Comparatives in Context

Steffen Staab & Udo Hahn

CLIF Computational Linguistics Research Group &
Graduate Programm on "Human and Artificial Intelligence"
Freiburg University
Platz der Alten Synagoge 1, D-79085 Freiburg, Germany
{staab,hahn}@coling.uni-freiburg.de

Abstract

We propose a model of semantic interpretation of comparatives which is based on a mechanism for semantic copying. Besides common phrasal and clausal forms of comparatives, our model also incorporates the analysis of referential and textual phenomena that interact with the interpretation of comparatives, *viz.* metonymies and omitted complements. In order to allow for efficient processing, guidance from syntactic, semantic, contextual and world knowledge sources is supplied.

Introduction

Comparatives form a class of linguistic phenomena that, up to now, has been given far less attention than, e.g., anaphoric reference, quantification or coordination. However, for expository texts in particular many of the crucial, "interesting" propositions are expressed by comparative utterances, e.g., differences between competing products, changing conditions of patients under medical treatment, etc. Despite their relevance for real-world text understanding, comparatives have been almost exclusively studied from a purely linguistic perspective. Though usually embedded in a Montague-style description framework (von Stechow, 1984; Lerner & Pinkal, 1992), syntactic criteria prevail, while semantic information is brought into play only to a small extent, if at all. This observation also holds for computational models of comparative interpretation as, e.g., proposed by Friedman (1989), Rayner & Banks (1990) and Alshawi & Pulman (1992). As a consequence, a large number of implausible readings are generated.[1] Even worse, syntax-oriented approaches systematically fail to account for interpretations that depend entirely on semantic or conceptual criteria.

In this paper, we propose to augment the methodology of comparative interpretation in three ways. First, in order to constrain emerging ambiguities, semantic and conceptual criteria are formulated to complement syntactic criteria.[2] Second, with the availability of conceptual criteria we provide analyses for comparatives which rely solely on semantic or conceptual criteria and have <u>no</u> syntactic triggering condition at all. Third, we substantially enhance the range of possible interpretations by encompassing the text-level of analysis. Actually, comparatives are often intertwined with referential and textual phenomena in a way that ultimately changes their interpretation. The influence of the context on determining which entities are to be contrasted can be illustrated by the following constructions:

- *Omitted complements* (cf. (1)) constitute a hard problem for interpretation mechanisms, since one of the correlates to which the comparison actually refers has been zeroed:

(1) *The LED line of the printer 810 is clearly thicker [than the one of the 410].*[3]

- *Metonymies* are another kind of frequently occurring referential phenomenon. They extend the set of possible correlates, because they allow concepts to stand for related concepts. For instance, in sentence (2) the company name "HP" indirectly denotes a printer manufactured by HP, which was mentioned in the foregoing sentence(s).

(2) *The Oki-X11 is faster than the [printer manufactured by] HP.*

The overall goal of this paper is to provide an efficient, knowledge-based interpretation mechanism for comparatives that covers well-known phrasal and clausal phenomena but can also be extended to handle the referential and textual phenomena just mentioned. Likewise, this mechanism builds on the already established common ground of previous research on the formal semantics of comparatives. In order to accomplish this task, a basic model of comparative interpretation that abstracts from textual influences is presented in the next section. Subsequently, this core mechanism will be slightly extended to incorporate reference relations and textual structures, as far as they are relevant for the interpretation of comparatives. Throughout these sections, we illustrate our approach with several sample analyses and then turn to the results of an empirical evaluation.

[1]Rayner & Banks (1990, p. 102): "...our treatment of phrasal contrastives also appears to allow more readings than really exist. It would certainly be desirable to find rules to eliminate these, or at least heuristics to say which readings can be regarded as unlikely."

[2]See also Banks & Rayner (1988) who mention the use of type lists to disambiguate syntactic ambiguities.

[3]The brackets "[...]" mark passages that do not explicitly occur in the text but must be inferred by the reader from the context of the discourse.

Basic Model of Comparative Interpretation

In this section, we outline our model of semantic copying as a new methodology for comparative interpretation. It incorporates syntactic restrictions but more heavily relies upon conceptual criteria which, when applied to candidate correlates, select the most plausible pair(s). A standard embedding within a Davidsonian-style semantic representation is described. We then turn to the algorithmic description of the interpretation process and give a worked-out example.

