


Bridging Textual Ellipses

Udo Hahn

Michael Strube

Katja Markert

Freiburg University

 Computational Linguistics Lab

Europaplatz 1, D-79085 Freiburg, Germany

{hahn, strube, markert}@coling.uni-freiburg.de

Abstract

We present a hybrid text understanding methodology for the resolution of textual ellipsis. It integrates language-independent conceptual criteria and language-dependent functional constraints. The methodological framework for text ellipsis resolution is the centering model that has been adapted to constraints reflecting the functional information structure within utterances, i.e., the distinction between context-bound and unbound discourse elements.

1 Introduction

Text phenomena, e.g., textual forms of ellipsis and anaphora are a challenging issue for the design of parsers for text understanding systems, since lacking recognition facilities either result in referentially incoherent or invalid text knowledge representations. At the conceptual level, textual ellipsis (also called functional anaphora) relates a quasi-anaphoric expression to its extrasentential antecedent by conceptual attributes (or roles) associated with that antecedent (see, e.g., the relation between “*Ladezeit*” (*charge time*) and “*Akku*” (*accumulator*) in (1c) and (1b)). Thus, it complements the phenomenon of nominal anaphora, where an anaphoric expression is related to its antecedent in terms of conceptual generalization (as, e.g., “*Rechner*” (*computer*) refers to “*316LT*”, a particular notebook, in (1b) and (1a)). The resolution of text-level nominal (and pronominal) anaphora contributes to the construction of referentially valid text knowledge bases, while the resolution of textual ellipsis yields referentially coherent text knowledge bases. Both phenomena tend to interact, as evidenced by the examples below. “*Akku*” (*accumulator*) in (1b) is a nominal anaphor referring to “*Nickel-Metall-Hydride-Akku*” (*nickel-metal-hydride accumulator*) in (1a), which, when resolved, provides the proper referent for relating “*Ladezeit*” (*charge time*) in (1c) to it.

- (1) a. Der *316LT* wird mit einem *Nickel-Metall-Hydride-Akku* bestückt.
(The *316LT* is – with a *nickel-metal-hydride accumulator* – equipped.)
- b. Der *Rechner* wird durch diesen neuartigen *Akku* für 4 Stunden mit Strom versorgt.
(The *computer* is – because of this new type of *accumulator* – for 4 hours – with power – provided.)
- c. Darüberhinaus ist die *Ladezeit* mit 1,5 Stunden sehr kurz.
(Also – is – the *charge time* of 1.5 hours quite short.)

In the case of textual ellipsis, the missing conceptual link between two discourse elements occurring in adjacent utterances must be inferred in order to establish the local coherence of the discourse (for an early statement of that idea, cf. Clark (1975)). In sentence (1c) the information is missing that “*Ladezeit*” (*charge time*) links up with “*Akku*” (*accumulator*). This relation can only be made explicit if conceptual knowledge about the domain, *viz.* the relation *charge-time-of* between the concepts CHARGE-TIME and ACCUMULATOR, is available.

The solution we propose is embedded within the centering model (Grosz et al., 1995). In this approach, discourse entities linking successive utterances in a discourse segment are organized in terms of centers. Local coherence in discourse is established whenever a center element of the previous utterance is associated with an expression that has a valid semantic interpretation (i.e., is *realized*) in the following utterance. Textual ellipsis has only been given insufficient treatment within the centering model in terms of rather sketchy realization conditions as opposed to the more elaborated constraints for (pro)nominal anaphora. The heuristics we propose include language-independent conceptual criteria and language-dependent information structure constraints reflecting the context-boundedness or unboundedness of discourse elements within the considered utterances.

2 Constraints on Conceptual Linkage

This section provides a highly condensed exposition of the conceptual constraints underlying the resolution of textual ellipses. A much more detailed presentation is given by Hahn et al. (1996). The constraints we posit require a domain knowledge base to consist of concepts and conceptual roles linking these concepts. Concepts and roles are hierarchically ordered by subsumption (a terminological knowledge representation framework is assumed; cf. Woods & Schmolze (1992)).

