

A computational account of comparative implicatures for a spoken dialogue agent

Luciana Benotti, David Traum

► **To cite this version:**

Luciana Benotti, David Traum. A computational account of comparative implicatures for a spoken dialogue agent. Eighth International Conference on Computational Semantics - IWCS-8 2009, Jan 2009, Tilburg, Netherlands. 2009. <inria-00424703>

HAL Id: inria-00424703

<https://hal.inria.fr/inria-00424703>

Submitted on 16 Oct 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A computational account of comparative implicatures for a spoken dialogue agent*

Luciana Benotti
LORIA/INRIA, France
benottil@loria.fr

David Traum
ICT/USC, USA
traum@ict.usc.edu

Abstract

Comparative constructions are common in dialogue, especially in negotiative dialogue where a choice must be made between different options, and options must be evaluated using multiple metrics. Comparatives explicitly assert a relationship between two elements along a scale, but they may also implicate positions on the scale especially if constraints on the possible values are present. Dialogue systems must often understand more from a comparative than the explicit assertion in order to understand why the comparative was uttered. In this paper we examine the pragmatic meaning of comparative constructions from a computational perspective.

1 Introduction

It is a big challenge for computational semantics of dialogue that much of the meaning of an utterance is conveyed not just through the compositional meanings of the words themselves, but in relation to the situation in which the utterance is performed, including the background knowledge, common ground, goals, and ontologies of the participants. A number of pragmatic principles have been proposed to bridge this gap, including Grice's maxims (and the associated concepts of Implicature and Relevance) [Grice, 1975], Accommodation [Lewis, 1979] and Bridging [Clark, 1975].

While these principles can provide elegant explanations of how meaning can be conveyed between people, they often require fairly strong assumptions about the common knowledge between participants and the ontologies that this knowledge must be organized in. There are two general problems in developing computational accounts of these phenomena. First, it is sometimes difficult to specify the required knowledge and relationships in a computational way, such that a

*This work was sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM). Statements and opinions expressed do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

reasoner given the input and context can compute a particular, specific meaning as opposed to other possibilities that are not as appropriate. Second, even if the principles are sufficiently clear so that a computational account can be formulated, there may still be a problem providing a given computational dialogue system with the appropriate knowledge to carry out the inferences in a way that is congruent with human interpretations. For a hand-constructed limited domain, the system designer will often take shortcuts and represent only the knowledge that is necessary to carry out and understand tasks in that domain. These limitations often render the dialogue system unable to reason about the domain in as much detail as a knowledgeable human would, but often this characterization is sufficient for the purposes of the conversation.

In this paper we examine one aspect of computational non-compositional meaning: the pragmatic meaning of comparative constructions. Comparative constructions are common in dialogue, especially in negotiative dialogue where a choice must be made between different options, and options must be evaluated using multiple metrics. Comparatives explicitly assert a relationship between two elements along a scale, but they may also implicate positions on the scale especially if either information about the position of one of the compared items or constraints on the possible values are present. Dialogue systems must often understand more from a comparative than the explicit assertion in order to understand why the comparative was uttered; why it is relevant to the dialogue.

In the next section, we present linguistic background on comparatives and conversational implicature. In section 3, we review some issues of how implicatures play out in dialogue in which multiple participants are involved, and a listener can clarify the (lack of) understanding. In section 4, we briefly introduce the computational framework in which the present work is implemented. In section 5, we present the extensions to this framework which provide the computational agents with the ability to understand the implicatures arising from comparatives. In section 6 we evaluate this with respect to the specific scenario of SASO-EN [Traum *et al.*, 2008a], in which an Army Captain, a Doctor, and a town Elder discuss the best location for a medical clinic. Finally, we discuss some related issues, remaining problems, and future work in Section 7.

2 Linguistic background

In this section we review some of the previous work on comparatives constructions and conversational implicatures in order to establish the necessary theoretical basis for the discussion on comparative implicatures in the rest of the paper.

2.1 Comparative constructions

In the classical literature on semantics of natural language comparatives (see e.g., [Cresswell, 1976]), comparative constructions such as (1) are analyzed as a relation between two degrees as in (2).