Comparative Interpretation as Semantic Copying

Syntactic Restrictions and Conceptual Specificity Constraints. Current approaches to comparative interpretation usually depend on syntactic triggering conditions (Friedman, 1989; Rayner & Banks, 1990). However, a purely syntactic reconstruction might fail, e.g., if the contrasted entities are realized by phrases with different syntactic forms (cf. (3)[4]), or if one of them makes reference to an essentially semantic feature (cf. (4)).

(3) *Die LED-Zeile des PrinterA ist dicker als beim PrinterB.*
(The LED line – {GENITIVE: *of printerA*} *– is thicker than –* {PP: *in the case of printerB*}*.)*

(4) *Hans is richer than last year.*

We therefore propose a multi-layered interpretation mode. First, pairs of contrasted objects — we call them *quasi-coordinated* — are determined by syntactic criteria. In the next step, if the algorithm fails to identify a syntactically valid quasi-coordination that is also conceptually plausible (such as in (3) and (4)), the complement is contrasted with semantic entities that are available in the matrix clause of the comparative. Based on a description logics framework (cf. Woods & Schmolze (1992) for a survey), we define the conceptual plausibility for the quasi-coordination of two concepts by the predicate "PlausiblePair" (cf. Table 1). This predicate evaluates to TRUE *iff* their least common superconcept (denoted by "lcs") is not too general. For instance, in the example (4), "*Hans*" as a PERSON and "*last year*" as a TIMEINTERVAL with TOP as their least common superconcept render the quasi-coordination between them implausible, while the one between TIMEINTERVAL and the implicitly available tense feature PRESENT yields a conceptually plausible quasi-coordination. In addition, plausible quasi-coordinations are preferentially ordered according to their conceptual similarity as defined by the predicate $\prec_{\mathsf{simil}}$ (cf. Table 2). Therefore, the quasi-coordination of "*printerA*" and "*printerB*" in (3) will be preferred to that of "*printerB*" with "*LED line*".

$\mathsf{PlausiblePair}(a,b) :\Leftrightarrow \mathrm{lcs}(a,b) \neq \textsc{Top}$, where $\forall x\ \mathsf{Concept}(x) \rightarrow \mathsf{Isa}(x, \textsc{Top})$ and $\mathrm{lcs}(x,y)$ denotes the least common superconcept[5] of x and y.

Table 1: The Plausibility Predicate PlausiblePair

$(a,b) \prec_{\mathsf{simil}} (c,d)\ :\Leftrightarrow$ $\mathsf{Isa}(\mathrm{lcs}(a,b), \mathrm{lcs}(c,d)) \wedge \neg\mathsf{Isa}(\mathrm{lcs}(c,d), \mathrm{lcs}(a,b))$

Table 2: The Similarity Ordering $\prec_{\mathsf{simil}}$

Semantic Reconstruction. The logical form of the matrix clause (5) where the comparative is not yet interpreted is described in a common Neo-Davidsonian style by the expression (6). The comparative is assumed to carry a comparison operator that is still missing the description for the second degree[6] d'.

(5) *Tom is taller than Sue.*

(6) $\lambda D' \exists e, d[\, ToBe(e) \wedge Patient(e, tom) \wedge Pred(e, d) \wedge Height(d) \wedge \forall d'[D'(d') \rightarrow d > d']\,]$

Next, a semantic reconstruction process is triggered. It searches for a chain of semantic relations in the matrix clause from the quasi-coordinated head to the degree that is described by the comparative and uses this information to recover the missing semantic elements in the complement. Hence, and because no type restrictions of the semantic predicates are violated, (7) is reconstructed as the semantics of the complement part:

(7) $\lambda d' \exists e'[\, ToBe(e') \wedge Patient(e', sue) \wedge Pred(e', d') \wedge Height(d')\,]$

Consequently, (7) is applied to (6), which results in (8).

(8) $\exists e, d[\ ToBe(e) \ \wedge\ Patient(e, tom) \ \wedge\ Pred(e, d) \ \wedge\ Height(d) \ \wedge$
$\forall d'[\, \exists e'[\, ToBe(e') \wedge Patient(e', sue) \wedge Pred(e', d') \wedge Height(d')\,] \ \rightarrow\ d > d'\,]\,]$

The recovery process is completed by several substitutions, the notation of which is given in example (9), where the quasi-coordinated matrix element is substituted by the corresponding complement object:

(9) $[\, ToBe(e) \ \wedge\ Patient(e, tom) \ \wedge\ Pred(e, d) \ \wedge \ldots] \mid \{tom/sue\} =$
$[\, ToBe(e) \wedge Patient(e, sue) \wedge Pred(e, d) \wedge \ldots]$

To keep the following description as simple as possible and to keep focused, we assume that all variables in a given sentence are existentially quantified and will be reconstructed as new variables in the complement clause, if they are reconstructed at all.[7]

[4]The English examples (in parentheses) are phrase-to-phrase translations of the German ones.