In order to determine suitable conceptual links between an antecedent and an elliptic expression, we distinguish two modes of constraining the linkage between concepts via conceptual roles. In the process of *path finding* an extensive unidirectional search is performed in the domain knowledge base and formal *well-formedness* conditions holding for paths between two concepts are considered, *viz.* complete connectivity (compatibility of domains and ranges of the included relations), non-cyclicity (exclusion of inverses of relations) and non-redundancy (exclusion of including paths).

The mode of *path evaluation* incorporates empirically plausible criteria in order to select the *strongests* of the ensuing paths. Based on analyses of approximately 60 product reviews from the information technology domain and experimental evidences from several (psycho)linguistic studies (e.g., Chaffin (1992)), we state certain predefined *path patterns*. From those general path patterns and by virtue of the hierarchical organization of conceptual relations, concrete conceptual role chains can automatically be derived in a terminological reasoning system. As a consequence, we may distinguish between a subset of all types of well-formed paths, which is labeled “*plausible*”, another subset which is labeled “*metonymic*”, and all remaining paths which are labeled “*implausible*”. Examples of plausible paths are all paths of length 1 — they are explicitly encoded in the domain’s concept descriptions and are therefore “*plausible*”, by definition — or any series of transitive relations, e.g. *has-physical-part* relations. Following well-established typologies of metonymies (Lakoff, 1987) we include *producer-for-product*, *part-for-whole*, and *whole-for-part* patterns among the metonymic paths.

The computation of paths between an antecedent x and an elliptic expression y , however, may yield several types of well-formed paths, *viz.* “*plausible*”, “*metonymic*” or “*implausible*”. For proper selection we define a ranking on those path labels according to their intrinsic conceptual strength in terms of the relation “ $>_{c-str}$ ” (*conceptually stronger than*) (cf. Table 1).

As a consequence of this ordering, “*metonymic*”

“ <i>plausible</i> ” $>_{c-str}$ “ <i>metonymic</i> ” $>_{c-str}$ “ <i>implausible</i> ”
--

Table 1: Path Labels Ordered by Conceptual Strength

paths will be excluded from a path list iff “*plausible*” paths already exist, while “*implausible*” paths will be excluded iff “*plausible*” or “*metonymic*” paths already exist. At the end of this selection process, only paths of the strongest type are retained in the final path list.

All conceptual paths which meet the above linkage criteria for two concepts, x and y , are contained in the final list denoted by $CP_{x,y}$. As, in the case of textual ellipsis, we have to deal with paths leading from the elliptical expression to several alternative antecedents, we usually have to compare pairs of path lists $CP_{x,y}$ and $CP_{x,z}$, where x, y, z are concepts. Obviously, the criterion which ranks conceptual paths according to their associated path markers is applicable as all paths in a single CP list have the same marker. A function, $PathMarker(CP_{i,j})$, yields either “*plausible*”, “*metonymic*” or “*implausible*” depending on the type of paths the list contains. Hence, the same ordering of path markers as in Table 1 can be applied to compare two CP lists (cf. Table 2).

$StrongerThan(CP_{x,y}, CP_{x,z}) :\Leftrightarrow$ $PathMarker(CP_{x,y}) >_{c-str} PathMarker(CP_{x,z})$ $asStrongAs(CP_{x,y}, CP_{x,z}) :\Leftrightarrow$ $PathMarker(CP_{x,y}) = PathMarker(CP_{x,z})$
--

Table 2: Path Lists Compared by Conceptual Strength

3 Constraints on Centers

Conceptual criteria are of tremendous importance, but they are not sufficient for proper ellipsis resolution. Additional criteria have to be supplied in the case of equal strength of CP lists for alternative antecedents. We therefore incorporate into our model criteria which relate to the functional information structure of utterances using the methodological framework of the well-known *centering* model (Grosz et al., 1995). Accordingly, we distinguish each utterance’s backward-looking center ($C_b(U_n)$) and its forward-looking centers ($C_f(U_n)$). The ranking imposed on the elements of the C_f reflects the assumption that the most highly ranked element of $C_f(U_n)$ is the most preferred antecedent of an anaphoric (or elliptical) expression in U_{n+1} , while the remaining elements are ordered by decreasing preference for establishing referential links.

