Understanding "Each Other"

Claire Gardent
Computational Linguistics
University of the Saarland
Saarbrücken, Germany
claire@coli.uni-sb.de

Karsten Konrad
Computer Science
University of the Saarland
Saarbrücken, Germany
konrad@ags.uni-sb.de

Abstract

Although natural language is ambiguous, various linguistic and extra-linguistic factors often help determine a preferred reading. In this paper, we show that model generation can be used to model this process in the case of reciprocal statements. The proposed analysis builds on insights from Dalrymple et al. 98 and is shown to provide an integrated, computational account of the interplay between model theoretic interpretation, knowledge-based reasoning and preferences that characterises the interpretation of reciprocals.

1 Introduction

Although there is widespread agreement that inference is an essential component of natural language processing, little work has been done so far on whether existing automated reasoning systems such as theorem provers and model builders could be fruitfully put to work in the area of natural language interpretation.

In this paper, we focus on the inference problems raised by the reciprocal expression *each other* and show that model generation provides an adequate tool for modeling them.

The paper is structured as follows. Section 3 discusses the meaning of reciprocal statements and proposes a formal semantics for *each other*. Section 2 shows how model generation can be used to provide this semantics with a computational interpretation. Section 4 compares our approach with the account of reciprocals which inspired it in the first place namely, (Dalrymple et al., 1998). Section 5 concludes with pointers for further research.

2 The meaning of reciprocal statements

In the linguistic literature, the reciprocal expression *each other* is often taken to denote a dyadic quantifier over a first-order set, which we will call the **antecedent set**, and a binary first-order relation, which we will call the **scope relation**. In what follows, we assume an approach of this type and will use the symbol RCP for such reciprocal quantifiers so that the semantic representation of e.g. *Jon and Bill saw each other* will be:

(1) $\text{RCP}(\{jon, bill\})(\lambda y \lambda x\ saw(x, y))$

When antecedent sets of just two members are considered, each set member is required to stand in the scope relation to each other member. For larger sets however, research on reciprocal statements has uncovered a variety of logical contributions that the reciprocal can provide. Here are some examples.

(2) The students like each other.
$\forall x\ (std(x) \rightarrow \forall y\ (x \neq y \wedge std(y) \rightarrow like(x, y))$

(3) The students stare at each other in surprise.
$\forall x\ (std(x) \rightarrow \exists y\ (x \neq y \wedge std(y) \wedge stare\_at(x, y))$

(4) The students gave each other measles.
$\forall x\ (std(x) \rightarrow \exists y\ (x \neq y \wedge std(y) \wedge (gave\_measles(x, y) \vee gave\_measle(y, x))))$

We can accept (2) to be true only if for each pair x and y of two students it holds that x likes y. But an analogous interpretation would be invalid in the case of (3) and (4) where not all pairs in the antecedent set *the students* can consistently stand in the scope relation (one can only stare at most at one person at a time, and one can only get measles from at most one person). More generally, (Langendoen, 1978; Dalrymple et al., 1998) convincingly argues that different reciprocal statements can have very different truth conditions. The challenge to be addressed is thus the following: How can we determine a (computational) semantics for the reciprocal expressions *each other* that accounts for these multiplicity of meanings while predicting the specific meaning of a particular reciprocal statement?

Clearly knowledge based reasoning plays an important role: only those readings are possible that are consistent with our knowledge about the situation and the world. Specifically, knowledge based reasoning constrains the strength of the truth conditions of a reciprocal statement. Thus if we abstract away from the specific scope relations, the truth conditions of examples such as (2),(3) and (4) are ordered through entailment as follows (with A the antecedent set and R the scope relation):

$$\forall x\ (A(x)\ \rightarrow \forall y\ (A(y) \rightarrow R(xy))$$
$$\models\ \forall x\ (A(x)\ \rightarrow \exists y\ (A(y) \wedge R(xy))$$
$$\models\ \forall x\ (A(x)\ \rightarrow \exists y\ (A(y) \wedge (R(xy) \vee (R(yx)))$$

Specifically, example (2), which does not involve any strong knowledge based constraint, has the strongest truth-conditions of the three examples. By contrast in (3), the knowledge that one can stare at most at one person, forces a $\forall\exists$ reading while in (4), a weaker meaning still is imposed by knowledge based constraints: the *x gave y measles* relation is asymmetric hence the $\forall\forall$ reading is ruled out; moreover, since one cannot be infected twice, some students in the group will be infected but not pass on the disease to anyone. Hence the strongest truth conditions that can be assigned the sentence are the $\forall\exists$ disjunctive reading indicated in (4).