[5]Given multi-hierarchies, the uniqueness of the "lcs" can be insured by computing the conjunction of the different least common superconcepts.

[6]The decision for the logical structure of the comparative ($\forall d'[D'(d') \rightarrow d > d']$) has been made in line with considerations raised, e.g., by Pinkal (1989). This becomes crucial for the proper treatment of negative polarity items and modal contexts, which are not an issue in this paper. Nevertheless, they can be accounted for in our approach similar to proposals made by von Stechow (1984) and Pinkal (1989), who use quantifier scoping generation to produce different readings.

[7]For the problems that are connected with strict versus sloppy identity and scope parallelism in ellipsis reconstruction, cf., e.g., Crouch (1995) who presents a substitutional approach to ellipsis handling that goes well with the mechanism described here.

The Core Algorithm

The core algorithm for the semantic interpretation of comparatives can now be specified as follows:

1. *Input:*

- *Syntax:* c is the semantic complement object, which is syntactically determined, and $S_s := \{(a_1, c) \dots (a_n, c)\}$ is the set of pairs of contrasted semantic objects[8] that are syntactically allowed.
- *Ontology:* provides the predicate "PlausiblePair" and the partial order "$\prec_{\mathsf{simil}}$" as described in Tables 1 and 2, respectively.
- *Semantics:* $\lambda D' \exists d, \bar{o}[\, D(d, \bar{o}) \wedge \forall d'[D'(d') \rightarrow d > d']\,]$ is the standard semantics of the matrix clause, where $D(d, \bar{o})$ describes a conjunction (which we often view as a set of assertions) of unary and binary relations ($r_i(x_i)$ and $r_j(x_j, x_{j+1})$, respectively) on a number of objects $\bar{o} = (o_1, \dots, o_n)$, $n \in \mathbb{N}_0$, and the degree d.
$S_p := \{\, (b_i, c) \mid \exists r_i(b_i, b_{i+1})[r_i(b_i, b_{i+1}) \in D(d, \bar{o}) \wedge (b_i, c) \notin S_s] \,\}$[9] is the cross product of all semantic objects[10] b_i in the matrix clause and the semantic complement object c, but without the pairs in S_s.

2. *Choose and Order Correlates:*

- Collect all possible and reasonable quasi-coordinations: $S := \{\, (a_i, c) \mid (a_i, c) \in S_p \cup S_s \wedge \mathsf{PlausiblePair}(a_i, c)\}$
- A partial preference order $\prec_{\mathsf{pref}}$ on S is given by the transitive closure of "$\prec_{\mathsf{simil}}$" (denoted by " * ") restricted to the two subsets $S'_s := S_s \cap S$ and $S'_p := S_p \cap S$, respectively, and the condition that syntactically given quasi-coordinations are preferred over semantically given ones:

$$\prec_{\mathsf{pref}}|_S := (\prec_{\mathsf{simil}}|_{S'_p} \cup \prec_{\mathsf{simil}}|_{S'_s} \cup \{((a_i, c), (b_j, c)) | (a_i, c) \in S'_s \wedge (b_j, c) \in S'_p\})^*$$

3. *Handle the Quasi-Coordinations $(a_i, c) \in S$ with the Priority Ordering Given by "$\prec_{\mathsf{pref}}$":*

