Natural Language Inference as Triggered Submodel Search [*]

Daniel Hardt
Copenhagen Business School
dh@id.cbs.dk

Abstract

I propose a simple, general framework for the interaction of inference with natural language interpretation. First, inference is only available when triggered by the violation of highly ranked constraints. Second, inference is constrained to be a search for a minimal submodel. I show that this correctly captures facts concerning deaccenting, ellipsis, and reciprocal interpretation.

Introduction

While it is widely agreed that inference plays an important role in the interpretation of natural language, there is little agreement about when inference is in fact available, and how it is constrained. In this paper I propose a simple, general framework for the interaction of inference with natural language interpretation. I suggest that inference is only available when triggered by the violation of highly ranked constraints. Furthermore, I propose that inference is constrained to be a search for a minimal submodel.

According to this proposal, a given linguistic construction has a default interpretation that does not require inference. For example, the default interpretation of an elliptical expression is that the elided material is *identical* to the antecedent expression. The default interpretation of a reciprocal expression is the interpretation relating *all* pairs of individuals in the range of the quantifier. Inference is only permitted in cases where the default interpretation violates a highly ranked linguistic constraint, such as AGREE, AVOID-CONTRADICTION, or PARALLEL. When inference is triggered in this way, the only possible inferences are those whose minimal models are *submodels* of the model associated with the default interpretation.

In what follows, I first describe the general model of triggered, submodel inference. Next, I apply the model to a variety of phenomena in which inference has been observed to play a role in interpretation. First, I address phenomena involving deaccenting and the PARALLEL constraint as applied to quantifier scoping. Next, I look at ellipsis phenomena, and finally, the interpretation of reciprocal expressions.

In each case, the model uniformly captures subtle facts concerning the interaction of inference with interpretation, successfully predicting when inference is triggered, and constraining inference to produce precisely the observed interpretations.

A General Model of Natural Language Inference

Highly Ranked Constraints

I will assume in this paper that a given construction has a *default* interpretation that does not involve inference. I will also assume the existence of several highly ranked linguistic constraints. It is the violation of these constraints that triggers inference. In this paper, I appeal to the following constraints:

- **AGREE:** This covers number agreement, and rules out, e.g., "Martha danced together", while permitting "Martha and Irv danced together"

- **AVOID-CONTRADICTION:** This is a logical inconsistency arising from a discourse. (See (Hendriks & de Hoop 2001) for recent discussion of this as a preference principle.)

- **PARALLEL:** The requirement or preference for certain utterances to receive parallel interpretations has been widely studied (see (Fox 2000; Asher 1993), as well as references cited in those works). I will follow Rooth's Alternative Semantics theory (Rooth 1985) in implementing certain aspects of this requirement, according to which two utterances are required to have matching interpretations, apart from focused elements.

Submodels and Inference

We assume a logical form, (LF) for a given natural language utterance. A first order model of an LF is as follows.
First Order Model:

- **Domain** of model M, dom(M): a set of individuals.

- **Interpretation Function** of M, I(M): a function from relation symbols to sets of n-tuples of elements of the Domain. For the ordering on models, we view the Interpretation as a set of *assertions* about individuals in the domain.

Ordering on models: M1 $\leq$ M2 iff card(dom(M1)) $\leq$ card(dom(M2)) and I(M1) $\subseteq$ I(M2). (Gardent & Webber 2001; Gardent & Konrad 2000a; Konrad 2004)

Following (Gardent & Webber 2001; Gardent & Konrad 2000a; Konrad 2004), we can now define a minimal model.

Minimal Model: for any LF α, M ϵ Min-models(α) iff M satisfies α and no model M$'$ that satisfies α is below M in the model ordering.

If M1 $\leq$ M2, we say that M1 is a *submodel* of M2.

Simply put, we can say that a model $M1$ is a *submodel* of $M2$ if the assertions in $M1$ is a subset of the set of assertions in $M2$.

A General Model of Inference

Our proposal can now be stated as follows: for a given discourse D, we produce a *default* LF L. If L violates no constraints, it is the preferred interpretation. If L does violate one or more constraints, we perform inferences. An inferred L' is a potential interpretation of D if it avoids the constraint violations, and if it has a minimal model that is a submodel of a minimal model of L. If there are several such alternatives, those LF's *closest* to L in the submodel ordering are preferred.

