Interpreting Negatives in Discourse

Michael Kohlhase & Mandy Simons

February 28, 2002

Technical Report No. CMU-PHIL-127

Philosophy

Methodology

Logic

Carnegie Mellon

Pittsburgh, Pennsylvania 15213

Interpreting Negatives in Discourse

Michael Kohlhase, Mandy Simons
Carnegie Mellon University
{kohlhase+@cs|simons@andrew}.cmu.edu

February 28, 2002

Abstract

In recent work, tableau-based model generation calculi have been used as computational models of the reasoning processes involved in utterance interpretation. In this linguistic application of an inference technique that was originally developed for automated theorem proving, natural language understanding is treated as a process of generating Herbrand models for the logical form of an utterance in a discourse. This approach captures ambiguity by generating multiple models for input logical forms.

In this paper we apply the model generation approach to a particular case of ambiguity: the interpretation of negated sentences. Using model generation, we will demonstrate how the various possible readings of simple negated sentences are generated, and by what criteria an interpreter chooses among these possibilities. Our investigation of negated sentences will lead us to propose constraints on the model generation system which, we will suggest, represent broadly applicable principles of interpretation.

1 Introduction

The interpretation of natural language utterances minimally involves two distinct, if interrelated, processes. One of these processes is semantic composition, which is driven by purely linguistic knowledge. The other is a process of reasoning from the output of the semantic composition to a conclusion as to the speaker/writer's communicative intent.[1] Perhaps the clearest cases in which there is a gap between compositional meaning and communicative intent are those involving indirection, as in the case of, say, relevance implicatures. However, even the most apparently straightforward uses of language involve utterances whose interpretation is underdetermined by semantic content, even when issues of context dependency, and lexical and structural ambiguity, have been resolved. Utterances containing nominal anaphors are a very familiar example

[1] There is a great deal of debate, particularly in the more recent literature, as to where semantic composition ends and inferential processes begin. For simplicity, we adopt here a very conservative position, assuming that the output of semantic composition is a complete proposition. However, other views are compatible with the framework which we develop.

of this underdetermination. To interpret such utterances, a hearer must reason (perhaps implicitly) about which possible antecedent is most likely to be the intended one.

In recent work, tableau-based model generation calculi have been used as computational models of the reasoning processes involved in utterance interpretation (see [BK00, GK00, Kon00, KK01]). Model Generation is an inference procedure developed in the field of automated reasoning to solve the problem of finding a Herbrand model satisfying a given set of formulas of predicate logic. In the linguistic application, natural language understanding is treated as a process of generating models for the logical forms of utterances in a discourse. The system is applied to particular advantage in those cases where multiple interpretations are compatible with the semantically determined logical form of an utterance, as in the case of utterances containing anaphors with multiple possible antecedents. But there are many other cases in which the semantic content (logical form) of an utterance is compatible with multiple possible interpretations, or models. The goal of this paper is to apply the model generation approach to a central case: the interpretation of negated sentences. Using model generation, we will demonstrate how the various possible readings of simple negated sentences are generated, and by what criteria an interpreter chooses among these possibilities.

The paper builds on [KK01], which refines the general model generation idea by introducing resource constraints and salience marking into the model generation process. These constraints improve the plausibility of the system as a cognitive model of sentence processing. Our investigation of negated sentences will lead us to propose further constraints which, we will suggest, represent broadly applicable principles of interpretation.

1.1 Semantic Underdetermination in Negated Sentences

Consider the various models which are compatible with the truth conditions of the simple negated sentence:

(1) *John didn't run.*

One compatible model is one in which John does nothing and nobody runs. A model in which running events take place may also be compatible with the truth conditions of the sentence, as long as John is not the agent of any of these events. Similarly, we can construct a compatible model in which John does something, as long as he does not run. Each of these models represents a possible interpretation of the sentence. Thus, the first task we face is to ensure that our model generation calculus has the potential to generate all of these models from the input logical form. The second task is to predict which interpretation the sentence will receive in particular situations of utterance.

There are two factors which clearly constrain the choice of interpretation: discourse context, and focal stress. Discourse context constrains the choice because information already given, or information given in following utterances, is often simply incompatible with one or more of the potential interpretations. For example, suppose an interpreter is faced with one of the following sequences:

(2) *John didn't run. Bill did.*

(3) *Bill ran. John didn't run.*

In both cases, the interpreter clearly cannot select as a model for the negated sentence one in which no running events take place, as this is inconsistent with the information that Bill ran. This type of constraint is dealt with rather straightforwardly in the model generation framework, as we will show below in Section 3.2.

The effects of focal stress provide an interesting test for our account. As is well known, negation and focus interact, with negation tending to "associate" with focused constituents. If our sentence (1) is uttered with heavy stress on the subject *John*, the preferred reading for the sentence is that John did not run, but someone else did. If stress falls on the verb, the preferred reading is that John did something other than run. Intonation and discourse context, of course, interact. In the sequence (2), our target sentence is most likely to be uttered with stress on *John*, while in (4), below, stress would probably fall on the verb.

(4) *John didn't run. He gambolled.*

While we do not attempt here any general account of focus, we will offer an analysis of the interaction of focus with sentential negation.

Discourse context and focal stress are two observable factors which constrain the choice of interpretation of a negated sentence. But there appear to be additional factors which constrain this choice. In certain cases, a negated sentence has a default, or preferred, reading, even when presented in the absence of context and with no focal stress – admittedly, a rather artifical case. One such set of cases are sentences containing an optional argument, such as:

(5) *John didn't vote for Nader.*

The default interpretation of this sentence is that John voted, but not for Nader, although the sentence also allows for an interpretation in which John did not vote at all. The question is why one of the available interpretations should be preferred. In the model generation framework, this question can be framed as a question about the quality of models: Given two models for a given input formula, what criteria determine whether one model is better than the other? Consideration of this question, in light of intuitions about default readings of negated sentences, will lead us to formulate some general principles of interpretation (Section 4).

1.2 Event semantics

In our treatment, we adopt an event-based semantics ([Dav67, Par90]). We conclude our introductory section with a brief review of the assumptions involved.

The fundamental idea of an event-based semantics is that verbs are predicates of events or states. A verb contributes to the logical form of a sentence containing it a predicate with a free variable over events as its argument. Verbal modifiers introduce further restrictions on the event variable, while NPs in argument position identify the participants of the event. In the absence of other sources of event quantification (such as adverbs of quantification or generic operators), the free event variable is existentially bound. Thus, the simple sentence *John ran* is assigned the logical form below. (Here and throughout, we ignore the contribution of tense.)

$$\exists e.\mathsf{r}(e) \wedge \mathsf{ag}(e, \mathsf{j})$$

It will be useful to have a term for the predicate-argument elements in these logical forms. Following [Par90], let us call these **subatomic formulae** or just **subatoms**.

The two-place predicate ag denotes the agent relation between an individual and an event. In the formula above, it identifies John as the agent of a running event. All thematic roles are represented by such two-place predicates. Additional roles include: theme; goal; benefactive; instrument; experiencer.[2]

Our event-semantic representations will deviate from standard ones in one way. For the sake of consistency, we prefer that all subatomic formulae contain 2-place predicates. Hence, instead of treating verbs as one-place predicates over events, we assume that verbs introduce a two-place event-type predicate, ty, which takes as arguments a variable over events and an event-type. We thus represent the logical form of *John ran* as follows.

$$\exists e.\mathsf{ty}(e, \mathsf{r}) \wedge \mathsf{ag}(e, \mathsf{j})$$

With these preliminaries in place, we turn to a fuller explanation of the model generation system.

2 Model Generation with First-Order Tableaux

In this section we will briefly review the state of the art in tableau-based model generation calculi and the calculus RAMG, which we will employ in our analysis.

First-order tableaux were originally used as data structures in refutation-based procedures for automated theorem proving. There a theorem is proved by decomposing its negation into a tree of possible instantiations (the tableau). Each branch in this tree corresponds to a possible model of the formula at the root, so if all branches are closed (inconsistent), the root formula is unsatisfiable, and the theorem is valid. Conversely, model generation procedures [MB88, LP93, FL96, Bau98, Pel99] build a tableau for the root formula and analyze the open branches as possible models, showing that it is indeed satisfiable.

[2]The question of whether it is either possible or useful to provide an exhaustive and definitive characterization of possible thematic roles is undecided. We adopt only the weak assumption that events have participants, that different participants have different types of roles, and that this can be represented as it is here. We do not place much theoretical weight on the particular thematic role assignment of NPs.