- *Select Relevant Relations:*
  - Search for shortest path $D_1 \subseteq D$ from matrix correlate a_i to degree d:
  $D_1 := \min_{|P|}\{\, P \mid P \in 2^D \wedge \mathsf{Path}(P)\}$, where[11]
  $\mathsf{Path}(P) :\Leftrightarrow \forall j \in [1, n] : r_j(o_j, o_{j+1}) \in P \wedge o_1 = a_i \wedge o_{n+1} = d$,
  where n is the cardinality of P
  - Add ASP[12] fillers recursively:
  $D_2 := \mathrm{AddF}^*(D_1)$, where AddF is defined by:
  $\mathrm{AddF}(P) := P \cup \{\, r_k(o_k, o_{k+1}) \mid r_k(o_k, o_{k+1}) \in D \wedge \exists r_l(o_l, o_{l+1}) \in P[\, \mathsf{Isa}(o_l, \mathsf{ASP}) \wedge o_l = o_k] \,\}$
- *Coordinate Relations:*
  - Assume D_C to be the semantics of the unreconstructed complement.[13]
  - Substitute correlate: $D'_1 := (\, D_2 \mid \{a_i/c\}\,) \mid \{d/d'\}$
  - Substitute other complement objects:
  $D'_2 := \mathrm{Sub}^*(D'_1)$, and $\mathrm{Sub}(P) := P \mid \{\, o_{j+1}/o'_{j+1} : \exists r_j, o_j[\, r_j(o_j, o_{j+1}) \in P \wedge r_j(o_j, o'_{j+1}) \in D_C\,]\,\}$
  - Copy existentially quantified objects:
  $D'_3 := \mathrm{CopyE}^*(D'_2) \cup \bigcup_{j:o_j \in E}\{\mathsf{Type}_j(o'_j)\}$, where $\mathrm{CopyE}(P) := P \mid \{\, o_j/o'_j : o_j \in E\,\}$, and E is the set of all variables in D'_2, which are not the result of a previous substitution and not contextually bound; Type_j is the type of o_j in D and $\{o'_j\}$ are new variables.
  - Combine reconstructed and original complement semantics: $D' := D_C \cup D'_3$
- *Application:*
  - Interpret the set D' as a logical formula, where all unbound variables are existentially quantified:
  $D'(d') \equiv \lambda d' \exists \bar{o}'[r_{j+1}(o'_j, o'_{j+1}) \wedge \dots]$.
  - Apply D' to $\lambda D' \exists d, \bar{o}[D(d, \bar{o}) \wedge \forall d'[D'(d') \rightarrow d > d']]$

An Example of Semantic Interpretation

We now supply the step-wise interpretation of example sentence (3), *"Die LED-Zeile des PrinterA ist dicker als beim PrinterB."* (The LED line – {GENITIVE: of printerA} – is thicker than – {PP: in the case of printerB}):

1. *Input:*

- *Syntax:* $c = printerB$, $S_s = \{\}$.
- *Ontology:* A knowledge base covering the information technology domain with currently 550 concepts and 350 relations.
- *Semantics:* The semantics of the matrix clause is $\lambda D' \exists e, l, d[\, HasPhysPart(printerA, l) \wedge Led\text{-}Line(l) \wedge ToBe(e) \wedge Patient(e, l) \wedge Tense(e, present) \wedge Pred(e, d) \wedge Thickness(d) \wedge \forall d'[D'(d') \rightarrow d > d']\,]$.
$S_p = \{(printerA, printerB), (l, printerB), (e, printerB), (present, printerB), (d, printerB)\}$ contains all the possible quasi-coordinations.

2. *Choose and Order Correlates:*

- $S = \{(printerA, printerB), (l, printerB)\}$
- $(printerA, printerB) \prec_{\mathsf{pref}} (l, printerB)$

3. *Handle $(printerA, printerB)$ First, Because It Is the Preferred Pair of Correlates:*

- *Select Relevant Relations:*
  - $D_1 = \{HasPhysPart(printerA, l), Patient(e, l), Pred(e, d)\}$[9]
  - $D_2 = D_1 \cup \{Tense(e, present)\}$
- *Coordinate Relations:*
  - The unreconstructed complement semantics is: $D_C = \{\}$

---

[8] In our implementation of a dependency grammar parser (cf. Neuhaus & Hahn (1996)) this set is computed by a tree walk.

[9] Throughout this paper, we assume that for each $r(a, b)$ in a set M its inverse $r^{-1}(b, a)$ is in M, too. This convention avoids abundant verbosity in the formal descriptions.

[10] This includes semantic features, such as TENSE and MODALITY, which are required, e.g., for the proper interpretation of sentence (4).

[11] The existence of such a path is obvious. The uniqueness is given in our model through the (almost) one-to-one correspondence between syntactic dependency relations and semantic relations. In other schemes it can be achieved by search restrictions.

[12] ASP denotes the concept ACTION-OR-STATE-OR-PROCESS, which incorporates concepts that are lexicalized as verbs.

[13] Often, D_C will be empty at the start. It will be nonempty, if the complement object is complex, e.g., in a clausal comparison.