The main difference between Grosz et al.’s work and our proposal (see Strube & Hahn (1996)) concerns the criteria for ranking the forward-looking centers. While Grosz et al. assume that *grammatical* roles are the major determinant for the ranking on the

C_f , we claim that for languages with relatively free word order (such as German), it is the *functional information structure (IS)* of the utterance in terms of the context-boundedness or unboundedness of discourse elements. The centering data structures and the notion of context-boundedness can be used to redefine Daneš' (1974) trichotomy between *given information*, *theme* and *new information* (or *rheme*). The $C_b(U_n)$, the most highly ranked element of $C_f(U_{n-1})$ realized in U_n , corresponds to the element which represents the *given information*. The *theme* of U_n is represented by the preferred center $C_p(U_n)$, the most highly ranked element of $C_f(U_n)$. The *theme/rheme hierarchy* of U_n is determined by the $C_f(U_{n-1})$: the most rhematic elements of U_n are the ones not contained in $C_f(U_{n-1})$ (unbound discourse elements), they express the *new information* in U_n . The distinction between context-bound and unbound elements is important for the ranking on the C_f , since bound elements are generally ranked higher than any other non-anaphoric elements (cf. also Hajičová et al. (1992)).

bound element(s)	$>_{IS_{base}}$	unbound element(s)
anaphora	$>_{IS_{bound}}$	elliptical antecedent
	$>_{IS_{bound}}$	elliptical expression
nom head ₁	$>_{prec}$	nom head ₂ $>_{prec}$... $>_{prec}$ nom head _n

Table 3: Functional Ranking Constraints on the C_f

The constraints holding for the ranking on the C_f for German are summarized in Table 3. They are organized at three layers. At the top, " $>_{IS_{base}}$ " denotes the basic relation for the overall ranking of information structure (IS) patterns. The second relation, " $>_{IS_{bound}}$ ", denotes preference relations dealing exclusively with multiple occurrences of *bound elements* in the preceding utterance. Finally, " $>_{prec}$ " covers the preference order for multiple occurrences of the *same type* of any information structure pattern, e.g., the occurrence of two anaphora or two unbound elements (all nominal heads in an utterance are ordered by linear precedence in terms of their text position). Given these basic relations, we may formulate the composite relation " $>_{IS}$ " (Table 4). It summarizes the criteria for the ordering of the items on the C_f (x and y denote lexical heads).

$>_{IS} := \{ (x, y) \mid$ <i>if</i> x and y represent the same type of IS pattern <i>then</i> the relation $>_{prec}$ applies to x and y <i>else if</i> x and y represent different forms of bound elements <i>then</i> the relation $>_{IS_{bound}}$ applies to x and y <i>else</i> the relation $>_{IS_{base}}$ applies to x and $y \}$
--

Table 4: Information Structure Relation

4 Predicates for Textual Ellipsis

The grammar formalism we use (for a survey, cf. Hahn et al. (1994)) is based on dependency relations between lexical heads and modifiers. The dependency specifications allow a tight integration of linguistic (grammar) and conceptual knowledge (domain model), thus making powerful terminological reasoning facilities directly available for the parsing process.¹ The resolution of textual ellipses is based on two major criteria, a conceptual and a structural one. The conceptual strength criterion for role chains is already specified in Table 2. The structural condition is embodied in the predicate *isPotentialEllipticAntecedent* (cf. Table 5). A quasi-anaphoric relation between two lexical items in terms of textual ellipsis is here restricted to pairs of nouns. The elliptical phrase which occurs in the n -th utterance is restricted to be a definite NP and the antecedent must be one of the forward-looking centers of the preceding utterance.