Deaccenting and PARALLEL

There is a vast literature on the requirement that a discourse be coherent (see (Hobbs 1979; Asher 1993), as well as references cited in those works). Here, I will focus on one well-studied aspect of that discourse coherence, namely Parallelism. Following the discussion in (Fox 2000), Direct Parallelism involves satisfying the requirements of Alternative Semantics (Rooth 1985); parallel sentences must match, except possibly for focused elements. If Direct Parallelism fails for two sentences A and B, Indirect Parallelism is an option; Indirect Parallelism can be satisfied if an inference from A to A$'$ is possible, where A$'$ and B satisfy Direct Parallelism.

I will use the following examples to show that the Inference involved in Direct Parallelism must be minimal with respect to submodel search. [1]

(1)   A doctor saw every patient. A NURSE saw every patient, too.

(2)   A doctor saw every patient. A NURSE saw many patients, too.

---

[1] A similar argument is made in (Hardt forthcoming), where it is shown that model-based constraints on inference capture these facts, while other theories, such as that of (Fox 2000), do not. However, (Hardt forthcoming) proposes that inferences be restricted to the nearest minimal models. The current proposal is more restrictive, since it limits inferences to the nearest *submodels*. Furthermore, the proposal in (Hardt forthcoming) does not apply uniformly to the range of phenomena considered here.

As observed by many authors (see (Asher, Hardt, & Busquets 2001) and references therein) scope ambiguity should be resolved in parallel in both examples: if "a doctor" takes wide scope, "a nurse" also takes wide scope; if "a doctor" takes narrow scope, so does "a nurse". For (1), this follows from Direct Parallelism. But for (2), Direct Parallelism fails, since "many" does not match "every".

We focus on the case where "a doctor" takes wide scope. Since Direct Parallelism fails, we must apply Indirect Parallelism. In other words, an inference is required to establish parallelism. We can then infer A$'$ (*A doctor saw many patients*) from A (*A doctor saw every patient*). (Note that this only follows if there is a presupposition that there are many patients.) The inferred A$'$ allows parallelism to be established, and this is consistent with scope parallelism. The problem is that A$''$, with inverse scoping, can also be inferred from A. But this would then license a non-parallel reading. Here are the three relevant LF's:

A:
(A doctor x) (every patient y) x saw y.
A$'$: (Parallel)
(A doctor x) (many patients y) x saw y.
A$''$: (Non-parallel)
(many patients y) (A doctor x) x saw y.

It is necessary to rule out the inference giving rise to A$''$, while permitting the inference to A$'$. An inspection of minimal models shows that the proposed approach achieves this result.

Note first that I assume that there are at least four doctors, patients, and nurses, and that many means "at least 3". While this may seem arbitrary, I believe some such assumption is required here: the use of "every" presupposes the existence of a domain of some reasonable size. For example, it is a bit odd to say "every student asked a question" if there are just two students; in this case "both" is a more felicitous determiner.

So each minimal model will contain the following information:

doctor(d1), doctor(d2), doctor(d3), doctor(d4)
patient(p1), patient(p2), patient(p3), patient(p4)
nurse(n1), nurse(n2), nurse(n3), nurse(n4)

Where the three models differ is in the *saw* relation. Below, we give the *saw* relation in the minimal model of each LF, A, A$'$, and A$''$.

- **A:**
  (A doctor x) (every patient y) x saw y.

- **M$_A$:**
  saw(d1,p1), saw(d1,p2), saw(d1,p3), saw(d1,p4)

- **A$'$:** (Parallel)
  (A doctor x) (many patients y) x saw y.

- **M$_{A'}$:**
  saw(d1,p1), saw(d1,p2), saw(d1,p3)

- **A″:** (Non-parallel)
  (many patients y) (A doctor x) x saw y.

- **M$_{A″}$:**
  saw(d1,p1), saw(d2,p2), saw(d3,p3)

Now it is clear that the inference from A to A′ involves submodel search; in this case, a single literal, $saw(d1, p4)$ is eliminated from M_A to produce $M_{A′}$. $M_{A″}$ is not a submodel of A, and thus is correctly ruled out.

Ellipsis and AGREE

Consider now the following example of Verb Phrase Ellipsis, from (Webber 1978):

(3) Irv and Martha wanted to dance together, but Martha couldn't, because her husband was there.

Here, the VP is elided in the clause "Martha couldn't". We make the standard assumption that the default reading of an ellipsis is that it is identical to the antecedent. In this example, the antecedent VP is "dance together". So the default reading of the elliptical sentence is:

(4) Martha couldn't *dance together*.

Of course, this is not the desired reading in this case. In our view, this is because the default reading gives rise to an AGREE violation (the VP modifier *together* requires a plural subject). As argued by Webber, the desired reading arises from the following inference:

(5) Irv and Martha wanted to dance together $⇒$ Martha wanted to dance with Irv.

Again, this is consistent with the submodel proposal. Intuitively, we have the following minimal models.