- Substitute complement correlate: $D_1' = \{HasPhysPart(printerB, l), Patient(e, l), Pred(e, d'), Tense(e, present)\}$
- There are no other complement objects: $D_2' = D_1'$
- Substitute old variables and add types: $D_3' = \{HasPhysPart(printerB, l'), Patient(e', l'), Pred(e', d'), Tense(e', present)\} \cup \{Led\text{-}Line(l'), ToBe(e'), Thickness(d')\}$

• *Application:*

- $D' = D_3' \cup \{\} \equiv \lambda d' \exists e', l' [\, HasPhysPart(printerB, l') \wedge Led\text{-}Line(l') \wedge ToBe(e') \wedge Patient(e', l') \wedge Pred(e', d') \wedge Thickness(d') \wedge Tense(e', present)]$
- Now the combined semantics amounts to: $\exists e, l, d[\; HasPhysPart(printerA, l) \wedge Led\text{-}Line(l) \wedge ToBe(e) \wedge Patient(e, l) \wedge Tense(e, present) \wedge Pred(e, d) \wedge Thickness(d) \wedge \forall d'[\, \exists e', l' [\, HasPhysPart(printerB, l') \wedge Led\text{-}Line(l') \wedge ToBe(e') \wedge Patient(e', l') \wedge Tense(e', present) \wedge Pred(e', d') \wedge Thickness(d')\,] \rightarrow d > d' \,] \,]$.

A complete KL-ONE-style representation of this result is depicted in Fig. 1 (note that DEGREE.6 actually stands for the concept of *Thickness* in the expression above).

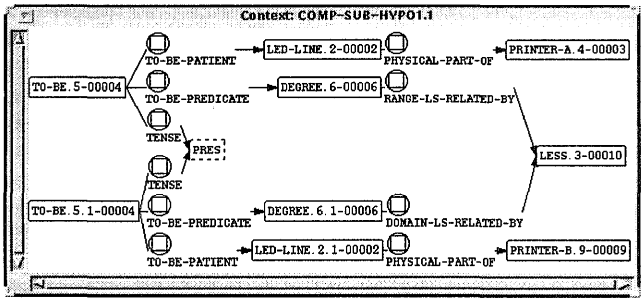

Figure 1: The Conceptual Representation of a Comparative

Referential and Textual Phenomena

Comparatives often cannot be viewed in isolation from other linguistic phenomena, reference relations in particular. The interaction of anaphora with comparatives becomes evident in the following two examples:

1. *Relative-clause-like* constructs (cf. (10) and (11)) require an anaphora-like reconstruction mechanism, as proposed by Pinkham (1985) and Lerner & Pinkal (1992).
   (10) *More guests than we invited [guests] visited us.*
   (11) *More guests than we expected [X guests to visit us] visited us.*
2. The difference between a narrow-reading construction (12), which derives from the standard interpretation proposed here, and the special *wide reading* in (13) results from the enforcement of reference resolution: the complement is either forced to be referential or to be embedded in a sentential construct.
   (12) *A taller man than John visited us.*
   (13) *More men [visited us] than women visited us.*

The examples (10)–(13), though vividly discussed in the semantic literature, occur far less often in our text corpus than referential and textual phenomena such as *comparatives without complement* (as in (1) or (14)) and *comparatives with metonymies* (as in (2) or (15)). Hence, we will subsequently concentrate on their analysis and incorporate additional conceptual and contextual criteria in order to extend the core model from the previous section. Harder linguistic constructs such as (10)–(13) can also be accounted for by just adding a few extra constraints to this core. Difficulties that arise for the interpretation of relative-clause-like and special wide-reading constructs, as well as omitted complements, do not multiply, because their analyses are triggered by mutually exclusive conditions.

Comparatives with Omitted Complements

A large number of omitted complements are essentially anaphorical and can therefore be handled with the help of an anaphor resolution mechanism based on the centering model (Grosz et al., 1995). In this model, for each utterance U_i an ordered list (the forward-looking centers $C_f(U_i)$) of the most plausible antecedents for referential expressions in the next utterance is computed. The possible antecedents in the C_f that have not been consumed by an anaphoric resolution process in the comparative matrix clause are considered to be possible complements for the comparative.