$isPotentialEllipticAntecedent(y, x, n) :\Leftrightarrow$ $y isa_C^* Nominal \wedge x isa_C^* Noun$ $\wedge \exists z: (x head z \wedge z isa_C^* DetDefinite)$ $\wedge x \in U_n \wedge y.r \in C_f(U_{n-1})$

Table 5: Potential Elliptic Antecedent

The predicate *PreferredConceptualBridge* (cf. Table 6) combines both criteria. A lexical item y is determined as the proper antecedent of the elliptic expression x iff it is a potential antecedent and if there exists no alternative antecedent z whose conceptual strength relative to x exceeds that of y or, if their conceptual strength is equal, whose strength of preference under the IS relation is higher than that of y .

$PreferredConceptualBridge(y, x, n) :\Leftrightarrow$ $isPotentialEllipticAntecedent(y, x, n)$ $\wedge \neg \exists z: isPotentialEllipticAntecedent(z, x, n)$ $\wedge (StrongerThan(CP_{x.c,z.c}, CP_{x.c,y.c})$ $\vee (asStrongAs(CP_{x.c,z.c}, CP_{x.c,y.c}) \wedge z >_{IS} y))$
--

Table 6: Preferred Conceptual Bridge

5 Resolution of Textual Ellipsis

The actor computation model (Agha & Hewitt, 1987) provides the background for the procedural interpretation of lexicalized grammar specifications, as those

¹We assume the following conventions to hold: $\mathcal{C} = \{\text{Word, Nominal, Noun, PronPersonal, ...}\}$ denotes the set of word classes, and $isa_C = \{(\text{Nominal, Word}), (\text{Noun, Nominal}), (\text{PronPersonal, Nominal}), \dots\} \subset \mathcal{C} \times \mathcal{C}$ denotes the subclass relation which yields a hierarchical ordering among these classes. Furthermore, *object.r* refers to the instance in the text knowledge base denoted by the linguistic item *object* and *object.c* refers to the corresponding concept class C . *Head* denotes a structural relation within dependency trees, viz. x being the head of modifier y .

given in the previous section, in terms of so-called word actors. Word actors communicate via asynchronous message passing; an actor can only send messages to other actors it knows about, its so-called acquaintances. The arrival of a message at an actor triggers the execution of a method, a program composed of grammatical predicates.

The resolution of textual ellipses depends on the results of the foregoing resolution of nominal anaphors (Strube & Hahn, 1995) and the termination of the semantic interpretation of the current utterance. It will only be triggered at the occurrence of the definite noun phrase *NP* when *NP* is not a nominal anaphor and (the referent of the) *NP* is only connected via certain types of relations (e.g., *has-property*, *has-physical-part*)² to referents denoted in the current utterance at the conceptual level.

The protocol level of text analysis encompasses the procedural interpretation of the grammatical constraints from Section 4. We will illustrate the protocol for text ellipsis resolution (cf. Fig. 1), referring to the already introduced text fragment (1) which is repeated at the bottom line of Fig. 1.

(1c) contains the definite noun phrase “*die Ladezeit*”. Since “*Ladezeit*” (*charge time*) does not subsume any word at the conceptual level in the preceding text, the anaphora test fails; the definite noun phrase “*die Ladezeit*” has also not been integrated in terms of a significant relation into the conceptual representation of the utterance as a result of its semantic interpretation. Consequently, a *SearchTextEllipsisAntecedent* message is created by the word actor for “*Ladezeit*”. Message passing consists of two phases:

1. In *phase 1*, the message is forwarded from its initiator “*Ladezeit*” to the *forward-looking centers* of the *previous sentence* (an acquaintance of that sentence’s punctuation mark), where its state is set to *phase 2*.