- (1) Downtown is safer than the market.
- (2) Degree to which downtown is safe $>$ Degree to which the market is safe.

The problem now, of course, is what a *degree* is and what are the properties of the relation $>$. Abstractly, a degree can be considered just as a collection of objects – the collection of those objects that share the same degree with respect to a given property P . These collections are defined as the equivalence classes of the equivalence relation $=_P$, which in turn is defined in terms of an order $>_P$ among objects. The relation $>$ among degrees is defined as a lifting of $>_P$ to the set of equivalent classes of $=_P$. Finally, scales are defined in terms of degrees, a *scale* S_P is a sequence of degrees of P ordered by the relation $>$.

Summing up then, once we know what $>_P$ is for a property P we know what degrees of P are and how to compare them on scale S_P . All is good and well, but this approach assumes the relation $>_P$ as given. Such a strong assumption was already criticized by [Klein, 1980] and certainly cannot be made in a dialogue system where information about the domain of discourse (in particular, any order $>_P$ for a given property P) is incomplete and is constructed (and negotiated) during the dialogue. As dialogue system builders, the issue that interests us is, not so much how to determine the truth value of a particular comparative utterance, but mainly how comparatives contribute to the construction of the information about the domain. So, for our task it is crucial to figure out “where do scales come from?”

2.2 Conversational implicatures

Modelling how listeners draw inferences from what they hear is a basic problem for the theories of understanding natural language. An important part of the information conveyed is inferred, as in the following classical example by Grice [Grice, 1975]:

- (3) A: I am out of petrol.
 B: There is a garage around the corner.
 \rightsquigarrow B thinks that the garage is open and has petrol to sell.

B’s answer *conversationally implicates* (\rightsquigarrow) information that is relevant to A. In Grice’s terms, B made a relevance implicature, he would be flouting Grice’s maxim of relevance unless he believes that the garage is open. A conversational implicature (CI) is different from an entailment in that it is *cancelable* without contradiction (B can append material that is inconsistent with the CI – “but I don’t know whether it’s open”). Since the CI can be cancelled, B knows that it does not necessarily hold and then both B or A are able to *reinforce* it (or negotiate it) without repetition. CIs are *non-conventional*, they are not part of the conventional meaning of the words but *calculable* from their utterance in context given the nature of conversation as a goal-oriented enterprise.

Scalar implicatures are a type of CI that are particularly relevant to this paper since they involve the use of scales, as comparatives do. These implicatures

are inferred on the basis of the assumption that the speaker is trying to make his utterance sufficiently informative for the current purposes of the exchange. A typical example is:

- (4) A is hungry (and B knows it).
A: Did somebody eat the brownies that I bought this morning?
B: Fred ate some of them.
 \leadsto B thinks that Fred didn't eat all of them, there are some left.

Theories of scalar implicature have been deeply influenced by Horn's dissertation work [Horn, 1972]. A Horn scale is an ordered sequence of expressions such as $\langle \textit{some}, \textit{many}, \textit{most}, \textit{all} \rangle$ and $\langle \textit{warm}, \textit{hot}, \textit{scalding} \rangle$. The recipe to calculate the scalar implicature is the following. The use of a weaker word in the scale (such as *some*) implicates that (the speaker believes that) the stronger form (such as *all*) is not the case, as exemplified in (4). However, there are cases in which such a recipe does not apply, such as in (5).

- (5) A realizes that the brownies were injected with strychnine and suspects that somebody may have eaten from them (and B knows it).
A: Did somebody eat the brownies that I bought this morning?
B: Fred ate some of them.

In this example the scalar implicature is (at the very least) less likely to arise, it's not *relevant* for the current purposes of the exchange; it doesn't matter how many brownies Fred ate, he needs medical attention. Comparing (4) and (5) we can see that conversational implicatures are affected by the conversational goals. We believe this to be an starting point for answering the question left open in Section 2.1: "Where do scales come from?" and we will investigate it in next section.

