{X} \subseteq \mathfrak{R}$ is a set of **discourse referents** (which we denote by $U, V, W \ldots$) and a **condition** $\mathbf{C}$, which can be an atom, a negated DRS, a disjunction or an implication of DRSes.

[1] Dept. of Computer Science, Universität des Saarlandes, Saarbrücken, Germany, email: kohlhase@cs.uni-sb.de

[2] Especially in discourse (or dialogue) processing applications, where the numbers readings of the sentences *multiply* to that of the whole discourse.

[3] We take the liberty to write DRSes in this linear notation instead of using boxes, since it makes the inference procedures simpler to write down. Furthermore, we will write $\delta\{U^1, \ldots, U^n\}.\mathbf{C}$ as $\delta\overline{U^n_{\mathbb{A}_n}}$ or as $\delta U^1.\delta U^2.\ldots.\delta U^n.\mathbf{C}$ or even as $\delta U^1.\mathbf{D}$, where $\mathbf{D} = \delta U^2.\ldots.\delta U^n.\mathbf{C}$.

Conditions:	$C ::=$	$p(U^1_{\mathbb{A}_1},\ldots,U^n_{\mathbb{A}_n}) \mid C_1 \wedge C_2$ $\mid \neg D \mid D_1 \vee D_2 \mid D_1 \Rrightarrow D_2$
DRSes:	$D ::=$	$\delta\{U^1_{\mathbb{A}_1},\ldots,U^n_{\mathbb{A}_n}\}.C \mid D_1 \mathbin{;;} D_2$
τ-equality:	$\delta\mathcal{X}.C_1 \mathbin{;;} \delta\mathcal{Y}C_2 \longrightarrow_\tau \delta\mathcal{X}\cup\mathcal{Y}.C_1 \wedge C_2$	

We assume that discourse referents in DRT$^+$ are **sorted**, i.e. that there is a sort hierarchy (a set $\mathcal{S} = \mathbb{A}, \mathbb{B}, \mathbb{C}\ldots$ of **sorts** and a partial ordering relation $\prec$ on $\mathcal{S}$) which is given a priory (see [19] for an introduction to sorted FOL). We annotate the sorts of discourse referents in the subscript, when they are not clear from the context. In our examples, we will use the sorts $\mathbb{M}, \mathbb{F}, \mathbb{N}$ for genders "male, female" and "neuter" and $\mathbb{H}$ for humans (of course we have $\mathbb{M} \prec \mathbb{H}$, $\mathbb{F} \prec \mathbb{H}$).

Traditionally, the semantics of DRSes is given by the following relativisation mapping into FOL:

$$\left(\begin{array}{|c|}\hline \overline{U^n_{\mathbb{A}_n}} \\ \hline \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_m \\ \hline \end{array}\right)^{fo} = \exists\overline{U^n_{\mathbb{A}_n}}.(\mathbf{C}_1^{fo} \wedge \cdots \wedge \mathbf{C}_m^{fo})$$

$$(p(U^1_{\mathbb{A}_1},\ldots,U^n_{\mathbb{A}_n}))^{fo} = p(U^1_{\mathbb{A}_n},\ldots,U^n_{\mathbb{A}_n})$$

$$(U_{\mathbb{A}} = V_{\mathbb{B}})^{fo} = U_{\mathbb{A}} = V_{\mathbb{B}}$$

$$(\neg\mathcal{D})^{fo} = \neg\mathcal{D}^{fo}$$

$$(\mathbf{C}_1 \vee \mathbf{C}_2)^{fo} = \mathbf{C}_1^{fo} \vee \mathbf{C}_2^{fo}$$

$$\left(\begin{array}{|c|}\hline \overline{U^n_{\mathbb{A}_n}} \\ \hline \mathbf{C}_1 \\ \vdots \\ \mathbf{C}_m \\ \hline \end{array} \Rrightarrow \mathbf{D}\right)^{fo} = \forall\overline{U^n_{\mathbb{A}_n}}.(\mathbf{C}_1^{fo} \wedge \cdots \wedge \mathbf{C}_m^{fo} \Rightarrow \mathbf{D}^{fo})$$

An (equivalent) alternative is to give a direct denotational semantics of DRT. For this, we presuppose the notion of a sorted FO model $\mathcal{M} = (\mathcal{U}, \mathcal{I})$, where $\mathcal{U} = \bigcup_{\mathbb{A}\in\mathcal{S}} \mathcal{U}_{\mathbb{A}}$ is the **universe of discourse** such that $\mathcal{U}_{\mathbb{A}} \subseteq \mathcal{U}_{\mathbb{B}}$, iff $\mathbb{A} \prec \mathbb{B}$ and $\mathcal{I}$ an interpretation of constants such that $\mathcal{I}(c_{\mathbb{A}}) \in \mathcal{U}_{\mathbb{A}}$.

Definition 1 (State) Let $\mathcal{M} := (\mathcal{U}, \mathcal{I})$ be a sorted FO model, then we call a referent assignment $\varphi: \mathfrak{R} \to \mathcal{U}$ a **state**, iff it is well-sorted ($\varphi(U_{\mathbb{A}}) \in \mathcal{U}_{\mathbb{A}}$). We write $\varphi[\mathcal{X}]\psi$, if $\varphi(U) = \psi(U)$ for all $U \notin \mathcal{X}$.