For omitted complements, the order in $C_f := C_f(U_{i-1})$ plays a crucial role, even though the conceptual similarity of two elements that are to be contrasted is still more important. Therefore, the specification of comparatives with omitted complements contains the following assumptions:

1. *Input:*

• *Text Grammar:* $C_f :=$ [center$_1 \prec_C \ldots \prec_C$ center$_n$] with the referential preference ordering "$\prec_C$"

• *Ontology:* yields the ordering "$\prec_{\mathsf{simil}}$" and the predicate[14] "PlausiblePair"

• *Semantics:* $\lambda D' \exists d, \bar{o} [\, D(d, \bar{o}) \wedge \forall d' [D'(d') \rightarrow d > d']]$

2. *Choose and Order Correlates:*

• Collect all possible and reasonable quasi-coordinations: $S := \{\, (a_i, b_i) \mid \exists r_i(a_i, a_{i+1}) \in D(d, \bar{o}) \wedge b_i \in C_f \wedge \mathsf{PlausiblePair}(a_i, b_i)\}$

• A partial preference order "$\prec_{\mathsf{pref}}$" on S is given by the partial order "$\prec_{\mathsf{simil}}$", which is further constrained by "$\prec_C$": $\prec_{\mathsf{pref}} := \prec_{\mathsf{simil}} \cup \{((a_i, b_i), (a_j, b_j)) \mid (a_i, b_i) \in S \wedge (a_j, b_j) \in S \wedge \neg((a_j, b_j) \prec_{\mathsf{simil}} (a_i, b_i)) \wedge b_i \prec_C b_j\}$

3. *Proceed as in Step 3 of the Core Comparative Interpretation Procedure.*

[14]Lacking complements, as a matter of fact, open up a wider range of semantic interpretations than constructions in which complements are lexicalized. This range must then be constrained by requiring permitted correlates to converge on the level of a conceptually plausible common denominator. We therefore restrict comparisons to be made only at the level of *basic categories* and concepts subsumed by them. Formally defining this level, however, is notoriously difficult (Lassaline et al., 1992). But we have some evidence that simple heuristics seem to be sufficient.

An Example for Omitted Complements

The sample text (14) contains a simple real-world example for a comparative without a complement in the cotext (the corresponding ordering of centers is stated as $C_f(U_{(i-1)})$):

(14) $U_{(i-1)}$: *The spaceball is as inconvenient as such a trackpoint.*
$C_f(U_{(i-1)})$ = [SPACEBALL $\prec_C$ TRACKPOINT]
U_i : *With a common mouse, one may move faster from one corner of the screen to the other [than by using the spaceball].*

The interpretation process proceeds as follows:

1. *Input:*

- *Text Grammar:* SPACEBALL $(= s') \prec_C$ TRACKPOINT $(= t')$
- *Ontology:* Knowledge base on information technology
- *Semantics:* $\lambda D' \exists e, m, p, r_1, s, r_2, d[\text{Move}(e) \wedge \text{Instr}(e,m) \wedge \text{Mouse}(m) \wedge \text{Agent}(e,p) \wedge \text{Person}(p) \wedge \text{StartLoc}(e,r_1) \wedge \text{Region}(r_1) \wedge \text{On}(s,r_1) \wedge \text{Screen}(s) \wedge \text{DestLoc}(e,r_2) \wedge \text{Region}(r_2) \wedge \text{On}(s,r_2) \wedge \text{Vel}(e,d) \wedge \text{Velocity}(d) \wedge \forall d'[D'(d') \rightarrow d > d']]$

2. *Choose and Order Correlates:*

- Collect all possible and reasonable quasi-coordinations (NB: all other pairs are filtered out by conceptual restrictions): $S = \{(m,t'), (m,s'), (p,t'), (p,s'), (s,t'), (s,s')\}$
- Determine "$\prec_{\text{pref}}$": The least common superconcept of the latter four pairs is PHYSICALOBJECT and HARDWARE, respectively, and, therefore, they are less similar than the first two (POINTINGDEVICE). (m,s') is preferred to (m,t') because of the ordering on $C_f(U_{(i-1)})$.

3. *Handle the preferred* (m,s').

Metonymies in the Complement

As was shown with example (2), not even complete syntactic and semantic information is always sufficient to interpret a comparative in the correct way. While metonymies in the matrix clause can be handled by standard metonymy resolution mechanisms (cf. Fass (1991)), metonymies in the complement cannot be treated along the same line. This is due to the fact that the common premiss, *viz.* the violation of a selection restriction, needs a semantic embedding first, while the comparative interpretation needs the *resolved* concept to establish the semantic relations for the complement. As a way out of this dilemma we suggest to extend the mechanism that determines the quasi-coordinated entities.