²Associated with the set of conceptual roles is the set of their inverses. This distinction becomes crucial for already established relations like *has-property* (subsuming *charge-time*, etc.) or *has-physical-part* (subsuming *has-accumulator*, etc.). These relations do not block the triggering of the resolution procedure for textual ellipsis (e.g., ACCUMULATOR – *charge-time* – CHARGE-TIME), whereas instantiations of their inverses, we here refer to as *POF-type relations*, e.g., *property-of* (subsuming *charge-time-of*, etc.) and *physical-part-of* (subsuming *accumulator-of*, etc.), do (e.g., ACCUMULATOR – *accumulator-of* – 316LT). This is simply due to the fact that the semantic interpretation of a phrase like “*the charge time of the accumulator*” already leads to the creation of the *POF-type* relation the resolution mechanism for textual ellipsis is supposed to determine. This is opposed to the interpretation of its elliptified counterpart “*the charge time*” in sentence (1c), where the genitive object “[*of the accumulator*]” is zeroed and, thus, the role *charge-time-of* remains uninstantiated.

2. In *phase 2*, the *forward-looking centers* of the previous sentence are tested for the predicate *PreferredConceptualBridge*.

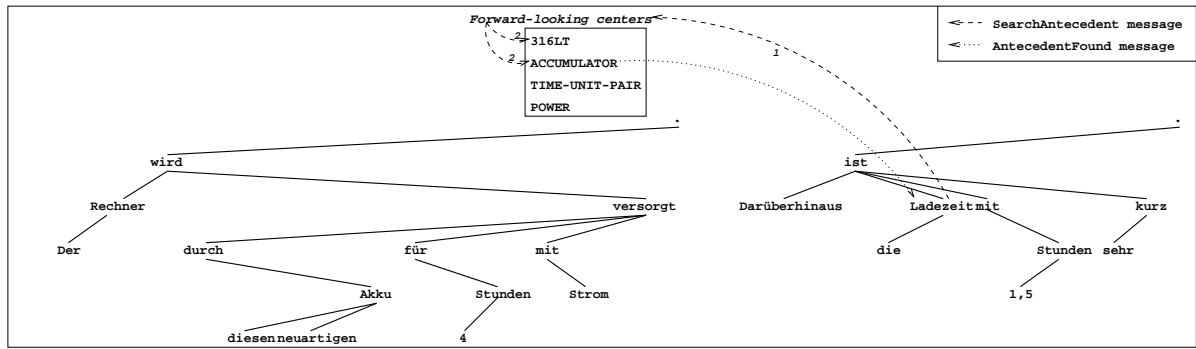
In this case, the instance 316LT (the conceptual referent of the nominal anaphor “*der Rechner*” (*the computer*), which has already been properly resolved) is related to CHARGE-TIME (the concept denoting “*Ladezeit*”) via a *metonymic path*, viz. (*charge-time-of accumulator-of*) indicating a *whole-for-part* metonymy, while the concept ACCUMULATOR is related to CHARGE-TIME via a *plausible path* (viz. *charge-time-of*). As plausible paths are the strongest type of conceptual paths, only an element which is more highly ranked in the centering list and is linked via a plausible path to the elliptical expression could be preferred as the elliptical antecedent of “*Ladezeit*” (*charge time*) over “*Akku*” (*accumulator*) according to the constraint from Table 6. As this can be excluded the remaining concepts associated with the current forward-looking centers (namely, TIME-UNIT-PAIR and POWER) need no longer be considered. Hence, “*Akku*” is determined as the proper elliptical antecedent³. As a consequence, a *TextEllipsisAntecedentFound* message is sent from “*Akku*” to the initiator of the *SearchAntecedent* message, viz. “*Ladezeit*”. An appropriate role filler update links the corresponding concepts via the role *charge-time-of* and, thus, local coherence is established at the conceptual level of the text knowledge base.