Definition 2 (Dynamic Interpretation) Let $\mathcal{M} = (\mathcal{U}, \mathcal{I})$ be a sorted FO model and φ a state, then we call $\mathcal{I}^\delta$ a **dynamic interpretation**[4] iff

1. $\mathcal{I}^\delta_\varphi(p(\overline{U^n})) = \mathtt{T}$, iff $(\overline{\varphi(U^n)}) \in \mathcal{I}(p)$.
2. $\mathcal{I}^\delta_\varphi(\mathbf{A} \wedge \mathbf{B}) = \mathtt{T}$, iff $\mathcal{I}^\delta_\varphi(\mathbf{A}) = \mathtt{T}$ and $\mathcal{I}^\delta_\varphi(\mathbf{B}) = \mathtt{T}$.
3. $\mathcal{I}^\delta_\varphi(\neg\mathbf{D}) = \mathtt{T}$, if $\pi_2(\mathcal{I}^\delta_\varphi(\mathbf{D})) = \emptyset$.
4. $\mathcal{I}^\delta_\varphi(\mathbf{D} \vee \mathbf{E}) = \mathtt{T}$, if $\pi_2(\mathcal{I}^\delta_\varphi(\mathbf{D})) \neq \emptyset$ or $\pi_2(\mathcal{I}^\delta_\varphi(\mathbf{E})) \neq \emptyset$.
5. $\mathcal{I}^\delta_\varphi(\mathbf{D} \Rightarrow \mathbf{E}) = \mathtt{T}$, if for every $\psi \in \pi_2(\mathcal{I}^\delta_\varphi(\mathbf{D}))$ there is a $\tau \in \pi_2(\mathcal{I}^\delta_\varphi(\mathbf{E}))$ with $\varphi[\pi_1(\mathcal{I}^\delta_\varphi(\mathbf{E}))]\tau$.
6. $\mathcal{I}^\delta_\varphi(\delta\mathcal{X}.\mathbf{C}]) = \langle\mathcal{X}, \{\psi: \varphi[\mathcal{X}]\psi \text{ and } \mathcal{I}^\delta_\varphi(\mathbf{C}) = \mathtt{T}\}\rangle$.
7. $\mathcal{I}^\delta_\varphi(\mathbf{D}_1 \mathbin{;;} \mathbf{D}_2) = \langle\pi_1(\mathcal{I}^\delta_\varphi(\mathbf{D})) \cup \pi_1(\mathcal{I}^\delta_\varphi(\mathbf{E})), \pi_2(\mathcal{I}^\delta_\varphi(\mathbf{D})) \cap \pi_2(\mathcal{I}^\delta_\varphi(\mathbf{E}))\rangle$

Conditions are evaluated to truth values; DRSes as pairs $\langle\mathcal{X}, \mathfrak{S}\rangle$, where $\mathcal{X} \subseteq \mathfrak{R}$ is a set of discourse referents and $\mathfrak{S}$ a set of states (we denote pair-projections by π_1 and π_2).

We will call a DRS $\mathbf{D}$ **valid** in $\mathcal{M}$, if $\pi_2(\mathcal{I}^\delta_\varphi(\mathbf{D})) \neq \emptyset$ and **satisfiable**, iff there is a model $\mathcal{M}$ where $\mathbf{D}$ is valid.

[4] This is a variant of Zeevat's semantics [25] for DRT, see [11] for details.

For instance the DRS (3 with anaphor resolution $U = V$) translates to (4) and has the direct semantics given in (5), assuming a sorted FO model $\mathcal{M} = (\mathcal{U}, \mathcal{I})$ with $\mathsf{a} \in \mathcal{U}_{\mathbb{H}}$.

(4) $\exists X_{\mathbb{H}}, Y_{\mathbb{H}}.\mathsf{man}(X) \wedge \mathsf{sleep}(X) \wedge \mathsf{snore}(Y) \wedge X \doteq Y$

(5) $\left\langle \{U_{\mathbb{H}}, V_{\mathbb{H}}\}, \{[\mathsf{a}/U], [\mathsf{a}/V]: \mathsf{a} \in \mathcal{I}(\mathsf{man}) \cap \mathcal{I}(\mathsf{snore}) \cap \mathcal{I}(\mathsf{sleep})\right\rangle$

It is easy to verify that the FO formula in (4) is satisfiable, iff there is a FO model, such that $\pi_2(5) \neq \emptyset$. As a consequence, validity and satisfiability of DRSes can in principle be checked by a translation approach: we can obtain FO models for a DRS by first translating it into FOL, and then using a traditional model building or refutation system (see [2] for an implementation and evaluation of this approach).

The main computational problem with this approach is that it is not incremental, since translation closes all dynamic contexts. The main difference between the semantical representations in (4) and (5) is that in the FOL translation[5], the scope of the existential quantifiers is fixed and the only possibility to add new sentences (with anaphors) is to adjoin them at the level of DRS representations and retranslate. Certainly, this is not practically feasible for larger discourses even though there are linear time translations [24].