But are there any other constraints on the interpretation process than these knowledge based constraints? And which meaning shall we assign a reciprocal expression? The computational semantics we will propose is inspired from (Dalrymple et al., 1998) and relies on the following observations.

First, we note that (Dalrymple et al., 1998) identifies a lower bound for the truth conditions of reciprocal sentences which they dub Inclusive Alternative Ordering (IAO). It is exemplified by sentence (4) above and corresponds to the following definition of RCP.

(5) $\text{RCP}_{IAO} \equiv \lambda P \lambda R\ (|P| \geq 2 \wedge \forall x\ (P(x) \Rightarrow \exists y\ P(y) \wedge x \neq y \wedge (R(x,y) \vee R(y,x))))$

This definition only adequately characterises examples such as (4). It does not cover the stronger meanings of the reciprocal in sentences such as (2) and (3). However, each known form of reciprocity entails RCP_{IAO}'s truth conditions, and RCP_{IAO} therefore provides us with a minimal semantics for reciprocals.

Further, we observe that given a particular reciprocal statement, there seems to be a preference for consistent interpretations where the number of pairs that are in the scope relation is as large as possible. For instance in (3), not every student can stare at every other student (one can stare at at most one person), but intuitively, the sentence requires that *every* student stares at some other student. While such an interpretation is much weaker than that of (2), this maximisation of the scope relation yields a reading that is also much stronger than the minimal IAO interpretation of (4). More generally, while IAO provides us with a lower bound for the interpretation of reciprocal statements, we will see in section 3 that the maximisation of the scope relation that is consistent with contextual knowledge yields the upper bound for the interpretation of a particular reciprocal statement i.e., its meaning.

Based on these observations, the principle determining the actual logical contribution of a reciprocal statement can be stated as follows:

> **Maximise Meaning Hypothesis (MMH)**: The valid interpretations of a reciprocal sentence S in a context Γ (where Γ includes knowledge about the previous discourse, the discourse situation and the world) are those which (a) are consistent both with the IAO form of reciprocity and the information provided by Γ, and (b) whose contributions to the scope relation are the strongest.

The MMH selects from the set of interpretations that are consistent with IAO and contextual knowledge, those that maximise the scope relation. Crucially, this view of reciprocals leads to an inference method that can actually compute the preferred interpretations of reciprocal sentences. We now turn to this.

3 Interpretation as Model Generation

In Discourse Representation Theory (DRT, (Kamp, 1981; Kamp and Reyle, 1993)), a sentence with semantic representation Φ is true with respect to a model **M** iff there is an embedding of Φ onto **M**. Intuitively, this requirement says that a sub-model **M'** of **M** must be found which satisfies Φ. So for instance, sentence (6a) is true in **M** iff there are two individuals *bugs* and *bunny* in **M** such that *bugs* and *bunny* stand in the *love* relation; or in other words, iff the partial model sketched in (6b) is part of **M**.

(6) a. Bugs likes Bunny.
    b. $\{\mathit{love}(\mathit{bugs}, \mathit{bunny})\}$

As shown in (Gardent and Konrad, To appear), model generators (i.e., programs that compute some of the models satisfying a finite set of logical formulas) can be used to provide DRT, and more generally model–theoretic approaches to natural language semantics, with a procedural interpretation: Given the semantic representation of a discourse and the relevant world knowledge Φ (i.e., a finite set of logical formulas), a model generator proves that Φ is satisfiable by generating some of its models.

Intuitively, satisfying models explain how discourses can be made true. They give an abstract representation of how (part of) the world should be for a discourse to be true. Concretely, satisfying models can be seen as capturing the meaning of discourses: data–bases that can be queried e.g. as part of a query/answer system or to interpret subsequent discourse. Satisfying models are also remininiscent of Johnson-Laird's *mental models* (Johnson-Laird and Byrne, 1991) and in essence, mental models are very much like the Herbrand models we are making use of here.