The two inference procedures mainly differ in strategy and share both the underlying data structure (the tableau) and the propositional rules in Figure 1. These inference rules act on tableaux have to be read read as follows: if the formulae over the line appear in a tableau branch, then the branch can be extended by the formulae or branches below the line. The rules in Figure 1 consist of two rules for each primary connective[3], and a branch closing rule that adds the special symbol $\perp$ (for unsatisfiability) to a branch. We will call a branch **closed**, iff it contains $\perp$, and **open** otherwise.

$$\frac{\mathbf{A}\land\mathbf{B}}{\begin{array}{c}\mathbf{A}\\ \mathbf{B}\end{array}}\mathcal{T}(\land) \qquad \frac{\neg(\mathbf{A}\land\mathbf{B})}{\neg\mathbf{A}\mid\neg\mathbf{B}}\mathcal{T}(\lor) \qquad \frac{\begin{array}{c}\mathbf{A}\\ \neg\mathbf{A}\end{array}}{\perp}\mathcal{T}(\perp) \qquad \frac{\begin{array}{c}\neg\mathbf{A}\\ \mathbf{A}\end{array}}{\perp}\mathcal{T}(\perp)$$

Figure 1: Propositional Tableau Rules

Model generation uses the quantifier rules in Figure 2 below, which are well-suited to the first-order logic without function symbols that we consider in this paper. The $\mathcal{RM}(\forall)$ rule tests the scope on all members of the Herbrand universe $\mathcal{H}$ of the current branch; it must be applied exhaustively to obtain a saturated branch. The $\mathcal{RM}(\exists)$ rule reuses constants that occur in the current branch and alternatively introduces a new constant c^{new}. Note that in contrast to other automated reasoning procedures, these rules do not use Skolemization and unification, so that no function symbols are re-introduced and models can remain finite. In addition, model generation normally assumes a kind of unique name assumption called Herbrand equality: two individual constants are equal iff they are identical. Consequently, a branch can be closed (extended by $\perp$) if it contains a literal $c = d$.

$$\frac{\forall X.\mathbf{A} \quad \mathcal{H}=\{a^1,\ldots,a^n\}}{\begin{array}{c}[a^1/X]\mathbf{A}\\ \vdots\\ [a^n/X]\mathbf{A}\end{array}}\mathcal{RM}(\forall)$$

$$\frac{\neg(\forall X.\mathbf{A}) \quad \mathcal{H}=\{a^1,\ldots,a^n\}}{\neg([a^1/X]\mathbf{A}) \mid \ldots \mid \neg([a^n/X]\mathbf{A}) \mid \neg([c^{new}/X]\mathbf{A})}\mathcal{RM}(\exists)$$

Figure 2: Quantifier rules for model generation ($\mathcal{RM}$)

[3]We have only shown those for (positive and negative) negations and conjunctions. The other connectives are defined in terms $\land$ and $\neg$ as usual, and the obvious inference rules can be obtained as derived rules using these definitions; we will use them in our examples without mention.

Definition 2.1 (Model Generation with $\mathcal{RM}$) The model generation calculus $\mathcal{RM}$ consists of the propositional tableau rules (see Figure 1) and the model generation rules for the quantifiers (see Figure 2).

We will call (a branch in) a tableau saturated, iff no rule application to one of its nodes yields new formulae.

Note that the tableau expansion rules in $\mathcal{RM}$ are essentially deterministic: It is a fair strategy to apply the propositional rules and the rule $\mathcal{T}(\exists)$ only to the leaves of a tableau, and the rule $\mathcal{T}(\perp)$, whenever possible. But when extending the Herbrand universe of a branch by c^{new}, all $\mathcal{RM}(\forall)$ must be re-instantiated with respect to c^{new}. It is a crucial observation that with this strategy, we only have to maintain one tableau, and do not need to consider multiple tableaux for completeness.

The $\mathcal{RM}$ model generation calculus was originally developed for a certain form of non-monotonic reasoning called minimal entailment [Lor94]. $\mathcal{RM}$ is a refutation complete first-order tableau calculus where each open saturated branch corresponds a so-called **Herbrand model**, i.e. a first-order model where the universe of discourse is the set of first-order terms without variables (in this case individual constants). In Herbrand models the valuation function can be represented by a set of **literals** (atomic propositions possibly with a negation). In this sense, open saturated tableau branches are Herbrand models of the input formula.

[Lor94] 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.

2.1 Resource-Adaptive Model Generation

[KK01] refines the general model-generation idea by introducing resource constraints and saliences into the model generation process improving both its computational behavior and its cognitive plausibility. Resource bounds on the inference process serve as an upper bound for the permitted computational complexity (introducing incompleteness but avoiding combinatorial explosion) and also alleviate the omniscience problem induced by unbounded computation. Associating saliences with objects allows ordering and restriction of the universe of individuals and eliminates another source of computational complexity.

We will use the RAMG calculus (Resource-Adaptive Model Generation) from [KK01] as a computational model in this paper. The details of salience assignment will not, however, play any central role in our discussion, so, for simplicity, we omit this from our review of the calculus. In a more detailed, quantitative version of the analyses we offer, salience considerations must be taken into account in the calculus, since they are relevant to the calculation of processing cost and model quality. We will allude to these interactions where they play a role as we go along.

Definition 2.2 (**RAMG: Resource-Adaptive Model Generation**) The calculus RAMG consists of the propositional rules in Figure 1 augmented with the rules in Figure 3.

The RAMG existential quantification rule introduces a new witness constant, instead of making a case distinction as does the $\mathcal{RM}(\exists)$ rule from $\mathcal{RM}$. This does not affect theoretical completeness, since $w^{new} \in \mathcal{W}$ and thus we can apply $\mathcal{R}(id)$ followed by $\mathcal{R}(=)$ (repeatedly in the right branch) to the result to obtain the branches of $\mathcal{RM}(\exists)$. The difference is that $\mathcal{R}(\exists)$ allows a finer-grained modeling of the reasoning processes.

$$\frac{\forall X.\mathbf{A} \quad a \in \mathcal{H}}{[a/X]\mathbf{A}}\,\mathcal{R}(\forall) \qquad \frac{\neg(\forall X.\mathbf{A}) \quad w^{new} \notin \mathcal{H}}{\neg([w^{new}/X]\mathbf{A})}\,\mathcal{R}(\exists)$$

$$\frac{\begin{array}{c} a = b \\ \mathbf{A} \end{array} \quad a \in \mathcal{W},\ b \in \mathcal{H}}{[b/a]\mathbf{A}}\,\mathcal{R}(=) \qquad \frac{a \in \mathcal{H}}{a = a}\,\mathcal{R}(\mathit{refl})$$

$$\frac{a, b \in \mathcal{H}}{a = b \mid a \neq b}\,\mathcal{R}(id) \qquad \frac{a = b \quad a, b \in \mathcal{U}}{\bot}\,\mathcal{R}(\mathrm{una})$$

Figure 3: RAMG rules

The RAMG calculus differs from $\mathcal{RM}$ in that it does not have a built-in unique name assumption: $\mathcal{RM}$ does not have to treat equality, since the underlying Herbrand semantics regards different constants as necessarily different. In RAMG, we partition the Herbrand base $\mathcal{H} = \mathcal{U} \cup \mathcal{W}$ into subsets $\mathcal{U}$ for constants with a unique name assumption, and $\mathcal{W}$ without (the new witness constants w^{new} from $\mathcal{R}(\exists)$ end up here). In addition, RAMG adds explicit rules for treatment of equality in a tableau. Note that the $\mathcal{R}(=)$ rule (Figure 3) is directional; it only allows substitution for a constant without the unique name assumption. Finally, $\mathcal{R}(\mathrm{una})$ mechanizes the unique name assumption by allowing a branch to close if two different constants with unique names are claimed to be equal. As a consequence of the introduction of equality into the calculus, the notion of an Herbrand model has to be generalized to that of a canonical model, i.e. a model where the universe is a set of equivalence classes of closed first-order terms. As these models can also be represented by the set of literals on a branch, this change of semantics does not affect our use of model generation as an inference procedure (See [Fit90] for details).

2.2 A Tableau Machine for Model Generation

Note that in contrast to the $\mathcal{RM}$ calculus, RAMG is no longer deterministic, i.e. tableau expansion can no longer be governed by a simple strategy that only employs one tableau. In particular the rule $\mathcal{R}(id)$ can be applied at any point and is indeterministic in the choice of constants to equate. In order to avoid losing completeness, we need to look at all the generated tableaux in parallel. To accommodate this we employ the notion of a **tableau machine**, that encapsulates a set of tableaux that represent the current state of the model generation, including relevant subgoals, and an account of the resources available for the computation. This encapsulation also makes it simpler to describe the non-local control mechanisms that we will introduce as analysis principles in the next sections.

We will use the tableau machine as a cognitive model for discourse understanding. We treat this as an online process that receives as input the logical forms of the sentences of the discourse one by one, and maintains a tableau that represents the current set of alternative models for the discourse. Since we are interested in the internal state of the machine (the current tableau), we do not specify the output of the tableau machine. We also assume that the tableau machine has a mechanism for choosing a preferred model from a set of open branches and that it maintains a set of deferred branches that can be re-visited, if extension of the the preferred model fails.

Upon input, the tableau machine will append the given logical form as a leaf to the preferred branch. (We will mark input logical forms in our tableaux by enclosing them in a box.) The machine then saturates the current tableau branch until some termination criterion is met – typically until the costs of all possible rule applications outweigh the expected gain in model quality – thereby partially exploring the set of possible models for the sequence of input sentences. If the subtableau generated by this saturation process contains open branches, then the machine chooses one of them as the preferred model, marks some of the other open branches as deferred, and waits for further input. If the saturation yields a closed sub-tableau, then the machine backtracks, i.e. selects a new preferred branch from the deferred ones, appends the input logical form to it, saturates, and tries to choose a preferred branch. Backtracking is repeated until successful, or until some termination criterion is met, in which case discourse processing fails altogether.