First, we have to define the predicate PMF (Potential-MetonymyFor; cf. Table 3), which evaluates to TRUE *iff* b is a possible metonymy for a, and, simultaneously, has a given type, C.

$$\text{PMF}(a,b,C) :\Leftrightarrow \exists r \in \text{MR} : r(a,b) \wedge \text{type}(b) = C,$$
where MR denotes the set of metonymic relations {has-physical-part, physical-part-of, produces, ...}

Table 3: The Metonymy Predicate PMF

We may now reformulate the second step of the core mechanism to account for metonymies, too.

2. *Choose and Order Correlates:*

- Collect all pairs (a_i, c') such that c' is a potential metonymy for the complement object c, and such that c' has the same type as the syntactically contrasted element a_i from the matrix clause:
$S_m := \{ (a_i, c') \mid (a_i, c) \in S_s \wedge \text{PMF}(c, c', \text{type}(a_i)) \}$
- Exclude implausible pairs from further consideration:
$S := \{ (a_i, c) \mid (a_i, c) \in (S_s \cup S_m \cup S_p) \wedge \text{PlausiblePair}(a_i, c) \}$
- $S'_s := S_s \cap S$, $S'_m := S_m \cap S$, and $S'_p := S_p \cap S$
- Conceptually plausible and syntactically available quasi-coordinations are preferred over metonymic ones, which themselves are preferred over conceptually plausible, but only semantically available pairs:
$$\prec_{\text{pref}}|_S := (\prec_{\text{simil}}|_{S'_p} \cup \prec_{\text{simil}}|_{S'_s} \cup \{ ((a_i,c),(b_j,c)) \mid ((a_i,c) \in S'_s \wedge (b_j,c) \in (S'_m \cup S'_p)) \vee ((a_i,c) \in S'_m \wedge (b_j,c) \in S'_p) \})^*$$

This extension of the core algorithm for comparative interpretation with metonymic complements has also been integrated into the mechanisms which handle comparatives with omitted complements. Due to its modular design the scaling up to the level of textual analysis of comparatives was easy to achieve. As an example which contains a *disk*-for-*disk drive* metonymy and the simultaneous occurrence of an omitted complement consider the text fragment (15):

(15) *Very-high-density disks offer the advantage that... Disk drives with more capacity [than these very-high-density disks] will be available soon.*

Theoretical and Empirical Coverage

Table 4 summarizes the types of comparatives that are covered by the core model and its extensions. While the first six types characterize fairly standard phenomena (cf. Banks & Rayner (1988) or Friedman (1989)), the last four, as well as the difference between wide and narrow readings, have not been adequately covered before.[15]

Predicative	*John is taller than Mary.*
Adverbial	*Mary runs faster than John.*
Attributive	*John bought better books than Mary.*
Clausal	*Mary reads more than John writes.*
Rel.-clause-like	*More guests than we invited visited us.*
Between PPs	*John spent more money in Paris than in Rio.*
Narrow reading	*A taller man than John visited us.*
Wide reading	*More men than women visited us.*
Nested contrastives	*Mary's car has a wider bumper than John's.*
Metonymic	*The Oki-X11 is faster than the HP.*
With implicit feature	*Mary is richer than last year.*
Omitted complement	*One moves faster with a mouse.*

Table 4: Types of Comparatives Covered by the Model

[15] Alshawi & Pulman (1992) give a rough sketch about how their comparative ellipsis resolution mechanism might deal with omitted complements. They fail, however, to make any proposals how such a schema can be put into work in realistic environments.

We also performed an empirical study on 32 real-world texts (composed of approximately 30,000 words) from the information technology domain. In this sample, a total of 160 comparatives occurred, 13 of which were absolute comparatives (such as "after a *longer* pause the device switches into a power-off mode"). The semantics of these is more closely related to positive adjectives and quantifiers than to the relative comparatives focused on in this paper.

relative comparatives	147	100%	
with complement	57	39%	100%
analyzed	52		91%
without complement	90	61%	100%
analyzed	43		48%
metonymy as complement	14	9.5%	

Table 5: Empirical Data

Table 5 briefly summarizes our data. Even though the sample size is far too small to generalize about the scope of our approach, the following tendencies seem valid. The rate of correct analyses for comparatives with complement is quite high (91%). The lower one for those with omitted complements (48%) is biased by the fact that a considerable number of them (approx. 25%) are not anaphorical *per se*. By this we mean, e.g., references to pragmatic implications from antonymic pairs such as *"with"* vs. *"without"* in sentences like "With this optical device the disk drive stores data more efficiently." Also, the metonymy mechanism (covering cases of comparatives with and without complements) turned out to yield a high bonus in our evaluation (about 10%). Note that none of the competing approaches accounts for textual and referential phenomena at all. Hence, the model we propose raises the rate of comparatives being covered by more than 66% (52 + 43) as opposed to 57 comparatives that constitute the upper limit of other approaches, even granting them a 100% success rate.