3 Comparative implicatures in dialogue

In this Section we are going to introduce comparative implicatures and to develop two lines of thought introduced in the previous section, but now in the context of dialogue. First, we relate the cancellability and negotiability of CIs and clarification requests in dialogue. It's often controversial whether something is actually a CI or not (people sometimes have different intuitions). Dialogue provide us with an extra tools for identifying the participants' beliefs about implicatures: feedback and negotiation of CIs. Listeners can give both positive (e.g. acknowledgements) and negative (e.g. clarification requests) feedback about their understanding and the acceptability of utterances, which can shed light on how they interpret CIs. Moreover, speakers can negotiate whether something is really implicated, or whether an implicature is meant by the speaker. We also investigate the fact that conversational implicatures change when the conversational goals change. Conversational goals establish structure on the information that is being discussed, determining which distinctions are relevant and which are not for the a given interaction.

3.1 Clarifications

In dialogue, a useful way to spot content that was meant but not actually said is to look at the kinds of dialogue feedback that is given. Clarification requests are one important part of this - focusing on cases where the initial listener is not sure of what the speaker means. Here we use the phrase *clarification requests* in its broad sense, including all those questions that would not make sense without the previous utterance in the dialogue.

Many times the clarification requests make the implicatures explicit, this is illustrated by the fact that the clarification request in (6) can follow Grice's example in (3).

(6) A: and you think it's open?

Here we present clarifications as a test that helps us identify not only the potential conversational implicatures but also the presuppositions in a dialogue. (7) is an example of a well studied case of inferrable content: the presupposition triggered by definite descriptions. In this exchange, A might believe that the intended referent is salient for different reasons: because it was mentioned before, because it's bigger, etc. If such a piece of information is available to B then he will be able to infer the presupposition and add it to the conversational record. Otherwise, B might well signal A that the presupposition did not get through, like in (7).

(7) There are 100 cups on a table.

A: Pick up THE cup.

B: Which cup? (I don't know which one you are referring to.)

The presence of a conversational implicature must be capable of being worked out, so the hearer might well inquire about them. The speaker will have to answer and support the implicature if he wants to get it added to the common ground. In [Rodríguez and Schlangen, 2004], the authors present a study of the different kinds of clarification requests (CRs) found in a dialogue corpus. Their results show that the most frequent CRs in their corpus (51.74%) are due to problems with referring expressions (such as example (7)) and the second most common (22.17%) are examples like the following.

(8) A: Turn it on.

B: By pushing the red button?

Here, the CR made explicit a potential requirement for the performance of the requested action. By saying "Turn it on" A meant that B had to push the red button but this was not explicitly mentioned in the command uttered by A. Such an implicature can be inferred from knowledge about the task (or in Clark terms, from knowledge about the interaction level of joint action [Clark, 1996]). From this examples it should be clear that implicated content is a rich source of CRs.

3.2 Goals and Contextual Scales

As described in Section 2, comparative sentences explicitly claim an ordering of two items on a relevant scale. However, with more context, a comparative can implicate degrees on a scale as well. If we know the degree of one of the items then we have a fixed range for the other item. If this range contains only one possible degree, then the comparative also sets the degree for this item. In the simplest possible scale in which a comparative could be applied, there are only two degrees, and a comparative indicates the degrees of each item even without more knowledge of the degrees of either item. For example, consider (9) inspired by [DeVault and Stone, 2004]:

- (9) B is drawing and A is giving instructions to B:
A: Draw two squares, one bigger than the other.
B: Done.
A: Now, paint the small square blue and the big one red.
 \rightsquigarrow A thinks that one of the squares can be described as small and the other as big.

The question remains, however, of how the scales are selected? These are not *generalized implicatures* that always arise. Every time we say “Fred is taller than Tom” we don’t mean that “Fred is tall” and “Tom is short”. As discussed in Section 2, conversational implicatures are affected by conversational goals. We believe that this is the path that we have to follow in order to explain comparative implicatures as *particularized implicatures*.