The tableau machine we use allows us to make use of derived rules of inference, i.e. rules that are not theoretically necessary in the calculus, since their effect can also be achieved by the other rules. In particular, we will use the derived rules shown in Figure 4 in our analysis. The first rule introduces a case distinction over an arbitrary formula by appealing to the law of the excluded middle ("tertium non datur"). This is called a "cut" rule in [Fit90]. Since this rule can be invoked anywhere in the computation, it has to be controlled explicitly by the tableau machine. The rule $\mathcal{R}(\Rightarrow)$ encapsulates a very common reasoning pattern often called *chaining*, which combines existential and universal reasoning. In the tableau machine, such rules can be given resource constraints different from the original rules they replace. Thus certain inferences can be penalized or rewarded, leading to a cognitively more adequate model of the derivation. In the particular case, an application of the rule $\mathcal{R}(\Rightarrow)$ consumes far fewer resources than the equivalent primitive applications of $\mathcal{R}(\exists)$, $\mathcal{R}(\forall)$, and $\mathcal{T}(\vee)$.

$$\frac{\begin{array}{c}\forall X^1.\mathbf{A}^1 \Rightarrow \ldots \forall X^n.\mathbf{A}^n \Rightarrow \exists Y.\mathbf{B} \\ \mathbf{A}^1(a^1) \\ \ldots \\ \mathbf{A}^n(a^1,\ldots,a^n) \qquad\qquad c \notin \mathcal{H}\end{array}}{[a^i/X^i, c/Y]\mathbf{B}}\,\mathcal{R}(\Rightarrow) \qquad \frac{}{\mathbf{A} \mid \neg\mathbf{A}}\,\mathcal{R}(tnd)$$

Figure 4: The chaining rule

3 Using Model Generation to Derive Possible Readings

We now have in place the systems that we will use for developing our account, and can begin the analysis itself. In this section, we develop the basic strategy for interpreting negated sentences in the model generation framework. In order avoid compounding issues, we will here show only how the RAMG calculus can be used to generate the range of possible readings for negated sentences. We will leave the discussion of how to control the search for models and how to determine preferred readings for section 4. We will begin by building a tableau for our original simple sentence "John didn't run" in isolation. Then we consider in turn the effects of context (in 3.2) and focus (3.3). We will then deal with some issues arising from more complex sentences in section 3.4

3.1 Model Generation in a Simple Case

We assign to our input sentence *John didn't run* (see 1) the logical form

(6) $$\neg\exists e.\mathsf{ty}(e,\mathsf{r}) \wedge \mathsf{ag}(e,\mathsf{j})$$

This logical form represents the simplest possible assumption about the semantics of negation: it is treated as a sentential operator with wide scope. Now, if we push the negation inside the quantifier, we see that what we are actually dealing with is a universal statement:

(7) $$\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \vee \neg\mathsf{ag}(e,\mathsf{j})$$

We assume that universals are unhelpful input for the model generation system. Since universally quantified variables can be instantiated with any salient object, applications of $\mathcal{T}(\forall)$ are costly. (For further discussion of this issue, see section 4.1). In order to proceed, we make use of the fact that this universal statement is propositionally equivalent to:

(8)
$$\forall e.\mathsf{ty}(e,\mathsf{r}) \Rightarrow \neg \mathsf{ag}(e,\mathsf{j})$$

Now, recall that in a first-order formula of the form $\forall x.p(x) \Rightarrow \mathbf{B}$ or $\exists x.p(x) \wedge \mathbf{B}$, we can consider the predicate p to be the sortal restriction on the bound variable x. In our example (8), the restriction p corresponds to *being a running event*, which we can represent by slightly extending the formalism into a Montagovian setting as $p = \lambda e.\mathsf{ty}(e,\mathsf{r})$. We now proceed by using the $\mathcal{R}(tnd)$ rule to make a case distinction on whether the restriction of the quantification is non-empty. This case distinction gives rise to the following RAMG tableau:

(9)
$$\begin{array}{c|c} \multicolumn{2}{c}{\boxed{\forall e.\neg \mathsf{ty}(e,\mathsf{r}) \vee \neg \mathsf{ag}(e,\mathsf{j})}} \\ \exists e.\mathsf{ty}(e,\mathsf{r}) & \neg\exists e.\mathsf{ty}(e,\mathsf{r}) \\ \mathsf{ty}(f,\mathsf{r}) & \forall e.\neg\mathsf{ty}(e,\mathsf{r}) \\ \neg\mathsf{ag}(f,\mathsf{j}) & \boxed{\mathrm{b}} \\ \boxed{\mathrm{a}} & \end{array}$$

On the left branch we have applied the rule $\mathcal{T}(\exists)$ introducing a new event f, which can then be used to chain with the original input formula (recall that it is equivalent to (8)) using the chaining rule in Figure 4. In the right branch we have only re-formulated the negative existential formula as a universal quantification.

Of course, our choice of the restrictor $p = \lambda e.\mathsf{ty}(e,\mathsf{r})$ was completely arbitrary in the somewhat marked situation assumed, where we have no context or focal stress. So we should also consider the other possible choice for a restrictor: $p = \lambda e.\mathsf{ag}(e,\mathsf{j})$. This leads to the RAMG tableau in (10):

(10)
$$\begin{array}{c|c} \multicolumn{2}{c}{\boxed{\forall e.\neg \mathsf{ty}(e,\mathsf{r}) \vee \neg \mathsf{ag}(e,\mathsf{j})}} \\ \exists e.\mathsf{ag}(e,\mathsf{j}) & \neg\exists e.\mathsf{ag}(e,\mathsf{j}) \\ \mathsf{ag}(f,\mathsf{j}) & \forall e.\neg\mathsf{ag}(e,\mathsf{j}) \\ \neg\mathsf{ty}(f,\mathsf{r}) & \boxed{\mathrm{d}} \\ \boxed{\mathrm{c}} & \end{array}$$

Thus we have a total of four open branches in the two tableaux (9) and (10), giving us four models, which represent the following states of affairs:

(11)
a. *There is a running event f, but John is not the agent. ("Someone else ran.")*
b. *There is no running event at all. ("Nobody ran.")*
c. *There is an event f of which John is the agent, but which is not a running event. ("John did something else.")*
d. *There is no event of which John is the agent. ("John did nothing.")*

The first two models come from tableau (9) and the second two from (10). Note that the inference that someone else ran in the glossing of (11a) is derived from the assumption that every running event has an agent, and that John did something else in (11c) from the fact that every event has a type. (We discuss this issue more thoroughly in section 4.3.)

By examining this sentence in isolation, we have demonstrated the basic problem for interpreting negated sentences, and the strategy which we will use to resolve it. The basic problem is that negations are only minimally informative. But the goal of the interpreter is to maximize informativity, to use the content of the utterance to extend her model. What an interpreter must do, then, is to decide which elements of the input can be treated as given, and thus (in effect) as falling outside the scope of the negation. This is what we represent by the technical move of invoking a case distinction.

3.2 The simple sentence in context

Note that when we instantiated the existential in the left branch of (9), we introduced a new event, thus implicitly assuming that there were no salient running events in the branch above. Suppose, however, that there were such a salient event, as there would be if the previous utterance were the sentence *Bob ran.* This would give us the following tableau.

(12)
$$\begin{array}{c}
\boxed{\exists e.\mathsf{ty}(e,\mathsf{r}) \wedge \mathsf{ag}(e,\mathsf{b})} \\
\mathsf{ty}(g,\mathsf{r}) \wedge \mathsf{ag}(g,\mathsf{b}) \\
\mathsf{ty}(g,\mathsf{r}) \\
\mathsf{ag}(g,\mathsf{b}) \\
\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \vee \neg\mathsf{ag}(e,\mathsf{j})} \\
\neg\mathsf{ty}(g,\mathsf{r}) \vee \neg\mathsf{ag}(g,\mathsf{j}) \\
\begin{array}{c|c}
\neg\mathsf{ty}(g,\mathsf{r}) & \neg\mathsf{ag}(g,\mathsf{j}) \\
\bot &
\end{array}
\end{array}$$

Note that here the presence of the salient running event g allows for the cheap application of $\mathcal{T}(\forall)$, and so eliminates the need for a case distinction. The result is a tableau with only one open branch, corresponding to a model in which there is a single running event of which Bob but not John is the agent. Alternatively, we could do the case distinction anyway, which would give us the following tableau:

