

Controlling Plural Ambiguities in Attempto Controlled English (ACE)

Uta Schwertel

Department of Information Technology, University of Zurich

Winterthurer Str. 190, CH-8057 Zurich

uschwert@ifi.unizh.ch

<http://www.ifi.unizh.ch/~uschwert>

Abstract. Attempto Controlled English (ACE) is a controlled natural language specifically constructed to write software specifications. ACE specifications are computer-processable and can be unambiguously translated into a first-order logic language. The current version of ACE deals only with singular phenomena. We discuss how ACE can be extended with plurality. The extension proves to be difficult in that plurals are a frequent source of semantic ambiguities. ACE uses a *constructive*, i.e. completely controlled approach to deal with syntactic and semantic ambiguities. We outline how this strategy can be applied to reduce and eliminate plural ambiguities in ACE.

1 Introduction

The main purpose of controlled natural languages is to reduce ambiguity and complexity inherent in full natural language. The reduction of ambiguity is particularly important when natural language texts are to be automatically processed by a computer, since a computer does not have access to the complex knowledge humans use when interpreting and disambiguating natural language.

Attempto Controlled English (ACE) is a controlled natural language specifically designed to write software specifications (Fuchs et al. 1999a/b). ACE specifications are computer-processable and can be unambiguously translated into a first-order logic language thus facilitating automatic verification and validation of specifications. To guarantee that every sentence gets just one semantic representation, ambiguity processing in ACE uses a special, so-called *constructive* approach that we will present in this paper.

Ambiguity occurs on all levels of natural language processing (lexical, syntactic, semantic, pragmatic). ACE is fundamentally concerned with syntactic and semantic ambiguities while many other controlled languages (CLAW 1998) aim at controlling lexical ambiguity, e.g. by standardizing the vocabulary. Syntactic ambiguity means that a sentence gets two or more syntactic analyses. A typical example are PP-attachment ambiguities found e.g. in *The driver stops the train with the defective brake*. Semantic ambiguities occur when a sentence has just one syntactic analysis but allows for different semantic analyses. Typical examples are so-called scope-ambiguities found in sentences like *A device controls every pump*. A further frequent source of semantic ambiguities are plural constructions, e.g. *John and Mary have a computer* can mean that John and Mary have a computer together, or that each of them has a computer. This has become known as the collective/distributive ambiguity.

The current version of ACE only deals with singular phenomena. We call this version the *singular fragment* of ACE. This paper shows how ACE can be extended with plurals, and gives an outline

of necessary restrictions to control plural ambiguities.

In section 2.1 we give a short overview of the singular fragment of ACE; section 2.2 explains the disambiguation strategy of ACE and contrasts it to other strategies. The semantic ambiguities postulated for full natural language plural sentences are briefly discussed in section 3.1. We show that uncontrolled plurals can lead to a combinatorial explosion of semantic representations that are hard to manage for computers and sometimes even for humans. Section 3.2 discusses possible means for reducing or eliminating plural ambiguities and shows how these can be incorporated in ACE. The results presented are just a starting point for the treatment of plurals in ACE. Necessary extensions are addressed in section 3.3.

2 Treatment of Ambiguity in ACE

2.1 ACE in a Nutshell

Attempto Controlled English (ACE) is a controlled natural language specifically constructed to write specifications (Fuchs et al. 1999a/b, Schwitter 1998). ACE allows users to express specifications precisely, nonetheless using the familiar terms of the application domain. ACE specifications are computer-processable and can be unambiguously translated into first-order logic. This means that users can work solely with ACE without having to take recourse to the internal logic representation. Though ACE seems perfectly natural, it is a formal language with the semantics of the underlying logic language. As a formal specification language ACE has the advantage that it supports the automatic analysis of specifications such as consistency verification, and the option to validate specifications through execution (Fuchs 1992).