Predicates such as *tall* and *small* are vague in that they can refer to different ranges of a scale and in fact different scales. *Small* refers to size, which as a default we might think of as a continuous scale measure, however, as in the example above, we might prefer to use a simpler scale. This phenomena was already noticed by [Hobbs, To appear] who says:

There is a range of structures we can impose on scales. These map complex scales into simpler scales. For example, in much work in qualitative physics the actual measurement of some parameter may be anything on the real line, but this is mapped into one of three values – positive, zero, and negative.

How are such structures over scales defined? The intuitive idea is that the structure only distinguishes those values that are relevant for the goals of the agent; that is when the attribute takes a particular value, it plays a causal role in deciding whether some agent’s goal can be achieved or not. In the example above, we have only *small* and *big* as relevant values for size, and a comparative in this context will implicate that we are using this binary scale as well as the degrees that each of the items of comparison have.

4 Computational framework

The computational framework we have used for the implementation is the ICT Virtual Human dialogue manager [Traum, 2003; Traum *et al.*, 2008b]. This dialogue manager allows virtual humans to participate in bilateral or multi-party task-oriented conversations in a variety of domains, including team and non-team negotiation. This dialogue manager follows the information state approach [Larsson and Traum, 2000], with a rich set of dialogue acts at different levels. The dialogue manager is embedded within the Soar cognitive architecture [Laird *et al.*, 1987], and decisions about interpreting and producing speech compete with other cognitive and physical operations. When an utterance has been perceived (either from processing of Automatic Speech Recognizer and Natural Language Understanding (NLU) components which present hypotheses of context-independent semantic representations, or messages from self or other agents), the dialogue manager processes these utterances, making decisions about pragmatic information such as the set of speech acts that this utterance performs as well as the meanings of referring expressions. Updates are then performed to the information state on the basis of the recognized acts. The representations of semantic and pragmatic elements for questions and assertions are presented in [Traum, 2003]. The semantic structure derives from the task model used for planning, emotion and other reasoning. (10a) shows an example proposition in this domain, where propositions have an object id, an attribute, and a value (whose type is defined by the attribute). (10b) shows an assertion speech act, with this proposition as content.

- (10) a. ($\langle \text{prop1} \rangle \hat{\text{attribute}} \text{ safety } \hat{\text{object}} \text{ market } \hat{\text{type}} \text{ state } \hat{\text{value}} \text{ no}$)
b. ($\langle \text{da1} \rangle \hat{\text{action}} \text{ assert } \hat{\text{actor}} \text{ doctor } \hat{\text{addressee}} \text{ captain } \hat{\text{content}} \langle \text{prop1} \rangle$)

In the next section, we extend this framework to allow comparative propositions and speech acts arising from implicatures of comparatives.

5 Implementing comparative implicatures

We have added an ability for the ICT dialogue manager to handle comparative constructions and comparative implicatures. Some of the attributes in the task model for a given domain will be scalar, where there is some implied ordering of the possible values. This will not be the case for all attributes, for example, the *location* allows for different possible places that an entity can be, but there is no scale among them. For each scalar attribute, P , in the task model of our domain, we can create a comparative attribute $>_P$ that will compare objects and values according to the designated scale for P . (11a) means that X is higher on the P scale than Y . In our system, when a comparative of the form “ X is P -er than Y ” is uttered, the NLU generates a semantic frame that has the structure shown in (11a). The dialogue manager will then create an assertion speech act as represented in (11b) and then infer the conversational implicatures that arise.

Zeevat, 2007] which applies both to accommodation of content that has been lexically triggered (such as presuppositions) as well as content that has not (such as conversational implicatures).

The after approach: A first approach is to add the inferred assertion I as further input received by the dialogue manager once it finished processing the utterance U . The idea for the implementation is simple: just re-insert the frame of the inferred assertion(s) as an input to the dialogue manager, as if it had been uttered in the dialogue.

At the same time: A second approach is to consider that a single assertion performs multiple speech acts. The implementation of this option in the current computational framework is also quite straightforward because the framework already allows for multiple speech acts associated with an utterance and then the information state is correctly updated.