(13)
$$\begin{array}{c}
\boxed{\exists e.\mathsf{ty}(e,\mathsf{r}) \wedge \mathsf{ag}(e,\mathsf{b})} \\
\mathsf{ty}(g,\mathsf{r}) \wedge \mathsf{ag}(g,\mathsf{b}) \\
\mathsf{ty}(g,\mathsf{r}) \\
\mathsf{ag}(g,\mathsf{b}) \\
\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \vee \neg\mathsf{ag}(e,\mathsf{j})} \\
\begin{array}{c|c}
\begin{array}{c}
\exists e.\mathsf{ty}(e,\mathsf{r}) \\
\begin{array}{c|c}
\mathsf{ty}(g,\mathsf{r}) & \mathsf{ty}(f_{new},\mathsf{r}) \\
\neg\mathsf{ag}(g,\mathsf{j}) & \neg\mathsf{ag}(f_{new},\mathsf{j}) \\
\boxed{\mathsf{a}} & \boxed{\mathsf{b}}
\end{array}
\end{array} &
\begin{array}{c}
\neg\exists e.\mathsf{ty}(e,\mathsf{r}) \\
\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \\
\neg\mathsf{ty}(g,\mathsf{r}) \\
\bot
\end{array}
\end{array}
\end{array}$$

This development makes a finer distinction on the running events than the tableau above. Going from left to right, the two open branches correspond to the following two models:

(14) a. *The running event which has Bob as agent does not also have John as agent. ("John did not run with Bob.")*[4]

b. *There were two running events (g and f_{new}), and John was not the agent of f_{new}. ("Somebody other than John ran, but not with Bob")*

We will return to the interpretation of this sentence, and the choice between these two strategies, in section 4.2, where we discuss mechanisms for controlling the model generation system.

3.3 The simple sentence with focal stress

As already noted, the sentence *John didn't run* would not only normally occur in some discourse context, as illustrated in the previous section, but would also usually bear some phonological marking of focus. The presence of focal stress has a quite straightforward effect: it constrains the possible interpretations of the negated sentence.[5] Thus, the sentence:

(15) *JOHN didn't run.*

is naturally interpreted as saying that someone ran, but not John. (Here and throughout, we use capitals to indicate focal stress.) In contrast, the sentence:

(16) *John didn't RUN.*

is naturally interpreted as saying that John did something other than run. Recall that in our initial discussion of this example, we generated models representing each of these readings, each model produced by a different case distinction. The fact that the presence of focal stress forces a choice between these two possible readings suggests a straightforward way of modeling its effect: focus controls the choice of material for the case distinction. Specifically, where a negated sentence contains focal stress, all nonfocused material is taken to be in the restriction of the quantifier. Whatever is focused may not be incorporated into the restriction. Thus, where focus falls on the subject NP, we can generate

[4] Note that this interpretation presupposes a framework where events can have multiple agents (or contain plural objects). Since we are concerned here with the computation of meaning and not with the structure of events, but on the computation of meaning, we will take the liberty of making such assumptions as needed.

[5] On some views, it is assumed that all sentences have semantic focus, even in the absence of focal stress. We do not wish to enter into this debate, as it is tangential to our purposes. Our proposal predicts that if in fact all sentences are marked for focus, then focal constraints are always in effect. In the absence of focal stress (and indeed, sometimes even when it is present, as the position of focal stress does not usually fix the extent of focus marking) we predict multiple ambiguity which, in our framework, will turn out to be equivalent to the absence of focus.

only tableau (9) above, where the case distinction is performed on the existence of a running event. This gives rise to a model in which someone other than John ran. Where focus falls on the verb, we can generate only tableau (10), where the case distinction is performed on the existence of an event with John as agent. This gives rise to a model in which John did something other than run. (Each of these case distinctions also gives rise to a second model, which represents a dispreferred reading of the sentence. We return in section 4 to the choice between these models.)

Finally, consider the interpretation of the same sentence with focal stress on negation:

(17) *John DIDN'T run.*

Following the strategy just outlined, the interpretation of this sentence produces the following tableau:

(18)

$$\begin{array}{c}
\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \vee \neg\mathsf{ag}(e,\mathsf{j})} \\
\begin{array}{c|l}
\begin{array}{c}
\exists e.\mathsf{ty}(e,\mathsf{r}) \wedge \mathsf{ag}(e,\mathsf{j}) \\
\mathsf{ty}(f,\mathsf{r}) \\
\mathsf{ag}(f,\mathsf{j}) \\
\begin{array}{c|c}
\neg\mathsf{ty}(f,\mathsf{r}) & \neg\mathsf{ag}(f,\mathsf{j}) \\
\bot & \bot
\end{array}
\end{array}
&
\begin{array}{l}
\neg\exists e.\mathsf{ty}(e,\mathsf{r}) \wedge \mathsf{ag}(e,\mathsf{j}) \\
\\
\\
\\
\\
\end{array}
\end{array}
\end{array}$$

In this case, focus forces the content of both subject and predicate into the case distinction, resulting in a computation which is vacuous, in that it gives back the very information with which we started. This, however, is consistent with the intuitive reading of the sentence: focus on negation results in a sentence which tells us that John was not the agent of any running event, but seems to commit the speaker neither to the existence of a running event with some other agent, nor to the existence of some other contextually relevant event with John as agent. The relatively uninformativity of such an utterance perhaps explains why this type of utterance normally occurs only as a denial of a previous assertion.

This proposed treatment of focus is inspired by Elena Herberger's proposal in [Her00], and may indeed be seen as a computational implementation of it. In that work, too, focus is treated within an event semantics. The central idea is that the interpretation of focused sentences can be derived by incorporating the non-focused material in a sentence into the restrictor of the event quantifier. In the case of negated sentences, this has the effect of allowing non-focused material to escape from the scope of negation, an effect which we achieve through the device of the case distinction.

Herburger argues further that her proposal illuminates the intuition that non-focused material in a sentence is "backgrounded" or "presupposed." For her, this fact about non-focused material reduces to the more general phenomenon of backgrounding of material in the restriction of a quantifier. In our view, the model generation framework provides an intuitive way of understanding this fact, at least as far as universal quantification is concerned. As we will discuss further below (section 4.1), universal quantifications are informationally weak. Faced with such a weak input, the interpreter looks for ways of strengthening the informational content. One way to do this is to treat some of the information represented in the quantificational input as given, as background information relative to which the quantificational statement is to be evaluated. This is what is represented by the (positive branch of) the case distinction. The model constructed in this branch represents the inferences which can be drawn on the basis of the quantificational input, given the material in the restrictor as an additional assumed premise. Thus, the initial statement in the affirmative branch of the case distinction—which corresponds to the material which Herburger would add to the restrictor of the implicit quantifier—intuitively represents material which the interpreter uses as background in interpreting the utterance.

As one final remark, note that Herburger actually requires a second mechanism to derive the various readings of negated sentences, namely, scopal ambiguity of negation. In her account, negation sometimes has wide scope, but sometimes attaches to one or another subatom. Our treatment appears to have the advantage of achieving similar results without positing such an ambiguity. However, Herburger covers a variety of cases with more complex quantificational structure than we discuss here, so we do not claim in the least to have presented a complete comparison of the two approaches. Indeed, we do not claim here to have developed a novel theory of focus, but rather to have shown that our model-generation approach is sufficiently flexible and general to incorporate existing ideas about focus in a very natural way. The approach has the further advantage of enabling us to compute the predictions of the analysis and to test them against linguistic data.

3.4 John didn't vote for Nader

In the above analysis for *John did not run*, the model generation process was driven by a case distinction on the emptiness of the restriction on the universal event quantifier. However, in the discussion we glossed over the question of what actually constitutes the restriction, simply assuming that it must be one or the other of the two subatoms of the input formula. However, in addition to the equivalences noted above, the formula in (7) is also equivalent to

$$\forall e.\top \Rightarrow (\neg \mathsf{ty}(e,\mathsf{r}) \vee \neg \mathsf{ag}(e,\mathsf{j})) \qquad \text{or} \qquad \forall e.(\mathsf{ty}(e,\mathsf{r}) \wedge \mathsf{ag}(e,\mathsf{j})) \Rightarrow \top$$

So we could also have taken the restriction p to be either $\lambda e.\top$ or $\lambda x.\mathsf{ty}(e,\mathsf{r}) \wedge \mathsf{ag}(e,\mathsf{j})$. Each of these choices would have given us different case distinctions. Generalizing this point, wherever we have a sentence whose logical form has the structure $\forall e.\neg p_1(e) \vee \ldots \vee \neg p_n(x)$, any subset of the disjuncts $p_i(e)$ can in principle be treated as the restriction.

We address this problem by assuming that the strongest restriction which produces reasonable models is always preferred. Hence, we work from the strongest restriction (incorporating all subatoms) to weaker ones. Note that a case distinction with the very strongest restriction does not make sense computationally, as we would end up with the following (redundant) tableau (in fact we have already seen an instance of this in tableau (18)):

$$
\begin{array}{c|c}
\multicolumn{2}{c}{\boxed{\neg\exists e.\mathbf{A}}} \\
\exists e.\mathbf{A} & \neg\exists e.\mathbf{A} \\
\bot &
\end{array}
$$

Consequently, we can always start with the strongest-but-one restriction (except, as noted above, where the strongest restriction is forced on us by focal stress on negation). This is exactly what we did in the last section, where we only had two sub-atoms. We did not need to look at the weakest restriction (corresponding to the empty subset of sub-atoms), since this would only have given rise to another model that did not have any events (which is unlikely in any situation).