In order to illustrate our approach under slightly varying conditions, consider text fragment (2):

- (2) a. Der 316LT geht sparsam mit Energie um.
(The 316LT uses – sparingly – energy.)
- b. Netzunabhängig wird *er* für ca. 2 Stunden mit Strom versorgt.
(In a power supply-independent mode *it* is – for approximately 2 hours – with power – provided.)
- c. Wenn die *Taktfrequenz* herabgesetzt wird, reicht die Energie sogar für 3 Stunden.
(When the *clock frequency* – is reduced – suffices – the power – even for 3 hours.)

Here, the elliptical expression “*Taktfrequenz*” (*clock frequency*) can tentatively be related to three antecedents in the preceding sentence: “*er*” (*it*) (which is an anaphoric expression for “316LT”), “*Stunden*” (*hours*), and “*Strom*” (*power*). Thus, in the path finding mode paths from CLOCK-MHZ-PAIR (the conceptual representation for “*Taktfrequenz*”) to 316LT, TIME-UNIT-PAIR (representing “*Stunden*”),

³Note that only nouns and pronouns are capable of responding to the *SearchTextEllipsisAntecedent* message and of being tested as to whether they fulfill the required criteria for an elliptical relation.



(1b) Der Rechner wird durch diesen neuartigen Akku für 4 Stunden mit Strom versorgt. (1c) Darüberhinaus ist die Ladezeit mit 1,5 Stunden sehr kurz.
 (1b) The computer is because of this new type of accumulator for 4 hours with power provided. (1c) Also, is the charge time of 1.5 hours quite short.

Figure 1: Sample Parse for Text Ellipsis Resolution

and POWER, respectively, are searched. As only a single well-formed role chain from CLOCK-MHZ-PAIR to 316LT can be determined (viz. (*clock-mhz-pair-of-cpu-of-motherboard-of-central-unit-of*)), “316LT” is selected as the valid elliptic antecedent. Under these circumstances, conceptual linkage could not be established via a *plausible path*, but only via a *metonymic path*, corresponding to a *whole-for-part* metonymy. This is due to the fact that “*Taktfrequenz*” (*clock frequency*) (conceptualized as CLOCK-MHZ-PAIR) is a property of the CPU of COMPUTER-SYSTEM and, therefore, only a mediated property of computers as a whole (hence, the *whole-for-part* metonymy).

Evaluation. A small-scale evaluation experiment was conducted on a test set of 109 occurrences of textual ellipses in 5 different texts taken from our corpus. The evaluation used our knowledge base from the information technology domain, which consists of 449 concepts and 334 relations. Among 46 (42.2%) *false negatives* (no resolution triggered though textual ellipsis occurs), the ellipsis handler encountered 42 (38.5%) cases of lacking concept specifications (half of which were gaps that can easily be closed, the other half constituted by “soft” concepts (e.g., referring to spatial knowledge) which are hard to get hold of). In 4 of the 46 cases the conceptual model was adequate but the triggering conditions were inappropriate.

Among the 63 cases where the ellipsis handler started processing 60 were correctly analyzed (*recall* rate of 55.05%), 2 modelling bugs were encountered in the knowledge base, and one case was due to incorrect conceptual constraints. Considering the performance of the criteria we propose — disregarding effects that come from deficient knowledge engineering, i.e. restricting the evaluation to the 63 cases run by the ellipsis handler — the *precision* rate amounts to 95.2%.

With respect to *accuracy*, however, we still have to consider the actual number of textual ellipses processed including *false positives*, i.e., cases where the

ellipsis resolution is carried out although no textual ellipsis actually occurs. Altogether, the ellipsis handler was triggered 82 times, thus it was incorrectly triggered in 19 cases (23.2%). 12 of these 19 false positives were due to insufficiencies of the parsing process (it failed to create suitable semantic/conceptual representations blocking the triggering of the ellipsis handler). In 4 cases the anaphora resolution process failed to resolve an anaphor, thus leading to an incorrect call of the ellipsis handler, and in the 3 remaining cases the triggering condition was not restrictive enough. This condition gives an overall accuracy score of 73.2%.