What are the main components of ACE? ACE is a subset of standard English which uses a domain-specific vocabulary, consisting of predefined function words like determiners, prepositions and conjunctions, and user-defined content words like nouns, verbs, adjectives and adverbs. Users can extend and modify the lexicon via a simple interface requiring little more than basic grammar knowledge. ACE employs a restricted grammar in the form of a small set of construction and interpretation rules. Construction rules define the form of ACE sentences and state restrictions intended to remove imprecision and to restrain ambiguities. Interpretation rules control the semantic analysis of grammatically correct ACE sentences and resolve remaining ambiguities. The construction and interpretation rules are implemented as a unification-based phrase structure grammar that is used by the chart-parser of the Attempto system (Schwitter 1998). The parser deterministically translates ACE texts into discourse representation structures (Kamp and Reyle 1993) which are a syntactic variant of first-order predicate logic (FOL), and optionally into the standard and clausal form of FOL. Furthermore, a paraphrase is generated that shows the user how the Attempto system interprets the ACE input.

Following are the most important construction rules of the singular fragment of ACE:

- ACE specifications are sequences of anaphorically interrelated simple and composite sentences.
- Simple sentences have the form *subject + verb + complements + adjuncts*, where complements can be noun phrases or prepositional phrases. Adjuncts (adverbs, prepositional phrases) are optional. An example of a simple sentence is

- (1) A customer enters a card into a slot.
- Composite sentences are built from other sentences through coordination (*and, or*) or subordination (*if ... then, who/which/that*). Furthermore sentences containing negation (*does not, is not, no*) or universal quantifiers (*every, for every*) are classified as composite sentences.
- (2) Every card has a code that is valid.
- ACE sentences can be interrelated by anaphora, i.e. personal pronouns (*he/she/it*), definite noun phrases like *the card*, and so-called dynamic names like *T* in the third sentence of (3).
- (3) John is a customer. He has a card that he enters into an automated teller T.
If a customer enters a card into T then T checks the card.
- Verbs are only used in the simple present tense, the active voice, the indicative mood, and the third person.
- Modal verbs (*can, must* etc.), intensional verbs (*hope, know* etc.), and modal adverbs (*possibly, probably* etc.) are not allowed.

Interpretation rules control the semantic analysis of grammatically correct ACE sentences. They resolve ambiguities that are not eliminated by the constructions rules. Following are the most important interpretation rules:

- Verbs denote events or states; the textual order of verbs determines the default temporal order of the associated events and states.
- Prepositional phrases in adjunct position always modify the verb, while relative sentences modify the immediately preceding noun phrase.
- The textual order of a quantifier opens its scope that extends to the end of the sentence; thus any following quantifier is within the scope of the preceding ones.
- The antecedent of anaphoric reference is always the most recent accessible noun phrase that agrees in number and gender.

For further details of the Attempto system see Fuchs et al. (1999a/b), Schwitter (1998).

2.2 Constraining Ambiguities in ACE

Natural language processing systems that deal with uncontrolled natural language rely on different disambiguation strategies, e.g. *Generate and Test* and *Underspecified Representations* (FRACAS 1994).

The *Generate and Test* approach first generates for every possible reading a semantic representation and in a second step eliminates the implausible ones. The elimination can be done by presenting all interpretations to the user who selects the intended one (Macias and Pulman 1995), or by formalizing relevant contextual knowledge (world knowledge, linguistic context, lexical semantics, pragmatics etc.) to determine the most fitting interpretation (Hindle and Rooth 1993). The disadvantage of this approach is that complex sentences tend to have a large number of readings and it is inefficient to first generate all representations and then to eliminate most of them. Manually selecting one of the many representations is a burden on the user. Furthermore, formalizing relevant contextual knowledge to filter the readings has proven to be very difficult, if not impossible.