Formally, a **model** is a mathematical structure that describes how the symbols of a logical theory are interpreted. Given a first-order language $\mathcal{L}$, a model is a pair $\langle I, D\rangle$ with D a non-empty set of entities (the **domain of individuals**) and I an interpretation function which maps relation symbols in $\mathcal{L}$ to relations of appropriate arity in D and constant symbols in $\mathcal{L}$ to elements of D. Here we identify these models with sets of positive assumptions that unambiguously define the interpretation of the relation symbols and fix the interpretation of terms to first-order entities that carry a unique name. These are known in the literature as *Herbrand models.*

The set (7c) is such a model for the logical form (7b) which is a semantic representation of the sentence (7a).

(7) a. Jon likes his cousin.
    b. $\exists x\ \mathit{cousin\_of}(x, jon) \wedge \mathit{like}(jon, x)$
    c. $\mathcal{M}_1 = \{\mathit{cousin\_of}(c_1, jon),$ $\mathit{like}(jon, c_1)\}$

The model $\mathcal{M}_1$ defines an interpretation of the predicates *cousin* and *like* over the universe of discourse $\mathcal{D} = \{jon, c_1\}$. It can also be taken as a valid interpretation of (7a).There are, however, infinitely many models for (7b) that do not correspond to such interpretations e.g.

(8) $\mathcal{M}_2 = \{\mathit{cousin\_of}(jon, jon), \mathit{like}(jon, jon)\}$

(9) $\mathcal{M}_3 = \{\mathit{cousin\_of}(c_1, jon), \mathit{like}(jon, c_1),$ $\mathit{like}(c_1, jon)\}$

The model $\mathcal{M}_2$ explains the truth of (7a) by declaring Jon as his own cousin. This is a result of the inappropriate semantic representation (7b) which fails to specify that the relation expressed by the noun *cousin* is irreflexive. In the case of $\mathcal{M}_3$, the model contains superfluous information. While it is consistent to assume $\mathit{like}(c_1, jon)$ it is not necessary for explaining the truth of the input.

3.1 Minimality

For applications to natural-language, we are interested in exactly those models that capture the meaning of a discourse, or at least capture the preferred interpretations that a hearer associates with it. As discussed in (Gardent and Webber, January 2000), obtaining only these models requires eliminating both models that are "too small" (e.g. $\mathcal{M}_2$) and models that are "too big" (e.g. $\mathcal{M}_3$).

Models such as $\mathcal{M}_2$ can be eliminated simply by using more appropriate truth conditions for NL expressions (e.g. $\exists x\ \mathit{cousin}(x) \wedge \mathit{of}(x, jon) \wedge x \neq jon \wedge \mathit{like}(jon, x)$ for (7a)). In general however, eliminating models that are "too small" is a non-trivial task which involves the interaction of model-theoretic interpretation not only with world knowledge reasoning but also with syntax, prosody and pragmatics. The issue is discussed at some length (though not solved) in (Gardent and Webber, January 2000).

To eliminate models that are "too big", some notion of minimality must be resorted to. For instance, (Gardent and Konrad, 1999; Gardent and Konrad, To appear) argues that *local minimality* is an adequate form of minimality for interpreting definite descriptions. Local minimality is defined as follows.

Local Minimality: *Let Φ be a set of first-order formulas and $\mathbf{D}$ be the set of Herbrand models of Φ that use some finite domain D whose size is minimal. Then a model $\langle I, D\rangle \in \mathbf{D}$ is* ***locally minimal*** *iff there is no other model $\langle I', D'\rangle \in \mathbf{D}$ such that $I' \subseteq I$.*

Locally minimal models are models that satisfy some input Φ within a minimal domain $\mathcal{D}$ of individuals *and* are subset-minimal with respect to all other domain minimal models. These models are the simplest in the sense of Occam's Razor and often the best explanation for the truth of an observation. In particular, if we assume that $\mathcal{M}_2$ is ruled out by a more appropriate semantics for the word *cousin*, local minimality rules out $\mathcal{M}_3$ as non locally minimal and therefore $\mathcal{M}_1$ is correctly identified as giving the preferred interpretation for example (7).

3.2 The MMH as a Minimality Constraint

In the case of reciprocals, local minimality is clearly not a characterisation of preferred interpretations. Our semantic representation RCP_{IAO} will only capture a reciprocal's meaning if the reciprocal group has exactly two members or if the input meets IAO, the weakest form of reciprocity. For instance, the locally minimal model (10c) of formula (10b) is too weak to constitute an acceptable interpretation of (10a). Instead, the model capturing the meaning of (10a) is the model given in (10d).