To get a feeling for the procedure, let us now turn to our modification example, repeated below:

(19) *John didn't vote for Nader.*

The logical form of this sentence has three sub-atoms:

(20) $$\forall e.\neg\mathsf{ty}(e,\mathsf{v}) \vee \neg\mathsf{ag}(e,\mathsf{j}) \vee \neg\mathsf{ben}(e,\mathsf{n})$$

The principle of working from the strongest reasonable restriction still leaves us with three possibilities for the case distinction, each of which would give rise to a different set of possible readings. All of the readings which would arise are possible readings, and all are readings which could be enforced by introducing focal stress. In the absence of focal stress, other factors may conspire to fix a preferred case distinction. In our example, it is clear that the optional argument is the best candidate to be dropped. This is presumably a consequence of the syntactic structure itself. It is reasonable for an interpreter to infer that if the speaker has gone to the effort of including in her utterance a syntactically optional element, then the content of that element should not be backgrounded, as a first choice of interpretation.[6]

The tableau which results from this choice of restriction for sentence (19) is given below. Note that in constructing this tableau, we repeat the process of performing case distinctions in the process of expanding the right branch, because here too the initial formula is a universal quantification.

(21)
$$
\begin{array}{c|c|c}
\multicolumn{3}{c}{\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{v}) \vee \neg\mathsf{ag}(e,\mathsf{j}) \vee \neg\mathsf{ben}(e,\mathsf{n})}} \\
\exists e.\mathsf{ty}(e,\mathsf{v}) \wedge \mathsf{ag}(e,\mathsf{j}) & \multicolumn{2}{c}{\forall e.\neg\mathsf{ty}(e,\mathsf{v}) \vee \neg\mathsf{ag}(e,\mathsf{j})} \\
\mathsf{ty}(f,\mathsf{v}) & \exists e.\mathsf{ty}(e,\mathsf{v}) & \neg\exists e.\mathsf{ty}(e,\mathsf{v}) \\
\mathsf{ag}(f,\mathsf{j}) & \mathsf{ty}(g,\mathsf{v}) & \boxed{\mathrm{c}} \\
\neg\mathsf{ben}(f,\mathsf{n}) & \neg\mathsf{ag}(g,\mathsf{j}) & \\
\boxed{\mathrm{a}} & \boxed{\mathrm{b}} &
\end{array}
$$

[6] Ultimately, we would like to say more about the factors affecting the choice of the case distinction, including such syntactic considerations.

In branch a we have chained with the input formula to populate the branch after introducing the witness event f, in branch b with the case restriction formula. Thus we end up with three models:

(22) a. *"John voted but not for Nader."*

b. *"Someone voted, but not John."*

c. *"Nobody voted."*

It should be observed here that the middle model, b. does not, as it stands, represent an actual reading of the sentence. We return to this issue in 4.3.

So far, we have considered the interpretation of this sentence in the absence of any focal stress. Consider now the effect of focal stress on *John*. We assume, as before, that all nonfocused material is incorporated into the restriction, while what is focused is left out. This requires us to perform the case distinction on the formula:

(23) $$\exists e.\mathsf{ty}(e,\mathsf{v}) \wedge \mathsf{ben}(e,\mathsf{n})$$

And this gives us the following tableau:

(24)
$$\begin{array}{c}
\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{v}) \vee \neg\mathsf{ag}(e,\mathsf{j}) \vee \neg\mathsf{ben}(e,\mathsf{n})} \\
\begin{array}{c|c}
\begin{array}{c}
\exists e.\mathsf{ty}(e,\mathsf{v}) \wedge \mathsf{ben}(e,\mathsf{n}) \\
\mathsf{ty}(f,\mathsf{v}) \\
\mathsf{ben}(f,\mathsf{n}) \\
\neg\mathsf{ag}(f,\mathsf{j})
\end{array}
&
\begin{array}{c}
\forall e.\neg\mathsf{ty}(e,\mathsf{v}) \vee \neg\mathsf{ben}(e,\mathsf{n}) \\
\begin{array}{c|c}
\begin{array}{c}
\exists e.\mathsf{ty}(e,\mathsf{v}) \\
\mathsf{ty}(g,\mathsf{v}) \\
\neg\mathsf{ben}(g,\mathsf{n})
\end{array}
&
\begin{array}{c}
\neg\exists e.\mathsf{ty}(e,\mathsf{v}) \\
\\
\\
\end{array}
\end{array}
\end{array}
\end{array}
\end{array}$$

The leftmost branch gives us a model in which someone voted for Nader, but not John. This corresponds to the most salient reading of the sentence, given the assumed focal stress. What we must now do is provide an explanation in terms of the model generation system as to why this model should be preferred as an interpretation over the others in the tableau. The same question arises, *mutatis mutandis*, for all of the other tableaux which we have generated. It is to this question that we turn in the next section.

4 Controlling RAMG and Selecting Models

In the previous section we demonstrated the basic mechanism for generating the range of possible readings of negated sentences, and also showed two ways in which this range can be restricted: by contextual information, and by focus. In this section, we address further the issue of controlling the model generation process and of determining a preference ordering among the tableau branches. This preference ordering will give us predictions as to preferred readings. The work of controlling generation and of ordering models will be done by two general principles: the Full Representation Principle and the Safe Commitment Principle. The principles are, of course, formulated in model generation terms. However, we conceive of them as general principles of interpretation, whose implementation in this framework increases its viability as a cognitive model of discourse understanding.

4.1 Universals and the Full Representation Principle

A crucial assumption in our treatment of all of the examples is that their logical forms are those of universal quantifications over events, and that such quantifications induce case distinctions, that is, applications of $\mathcal{R}(tnd)$. We want here to justify that assumption.

Generally in RAMG, universal quantification poses a control problem. In a situation where there are no individuals that are much more salient than the rest, it is unclear whether it is better to continue tableau expansion or to stop with the current state, given that the computational resources are bounded. Instantiating universal quantifiers with non-salient individuals is expensive, and therefore dispreferred. Moreover, if there are many salient individuals, then the sum of the (lesser costs) of instantiation may be prohibitive.[7] This is especially problematic for quantification over events, as there are typically many of these around. Hence, it would be highly dispreferred to proceed in this situation by applying $\mathcal{T}(\forall)$.

But what prevents us from simply stopping the model generation with the universal statement itself? The intuitive idea is that there is something highly unsatisfactory about stopping the model generation without extracting some information from the content which is explicitly represented in the input sentence. In informal terms, we can put it like this: A hearer will recognize that the speaker has gone to the effort of saying something about John, about his possible role as agent in an event of a particular kind, and so on. Consequently, she expects to extract from what has been said some information about John, some information about the agent of this kind of event, and so on. Extracting information, in model generation terms, is a matter of extending the model. We express the idea in the following principle:

Principle 4.1 (Full Representation)
Interpretations which maximize the contribution of information explicitly realized in the input are preferred.

We intentionally formulate this principle in very broad terms, since we claim that its application goes beyond the use we make of it here. However, we can interpret it more precisely in the context of our model-generation approach.

Our initial assumption is that what counts as explicitly realized information in the input is at least that information represented by non-logical constants. At this point, we treat all such information as equivalent, allowing for the possibility that further investigation will show that Full Representation is more sensitive to the inclusion of some types of information than others. The more difficult question is what counts as a non-trivial contribution to the extension of the model. It is crucial to our understanding of this principle that the quality of model extensions is graded: some extensions are better than others. The best

[7] As noted in section 2, the model generation framework presented here can be enhanced with explicit marking of salience. To make these claims about processing costs more precise we would need to incorporate this marking, see [KK01] for details.

kind of contribution information can make is to result in extension of the model with a new literal. This is the most informative kind of extension possible[8]. Extension with a disjunction of literals is less good (because less informative), but this is still better than extension with a universally quantified formula.

Our Full Representation Principle is intentionally reminiscent of Chomsky's Principle of Full Interpretation [Cho86, Cho95], which disallows superfluous symbols in representations at any level, from PF (phonetic form) to LF (logical form). Applied to the syntactic level of LF, it requires that every LF element have a non-vacuous role in the final interpretation of the syntactic representation. Chomsky's discussion of the principle makes clear that one of its intended consequences is to rule out the possibility of interpretations which ignore syntactically represented elements – at least, those elements represented at LF. Thus, Full Interpretation is to rule out vacuous quantification, and also to ensure that in a sentence such as *John left town at noon*, *at noon* is predicated of some element [Cho86]. Our principle of Full Representation is a natural extension of Full Interpretation. Full Interpretation, in one sense, serves to guarantee that no information is lost in the move from (syntactic) LF to the semantic representation. Full Representation, in turn, guarantees that information cannot be lost in the move from the semantic representation to model construction.

Our principle of Full Representation is, however, different from Full Interpretation in that it is a soft constraint. We can generate models which violate the principle, but such models will always be dispreferred relative to models which satisfy it. The principle thus provides us with both a motivation for continuing model generation when faced with a universal quantification, and also with a criterion for choosing among the resulting models.