In the *Underspecification* approach a semantically ambiguous sentence gets just one representa-

tion which is derivable from the syntactic form of the sentence leaving certain aspects of the meaning not specified (Alshawi 1992, Reyle 1993/1994). Fully specified readings are obtained by filling in material that is derived from contextual knowledge. This approach tackles the efficiency problem of the *Generate and Test* method. However, the drawback remains that it is very hard to formalize the contextual knowledge that leads to more specific interpretations. Furthermore, there are currently no well-developed reasoning mechanisms that perform reasoning directly on underspecified representations. Existing reasoning techniques mostly rely on listing all the representations that the underspecified representation comprises. But then there is no efficiency gain compared to *Generate and Test*.

Since ACE is a formal language every sentence has to get just one and only one representation. This representation should be very reliable, i.e. the same syntactic construct must lead to the same semantic interpretation in all cases. Recourse to context would make this task very difficult. For this reason ACE chooses a different approach to handle ambiguity. We call our approach *constructive*. The *Constructive Approach* consists of three simple means:

- Some ambiguous constructs are excluded from the language; unambiguous alternatives are available in their place.
- All remaining ambiguous constructs are interpreted deterministically on the basis of a small number of interpretation rules that use syntactic information only; the interpretations are reflected in a paraphrase.
- Users can either accept the assigned interpretation, or they must rephrase the input to obtain another one.

For a sentence like

(4) John enters a card with a code.

ACE generates the paraphrase

(5) John {enters a card with a code}.

that reflects ACE's interpretation rule that a prepositional phrase always modifies the verb. Since this is probably not the intended interpretation the user can reformulate the sentence using the interpretation rule that relative sentences always modify the immediately preceding noun phrase

(6) John enters a card that has a code.

yielding the paraphrase

(7) John enters {a card that has a code}.

The singular fragment ACE has about a dozen interpretation rules to handle ambiguity.

3 Towards a Plural Fragment of ACE

3.1 Postulated Plural Ambiguities

Plurals in full natural language are a notorious source of semantic ambiguities. The sentence

(8) Three men lift a table.

can mean that the three men lift the table in one collective act (*collective reading*) or that each of

the three men lifts a (possibly different) table in an individual act (*distributive reading*), i.e. three tables can be involved. Further readings have been discussed (cf. van den Berg 1994).

If sentences contain more than one plural noun phrase possible readings multiply even further. Assuming only the collective/distributive distinction the sentence

(9) Four men lift three tables.

is claimed to have seven readings (Link 1991). Since there are two plural argument places, four readings come from the collective/distributive distinction; a possible distinction in scope multiplies the readings by two. In the doubly collective case the scope distinction is not relevant reducing the readings by one.

In addition, Scha (1981) and others suggest that there is yet another reading, the so-called *cumulative reading* prominent in sentences with two (or more) numeric noun phrases:

(10) 600 Dutch firms bought 5000 American computers.

Scha correctly claims that (10) can be used truthfully if there are a total of 600 Dutch firms involved in buying American computers, and 5000 American computers which were bought by Dutch firms. Furthermore, Scha says that the computers can be bought in arbitrary groups and the firms could form arbitrary collections to buy the computers, so long as all firms and all computers were involved.

Figure 1 gives a summary of the eight proposed readings of two-place plural sentences.

It is obvious, that this combinatorial explosion of readings cannot be a reasonable basis for a controlled unambiguous use of plurals in ACE, nor is this classification suitable for automatic processing. Moreover, the proposed classification is not intuitive since many readings are generated that humans hardly perceive unless they are trained in formal semantics.