(10) a. Jon, Bill and Dan like each other.
b. $\text{RCP}_{IAO}(\{jon, bill, dan\})(\lambda y \lambda x\ like(x,y))$
c. $\{like(jon,bill), like(bill,dan)\}$
d. $\{like(jon,bill), like(jon,dan), like(bill,dan), like(bill,jon), like(dan,bill), like(dan,jon)\}$

Since the MMH maximises rather than minimises the logical contribution of formulas, it seems at first sight incompatible with local minimality. However, a simple method to combine the MMH and model minimality is to consider the maximisation of reciprocal relations as a minimisation of their complement sets. After all, the difference in acceptability between (10c) and (10d) as models for (10a) is due to exactly those pairs $\langle x,y\rangle$ (with $x \neq y$) that are *not* in the *like* relation. To capture this intuition, we introduce a special predicate *\$R* that indicates assumptions whose truth is considered "costly". In our case, these assumptions correspond to the pairs of individuals that are not in the scope relation. The semantic representation of reciprocal *each other* is then as follows.

(11) $\text{RCP} \equiv \lambda P \lambda R\ (\text{RCP}_{IAO}(P)(R) \wedge \forall x \forall y\ (P(x) \wedge P(y) \wedge x \neq y \wedge \neg R(x,y) \Leftrightarrow \$R(x,y)))$

The first conjunct says that a reciprocal sentence has as weakest possible meaning an IAO reading. Since IAO is entailed by other identified meaning for reciprocal statements, this is compatible with the fact that reciprocal sentences can have other, stronger meanings. The second conjunct says that each pair $\langle x,y\rangle$ (with $x \neq y$) that is **not** in the *like* relation is in the *\$R* relation. This encoding leads to models like (12b) and (12c) for (12a). We say that model (12b) has a *\$R*-cost of 4 (*\$R4*), while model (12c) has a cost of 0.

(12) a. $\text{RCP}(\{jon, bill, dan\})(\lambda y \lambda x\ like(x,y))$
b. $\{like(jon,bill), like(jon,dan), \$R(bill,dan), \$R(bill,jon), \$R(dan,bill), \$R(dan,jon)\}$
\$R4
c. $\{like(jon,bill), like(jon,dan), like(bill,dan), like(bill,jon), like(dan,bill), like(dan,jon)\}$
\$R0

We now introduce a new form of minimality whose definition is as follows.

Conservative Minimality: *Let Φ be a set of first-order formulas and $\mathbf{D}$ be the set of Herbrand models of Φ with a minimal domain $\mathcal{D}$. Then $\mathbf{D}$ has a subset $\mathbf{C}$ of models that carry a minimal cost. A model $\langle I, D\rangle \in \mathbf{C}$ is conservative minimal iff there is no other model $\langle I', D'\rangle \in \mathbf{C}$ such that $I' \subseteq I$.*

Conservative minimality is a conservative extension of local minimality: if there are no costs at all, then all local minimal models are also conservative models. Conservative minimality is a combination of local minimality and cost minimisation that correctly identifies the preferred interpretation of reciprocal sentences. For instance since (12c) carries a minimal cost, it is a conservative minimal model for (12a) whereas (12b) isn't. Intuitively the approach works as follows: the more pairs there are that do not stand in the scope relation of the reciprocal, the bigger the *\$R* predicate and the more costly (i.e. the least preferred) the model. That is, the combined use of a *\$R*-predicate and of conservative minimality allows us to enforce a preference for interpretations (i.e. models) maximising R.

3.3 The System

KIMBA (Konrad and Wolfram, 1999) is a finite model generator for first-order and higher-order logical theories that is based on a translation of logics into constraint problems over finite-domain integer variables. KIMBA uses an efficient constraint solver to produce solutions that can be translated back into Herbrand models of the input.

We have tailored KIMBA such that it enumerates the conservative models of its input. Intuitively, this works as follows. First, KIMBA searches for some arbitrary model of the input that mentions a minimum number of individuals. Then, it takes the *\$R*-cost of this model as an upper bound for the cost of all successor models and further minimises the cost as far as possible by branch-and-bound search. After KIMBA has determined the lowest cost possible, it restarts the model search and eliminates those models from the search space that have a non-minimal cost. For each model $\mathcal{M}$ that it identifies as a cost-minimal one, it proves by refutation that there is no other cost-minimal model $\mathcal{M}'$ that uses only a subset of the positive assumptions in $\mathcal{M}$. Each succesful proof yields a conservative minimal model.