The before approach: The third approach is to interrupt the processing of the explicit assertion U to update the information state first with the implicature I . Such an implementation requires significant changes to the dialogue manager to interrupt processing on the current assertion and first interpret the inferred speech acts before re-interpreting.

The crucial difference among the three implementations is what content will be available in the information state when the rest of the content has to be interpreted. In the *same time* approach all the updates are independent, the information state does not contain the implicatures, nor the explicit assertion when interpreting any of them. In the *after* approach, U is available in the information state when interpreting I ; by contrast, in the *before* approach I is available when interpreting U .

These differences are relevant when they interact with other parts of the interpretation process, such as reference resolution. Consider for instance the following utterance:

- (15) A: I went to the hospital Saint-Joseph, the British clinic was too far.
A: The doctor gave me some medicine

The utterance “The doctor gave me some medicine” would normally be interpreted as implicating that A saw a doctor in the hospital Saint-Joseph. In order to properly interpret this utterance, it’s important that the implicature is in the context before resolving the referring expression “the doctor”. We leave for future work the interaction with other aspects of interpretation, such as word sense disambiguation.

6 A case-study

We have evaluated the implementation in the context of the SASO-EN scenario [Traum *et al.*, 2008a], where both comparatives and binary scales are

present in the task model. This domain contains over 60 distinct states that the agents consider as relevant for plan and negotiation purposes. There are currently 11 attributes used for this domain, 7 of which are binary scales and 4 of which are non-scalar. Adding the comparative implicature rules from the previous section allows understanding of additional arguments that were not previously dealt with adequately by the system. Consider the following fragment of a dialogue among the Captain (a human agent), the Elder and the Doctor (two virtual agents) about the location of a clinic.

- (16) Captain: Doctor would you be willing to move the clinic downtown?
 Doctor: It is better to keep the clinic here in the marketplace.
 Captain: Well, downtown is safer than the market
 Elder: Why do you think that the market is not safe?

During the interpretation of the comparative uttered by the Captain, the dialogue manager receives the following semantic frame:

- (17) (<prop> ^attribute safer ^object downtown ^type state ^value market)

Then the inferred rules developed in Section 5 are applied and the information state is updated not only with an assertion of (17) but also with the following two assertions:

- (18) (<prop1> ^attribute safety ^object downtown ^type state ^value yes)
 (<prop1> ^attribute safety ^object market ^type state ^value no)

Since these propositions assert the fact that the market is not safe before the Elder generates a response, he can directly address and query the reason for one of these implicatures with “why do you think that the market is not safe?”

In the general case, we might have to reason about the appropriate scale to use for this attribute, but in our domain, only one scale is relevant, so this additional inference is not needed.

Without the comparative implicature rules, neither the elder nor the doctor would recognize that the captain is asserting something about the safety of each of the locations and will not be able to properly assess the argument about desirability of moving the clinic.

7 Discussion

As mentioned in Section 3, the relevant information sources of a dialogue, such as the task model, shed light on how the scales that are relevant for calculating implicatures should be constructed. However, once comparative implicatures are inferred they interact with the structure of the dialogue in relevant ways. When implicated assertions are already in the context, because they have been uttered before, the information state does not need to be updated again with them.

If A says ϕ and implicates ψ but A said ψ before in the dialogue then the context of the dialogue does not have to be updated again with ψ . In this case, we say that the implicature has been *bound* (to use standard terminology in the area) and the information state is not modified by it. The result is a coherent dialogue, a dialogue that gives the intuition of “continuing on the same topic or stressing the same point” such as in the following example:

- (19) Captain: the market is not safe
Captain: downtown is safer than the market

The point can be better illustrated when the two contributions are not made by the same speaker. In this case, A says ψ and B says ϕ which implicates ψ , then ϕ has the effect of accepting ψ and adding it to the common ground. This implements the intuition that, in (20), the elder seems to be supporting the captain in his claim that the market is not safe.

- (20) Captain: the market is not safe
Elder: downtown is safer than the market

Finally, *implicature cancellation* is implemented in a simplistic way in the current framework. When A says ϕ and implicates ψ but $\neg\psi$ is already in the common ground, the implicature is simply ignored.