- **A:** *Irv and Martha wanted to dance together*

- **M$_A$:**
  want(M,dance(M,I))
  want(I,dance(M,I))

- **A′:** *Martha wanted to dance with Irv*

- **M$_A$′:**
  want(M,dance(M,I))

It is clear that this respects the submodel constraint, since the model of $A′$ is a submodel of A.

While this representation is perspicuous, it goes beyond first order logic, in that the proposition $dance(M, I)$ is an argument to $want$. There are a variety of proposals in the literature for treating such predications with first order representations. Here, we will follow (Hobbs 1985): we assume a set of situation[2] variables, $s_1 \ldots s_n$, which appear as extra arguments to all predications. Thus we can write that there is a situation s in which Martha and Irv are dancing, and that s is a situation which Martha wants. No higher order terms are necessary.

- **A:** *Irv and Martha wanted to dance together*

[2]Hobbs called these *events*, following (Davidson 1980).

- $\exists s1.(want(M, s1) \quad \wedge \quad dance(M, I, s1)) \quad \wedge$
  $\exists s2.(want(I, s2) \wedge dance(M, I, s2))$

- **M$_A$:**
  want(M,s), dance(M,I,s), want(I,s)

- **A′:** *Martha wanted to dance with Irv*

- $\exists s1.(want(M, s1) \wedge dance(M, I, s1))$

- **M$_A$′:**
  want(M,s), dance(M,I,s)

It is clear that the desired inference from A to A′ is consistent with submodel search; we simply delete the assertion $want(I, s)$. Furthermore, the inferred $A′$ removes the AGREE violation.

The proposal here follows the basic observation in (Webber 1978) concerning the need for inference in constructing the desired meaning. However, Webber does not propose any constraints on the interaction of inference with interpretation of ellipsis.

The triggering and submodel constraints play a crucial role in correctly determining the role of inference in ellipsis interpretation, and preventing unwanted inference-based readings. Note that there are many other inferences that could be drawn from *Irv and Martha wanted to dance together*, other than the desired *Martha wanted to dance with Irv*. For example, one could infer *Martha wanted to dance with someone who wants to dance*. This is a valid inference, and, arguably, just as easy an inference as the desired inference. However, (3) cannot mean *Martha can't dance with someone who wants to dance*. This inference is ruled out by the submodel constraint.

The triggering constraint also rules out unwanted inference-based readings. For example, consider the following variant of (3):

(6) Irv and Martha wanted to dance together, but Tom and Susan didn't want to.

Here, the only possible reading is *Tom and Susan didn't want to dance together*. On our proposal, this is because there is no AGREE violation to trigger inference here. Without this triggering constraint, it's not clear how the inference-based reading *Tom and Susan didn't want to dance with Irv* is to be ruled out.

Reciprocals and AVOID-CONTRADICTION

We turn now to the interpretation of reciprocals, as illustrated by (7):

(7) The students like each other.

Here, the preferred reading is that every pair of students participates in the *like* relation. Following the discussion in (Gardent & Konrad 2000b), we can describe a reciprocal as an operator RCP, applying to a set S and relation R. We define the default reading for a reciprocal (RCP(S,R)) as follows:

$$\forall x, y.(x \epsilon S \wedge y \epsilon S \wedge x \neq y) \rightarrow R(x, y)$$

As discussed by (Dalrymple *et al.* 1998) other, weaker readings are sometimes observed, as in the following example:

(8)   The students stare at each other in surprise.

Here, the default meaning is this:

$\forall x, y(student(x) \land student(y) \land (x \neq y) \rightarrow stare\_at(x, y)))$

(Every student stared at every other student.)

However, this is not the preferred reading here, as discussed by (Dalrymple *et al.* 1998). They point out that this reading contradicts world knowledge: it is not possible to stare at more than one person at a time.

- **A:**
  $\forall x, y(stare\_at(x, y) \rightarrow \forall z(z \neq y) \rightarrow not(stare\_at(x, z)))$

The preferred reading is, rather, the following:

- **A′:**
  $\forall x(student(x) \rightarrow \exists y(x \neq y \land student(y) \rightarrow stare\_at(x, y)))$

This result follows directly from our proposed framework: the default reading A produces a highly ranked violation, AVOID-CONTRADICTION. The desired reading, A', repairs the violation, and is consistent with the submodel constraint. To see that A' is consistent with the submodel constraint, consider the following model for the default meaning A, with three students:

- **M$_A$:**
  student(s1),student(s2),student(s3)
  stare_at(s1,s2), stare_at(s2,s3), stare_at(s1,s3)
  stare_at(s2,s1),stare_at(s3,s1), stare_at(s3,s2)

The desired reading A' has several minimal models, such as the following:

- **M$'_A$:**
  student(s1),student(s2),student(s3)
  stare_at(s1,s2), stare_at(s2,s3), stare_at(s3,s2)

It is clear that this is a submodel of M_A, since it results from removing the stare_at literals, *stare_at(s1,s3), stare_at(s2,s1),stare_at(s3,s1)*.