3 (Automated) Deduction for Discourse Logics

There have been several attempts to mechanize dynamic logics, i.e. to develop calculi and inference procedures for the satisfiability, validity and entailment problems for dynamic logics. [20, 18, 17, 9, 23, 16] give deductive calculi that operate either on DRSes or on FO formulae with dynamic (DPL) semantics.

[20, 17, 9] present calculi for the *validity problem* in DRT and [16] for that in DPL. In FOL, it is sufficient to study the validity problem, since it subsumes the entailment problem: FOL admits a deduction theorem, so $\mathbf{A}_1, \ldots, \mathbf{A}_n \models \mathbf{C}$ iff $\mathbf{A}_1 \wedge \ldots \wedge \mathbf{A}_n \Rightarrow \mathbf{C}$ is valid. Classical DRT [8] does not admit a deduction theorem, since the symmetric merge operator $\otimes$ which is the dynamic analogue of conjunction does not have the necessary accessibility relation.

The sequential merge operator $;;$ of DPL, does[6]; and as a consequence [16] is currently the only calculus that can be used to check for dynamic entailment. Unfortunately, it suffers from the same lack of incrementality as the translation approach. To test that a discourse $\mathbf{A}_1 \mathbin{;;} \ldots \mathbin{;;} \mathbf{A}_n$ does not entail a new utterance $\mathbf{C}$ (e.g. to check for informativity) Monz and De Rijke's calculus needs to transform $\mathbf{A}_1 \wedge \ldots \wedge \mathbf{A}_n \Rightarrow \mathbf{C}$ to a *dynamic* clause normal form that can be refuted by a variant of the resolution calculus. Since dynamic conjunction is not symmetric, it is impossible to reuse the computed clause normal form for subsequent informativity checks ($\mathbf{CNF}(\neg(\mathbf{A}_1 \wedge \ldots \wedge \mathbf{A}_n \Rightarrow \mathbf{C}))$ and $\mathbf{CNF}(\neg(\mathbf{A}_1 \wedge \ldots \wedge \mathbf{A}_n)) \cup \mathbf{CNF}(\neg\mathbf{C})$ differ).

Saurer's natural deduction calculus for DRT [18] which *is* incremental is only sound for checking validity and static entailment (which does not take into account anaphoric binding). Jan van Eijck's sequent-based approach [23] directly addresses the *entailment problem* but has not been developed for mechanization in an automated theorem prover. Therefore the incrementality issue is hard to judge.

In this situation, we will generalize an inference technique from FOL that is inherently incremental, namely that of *model generation*.

[5] Note that this only holds for classical FOL. In DPL [6] that assumes a dynamic semantics similar to ours but keeps classical FO syntax, the scope of quantifiers is governed by similar principles as the DRT accessibility relation. Therefore it is possible to simply adjoin the semantics of new sentences by (dynamic) conjunction (which corresponds to $;;$).

[6] This is the reason, why we use $;;$ for sentence composition in DRT$^+$.

This inference approach is dual to that of refutation theorem proving: instead of trying to find a refutation showing unsatisfiability of the negation of the formula to be proven, model generation tries to show satisfiability by constructing a model.

In the next section, we will generalize the notion of Herbrand models used in FO model generation to DRT and then generalize the $\mathcal{RM}$ model generation calculus [14, 13] accordingly. Then we will show in section 5 that the calculus can be used to account for a variety of linguistic phenomena.

4 Model Generation for DRT

In this section, we will develop a "model generation calculus" for DRT. In contrast to FOL, the scope of discourse referents is not governed by the term structure, but by the DRT accessibility relation. The truth definition with respect to $\mathcal{I}^\delta$ crucially depends on the current state φ, therefore, "model generation" for dynamic logics must also generate states and dynamic interpretations along the way.

FO model generation relies on the well-known Herbrand theorem that singles out Herbrand models as canonical representatives of models (if a FO theory is satisfiable at all, then it must be satisfiable in a Herbrand model).

Definition 3 (First-Order Herbrand Model) Let $\mathcal{L}$ be a (sorted) FO language, then the set $\mathcal{H}^\mathcal{L} = \bigcup_{\mathbb{A}\in\mathcal{S}} \mathcal{H}^\mathcal{L}_\mathbb{A}$ of closed terms (of sort $\mathbb{A}$) in $\mathcal{L}$ is called the **Herbrand universe** of $\mathcal{L}$. Let $\mathcal{M} = (\mathcal{U}, \mathcal{I})$ be a (sorted) FO model, then we call $\mathcal{M}$ a **Herbrand Model**, iff $\mathcal{U}_\mathbb{A} \subseteq \mathcal{H}^\mathcal{L}_\mathbb{A}$ and $\mathcal{I}(t) = t \in \mathcal{U}$ for all ground terms $t \in \mathcal{H}^\mathcal{L}$.

In a Herband model, only the interpretation of predicate symbols must be specified, e.g. by giving values on the closed literals of $\mathcal{L}$. Thus any Herbrand model $\mathcal{M}$ can be uniquely represented (under a closed-world assumption) by the set of closed atoms it makes true (the so-called **Herbrand base** $\mathcal{AT}(\mathcal{M})$).