Returning now to the interpretation of the sentence *John didn't run*, it is clear why Full Representation disprefers stopping without expanding the (complex) input formula. If we did so, the tableau would not be extended with any new literals, and as a consequence, the current canonical model of the branch (which consists of the set of literals on this branch) would not be changed.

In this situation, the only rule available for continuing the tableau expansion is $\mathcal{R}(tnd)$; we have already argued against random application of the $\mathcal{T}(\forall)$ rule. The question which then arises is which formula should be used in applying this rule. In principle, any formula whatsoever could be used. However, in order to guarantee at least minimal satisfaction of Full Representation, we need to trigger inferences pertaining to running events, to the agent role, and to John. This is achieved by using information from the input sentence itself to construct the cut formula, as we have done in the examples so far. It now also becomes clear why the preferred cut formula is the strongest one which produces non-trivial results: our aim is to maximize the use we make of explicitly represented

[8]We have not explored whether the polarity of literals should play a role in determining adherence to the Full Representation principle. It seems plausible that only those literals that coincide in polarity with the subterm they derive from (i.e. the number of dominating negations, antecedents of implications, ... modulo 2) should count. But such reasoning would also apply to other modalities like belief, knowledge, etc. At the moment, it is not clear how to define a suitable notion of polarity here.

information.

We have thus provided a justification both for the assumption that the universal quantification induces a case distinction, and for our choice of formula for use in applying this rule. Let us now demonstrate the use of Full Representation as a criterion for ranking models. Here again is the tableau which we generate for the sentence *John didn't run.* For simplicity, we consider here only one of the possible case distinctions. (Recall that there are two possibilities which are compatible with our assumptions, a point to which we return.)

(25)
$$\begin{array}{c|c} \multicolumn{2}{c}{\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \vee \neg\mathsf{ag}(e,\mathsf{j})}} \\ \exists e.\mathsf{ag}(e,\mathsf{j}) & \neg\exists e.\mathsf{ag}(e,\mathsf{j}) \\ \mathsf{ag}(f,\mathsf{j}) & \forall e.\neg\mathsf{ag}(e,\mathsf{j}) \\ \neg\mathsf{ty}(f,\mathsf{r}) & \boxed{\mathrm{b}} \\ \boxed{\mathrm{a}} & \end{array}$$

Although we have chosen one particular case distinction, the tableau still gives us two open branches. But the left branch satisfies Full Representation more fully than does the right branch. We might attempt to improve the quality of the right hand model, but doing so would require continuing tableau expansion, a process that would quickly become more resource-expensive than the left branch. Consequently, we predict that the interpretation represented by the left branch – in which John did something other than run – is preferred over the interpretation represented by the right branch – in which John is the agent of no (relevant) events. But whether this would be the preferred reading for the sentence in a given situation would depend on additional factors, to which we turn in the next section.

First, however, let us consider briefly how Full Representation applies to the interpretation of our more complex sentence:

(26) *John didn't vote for Nader.*

the tableau for which is given in (21) on page 15. In this tableau, the leftmost model is the best satisfier of Full Representation, as each predicate constant in the input appears in a new literal (either positive or negative) in the model. Thus, we predict, correctly, that this model represents the default reading of the sentence.

The same prediction might also be made on the basis of standard Gricean assumptions.[9] A speaker who utters (26) intending to convey, say, that John did not vote, would be in violation of at least two maxims: Manner (for lack of brevity and for choosing an ambiguous formulation when an unambiguous one is available) and Quantity 1 (for making a weaker assertion than he is in a position to make). In the absence of special discourse considerations, a cooperative speaker should not use (26) to express a proposition which could be expressed less ambiguously and with a simpler form. This requirement, however,

[9]Thanks to Barbara Kaup (p.c.) for raising this point.

falls out as a speaker corollary of Full Representation.[10] For a speaker must be aware that an addressee will prefer whatever interpretation makes maximal use of the information explicitly represented in the input. Hence, a cooperative speaker should avoid explicitly representing information which in fact has no role to play in the intended interpretation. Our predictions thus converge with the Gricean view not only in selection of interpretation but also with respect to the underlying intuitions which drive the analysis.

4.2 The Safe Commitment Principle

Recall once more that in the absence of focal stress or contextual information, the interpretation of the sentence *John didn't run* can proceed in two different ways, depending on which subatom of its logical form is used in constructing the case distinction. One possibility considered above was that in interpreting this sentence, one should simply do both case distinctions in parallel. This would give rise to the following two tableaux, giving four models to be compared.

(27)
$$\begin{array}{c|c} \multicolumn{2}{c}{\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \vee \neg\mathsf{ag}(e,\mathsf{j})}} \\ \exists e.\mathsf{ag}(e,\mathsf{j}) & \neg\exists e.\mathsf{ag}(e,\mathsf{j}) \\ \mathsf{ag}(f,\mathsf{j}) & \forall e.\neg\mathsf{ag}(e,\mathsf{j}) \\ \neg\mathsf{ty}(f,\mathsf{r}) & \boxed{\mathrm{b}} \\ \boxed{\mathrm{a}} & \end{array} \qquad \begin{array}{c|c} \multicolumn{2}{c}{\boxed{\forall e.\neg\mathsf{ty}(e,\mathsf{r}) \vee \neg\mathsf{ag}(e,\mathsf{j})}} \\ \exists e.\mathsf{ty}(e,\mathsf{r}) & \neg\exists e.\mathsf{ty}(e,\mathsf{r}) \\ \mathsf{ty}(g,\mathsf{r}) & \forall e.\neg\mathsf{ty}(e,\mathsf{r}) \\ \neg\mathsf{ag}(g,\mathsf{j}) & \boxed{\mathrm{d}} \\ \boxed{\mathrm{c}} & \end{array}$$

In this tableau, branches a and c win out over b and d as far as satisfaction of Full Representation is concerned: the former two branches have been extended by the addition of new literals, while the latter two have not. However, there seems no way to choose between a and c. The two models differ, but they satisfy Full Representation to exactly the same degree. What happens in this situation?

Intuitively, it seems to us that if this sentence really is taken completely "out of the blue," and is uttered with no focal stress, then it does not allow for expansion of the model beyond the quantificational input. We understand the utterer of this sentence to have committed to the absence of any running events with John as agent, and no more; at least, not until some further information is provided in the continuation of the discourse. What seems to be at work here is a principle which works in opposition to Full Representation, a principle which prevents the interpreter from making random selections among equally satisfactory interpretations. We call this principle Safe Commitment, and define it as follows:

Principle 4.2 (Safe Commitment Principle)
The processing of an input sentence should eventually commit to a safe interpretation.

[10] Compare with [Lev00], who formulates a speaker corollary for each of his interpretative heuristics.

This principle has two components. It requires commitment to some particular interpretation; but it further requires that commitment to be safe. In our system, to commit to an interpretation is to choose one model as the preferred interpretation. Committing to an interpretation always involves the pruning – or at least postponing the exploration of – dispreferred branches. A commitment is safe if there are clear reasons (of model quality) to prefer the chosen model over other options. This rules out arbitrary selection of a candidate interpretation. There is still no guarantee, of course, that the chosen interpretation will be that intended by the speaker; the speaker might, for example, have expected the addressee to make use of some additional background information in the interpretation. So even a safe choice may have to be rejected in terms of later evidence, requiring the interpreter to backtrack. What the principle guarantees is that the interpreter will commit to an interpretation, but will only commit to a highly specified interpretation if this is warranted.

The effect of this principle for our approach will be this: If after a certain amount of processing no clear preference among models emerges, we discard the results of the computation on the input sentence and return to a state of the tableau machine where there is a clear choice between branches. In (27), the only way to achieve this is to return to the initial tableau that only contains the input representation.

$$\forall e.\neg \mathsf{ty}(e,\mathsf{r}) \vee \neg \mathsf{ag}(e,\mathsf{j})$$

Clearly, this is a state which has one tableau branch, and avoids unnecessary model specialization and is therefore "safe". Note that by discarding the results of the computation we have not precluded any more specific interpretations, since the input sentence is in effect an underspecified representation of the four possible models. If in the ensuing discourse, information becomes available that favors one of the models, the same computation can be carried out again to re-specialize the models (discarding the other branches in the process).

Another example where the safe commitment principle helps avoid unmotivated over-specialization is in tableau (13) on page 11 above. Models a and b, although distinct, are not distinguished with respect to Full Representation. Hence, an interpreter cannot be justified in choosing one over the other; and so the case distinction is a waste of processing effort. Indeed, the sentence does not seem to support a distinction between these two readings in the absence of further discourse information.

Clearly, the Safe Commitment Principle works in contrast to most machine-oriented strategies and implementations of model generation calculi, as it it prevents the method from being complete in a model-theoretic sense. As we want to use model generation as cognitive model for discourse understanding, this is an advantage of the analysis rather than a problem.[11]

[11] One issue which should certainly be resolved empirically is what units of discourse Full Representation applies to. Here, we have assumed that Safe Commitment must be satisfied for each input sentence independently. But it is quite plausible that the principle need only be satisfied for larger (or at least different) discourse segments.