All the examples discussed in this paper have been tested on Kimba and can be tried out at:

`http://www.coli.uni-sb.de/cl/projects /lisa/kimba.html`

3.4 A spectrum of possible meanings

Let us see in more detail what the predictions of our analysis are. As we saw in section 2, reciprocal statements can have very different truth conditions. Intuitively, these truth-conditions lie on a spectrum from the weakest IAO interpretation (A is the antecedent set and R the scope relation):

$$|A| \geq 2 \wedge \forall x \in A(x)\ \exists y\ (A(y) \wedge x \neq y \wedge (R(x,y) \vee R(y,x))$$

to the strongest so-called Strong Reciprocity (SR) interpretation namely:

$$|A| \geq 2 \wedge \forall x\ A(x) \forall y\ A(y)(x \neq y \Rightarrow R(x,y))$$

We now see how the MMH allows us to capture this spectrum.

Let us start with example (2) whose truth-conditions are the strongest Strong Reciprocity conditions: every distinct x and y in the antecedent set are related to each other by the scope relation. In this case, there is no constraining world knowledge hence the content of the *like* relation can be fully maximised. For instance if there are five students, the cheapest model is one in which the cardinality of *like* is twenty (and consequently the cardinatity of *\$R* is zero).

(13) $\{like(s1,s2), like(s1,s3), like(s1,s4), like(s1,s5), like(s2,s1), like(s2,s3), like(s2,s4), like(s2,s5), like(s3,s1), like(s3,s2), like(s3,s4), like(s3,s5), like(s4,s1), like(s4,s3), like(s4,s2), like(s4,s5), like(s5,s1), like(s5,s3), like(s5,s2), like(s5,s4)\}$ *\$R0*

By contrast, example (3) has a much weaker meaning. In this case there is a strong world knowledge constraint at work, namely that one can stare at only one other person at some time. The cheapest models compatible with this knowledge are models in which *every* student stare at exactly *one* other student. Thus in a universe with five students, the preferred interpretations are models in which the cardinality of the scope relation *x stares at y in surprise* is five. The following are example models. For simplicity we ommit the *\$R* propositions and give the cost of the model instead (i.e. the cardinality of the complement set of the scope relation).

(14) $\{stare\_at(s1,s2), stare\_at(s2,s3), stare\_at(s3,s4), stare\_at(s4,s5), stare\_at(s5,s3)\}$ *\$R15*

(15) $\{\mathit{stare\_at}(s1, s2), \mathit{stare\_at}(s2, s3), \mathit{stare\_at}(s3, s4), \mathit{stare\_at}(s4, s5), \mathit{stare\_at}(s5, s1)\}$ *$R15*

Sentence (4) illustrates an intermediate case with respect to strength of truth conditions. World knowledge implies that the scope relation *x give y measles* is assymetric and further that every individual is given measles by at most one other individual. Given a set of five students, model (16) and (17) are both acceptable interpretations of (4), (16) being the preferred interpretation.

(16) $\{\mathit{gave\_measles}(s1, s2), \mathit{gave\_measles}(s1, s3), \mathit{gave\_measles}(s2, s4), \mathit{gave\_measles}(s3, s5)\}$ *$R16*

(17) $\{\mathit{gave\_measles}(s1, s2), \mathit{gave\_measles}(s2, s4), \mathit{gave\_measles}(s3, s5)\}$ *$R17*

In short, these examples show the MMH at work. They show how given a *single* semantic representation for reciprocals, a variety of meanings can be derived as required by each specific reciprocal statement. Two elements are crucial to the account: the use of model building, and that of minimality as an implementation of preferences. Model building allows us to compute all the finite interpretations of a sentence that are consistent with contextual knowledge and with an IAO interpretation of the reciprocal expression. Preferences on the other hand (i.e. the use of the cost predicate *$R* and the search for conservative mininal models), permits choosing among these interpretations the most likely one(s) namely, the interpretation(s) maximising the scope relation.

4 Related Work

(Dalrymple et al., 1998) (henceforth DKKMP) proposes the following taxonomy of meanings for reciprocal statements (A stands for the antecedent set and R for the scope relation):

Strong Reciprocity (SR)
$\forall x, y \in A(x \neq y \rightarrow xRy)$.