3.2 Controlling Plural Ambiguities in ACE

3.2.1 Disambiguation Information

Two criteria are important for the formulation of disambiguation rules in ACE. On the one hand the rules should be intuitive, i.e. correspond to a natural use and interpretation of English, on the other hand the rules must be systematic so that ACE texts can be deterministically processed by a computer. In any case the proposed representations should support the desired inferences. These requirements are sometimes hard to reconcile since disambiguation is often triggered by real world-knowledge, by linguistic context, by lexical knowledge, intonation or other factors – factors that are not accessible to the Attempto system (nor to other natural language processing tools). As in the singular case we therefore looked for "accessible" plural disambiguation patterns. With "accessible" we mean that the information consists of structural and lexical information that is available to every user and to the Attempto system, and that does not rely on contextual knowledge. For plural disambiguation we use the following accessible information sources:

- information about the verb
(arity, event or state, copula or full verb, verbs requiring plural arguments)

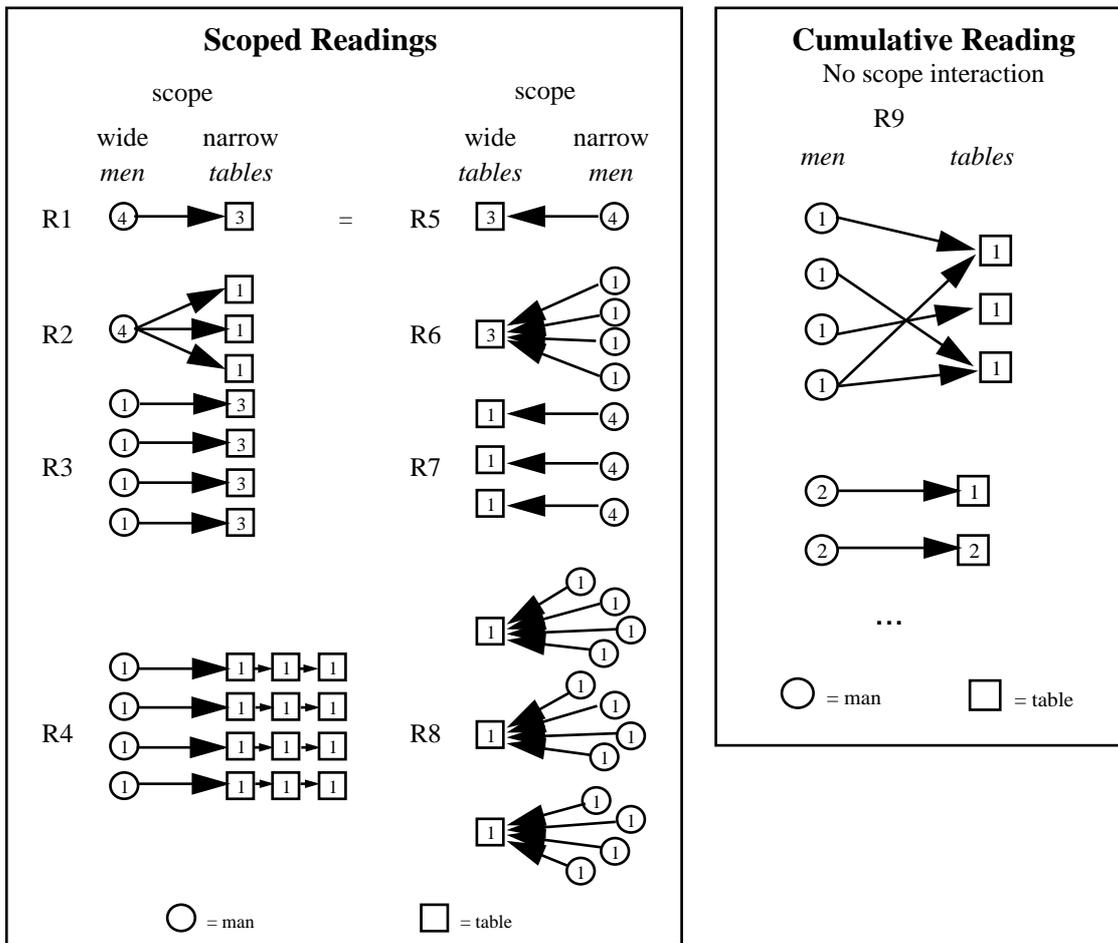


Figure 1 Eight Readings of *Four men lift three tables*.