A Herbrand model $\mathcal{M}$ is called **finite** if its universe of discourse $\mathcal{U}$ is finite and **minimal** if for all $\mathcal{M}'$ the following holds: $\mathcal{AT}(\mathcal{M}') \subseteq \mathcal{AT}(\mathcal{M}) \Rightarrow \mathcal{M}' = \mathcal{M}$. It is called **domain minimal** if $|\mathcal{U}(\mathcal{M})| \leq |\mathcal{U}(\mathcal{M}')|$.

The tableaux-based model generation procedure $\mathcal{T}_\delta$, which we will introduce in this section constructs a Herbrand model together with a state and a dynamic interpretation to verify that a given DRS is satisfiable. DRT is ideal for model generation applications since it does not contain function symbols, i.e. $\mathcal{H}^{\mathrm{DRT}^+}$ is a set of constants; as a consequence it is possible to generate finite Herbrand models.

The $\mathcal{T}_\delta$ model generation calculus is based on the $\mathcal{RM}$ calculus (see Definitions 4 and 5), which has been originally developed for a certain form of non-monotonic reasoning, called minimal entailment [14].

$$\frac{\neg\mathbf{A}^{\mathtt{T}}}{\mathbf{A}^{\mathtt{F}}}\mathcal{T}(\neg) \qquad \frac{\neg\mathbf{A}^{\mathtt{F}}}{\mathbf{A}^{\mathtt{T}}}\mathcal{T}(\neg) \qquad \frac{\begin{array}{c}\mathbf{A}^{\mathtt{T}}\\ \mathbf{A}^{\mathtt{F}}\end{array}}{\bot}\mathcal{T}(\bot)$$

$$\frac{(\mathbf{A}\wedge\mathbf{B})^{\mathtt{T}}}{\begin{array}{c}\mathbf{A}^{\mathtt{T}}\\ \mathbf{B}^{\mathtt{T}}\end{array}}\mathcal{T}(\wedge) \qquad \frac{(\mathbf{A}\wedge\mathbf{B})^{\mathtt{F}}}{\mathbf{A}^{\mathtt{F}} \mid \mathbf{B}^{\mathtt{F}}}\mathcal{T}(\vee) \qquad \frac{(\forall X_\mathbb{A}.\mathbf{A})^{\mathtt{T}}}{[a_1/X]\mathbf{A}^{\mathtt{T}}}\mathcal{T}(\forall)$$

$$\frac{(\forall X_\mathbb{A}.\mathbf{A})^{\mathtt{F}}}{[a_1/X]\mathbf{A}^{\mathtt{F}} \mid \ldots \mid [a_n/X]\mathbf{A}^{\mathtt{F}} \mid [c^{new}_\mathbb{A}/X_\mathbb{A}]\mathbf{A}^{\mathtt{F}}}\mathcal{T}(\exists)$$

Definition 4 (Static Model Generation) The static model generation calculus consists of the usual tableau rules for the connectives and the model generation rules for the quantifiers. The $\mathcal{T}(\forall)$ rule tests the scope on all members of the Herbrand universe $\mathcal{H}_\mathbb{A}$ of the current branch ($\mathcal{H}_\mathbb{A} = \{a_1, \ldots, a_n\}$); it must be applied exhaustively to obtain a saturated branch. The $\mathcal{T}(\exists)$ rule reuses constants that occur in the current branch and alternatively introduces a new constant $c^{new}_\mathbb{A}$. In this way, $\mathcal{T}(\exists)$ minimizes the size of the universe and also avoids Skolem functions which would introduce problematic function symbols into the Herbrand universe. When extending the Herbrand universe of a branch by $c^{new}_\mathbb{A}$, all $\mathcal{T}(\forall)$ must be re-instantiated with respect to $c^{new}_\mathbb{A}$.

To get a feeling for the model construction process, let us consider a simple (static) sentence *No man walks* in a situation including a man (say Peter). The logical form is $\neg(\exists X \mathsf{man}(X) \wedge \mathsf{walk}(X))$ (obtained e.g. as $\neg(\delta U.\mathsf{man}(U) \wedge \mathsf{walk}(U))^{fo}$) and we have the tableau

$$\begin{array}{c}\mathsf{man}(\mathsf{peter})^{\mathtt{T}}\\ (\exists X \mathsf{man}(X) \wedge \mathsf{walk}(X))^{\mathtt{F}}\\ (\mathsf{man}(\mathsf{peter}) \wedge \mathsf{walk}(\mathsf{peter}))^{\mathtt{F}}\\ \begin{array}{c|c}\begin{array}{c}\mathsf{man}(\mathsf{peter})^{\mathtt{F}}\\ \bot\end{array} & \mathsf{walk}(\mathsf{peter})^{\mathtt{F}}\end{array}\end{array} \tag{6}$$

$\mathcal{T}(\forall)$ converts the negative existential (interpreted as a universal via $\exists x.\mathbf{A} = \neg\forall X.\neg\mathbf{A}$) into a negative conjunction which is then split into two branches by $\mathcal{T}(\wedge)$. The left one is contradictory with the information already present in the model, so we obtain the minimal Herbrand model $\{\mathsf{man}(\mathsf{peter})^{\mathtt{T}}, \mathsf{walk}(\mathsf{peter})^{\mathtt{F}}\}$.