Intermediate reciprocity (IR)
$\forall x, y \in A\ \exists z_1, \ldots \exists z_m \in A(x \neq y \rightarrow xRz_1 \wedge \ldots \wedge z_mRy)$

One-way Weak Reciprocity (OWR)
$\forall x \in A\ \exists y \in A\ (xRy)$

Intermediate Alternative Reciprocity (IAR)
$\forall x, y \in A \exists z_1, \ldots \exists z_m \in A(x \neq y \rightarrow (xRz_1 \vee z_1Rx) \wedge \ldots \wedge (z_mRy \vee yRz_m))$

Inclusive Alternative Ordering (IAO)
$\forall x \in A\ \exists y \in A(xRy \vee yRx)$

To predict the meaning of a specific reciprocal sentence, DKKMP then postulate the *Strongest Meaning Hypothesis* which says that the meaning of a reciprocal sentence is the logically strongest meaning consistent with world and contextual knowledge.

The main difference between the DKKMP approach and the present approach lies in how the best reading is determined: it is the logically strongest of the five postulated meanings in DKKMP, whereas in our approach, it is that reading which maximises the scope relation of the reciprocal. This difference has both empirical and computational consequences.

Empirically, the predictions are the same in most cases because maximising the scope relation often results in yielding a logically stronger meaning. In particular, as is illustrated by the examples in section 2, the present approach captures the five meanings postulated by DKKMP. Thus model (13) exemplifies an SR reading, model (15) an IR reading and model (14) an OWR reading. Further, model (16) is an IAR interpretation while model (17) shows an IAO reading.

But as the examples also show there are cases where the predictions differ. In particular, in the DKKMP approach, sentence (3) is assigned the IR reading represented by model (15). However as they themselves observe, the sentence also has a natural OWR interpretation namely, one as depicted in model (14), in which some pairs of students reciprocally stare at each other. This is predicted by the present approach which says that models (14) and (15) are equally plausible since they both maximise the *stare at* relation to cardinality five.

On the other hand, the DKKMP account is more appropriate for examples such as:

(18) The students sat next to each other
  a. forming a nice cercle.

b. filling the bench.
c. some to the front and others to the back of the church.

An IR interpretation is predicted for (18) which is compatible with both continuation (18a) and continuation (18b). By contrast, the model generation approach predicts that the preferred interpretation is a model in which the students form a circle, an interpretation compatible with continuation (18a) but not with continuations (18b-c).

However, both approaches fail to predict the reading made explicit by continuation (18c) since this corresponds to the weaker OWR interpretation under the DKKMP account and to a model which fails to maximise the scope relation under the present approach. More generally, both approaches fail to capture the semantic vagueness of reciprocal statements illustrated by the following examples[1]:

(19) a. The students often help each other with their homework.
b. In the closing minutes of the game, the members of the losing team tried to encourage each other.

In both cases, the sentences can be true without maximising either the strength of its truth conditions (Strong Reciprocity) or the scope relation. This suggests that an empirically more correct analysis of reciprocals should involve prototypical and probabilistic knowledge – as it is essentially a computational approximation of the DKKMP approach, the present account does not integrate such information though it is compatible with it: just as we restrict the set of generated models to the set of conservative minimal models, we could restrict it to the set of models having some minimal probability.

Computationally, the difference between the DKKMP and the present approach is as follows. In essence, the DKKMP approach requires that each of the five possible readings (together with the relevant world knoweldge) be checked for consistency: some will be consistent, others will not. Since the first order consistency and validity problems are not decidable, we know that there can be no method guaranteed to always return a result. In order to implement the DKKMP approach, one must therefore resort to the technique advocated in (Blackburn et al., 1999) and use both a theorem prover and a model builder: for each possible meaning M_i, the theorem is asked to prove $\neg M_i$ and the model builder to satisfy M_i. M_i is inconsistent if the theorem prover succeeds, and consistent if the model builder does. Theoretically however, cases may remain where neither theorem proving nor model building will return an answer. If these cases occur in practice, the approach simply is not an option. Further, the approach is linguistically unappealing as it in essence requires the reciprocal *each other* to be five–way ambiguous.