- structural information
(internal vs. external argument, surface order of noun phrases)
- information about noun phrase
(definite, indefinite or quantifying, type of quantifier)
- predefined syntactic markers
(modifiers, floated-quantifier *each*)

The results are at the moment restricted to simple sentences containing one- or two-place verbs and definite and indefinite plural noun phrases as arguments. We assume that only definite and indefinite noun phrases can lead to a collective reading. Definite plural noun phrases in ACE are e.g. *the men*, *John and Mary*, indefinite noun phrases are e.g. *some men* and *two men*. ACE treats noun phrases introduced by *every* as quantifying. Quantifying noun phrases cannot lead to a collective reading. Other plural determiners that can under certain circumstances be read collectively are currently only partially covered in ACE (e.g. *at least three*, *more than three*, *not more than three*, *exactly three*, *at most three*). The plural quantifiers *no* and *all* will be treated as quantifiers. See Lønning (1987) for a discussion of this classification.

3.2.2 Reducing Plural Ambiguities

The following observations lead to a successive reduction and elimination of plural ambiguities in ACE.

Surface Order of Noun Phrases

Very often there is no scope distinction operative in sentences like

(11) Four men lift three tables.

The sentence is either about four men and three tables, or about four men and twelve tables. A wide scope interpretation of three tables would allow a situation with three tables and twelve men – a reading that is not plausible for (11).

In ACE, we therefore decided to interpret noun phrases in the order of their occurrence. As a consequence a noun phrase later in a sentence can never take scope over a preceding noun phrase. This correlates to the *Principle of Surface Order* stated for the singular fragment of ACE.

This rule reduces the readings of Figure 1 to the five readings R1-R4 and R9.

External and Internal Argument – Ambiguity vs. Indeterminacy

Furthermore, an asymmetry between the external and the internal argument of a transitive verb can be observed (Lønning 1987, Verkuyl 1999). The external argument of a transitive verb in the active voice is the subject, the internal argument is the direct object.

(12) Three girls lift a table.

(13) A girl lifts three tables.

The external plural argument in (12) leads to *ambiguity*, the internal plural argument in (13) to *indeterminacy*.

Ambiguity. In full natural language sentence (12) having a plural external argument is truly ambiguous between a reading where three girls each lift a table (three table lifting events) and a reading where three girls as a group lift a table (one table lifting event). Note that mixed readings where two girls together lift a table and a third girl lifts another table (i.e. two tables involved) are very unlikely.

In ACE the collective/distributive ambiguity of (12) has to be resolved by the user by adding the floated-quantifier *each* directly after the external argument noun phrase to get the distributive reading.

(14) Three girls each lift a table.

In absence of a cue word ACE interprets the sentence collectively:

(15) Three girls lift a table.

To prevent an inflation of the language we do not require users to add collectivizing markers explicitly to indicate the collective reading.

Indeterminacy. Sentence (13), however, is not considered to be ambiguous in its internal argument. Our claim is that it is not a matter of ambiguity but a matter of indeterminacy whether the tables are lifted one-by-one, as a whole or in any other constellation. In any case three tables get finally lifted in a possibly complex event. Sentence (13) fundamentally underinforms the hearer about how the actual situation was realized. This adds to the efficiency of natural language in that

the user is not forced to be more precise than necessary in a particular situation.

Thus (13) gets just one logical representation, no additional disambiguation markers are necessary in ACE. The representation comprises possible constellations. This is discussed as the partitional interpretation e.g. in Verkuyl (1999).

If the users want to express additional information and thus explicitly narrow down the possible constellations ACE allows to add the explicit modifiers *one-by-one*, *as a whole* or *as a group* directly after the internal argument noun phrase.

- (16) a. A girl lifts three tables.
b. A girl lifts three tables one-by-one.
c. A girl lifts three tables as a whole/as a group.

ACE does not yet allow for other modifiers like *in groups* or *in groups of two* which detail a more complex substructuring of the internal argument NP.