This example gives us the opportunity to compare the influence of sorts. If we had chosen to model the predicate man as a sort $\mathbb{M}an$, then the declaration peter::$\mathbb{M}an$ would be part of the signature and we would have interpreted the second sentence as the following tableau:

$$\begin{array}{c}(\exists X_{\mathbb{M}an}.\mathsf{walk}(X))^{\mathtt{F}}\\ \mathsf{walk}(\mathsf{peter})^{\mathtt{F}}\end{array} \tag{7}$$

It is easy to see that the introduction of sorts yielded a smaller initial representation and a more guided computation (without infertile branches for half the population). In particular, sorts make the model generation calculus less vulnerable to computational inefficiency induced by non-trivial but unstructured universes. Of course the simple sort system employed for our examples has to be extended to a more elaborate one for real-world applications, we will not pursue this here, and leave the integration of more expressive sort systems like terminological logics e.g. KL-ONE for future work.

$\mathcal{RM}$ is a refutation complete FO tableau calculus where each open saturated branch is a Herbrand model. [12] proves that $\mathcal{RM}$ is complete for finite satisfiability, i.e. $\mathcal{RM}$ is a decision procedure for theories that either are unsatisfiable or have a finite model. Additionally, $\mathcal{RM}$ is complete for finite minimal models that also are domain minimal: if a theory is finitely satisfiable, then one of the models generated by $\mathcal{RM}$ will be minimal with the smallest possible universe. These properties are inherited by the $\mathcal{T}_\delta$ calculus defined below, making it an ideal basis for the linguistic applications. The proofs are straightforward, we cannot exhibit the proofs here for space restrictions.

Definition 5 (Dynamic Model Generation) The $\mathcal{T}_\delta$ calculus extends static model generation by the inference rules in the box below. Like $\mathcal{T}(\exists)$, the rule $\mathcal{T}_\delta(\delta)$ is existential in nature, it introduces a witness constant c^{new} with the consequences for universal formulae discussed above. Furthermore, it extends the state represented by the current branch in all possible ways by *state* nodes of the form $[c/U]$. In this way it captures the accessibility relation of DRT. Sentence composition ($;;$) is mechanized by adding the respective DRS at all leaves (instantiated by the state represented in the respective branch).

Conditions are mechanized by translation, since they cannot change the current state, but only the Herbrand-representation of the current model.

$$\frac{\delta U_{\mathbb{A}}.\mathbf{A}^{\mathtt{T}} \quad \mathcal{H}_{\mathbb{A}} = \{a_1,\ldots,a_n\}}{\begin{array}{c|c|c|c} [a_1/U] & & [a_1/U] & [c_{\mathbb{A}}^{new}/U_{\mathbb{A}}] \\ [a_n/U]\mathbf{A}^{\mathtt{T}} & \cdots & [a_n/U]\mathbf{A}^{\mathtt{T}} & [c_{\mathbb{A}}^{new}/U_{\mathbb{A}}]\mathbf{A}^{\mathtt{T}} \end{array}} \mathcal{T}_{\delta}(\delta)$$

$$\frac{\neg\mathcal{D}}{(\neg\mathcal{D})^{fo}} \mathcal{T}_{\delta}(\neg) \quad \frac{\mathcal{D} \Rrightarrow \mathcal{D}'}{(\mathcal{D} \Rrightarrow \mathcal{D}')^{fo}} \mathcal{T}_{\delta}(\Rrightarrow) \quad \frac{\mathcal{D} \mathbin{\mathbb{W}} \mathcal{D}'}{(\mathcal{D} \mathbin{\mathbb{W}} \mathcal{D}')^{fo}} \mathcal{T}_{\delta}(\mathbb{W})$$

The dynamic interpretation induced by a given branch $\mathcal{B}$ in the tableau is determined by the positive literals $\mathcal{L}it^{\mathtt{T}}(\mathcal{B})$ and the state $\varphi(\mathcal{B})$ induced by the state nodes. It is the pair $\langle \mathbf{Dom}(\varphi), \{\varphi\colon \mathcal{L}it^{s}emtrue(\mathbf{B})\}\rangle$. Note that dynamic interpretation generalizes the induced Herbrand model by (dynamic) state information.

If we reconsider our example from above, we can see that we obtain the tableau in (6/7) resulting in the minimal dynamic Herbrand interpretation $\langle \emptyset, \{\emptyset\colon \mathsf{man}(\mathsf{peter})^{\mathtt{T}}, \mathsf{walk}(\mathsf{peter})^{\mathtt{F}}\}\rangle$. This is plausible, since our little discourse is static (does not have any anaphoric potential).

5 Linguistic Applications

We will now test the proposed model generation approach to discourse processing on some well-known examples from the literature.