Next, we turn to the following example:

(9)   The students gave each other measles.

The default reading A, is:

- **A:**
  $\forall x, y(student(x) \land student(y) \land (x \neq y) \rightarrow gave\_measles(x, y))$

This gives rise to the following minimal model:

- **M$_A$:**
  student(s1),student(s2),student(s3)
  gave_measles(s1,s2), gave_measles(s2,s3),
  gave_measles(s1,s3), gave_measles(s2,s1),
  gave_measles(s3,s1), gave_measles(s3,s2)

The problem here is that one cannot get measles from more than one person at a time. Again, we have a violation of AVOID-CONTRADICTION. As noted by (Dalrymple *et al.* 1998), the desired reading is

- **A′:**
  $\forall x(student(x) \rightarrow \exists y(student(y) \land (x \neq y) \land gave\_measles(x, y) \lor gave\_measles(y, x))$

(This is dubbed the *Inclusive Alternative Ordering* by (Dalrymple *et al.* 1998).)

The desired reading A' has several minimal models, such as the following:

- **M$'_A$:**
  student(s1),student(s2),student(s3)
  gave_measles(s1,s2), gave_measles(s2,s3)

Again, it is clear that the desired reading is an inference that can be characterized as submodel search.

Conclusions

In this paper, I have proposed that inference is only available when triggered by the violation of a highly ranked constraint. Furthermore, not all possible inferences are possible; only those consistent with a *sub-model* constraint. What this means is that an inference from A to B is only permitted if the set of assertions in B's minimal model is a subset of the set of assertions in A's minimal model.

I have considered three types of phenomena in which inference plays an important role: ellipsis, deaccenting, and reciprocals. I have shown that, in all three cases, the proposed framework gives the correct analysis. In all three cases, the default reading is preferred unless there is a violation, and in case of a violation, an inference is permitted only if it respects the submodel constraint and repairs the violation.

These successful results rely on the postulation of highly ranked constraints: in particular, AGREE, PARALLEL, and AVOID-CONTRADICTION. Of course, these claims are only convincing if the constraints have independent motivation. This must be established in future research. It is equally important to establish what other highly ranked constraints there are; our prediction would be that any highly ranked constraint will serve as a trigger for submodel inference.

References

Asher, N.; Hardt, D.; and Busquets, J. 2001. Discourse parallelism, ellipsis, and ambiguity. *Journal of Semantics* 18(1).

Asher, N. 1993. *Reference to Abstract Objects in English*. Dordrecht.

Dalrymple, M.; Kanazawa, M.; Kim, Y.; Mchombo, S.; ; and Peters, S. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy* 21(2):159–210.

Davidson, D. 1980. The logical form of action sentences. In *Essays on actions and events*. Oxford: Clarendon Press. 105–148.

Fox, D. 2000. *Economy and Semantic Interpretation*. MIT Press.

Gardent, C., and Konrad, K. 2000a. Interpreting definites using model generation. *Journal of Logic, Language and Information* 1(2):193–209.

Gardent, C., and Konrad, K. 2000b. Understanding each other. In *Proceedings of the First Annual Meeting of the North American Chapter of the Association for Computational Linguistics*.

Gardent, C., and Webber, B. 2001. Towards the use of automated reasoning in discourse disambiguation. *Journal of Logic, Language and Information* 10:487–509.

Hardt, D. (forthcoming). Ellipsis resolution and inference. *Acta Linguistica Hafniensis*.

Hendriks, P., and de Hoop, H. 2001. Optimality theoretic semantics. *Linguistics and Philosophy* 24:1–32.

Hobbs, J. 1979. Coherence and coreference. *Cognitive Science* 3:67–90.

Hobbs, J. 1985. Ontological promiscuity. In *Proceedings of the 23rd conference on Association for Computational Linguistics*.

Konrad, K. 2004. *Model Generation for Natural Language Interpretation and Analysis*. Springer.

Rooth, M. 1985. *Association with Focus*. Ph.D. Dissertation, University of Massachusetts-Amherst.

Webber, B. L. 1978. *A Formal Approach to Discourse Anaphora*. Ph.D. Dissertation, Harvard University. (Published 1979 by Garland Publishing, New York).