Our implementation captures the intuition that the relevant values of the properties in a domain are those that are causally related to the agent’s goals. This is used in order to construct or locate appropriate scales and infer useful implicatures that interact with the dialogue structure in different ways. However, not all the predictions achieved by such an implementation are explainable and further refinement is needed. Consider the following dialogue:

- (21) C: Downtown is safer than the market
C: The US base is safer than downtown

If we apply the inference rules developed in Section 5 to this exchange, which implicature gets through will depend on the order in which the utterances in (21) are said. If they are said in the order shown in (21) then the implicature that *downtown is safe* will get through; if they are uttered in the opposite order, then the implicature that *downtown is not safe* will get through. We leave the study of these kinds of interactions to further work.

It is clear that treating conversational implicatures in dialogue is a complex problem, and that it interacts in relevant ways with other key features of dialogue, namely positive and negative evidence of understanding. It is important that the theory of implicatures acknowledges the fact that conversational implicatures are a phenomena that arises in conversation and that then needs to be studied in its environment to be fully understood. This paper starts by motivating this big picture, and then relates it to the semantic theory of comparatives and the pragmatic theory of implicatures in order to address the practical problem of comparative implicatures that arises in an implemented dialogue agent.

References

- [Beaver and Zeevat, 2007] D. Beaver and H. Zeevat. Accommodation. In *The Oxford Handbook of Linguistic Interfaces*, pages 503–539. OUP, 2007.
- [Clark, 1975] H. Clark. Bridging. In *The 1975 Workshop on Theoretical issues in natural language processing*, pages 169–174. ACL, 1975.
- [Clark, 1996] H. Clark. *Using Language*. CUP, 1996.
- [Cresswell, 1976] M. Cresswell. The semantics of degree. In *Montague Grammar*, 1976.
- [DeVault and Stone, 2004] D. DeVault and M. Stone. Interpreting vague utterances in context. In *Proc. of COLING04*, pages 1247–1253, 2004.
- [Grice, 1975] H. Grice. Logic and conversation. In P. Cole and J. L. Morgan, editors, *Syntax and Semantics: Vol. 3: Speech Acts*, pages 41–58. AP, 1975.
- [Hobbs, To appear] J. Hobbs. Discourse and inference. To appear.
- [Horn, 1972] L. Horn. *On the semantic properties of logical operators in english*. PhD thesis, University of California Los Angeles (UCLA), 1972.
- [Klein, 1980] E. Klein. A semantics for positive and comparative adjectives. *Linguistics and Philosophy*, 4:1–45, 1980.
- [Laird *et al.*, 1987] J. Laird, A. Newell, and P. Rosenbloom. SOAR: an architecture for general intelligence. *AI*, 33(1):1–64, September 1987.
- [Larsson and Traum, 2000] S. Larsson and D. Traum. Information state and dialogue management in the TRINDI dialogue move engine toolkit. *Natural Language Engineering*, 6:323–340, September 2000.
- [Lewis, 1979] D. Lewis. Scorekeeping in a language game. *Journal of Philosophical Logic*, 8:339–359, 1979.
- [Rodríguez and Schlangen, 2004] K. Rodríguez and D. Schlangen. Form, intonation and function of clarification requests in german task oriented spoken dialogues. In *CATALOG04*, pages 101–108, 2004.
- [Traum *et al.*, 2008a] D. Traum, S. Marsella, J. Gratch, J. Lee, and A. Hartholt. Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In H. Prendinger, J. Lester, and M. Ishizuka, editors, *IVA*, volume 5208 of *LNCS*, pages 117–130. Springer, 2008.
- [Traum *et al.*, 2008b] D. Traum, W. Swartout, J. Gratch, and S. Marsella. A virtual human dialogue model for non-team interaction. In L. Dybkjaer and W. Minker, editors, *Recent Trends in Discourse and Dialogue*. Springer, 2008.
- [Traum, 2003] D. Traum. Semantics and pragmatics of questions and answers for dialogue agents. In *Proc. of IWCS-5*, pages 380–394, 2003.