6 Comparison with Related Approaches

As far as text-level processing is concerned, the framework of DRT (Kamp & Reyle, 1993), at first sight, constitutes a particularly strong alternative to our approach. The machinery of DRT, however, might work well for (pro)nominal anaphora, but faces problems when elliptical text phenomena are to be interpreted (though Wada (1994) has recently made an attempt to deal with restricted forms of textual ellipsis in the DRT context). This shortcoming is simply due to the fact that DRT is basically a semantic theory, not a full-fledged model for text understanding. In particular, it lacks any systematic connection to well-developed reasoning systems accounting for conceptual domain knowledge. Actually, the sort of constraints we considered seem much more rooted in encyclopedic knowledge than are they of a primarily semantic nature anyway.

As far as proposals for the analysis of textual ellipsis are concerned, none of the standard grammar theories (e.g., HPSG, LFG, GB, CG, TAG) covers this issue. This is not at all surprising, as their advocates pay almost no attention to the text level of linguistic description (with the exception of several forms of anaphora) and also do not seriously take conceptual criteria beyond semantic features into account. Hence their indeterminateness with respect to conceptually driven inferring in the context of text understanding.

Actually, only few systems exist which deal with textual ellipsis in a dedicated way. For example, the PUNDIT system (Palmer et al. (1986)) provides a fairly restricted solution in that only direct conceptual links between the concept denoted by the antecedent and the elliptical expression are considered (“plausible” paths of length l , in our terminology). A pattern-based approach to inferencing (including textual ellipsis) has also been put forward by Norvig (1989). The main difference to our work lies in the fact that these path patterns do not take the compositional properties of relations into account (e.g., transitive relations). Furthermore, numerical constraints like path length restrictions are posited without motivating their origin, whereas we state formal well-formedness and empirical criteria the evidence for which is derived from psycholinguistic studies. The abduction-based approach to inferencing underlying the TACITUS system (Hobbs et al. (1993)) also refers to weights and costs and, thus, shares some similarity with Norvig’s proposal (Hobbs et al., 1993, p. 122). Moreover, the crucial problem still unsolved in this logically very principled framework concerns a proper choice methodology for fixing appropriate costs for specific assumptions on which, among other factors, textual ellipsis resolution is primarily based. The approach reported in this paper also extends our own previous work on textual ellipses (Hahn, 1989) by the incorporation of an elaborated model of functional preferences on C_f elements.

7 Conclusion

We have outlined a model of textual ellipsis resolution. It considers conceptual criteria to be of primary importance and provides linkage constraints for role paths in a terminological knowledge base in order to assess the plausibility of possible antecedents as proper bridges (Clark, 1975) to elliptical expressions. In addition, functional information structure constraints contribute further restrictions on proper elliptical antecedents. The particular advantage of our approach lies in the integrated treatment of textual ellipsis within a single coherent grammar format that integrates linguistic and conceptual criteria in terms of general constraints.

The ellipsis handler has been implemented in Smalltalk as part of a comprehensive text parser for German. Besides the information technology domain, experiments with our parser have also been successfully run on medical domain texts, thus indicating that the heuristics we have been developing are not bound to a particular domain. The current lexicon contains a hierarchy of approximately 100 word class specifications with nearly 3.000 lexical entries and corresponding concept descriptions available from the LOOM knowledge representation system (MacGregor

& Bates, 1987) — 800 and 500 concept/role specifications for the information technology and medicine domain, respectively.

Acknowledgments. We would like to thank our colleagues in the *CLLF* group for fruitful discussions. This work has been funded by *LGFG Baden-Württemberg* (M. Strube). Katja Markert is supported by a grant from DFG within the Freiburg University Graduate Program on “*Human and Artificial Intelligence*”. We also gratefully acknowledge the provision of the LOOM system from USC/ISI.