Conclusion

In this paper, a conceptually constrained mechanism for comparative interpretation has been presented. It builds on parallel semantic relations between the matrix clause and the complement of the comparative resulting from a semantic copying process. In particular, domain-knowledge-based heuristics guide the analytic mechanism. This does not, of course, preclude the consideration of more refined syntactic criteria, e.g., concerning the relation of comparatives and quantifier scope (Lerner & Pinkal, 1992).

This core mechanism was then extended to include frequently occurring referential and textual phenomena, *viz.* comparatives with omitted complements and comparatives with metonymies. This coupling was recognized as an important problem several years ago (Friedman, 1989; Rayner & Banks, 1990) and has remained unsolved so far. Our solution extends standard model-theoretic methodologies from theoretical linguistics by contextual constructs (in terms of centering), conceptual criteria and taxonomic inference mechanisms from the field of AI.

An empirical study revealed that while the core mechanism is sound, the extension to comparatives without complements significantly increases the coverage of relative comparatives. There remains, however, a subclass of comparatives with omitted complements that should be related to pragmatic rather than to conceptual inferences.

Acknowledgments. S. Staab is a member of the Graduate Program on *Human and Machine Intelligence* at Freiburg University, which is funded by *DFG*. The work of U. Hahn is partially funded by a grant from *DFG* under the account *Ha 2097/5-1*. We thank our colleages, in particular Joe Bush, Katja Markert and Michael Strube, who read versions of this paper.

References

Alshawi, Hiyan & Stephen G. Pulman (1992). Ellipsis, comparatives, and generation. In Hiyan Alshawi (Ed.), *The Core Language Engine*, pp. 251–274. Cambridge, MA: MIT Press.

Banks, Amelie & Rayner Manny (1988). Comparatives in logic grammar: Two viewpoints. In Veronica Dahl & Patrick Saint-Dizier (Eds.), *Natural Language Understanding and Logic Programming, II*, pp. 131–137. Amsterdam: North-Holland.

Crouch, Richard (1995). Ellipsis and quantification: A substitutional approach. In *Proc. of the 7th Conference of the European Chapter of the ACL*, pp. 229–236.

Fass, Dan (1991). met*: A method for discriminating metonymy and metaphor by computer. *Computational Linguistics*, 17(1):49–90.

Friedman, Carol (1989). A general computational treatment of the comparative. In *Proceedings of the 27th Annual Meeting of the ACL*, pp. 161–168.

Grosz, Barbara J., Aravind K. Joshi & Scott Weinstein (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.

Lassaline, Mary E., Edward J. Wisniewski & Douglas L. Medin (1992). Basic levels in artificial and natural categories. In Barbara Burns (Ed.), *Percepts, Concepts, and Categories*, pp. 327–378. Amsterdam: Elsevier.

Lerner, Jean-Yves & Manfred Pinkal (1992). Comparatives and nested quantification. In M. Stokhof & P. Dekker (Eds.), *Proc. of the 8th Amsterdam Colloquium.*

Neuhaus, Peter & Udo Hahn (1996). Restricted parallelism in object-oriented lexical parsing. In *COLING '96: Proceedings of the 16th International Conference on Computational Linguistics*, Vol. 1, pp. 502–507.

Pinkal, Manfred (1989). Die Semantik von Satzkomparativen. *Zeitschrift für Sprachwissenschaft*, 8:206–256.

Pinkham, J. (1985). *The Formation of Comparative Clauses in French and English.* New York: Garland Press

Rayner, Manny & Amelie Banks (1990). An implementable semantics for comparative constructions. *Computational Linguistics*, 16(2):86–112.

von Stechow, Arnim (1984). Comparing semantic theories of comparison. *Journal of Semantics*, 3:1–77.

Woods, William A. & James G. Schmolze (1992). The KL-ONE family. *Computers & Mathematics with Applications*, 23(2-5):133–177.