Note, that these markers are not *disambiguation markers*, but what we call *precisification markers*. Precisification markers add detail to a logical representation, while the detailed logical representation still implies the original representation. The use of disambiguation markers like *each*, however, leads to two logically incompatible logical representations.

A thorough explanation for the asymmetry of external and internal arguments can be found in Verkuyl (1999).

In sum, the distinction between external and internal arguments collapses readings R1 and R2 of Figure 1 to one collective reading, whereas R3 and R4 are reduced to one distributive reading, leaving us for the moment with three readings.

One-place Verbs

There is a large class of one-place verbs where the subject, i.e. the only argument of the verb, behaves like the internal argument of a transitive verb.

- (17) Five books disappear.
(18) The trains depart.

In these sentences we claim that it is not adequate to distinguish between a collective or a distributive reading. The sentences underinform the hearer about possible constellations in which the books disappear or the trains depart. Again, we consider these examples not as ambiguous but as indeterminate and assign one representation comprising possible constellations.

As in the two-place case the additional markers *one-by-one*, *as a whole* or *as a group* can be added. In the one-place case these precisification markers can either be put directly after the plural noun phrase (which sounds less natural) or directly after the verb. The markers then relate to the closest preceding plural noun phrase.

- (19) a. Five books disappear.
b. Five books disappear one-by-one.
c. Five books disappear as a whole/as a group.

State Verbs

In ACE users have to classify verbs either as event verbs or as state verbs. Event verbs (*arrive*, *depart*, *succeed*, *disappear*, *walk*, *drive*) express progress in time whereas state verbs (*hang*, *sit*,

stand, sleep, have, own, watch) do not express a change in time. The examples in the previous sections all contained event verbs and the results have to be slightly modified to deal with state verbs properly. State verbs behave similar to event verbs in that plural noun phrases as external arguments require disambiguation

- (20) a. Five students have an email account.
- b. Five students each have an email account.

whereas plural noun phrases as internal arguments do not

- (21) A man owns five houses.
- (22) Mary has three email accounts.

There are, however, differences to event verbs in that states do not induce a substructuring of the internal argument (Verkuyl 1999). In (21) it does not make sense to talk of sub-states of the state of owning five houses, i.e. compared to event verbs state verbs lack the indeterminacy in the internal argument. This also explains why precisification markers like *one-by-one* are infelicitous.

Two Plural Noun Phrases

With the restrictions stated above sentences containing two plural noun phrases like

- (23) Four men lift three tables.

no longer lead to a proliferation of readings.

In ACE we can express a distributive reading by adding *each* after the external plural argument

- (24) Four men each lift three tables.

The collective reading does not require a marker.

- (25) Four men lift three tables.

Additionally, precisification markers can be added that lead to a detailing of the part structure of the internal argument.

- (26) a. Four men each lift three tables one-by-one.
- b. Four men lift three tables as a whole.

Cumulative Reading

As introduced in section 3.1 it has been argued that sentences containing two numerical noun phrases like

- (27) Four men lift three tables.
- (28) 600 Dutch firms own 5000 American computers.

can express a so-called cumulative reading (Scha 1981). For example, sentence (27) can be true in a situation where two men lift two tables and the other two men together lift one table. We follow Roberts (1987), Lønning (1987), Link (1991), Verkuyl (1999) and others in that the cumulative case does not constitute a separate reading but is a special case of the collective reading. In (27) there is a group of four men and a group of three tables and the two groups stand in the lifting relation, i.e. there is something which brings the four men together in one collection. Similar to the indeterminacy of an internal plural argument sentences like (27) and (28) underinform the hearer of what actually happened. Whatever more can be said about the possible constellations depends on many factors, particularly on the meaning of the verb which is beyond the accessible information of ACE.

Altogether this leaves us with two readings: the distributive and the non-distributive. Readings R3 and R4 of Figure 1 are instances of the distributive reading, whereas R1, R2 and R9 constitute the non-distributive cases.