References

- Agha, Gul & Carl Hewitt (1987). Concurrent programming using actors. In A. Yonezawa & M. Tokoro (Eds.), *Object-Oriented Concurrent Programming*, pp. 37–53. Cambridge, Mass.: MIT Press.
- Chaffin, Roger (1992). The concept of a semantic relation. In A. Lehrer & E.F. Kittay (Eds.), *Frames, Fields and Contrasts*, pp. 253–288. Hillsdale, N.J.: Lawrence Erlbaum.
- Clark, Herbert H. (1975). Bridging. In *Proc. of the Conference on Theoretical Issues in Natural Language Processing; Cambridge, Mass., June 1975*, pp. 169–174.
- Daneš, František (1974). Functional sentence perspective and the organization of the text. In F. Daneš (Ed.), *Papers on Functional Sentence Perspective*, pp. 106–128. Prague: Academia.
- Grosz, Barbara J., Avarind K. Joshi & Scott Weinstein (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.
- Hahn, Udo (1989). Making understanders out of parsers: Semantically driven parsing as a key concept for realistic text understanding applications. *International Journal of Intelligent Systems*, 4(3):345–393.
- Hahn, Udo, Katja Markert & Michael Strube (1996). A conceptual reasoning approach to textual ellipsis. In *Proc. of the 12th European Conference on Artificial Intelligence; Budapest, Hungary, 10–14 August 1996*, pp. 572–576.
- Hahn, Udo, Susanne Schacht & Norbert Bröker (1994). Concurrent, object-oriented dependency parsing: The ParseTalk model. *International Journal of Human-Computer Studies*, 41(1/2):179–222.
- Hajičová, Eva, Vladislav Kuboň & Petr Kuboň (1992). Stock of shared knowledge: A tool for solving pronominal anaphora. In *Proc. of the 14th International Conference on Computational Linguistics; Nantes, France, 23–28 August 1992*, Vol. 1, pp. 127–133.
- Hobbs, Jerry R., Mark E. Stickel, Douglas E. Appelt & Paul Martin (1993). Interpretation as abduction. *Artificial Intelligence*, 63:69–142.
- Kamp, Hans & Uwe Reyle (1993). *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht, Netherlands: Kluwer.
- Lakoff, George (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago, Ill: Chicago University Press.
- MacGregor, Robert & Raymond Bates (1987). *The LOOM Knowledge Representation Language*. (ISI/RS-87-18) USC/ISI.
- Norvig, Peter (1989). Marker passing as a weak method for inferencing. *Cognitive Science*, 13(4):569–620.

- Palmer, Martha S., Deborah A. Dahl, Rebecca J. Schiffman, Lynette Hirschman, Marcia Linebarger & John Dowding (1986). Recovering implicit information. In *Proc. of the 24th Annual Meeting of the Association for Computational Linguistics; New York, NY, 10–13 June 1986*, pp. 10–19.
- Strube, Michael & Udo Hahn (1995). *ParseTalk* about sentence- and text-level anaphora. In *Proc. of the 7th Conference of the European Chapter of the Association for Computational Linguistics; Dublin, Ireland, 27–31 March 1995*, pp. 237–244.
- Strube, Michael & Udo Hahn (1996). Functional centering. In *Proc. of the 34th Annual Meeting of the Association for Computational Linguistics; Santa Cruz, Cal., 23–28 June 1996*, pp. 270–277.
- Wada, Hajime (1994). A treatment of functional definite descriptions. In *Proc. of the 15th International Conference on Computational Linguistics; Kyoto, Japan, 5–9 August 1994*, Vol. 2, pp. 789–795.
- Woods, William A. & James G. Schmolze (1992). The KL-ONE family. *Computers & Mathematics with Applications*, 23(2-5):133–177.