4.3 Full Representation and Model Population

In the last section we saw how the two control principles work together to drive the exploration of a tableau that represents the possible interpretations of the input. In the examples so far, we have not made use of our knowledge about the world, and so have not yet utilized one of the strongest features of the model generation approach. In this section, we will look at how a particular type of world knowledge is used in interpretation: knowledge about the structure of events, which lets us populate models to partially fulfill the Full Representation Principle.

To keep the discussion simple, we will re-use our simple example *John didn't run* (1). We saw in section 3.1 that this input sentence gives us the tableau 9, which we repeat here:

$$
\begin{array}{c|c}
\multicolumn{2}{c}{\boxed{\forall e.\neg \mathsf{ty}(e,\mathsf{r}) \vee \neg \mathsf{ag}(e,\mathsf{j})}} \\
\exists e.\mathsf{ty}(e,\mathsf{r}) & \neg\exists e.\mathsf{ty}(e,\mathsf{r}) \\
\mathsf{ty}(f,\mathsf{r}) & \forall e.\neg\mathsf{ty}(e,\mathsf{r}) \\
\neg\mathsf{ag}(f,\mathsf{j}) & \boxed{\mathrm{b}} \\
\boxed{\mathrm{a}} &
\end{array}
$$

In the earlier discussion, we glossed the left branch as "Someone else ran," even though strictly speaking no agent of the event f is present in the branch. In effect, the inference that there is an agent of f was not made by the model generation procedure, and so the gloss was in fact fallacious. What we need to do is to provide the model generation procedure with the same information which we, as interpreters, use to draw the inference in question. We can do this by invoking a meaning postulate concerning running events. The postulate we need is this: $\forall e.\mathsf{ty}(e,\mathsf{r}) \Rightarrow \exists x.\mathsf{ag}(e,x)$, i.e. every running event has an agent (required role). Given this postulate, the chaining rule in Figure 4 allows us to infer the existence of an (unspecified) agent from the existence of the running event. The complete tableau has the following form (which supports the gloss "someone else (c) ran").[12]

$$
\begin{array}{c|c}
\multicolumn{2}{c}{\forall e.\mathsf{ty}(e,\mathsf{r}) \Rightarrow \exists x.\mathsf{ag}(e,x)} \\
\multicolumn{2}{c}{\boxed{\forall e.\neg \mathsf{ty}(e,\mathsf{r}) \vee \neg \mathsf{ag}(e,\mathsf{j})}} \\
\exists e.\mathsf{ty}(e,\mathsf{r}) & \neg\exists e.\mathsf{ty}(e,\mathsf{r}) \\
\mathsf{ty}(f,\mathsf{r}) & \forall e.\neg\mathsf{ty}(e,\mathsf{r}) \\
\neg\mathsf{ag}(f,\mathsf{j}) & \boxed{\mathrm{b}} \\
\exists x.\mathsf{ag}(f,x) & \\
\mathsf{ag}(f,c) &
\end{array}
$$

The meaning postulates we invoke here are, in essence, the thematic or theta grids postulated in Government Binding theory as part of the lexical representation of verbs and other predicates which have an argument structure (see [Hae94]). Theta grids contain, at least, the information as to which thematic roles are assigned by the predicate: the very information contained in our

[12] We follow established practice and add world knowledge at the root of the tableau.

meaning postulates. Thus, the meaning postulates that we require are independently motivated.

We can now make good on a promise made with respect to tableau (21) on page 15. We saw in section 4.1 above that, given Full Representation, this tableau makes correct predictions about the preferred reading of the sentence *John didn't vote for Nader* in the absence of focal stress or prior discourse. But we noted in our original discussion of this tableau that one of the models generated did not correspond to any possible interpretation of the sentence. This is undesirable. We would like to know that if the preferred reading is overridden by later information, the system will provide an intuitively correct alternative.

But the problem is resolved by the possibility of model population. Let us assume a postulate according to which every voting event has an agent and a beneficiary.[13] Below we give an expanded version of the original tableau, incorporating this postulate. The resulting open models all correspond to possible interpretations. (We have closed the central branch under the assumption, not represented, that John is not a candidate for beneficiary.) However, the leftmost model continues to best satisfy Full Representation with the least use of resources.

(28)

$$
\begin{array}{c}
\boxed{\forall e.\mathsf{ty}(e,\mathsf{v}) \Rightarrow \exists x.\mathsf{ag}(e,x) \wedge \exists(.y)\mathsf{ben}(e,y)} \\
\forall e.\neg\mathsf{ty}(e,\mathsf{v}) \vee \neg\mathsf{ag}(e,\mathsf{j}) \vee \neg\mathsf{ben}(e,\mathsf{n}) \\
\begin{array}{c|c}
\begin{array}{c}
\exists e.\mathsf{ty}(e,\mathsf{v}) \wedge \mathsf{ag}(e,\mathsf{j}) \\
\mathsf{ty}(f,\mathsf{v}) \\
\mathsf{ag}(f,\mathsf{j}) \\
\neg\mathsf{ben}(f,\mathsf{n})
\end{array}
&
\begin{array}{c}
\forall e.\neg\mathsf{ty}(e,\mathsf{v}) \vee \neg\mathsf{ag}(e,\mathsf{j}) \\
\begin{array}{c|c}
\begin{array}{c}
\exists e.\mathsf{ty}(e,\mathsf{v}) \\
\mathsf{ty}(g,\mathsf{v}) \\
\neg\mathsf{ag}(g,\mathsf{j}) \\
\exists x.\mathsf{ben}(g,x) \\
\mathsf{ben}(g,c) \\
\begin{array}{c|c}
\begin{array}{c}
c = n \\
\mathsf{ben}(g,n)
\end{array}
&
\begin{array}{c}
c \neq n \\
\begin{array}{c|c}
\begin{array}{c}
c = j \\
\mathsf{ben}(g,j) \\
*
\end{array}
&
c \neq j
\end{array}
\end{array}
\end{array}
\end{array}
&
\neg\exists e.\mathsf{ty}(e,\mathsf{v})
\end{array}
\end{array}
\end{array}
\end{array}
$$

5 Extensions

Before concluding, we would like to speculate briefly about some possible extensions of this proposal. These are the treatment of focus, and the treatment of certain cases of presupposition projection.

As noted at the outset, our treatment of focus here is very limited, and we would not suggest that there is any obvious and simple extension of it which would cover all focus phenomena. Nonetheless, it is tempting to think that we

[13] The presence of an NP filling the beneficiary role is syntactically optional; but nonetheless one simply cannot have a (semantic) voting event which does not involve a vote for *someone*. Casting an empty ballot doesn't count as voting.

have here the barest bones of a more extensive treatment of focus within the model generation framework. Focus is, fundamentally, a mechanism provided by natural language for distinguishing foreground and background material, for marking which information conveyed by an utterance is to be treated as new and newsworthy and which is to be treated as given (in some sense of this very difficult term). In our treatment of negated sentences, the case distinction mechanism provides us with a means of distinguishing between foreground and background. In essence, material which enters into a case distinction is treated as background. We are thus able to model the effects of focus in these cases straightforwardly, as enforcing a particular choice of material for the case distinction. This suggests that it may be possible to model the effects of focus more generally using the case distinction mechanism.

The proposal offered here also has interesting implications for the treatment of certain cases of projection of presuppositions over entailment-canceling operators. In the case of presupposition projection, one entailment (or more) of an affirmative declarative sentence "survives" embedding under an entailment-canceling operator, such as negation. This surviving entailment – the presupposition – is generally felt to be backgrounded, or treated as taken for granted. Two standard examples are given below:

(29) *Jane didn't stop laughing.*
Presupposes: *Jane was laughing immediately prior to the reference time.*

(30) *Jane wasn't late.*
Presupposes: *Jane had an appointment*

As readers may have noted, the interpretation of the simple sentences considered here involves something analogous: the preferred reading of *John didn't vote for Nader* is one in which several entailments of the embedded clause *John voted for Nader* "survive" as background, "projecting" over the negation. Moreover, just as the Nader sentence has other possible interpretations, so too do presuppositional sentences like (29) and (30) above. Both allow, in addition to the default presuppositional reading, a presupposition-canceling reading, often induced by a continuation which denies the presupposition, as in:

(31) *Jane didn't stop laughing. She wasn't laughing to begin with.*

(32) *Jane wasn't late for that meeting. She wasn't supposed to be there at all.*

Standardly, presuppositions are treated by identifying a specific trigger (either a lexical item or, in some cases, a construction), and associating with that trigger a presuppositional content distinct from its ordinary truth conditional content. The survival of the presupposition is then attributed to the special properties of this kind of content. The similarities between the cases considered here and the more familiar presupposition projection cases suggest, however, that the mechanism proposed here could be used to treat certain cases of supposed presupposition projection without postulating any special presuppositional content.