By contrast, the model generation approach assigns a single semantic representation to *each other*. The approach strengthens the logical contribution of the weak semantic representation as a process based on *computational* constraints on a set of effectively enumerable models. As a result, we will never encounter undecidable logical problems as long as the represented discourse is consistent. The model generator is the only computational tool that we need for determining preferable readings, and our experiment shows that for the examples discussed in this paper, it returns preferred readings in a few seconds on standard PCs as long as the background theory and the size of the domain remain managably small.

5 Conclusion

We have argued that model building can be used to provide a computational approximation of DKKMP's analysis of reciprocals.

One crucial feature of the account is that it permits building, comparing and ranking of natural-language interpretations against each other. In the case of reciprocals, the ranking is given by the size of the scope relation, but other ranking criteria have already been identified in the literature as well. For instance, (Gardent and Konrad, To appear) shows that in the case of definite descriptions, the ranking defined by local minimality permits capturing the preference of binding over bridging, over accomodation. Similarly (Baumgartner and Kühn, 1999) shows that a predicate minimisation together with a preference for logically consequent reso-

[1] I am thankful to an anonymous NAACL referree for these examples.

lutions can be used to model the interpretation of pronominal anaphora.

This suggests that one of the most promising application of model generators is as a device for developing and testing *preference systems* for the interpretation of natural language. Inference and knowledge based reasoning are needed in NLP not only to check for consistency and informativity (as illustrated in e.g. (Blackburn et al., 1999)), but also to express preferences between, or constraints on, possible interpretations. For this, finite model builders are natural tools.

Another area that deserves further investigation concerns the use of minimality for disambiguation. In this paper, conservative minimality is used to choose among the possible interpretations of a particular reciprocal statement. On the other hand, (Gardent and Webber, January 2000) shows that minimality is also an important tool for disambiguating noun-compounds, logical metonymy and definite descriptions. As the paper shows though, many questions remains open about this use of minimality for disambiguation which are well worth investigating.

In further work, we intend to look at other ambiguous natural language constructs and to identify and model the ranking criteria determining their preferred interpretation. Plurals are a first obvious choice. But more generally, we hope that looking at a wider range of data will unveil a broader picture of what the general biases are which help determine a preferred reading — either in isolation, as here, or in context, as in (Gardent and Webber, January 2000) — and of how these biases can be modelled using automated reasoning systems.

Acknowledgements

We are grateful to audiences from ITRI-Brighton, the Edinburgh School of Cognitive Science, the Paris VI TALANA seminar and the Amsterdam DIP colloquium for helpful comments and discussion on the material presented here as well as to the three NAACL anonymous referrees for constructive feedback. This work was partially supported by the Project C2 (LISA) in SFB-378, grant by the Deutsche Forschungsgemeinschaft to the University of Saarbrücken.

References

Peter Baumgartner and Michael Kühn. 1999. Abductive coreference by model construction. In *ICoS-1 Inference in Computational Semantics*, Institute for Logic, Language and Computation, University of Amsterdam, August.

P. Blackburn, J. Bos, M. Kohlhase, and H. de Neville. 1999. Inference and Computational Semantics. In *Third International Workshop on Computational Semantics (IWCS-3)*, Tilburg, The Netherlands.

Mary Dalrymple, Makoto Kanasawa, Yookyung Kim, Sam Mchombo, and Stanley Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 21(2):159–210, April.

Claire Gardent and Karsten Konrad. 1999. Definites and the proper treatment of rabbits. In *Proceedings of ICOS*. Also CLAUS Report 111, http://www.coli.uni-sb.de/claus/.

Claire Gardent and Karsten Konrad. To appear. Interpreting Definites using Model Generation. *Journal of Language and Computation*.

Claire Gardent and Bonnie Webber. January 2000. Automated deduction and discourse disambiguation. Submitted for Publication. Also CLAUS Report 113, http://www.coli.uni-sb.de/claus/.

P.N. Johnson-Laird and Ruth M.J. Byrne. 1991. *Deduction*. Lawrence Erlbaum Associates Publishers.

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

Hans Kamp. 1981. A theory of truth and semantic representation. In J. Groenendijk, Th. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*, pages 277 - 322. Mathematisch Centrum Tracts, Amsterdam.

Karsten Konrad and D. A. Wolfram. 1999. Kimba, a model generator for many-valued first-order logics. In *Proc., 16th International Conference on Automated Deduction, CADE 99*, LNCS, forthcoming, Trento, Italy. Springer.

D. Terence Langendoen. 1978. The logic of reciprocity. *Linguistic Inquiry*, 9(2):177–197.