Anaphora resolution is just a simple consequence of the search for minimal models. Consider for instance the discourse (2). Then we obtain the following tableau:

$$\begin{array}{c} \delta U_{\mathbb{M}}.\mathsf{man}(U) \wedge \mathsf{sleep}(U)^{\mathtt{T}} \\ [c^1_{\mathbb{M}}/U_{\mathbb{M}}] \\ \mathsf{man}(c^1_{\mathbb{M}})^{\mathtt{T}} \\ \mathsf{sleep}(c^1_{\mathbb{M}})^{\mathtt{T}} \\ \delta V \mathbb{M}.\mathsf{snore}(V)^{\mathtt{T}} \\ \begin{array}{c|c} [c^1_{\mathbb{M}}/V_{\mathbb{M}}] & [c^2_{\mathbb{M}}/V_{\mathbb{M}}] \\ \mathsf{snore}(c^1_{\mathbb{M}})^{\mathtt{T}} & \mathsf{snore}(c^2_{\mathbb{M}})^{\mathtt{T}} \end{array} \end{array}$$

which leads to the two dynamic interpretations
$\langle\{U_{\mathbb{M}}, V_{\mathbb{M}}\}, \{[c^1_{\mathbb{M}}/U_{\mathbb{M}}], [c^1_{\mathbb{M}}/V_{\mathbb{M}}]\colon \mathsf{man}(c^1_{\mathbb{M}}), \mathsf{sleep}(c^1_{\mathbb{M}}), \mathsf{snore}(c^1_{\mathbb{M}})\}\rangle$
$\langle\{U_{\mathbb{M}}, V_{\mathbb{M}}\}, \{[c^1_{\mathbb{M}}/U_{\mathbb{M}}], [c^2_{\mathbb{M}}/V_{\mathbb{M}}]\colon \mathsf{man}(c^1_{\mathbb{M}}), \mathsf{sleep}(c^1_{\mathbb{M}}), \mathsf{snore}(c^2_{\mathbb{M}})\}\rangle$

We can see that the anaphor resolution is a direct consequence of the $\mathcal{T}_{\delta}(\delta)$ rule. Both possible interpretations (one where *He* refers to the sleeping man introduced before, and also the deictic use of *He* that does not need an antecedent or accommodates (infers) one) have been derived. However, only the first one is minimal, and leads to the preferred interpretation.

The particular computation in this example only relies on the fact that there are no men in the context (more would have lead to more interpretations). The number of e.g. women is irrelevant due to the presence of sorts: in both applications of the $\mathcal{T}_{\delta}(\delta)$ rule, the discourse referents could only by assigned to constants of sort $\mathbb{M}$. Even if we choose not to represent gender by sorts but by the unary predicates male, female (say, since we are in a context, where genders can change), the interpretation process works, only that with n women we would get $2n$ additional closed branches as in the tableau (6) compared to the one in (7)

Traditional approaches to anaphora resolution would have obtained the same behavior, but on different grounds. There, the information about gender would have been treated on a *syntactic* basis, making the (reasonable) assumption that the world knowledge that "men are not women" is hard-wired into the grammar. While this is reasonable for syntactically marked properties like gender, the inference-based approach also generalizes to other sorts.

Let us now consider an example, where real world knowledge comes into play. To resolve the pronouns and the implicit reference in *her husband* in (8), we need to know that *if a female X is married to a male Y, then Y is X's only husband*, which is encoded in (10).

(8) *Mary is married to Jeff. Her husband is not in town.*

(9) $\delta U_{\mathbb{F}}, V_{\mathbb{M}}.U = \mathsf{mary} \wedge \infty(U,V) \wedge V = \mathsf{jeff}$;; $\delta W_{\mathbb{M}}, W'_{\mathbb{F}}.\mathsf{hub}(W,W') \wedge \neg\mathsf{intn}(W)$

(10) $\forall X_{\mathbb{F}}, Y_{\mathbb{M}}.\infty(X,Y) \Rightarrow (\mathsf{hub}(Y,X) \wedge \forall Z.\mathsf{hub}(Z,X) \Rightarrow Z \doteq Y)$

Model generation using $\mathcal{T}_{\delta}$ yields a (rather large) tableau, whose branches all contain the interpretation

$$\left\langle \{U,V,W,W'\}, \left\{ \begin{array}{l} [\mathsf{mary}/U], [\mathsf{jeff}/V], [\mathsf{jeff}/W], [\mathsf{mary}/W'] \\ :\ \infty(\mathsf{mary},\mathsf{jeff}), \mathsf{hub}(\mathsf{jeff},\mathsf{mary}), \neg\mathsf{intn}(\mathsf{jeff}) \end{array} \right\} \right\rangle$$

and only differ in some additional negative facts that have been created by the world knowledge, for instance $\infty(\mathsf{mary},\mathsf{mary})^{\mathtt{F}}$.

The figure below shows one branch of the model generation *without* using the world knowledge (10). In the associated reading, we get a new discourse referent $c^1_{\mathbb{M}}$ that denotes Mary's husband although we already have given the information that Mary's husband is Jeff.