On this account, information has the status of a presupposition when it is used in constructing the proposition on which a case distinction is performed. Thus, to extend the account to examples like (31) and (32), we would have to make a plausible case that the presuppositional content is available to the model generation system for this purpose. There are two ways in which this could be done. One way would be to assume that the logical structure of the sentences in question is more complex than revealed by surface form, and that the presuppositional content is actually represented in the logic form of the sentence. For example, one might posit something like the following as the logical form of *Jane stopped laughing.*

$$\exists d.\mathsf{ty}(d,\mathsf{l}) \wedge \mathsf{ag}(d,\mathsf{j}) \wedge \exists e.\mathsf{ty}(e,\mathsf{end}(d))$$

This representation, based very loosely on a proposal in [Piñ97], decomposes the predicate *stop laughing* into two separately specified events: a laughing event, and an (instantaneous) event of ending. Given such a representation, the interpretation strategy we have argued for above allows for the generation of the following tableau:

(33)
$$\begin{array}{c|c}
\multicolumn{2}{c}{\boxed{\forall d.\neg\mathsf{ty}(d,\mathsf{l}) \vee \neg\mathsf{ag}(d,\mathsf{j}) \vee \neg\exists e.\mathsf{ty}(e,\mathsf{end}(d))}} \\
\exists d.\mathsf{ty}(d,\mathsf{l}) \wedge \mathsf{ag}(d,\mathsf{j}) & \forall d.\neg\mathsf{ty}(d,\mathsf{l}) \vee \neg\mathsf{ag}(d,\mathsf{j}) \\
\mathsf{ty}(f,\mathsf{l}) & \begin{array}{c|c} \exists d.\mathsf{ty}(d,\mathsf{l}) & \neg\exists d.\mathsf{ty}(d,\mathsf{l}) \\ \mathsf{ty}(g,\mathsf{l}) & \\ \neg(\mathsf{ag}(g,\mathsf{j})) & \end{array} \\
\mathsf{ag}(f,\mathsf{j}) & \\
\neg\exists e.\mathsf{ty}(e,\mathsf{end}(f)) &
\end{array}$$

The leftmost model in this tableau, which satisfies the Full Representation principle best given the assumed input, represents the expected presuppositional reading of the sentence.

The representation suggested here for sentence (29) is offered only for illustration, and may turn out not to be an adequate analysis of this type of sentence. However, it is quite plausible that change of state predicates such as *stop V-ing* do have complex internal structure along these lines, as [Piñ97] argues.[14] To avoid committing to a complex logical form for the input sentence, one could achieve precisely the same effects by introducing a meaning postulate that introduces an end event or subatoms that constrain the timing of events. Whichever strategy one prefers, a treatment along these lines seems promising for the very many cases where a lexical entailment (i.e. an entailment due to lexical content) shows projection behavior.

We would also like to look further at the projection behavior of entailments induced by various syntactic structures. In this paper, we have looked in detail at the interpretation of a sentence containing a syntactically optional modifier: the PP *for Nader* in the sentence *John didn't vote for Nader.* The default reading for the sentence can be characterized in presupposition projection terms: on

[14]We should emphasize that although we borrow here from Piñón, the representation given both simplifies and modifies his proposal.

this reading the proposition that John voted (for someone), which is of course entailed by the affirmative version of the sentence, "projects" over negation, just like a presupposition. Similar projection behavior occurs with negations of sentences with secondary predicates, like (34), and even, in some cases, with adjectival modification, as in (35):

(34) *Jane didn't come home drunk.*
Default interpretation: *Jane came home, but not drunk.*

(35) *I didn't read a boring book all summer.*
Default interpretation: *I read books, but no boring ones.*

However, the default readings for negated sentences with no optional constituents are not "presuppositional." We saw this with the simple sentence *John ran*, which shows projection behavior only given focal stress on one of its elements. The same holds with sentences with transitive verbs. The default interpretation of:

(36) *Jane didn't wash the windows.*

is simply that there is no window-washing event with Jane as agent.

These observations reinforce the suggestion made above that syntactic information plays a role in determining the proposition to be used in the case distinction, and perhaps even in determining whether the case distinction should be performed. Full Representation, too, may be sensitive to syntactic structure: information introduced by a syntactically optional constituent may be weighted more heavily with respect to satisfaction of the principle than information whose inclusion is syntactically required. We intend in future work to extend our treatment of negation to additional structures. Such an extension will, we hope, provide us with a better understanding of the workings of the case distinction mechanism and also of the two pragmatic principles which constrain the model generation system: Full Representation, and Safe Commitment.

6 Conclusion

In this paper we have applied the recently developed technique of resource-adaptive model generation to the interpretation of monoclausal negated sentences. We have adopted the position that these sentences are semantically unambiguous, but that their actual interpretation in a given instance is underdetermined by their semantics. The project undertaken here was thus to model the pragmatic inferencing which guides the selection of a specific interpretation.

The fundamental idea is that, through the mechanism of the case distinction (the $\mathcal{R}(tnd)$ rule), the system generates multiple models, each of which is compatible with the content of the negated sentence. The preferred interpretation is then predicted to be given by that model which best satisfies certain quality measures. Here, we have argued that a crucial criterion for model selection is degree of satisfaction of the Full Representation principle (4.1), which leads to a preference for models that maximize the effects of information explicitly represented in the input sentence. But the application of Full Representation is moderated by the workings of the Safe Commitment principle (4.2), which precludes unmotivated overspecification, and may result in adoption of a less specified, but safe, interpretation.

It seems to us that the application of these principles should go well beyond their use here. We propose these as fundamental principles of interpretation, which are applied systematically in cases where an interpreter must choose among competing possible interpretations of a given input. But of course the principles as formulated here leave open a number of questions, which we hope to address in future work. Some of these open issues were raised in the previous section. In summary, we need to explore further at least the following questions:

- What counts as explicitly represented information? Do some types of information "count" more for satisfaction of Full Representation than others?
- What is the contribution of syntactic information to the workings of Full Representation?
- How can degree of satisfaction of Full Representation be more precisely quantified?
- What are the units of discourse to which Safe Commitment applies?

Future work should also try to uncover additional principles guiding interpretation. And finally, we would hope that the claims made here can be tested experimentally, to evaluate the degree to which the model generation system, constrained by the proposed principles, provides a working model of the actual process of language interpretation by human speakers.

References

[Bau98] Peter Baumgartner. Hyper Tableaux — The Next Generation. In Harry de Swaart, editor, *Automated easoning with Analytic Tableaux and Related Methods*, Lecture Notes in Aritificial Intelligence. Springer, 1998.

[BK00] Peter Baumgartner and Michael Kühn. Abducing Coreference by Model Construction. *Journal of Language and Computation*, 1(2):175–190, 2000.

[Cho86] Noam Chomsky. *Knowledge of Language.* Praeger, New York, 1986.

[Cho95] Noam Chomsky. *The Minimalist Program.* MIT Press, Cambridge, MA, 1995.

[Dav67] Donald Davidson. The logical form of action sentences. In N. Rescher, editor, *The logic of decision and action.* University of Pittsburgh Press, Pittsburgh, 1967.

[Fit90] Melvin Fitting. *First-Order Logic and Automated Threorem Proving.* Springer Verlag, 1990.

[FL96] Christian Fermüller and Alexander Leitsch. Hyperresolution and automated model building. *Journal of Logic and Computation*, 6(2):173–230, 1996.

[GK00] Claire Gardent and Karsten Konrad. Interpreting definites using model generation. *Journal of Language and Computation*, 1(2):193–209, 2000.

[Hae94] L. Haegeman. *Introduction to Government and Binding Theory.* Blackwell, Cambridge, USA, 2 edition, 1994.

[Her00] Elena Herburger. *What Counts: Focus and Quantification.* Number 36 in Linguistic Inquiry Monograph. MIT Press, Cambridge, Mass., 2000.

[KK01] Michael Kohlhase and Alexander Koller. Resource-adaptive model generation as a performance model. submitted to Journal of Language and Computation, 2001.

[Kon00] Karsten Konrad. *Model Generation for Natural Language Semantics.* PhD thesis, Universität des Saarlandes, 2000.

[Lev00] Stephen C. Levinson. *Presumptive Meanings.* MIT Press, Cambridge, MA, 2000.

[Lor94] Sven Lorenz. A tableau prover for domain minimization. *Journal of Automated Reasoning*, 13:375–390, 1994.

[LP93] Shie-Jue Lee and David A. Plaisted. Problem solving by searching for models with a theorem prover. *Artificial Intelligence*, 69(1-2):205–233, 1993.

[MB88] Rainer Manthey and François Bry. SATCHMO: A theorem prover implemented in Prolog. In Ewing L. Lusk and Ross A. Overbeek, editors, *Proc. of the 9th Conference on Automated Deduction*, number 310 in LNCS, pages 415–434, Argonne, Illinois, USA, 1988.

[Par90] Terence Parsons. *Events in the semantics of English.* MIT Press, 1990.

[Pel99] N. Peltier. Pruning the search space and extracting more models in tableaux. *Logic Journal of the IGPL*, 7(2):217–251, 1999.

[Piñ97] Christopher Piñón. Achievements in event semantics. *Semantics and Linguistic Theory*, 7:276–293, 1997.