$$\begin{array}{c} (\delta U_{\mathbb{F}}, V_{\mathbb{M}} U \doteq \mathsf{mary} \wedge \infty(U,V) \wedge V \doteq \mathsf{jeff})^{\mathtt{T}} \\ [\mathsf{mary}/U] \\ (\delta V\ \mathsf{mary} \doteq \mathsf{mary} \wedge \infty(\mathsf{mary},V) \wedge V \doteq \mathsf{jeff})^{\mathtt{T}} \\ [\mathsf{jeff}/V] \\ (\mathsf{mary} \doteq \mathsf{mary} \wedge \infty(\mathsf{mary},\mathsf{jeff}) \wedge \mathsf{jeff} \doteq \mathsf{jeff})^{\mathtt{T}} \\ \mathsf{mary} \doteq \mathsf{mary}^{\mathtt{T}} \\ \infty(\mathsf{mary},\mathsf{jeff})^{\mathtt{T}} \\ \mathsf{jeff} \doteq \mathsf{jeff}^{\mathtt{T}} \\ (\delta W_{\mathbb{M}} W'_{\mathbb{F}}.\mathsf{hub}(W,W') \wedge \neg\mathsf{intn}(W))^{\mathtt{T}} \\ [c^1_{\mathbb{M}}/W] \\ (\delta W'\ \mathsf{hub}(c^1_{\mathbb{M}},W) \wedge \neg\mathsf{intn}(c^1_{\mathbb{M}}))^{\mathtt{T}} \\ [\mathsf{mary}/W'] \\ (\mathsf{hub}(c^1_{\mathbb{M}},\mathsf{mary}) \wedge \neg\mathsf{intn}(c^1_{\mathbb{M}}))^{\mathtt{T}} \\ \mathsf{hub}(c^1_{\mathbb{M}},\mathsf{mary})^{\mathtt{T}} \\ \mathsf{intn}(c^1_{\mathbb{M}})^{\mathtt{F}} \end{array}$$

Let us now turn to a phenomenon, called bridging. Concretely, we will analyze the utterance

(11) *The Boston office called.*

(12) $\delta U_{\mathbb{H}}, V_{\mathbb{N}}, W_{\mathbb{N}}.\mathsf{boston}(U) \wedge \mathsf{office}(V) \wedge \mathsf{called}(W) \wedge \mathsf{rel}(U,V) \wedge \mathsf{rel}(V,W)$

introduced by Hobbs et al. in [7]. This sentence has at least three pragmatic problems that need world knowledge: resolving the reference of "the Boston office", expanding the metonymy to "[Some person at] the Boston office called", and determining the implicit relation between Boston and the office.

We will do this model generation from roughly the same world knowledge as in [7]:

(13) $\mathsf{boston}(b), \mathsf{office}(o),$

(14) $\forall X_{\mathbb{H}}, Y_{\mathbb{N}}.\mathsf{employed}(X,Y) \Rightarrow \mathsf{rel}(X,Y)$

(15) $\forall X_{\mathbb{N}}, Y_{\mathbb{N}}.\mathsf{intn}(X,Y) \Rightarrow \mathsf{rel}(X,Y)$

Given this input, $\mathcal{T}_{\delta}$ will (among others) generate the following tableau branch:

$$\begin{array}{c}(12)\\ [b/U],[o/V],[c^1_{\mathbb{H}}/W]\\ \mathsf{called}(c^1_{\mathbb{H}})^{\mathtt{T}}\\ \mathsf{office}(o)^{\mathtt{T}}\\ \mathsf{boston}(b)^{\mathtt{T}}\\ \mathsf{intn}(o,b)^{\mathtt{T}}\\ \mathsf{employed}(c^1_{\mathbb{H}},o)^{\mathtt{T}}\end{array}$$

This model is minimal, since the last five literals are already present in any Herbrand model, since they are entailed by the world knowledge. If we have additional knowledge, such as $\mathsf{employed}(\mathsf{harry}, o)$, then we obtain additional minimal models, in this case the one, where $c^1_{\mathbb{H}}$ is replaced by harry.

Naturally, a more thorough analysis of the example would also take into account the uniqueness presupposition induced by *the*, or the salience of relatedness. In [10], we have a $\mathcal{RM}$-like model generation calculus with saliences and a system for weighting inferences like that the weighted abduction system introduced in [7]. This leads to a more flexible notion of minimality of interpretations and thus to better predictions about preferred interpretations. Furthermore, the possibility to to resource-bounded best-first search helps control search spaces involved in model generation. It will be a logical next step to transport these methods to $\mathcal{T}_\delta$, to combine the advantages.

6 Conclusion

We have presented a model-generation calculus for DRT and proven it sound and (refutationally and minimal model) complete with respect to a natural dynamic (state-based) semantics for DRT. In contrast to other calculi for dynamic discourse logics, our approach offers an *incremental* inference procedure that allows to integrate world knowledge into the natural language understanding process.

We have exhibited a variety of examples that suggest that the incremental, dynamic model generation procedure can serve as a plausible analysis for natural language understanding. The underlying model-generation-based approach has been validated for the static case in e.g. [4, 1, 12]. While our examples support this claim at a technical level, the psycholinguistic literature supports the model-based analysis from a conceptual and scientific point:

Numerous psycholinguistic studies have shown that during discourse comprehension readers or listeners not only represent the logical form of a text but also construct a representation of the states of affairs described by the text, i.e. a representation, the elements of which are mental tokens standing for the referents of linguistic expressions (for an overview see [26]). These representations are constructed on-line during discourse comprehension in an incrementally manner (e.g. [3], they are enriched by a large amount of world knowledge (cf. [21]) and their major function is to provide the basis for anaphor resolution (e.g. [5])

An interesting question that remains to be answered is how the dynamics inherent in the model construction process (e.g. new witness constants are introduced into the Herbrand universe) and the dynamics explicit in discourse logics like DRT [8] or DPL [6] interact. The analysis in this paper suggests that they can happily coexist, and even more that for inference purposes, model generation can be harnessed to implement an adequate inference procedure for dynamic/discourse logics. To determine whether this effect can be extended to the whole field of dynamic semantics we will leave to further research.

REFERENCES

[1] Peter Baumgartner and Michael Kühn, 'Abducing coreference by model construction', in *Proceedings of Inference in Computational Semantics*, eds., Christof Monz and Maarten de Rijke, pp. 21–39, Amsterdam, (1999).

[2] Patrick Blackburn, Johan Bos, Michael Kohlhase, and Hans de Nivelle, 'Inference and computational semantics' International Workshop on Computational Semantics IWCS-3, Tilburg, 1999.

[3] M. de Vega, 'Backward updating of mental models during continuous reading of narratives', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 373–385, (1995).

[4] Claire Gardent and Karsten Konrad, 'Definites and the proper treatment of rabbits', in *Proceedings of Inference in Computational Semantics*, eds., Christof Monz and Maarten de Rijke, pp. 53–69, Amsterdam, (1999). ILLC.

[5] A. M. Glenberg, M. Meyer, and K. Lindem, 'Mental models contribute to foregrounding during text comprehension', *Journal of Memory and Language*, **26**, 69–83, (1987).

[6] Jeroen Groenendijk and Martin Stokhof, 'Dynamic predicate logic', *Linguistics & Philosophy*, **14**, 39–100, (1991).

[7] J. Hobbs, M. Stickel, D. Appelt, and P. Martin, 'Interpretation as abduction', *Artificial Intelligence*, **63**, 69–142, (1993).

[8] Hans Kamp and Uwe Reyle, *From Discourse to Logic*, Kluwer, Dordrecht, 1993.

[9] Hans Kamp and Uwe Reyle, 'A calculus for first order discourse representation structures', *Journal of Logic, Language and Information*, **5**(3-4), 297–348, (1996).

[10] Michael Kohlhase and Alexander Koller, 'Towards A Tableaux Machine for Language Understanding', in *Proceedings of Inference in Computational Semantics ICoS-2*, eds., J. Bos and M. Kohlhase, (2000) forthcoming.

[11] Michael Kohlhase, Susanna Kuschert, and Manfred Pinkal, 'A type-theoretic semantics for λ-DRT', in *Proceedings of the 10th Amsterdam Colloquium*, eds., P. Dekker and M. Stokhof, pp. 479–498, Amsterdam, (1996). ILLC.

[12] Karsten Konrad, *Model Generation for Natural Language Semantics*, Ph.D. dissertation, Universität des Saarlandes, 2000.

[13] Karsten Konrad and D. A. Wolfram, 'Kimba, a model generator for many-valued first-order logics', in *Proceedings of the 16th Conference on Automated Deduction*, ed., Harald Ganzinger, number 1632 in LNAI, pp. 282–286, Trento, Italy, (1999). Springer Verlag.

[14] Sven Lorenz, 'A tableau prover for domain minimization', *Journal of Automated Reasoning*, **13**, 375–390, (1994).

[15] Rainer Manthey and François Bry, 'SATCHMO: A theorem prover implemented in Prolog', in *Proceedings of the 9th Conference on Automated Deduction*, eds., Ewing L. Lusk and Ross A. Overbeek, number 310 in LNCS, pp. 415–434, Argonne, Illinois, USA, (1988).

[16] C. Monz and M. de Rijke, 'A tableaux calculus for ambiguous quantification', in *Automated Reasoning with Analytic Tableaux and Related Methods*, TABLEAUX'98, ed., H. de Swart, LNAI 1397, pp. 232–246. Springer, (1998).

[17] Uwe Reyle and Dov M. Gabbay, 'Direct deductive computation on discourse representation structures', *Linguistics & Philosophy*, **17**, 343–390, (1994).

[18] Werner Saurer, 'A natural deduction system for discourse representation theory', *Journal of Philosophical Logic*, **22**, (1993).

[19] Manfred Schmidt-Schauß, *Computational Aspects of an Order-Sorted Logic with Term Declarations*, volume 395 of *LNAI*, Springer Verlag, 1989.

[20] C. Sedogbo and M. Eytan, 'A tableau calculus for DRT', Technical report, Bull Report, France, (1987).

[21] M. Singer, 'Discourse inference processes', in *Handbook of Psycholinguistics*, ed., M. A. Gernsbacher, 479–515, Academic Press, San Diego, (1994).

[22] Rob A. Van der Sandt, 'Presupposition projection as anaphora resolution', *Journal of Semantics*, **9**(4), 333–377, (1992).

[23] Jan van Eijck, 'Axiomatising dynamic logics for anaphora', *Journal of Language and Computation*, **1**(1), 103–126, (1999).

[24] Jan Van Eijck and Fer-Jan De Vries, 'Dynamic interpretation and hoare deduction', *Journal of Logic, Language and Information*, **1**(1), 1–44, (1992).

[25] Henk Zeevat, 'A compositional approach to DRT', *Linguistics & Philosophy*, **12**, 95–131, (1989).

[26] R. A. Zwaan and G. A. Radvansky, 'Situation models in language comprehension and memory', *Psychological Bulletin*, **123**, 162–185, (1998).