Distributivity vs. Totalization

ACE users thus only have to distinguish between two readings when plural noun phrases are involved: the *distributive reading* and what – in the light of the above discussions – we now call the *totalizing reading* (Verkuyl 1999). The distributive reading is triggered by the disambiguation marker *each* directly after the external argument noun phrase. The totalizing (= non-distributive) reading is triggered in absence of markers. The totalizing reading subsumes the sentences that are *collective in a narrow sense* (the predicate applies to a plural object as a group and not to any sub-part of it), and sentences that are *collective in a broad sense* where the mixed cases, e.g. the cumulative reading, are included. The narrowly collective reading is triggered by precisification markers *as a whole* or *as a group*. In absence of any markers the totalizing reading in general is assumed.

Figure 2 summarizes the readings of simple two-place plural sentences in ACE.

3.2.3 Problems

Our mechanisms to reduce plural ambiguities rely on a very coarse classification of verbs. Of course, the mechanisms in full natural language are not so easily generalized. Examples like

(29) The tables are heavy.

indicate that there are one-place stative predicates that are sensitive to a collective-distributive distinction. The class of adjectives (*light, dense, thin, long* etc.) having this property is rather small. It is, however, not realistic to burden the user with classifying these adjectives separately.

One could argue that sentence (29) should be classified as a two-place sentence where the subject corresponds to the external argument which would predict the ambiguity. After all, this misses a generalization found in most other constructions involving copula plus adjective:

(30) The books are red.

(31) Three cards are valid.

These constructions are not ambiguous. They are not even indeterminate. The predicates denote states and do not involve a partition on the denotation of the subject noun phrase that is treated like the internal argument. Furthermore, most adjectives allow the inference that the predicate distributes down to the individual members of the subject denotation, a generalization we want to maintain in ACE.

It is no problem for ACE to express the narrowly collective reading of (29) by adding the markers *as a group* or *as a whole*.

(32) The tables as a whole are heavy.

To get the distributive interpretation the user has to reformulate the sentence. ACE users can either use the universal quantifier *every*

(33) Every table is heavy.

however, losing the anaphoric potential of the noun phrase *the tables*. The construction

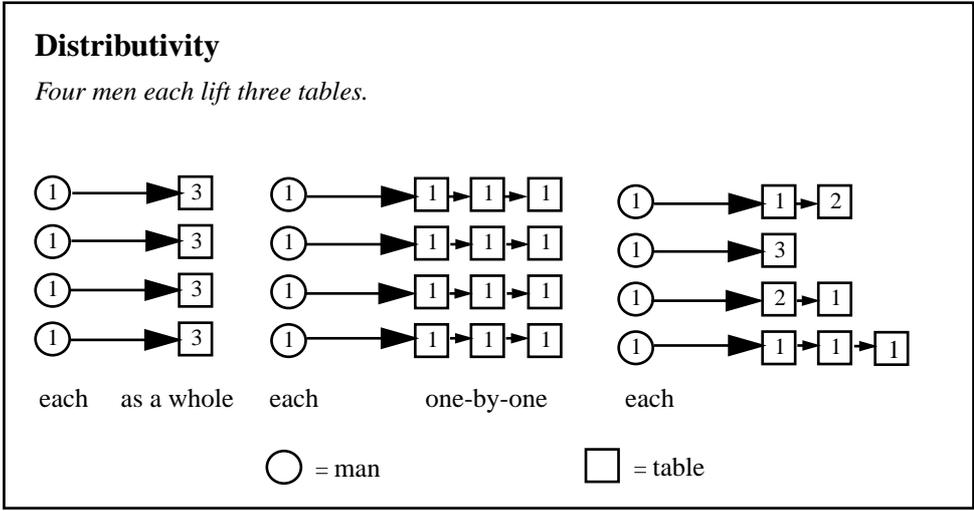
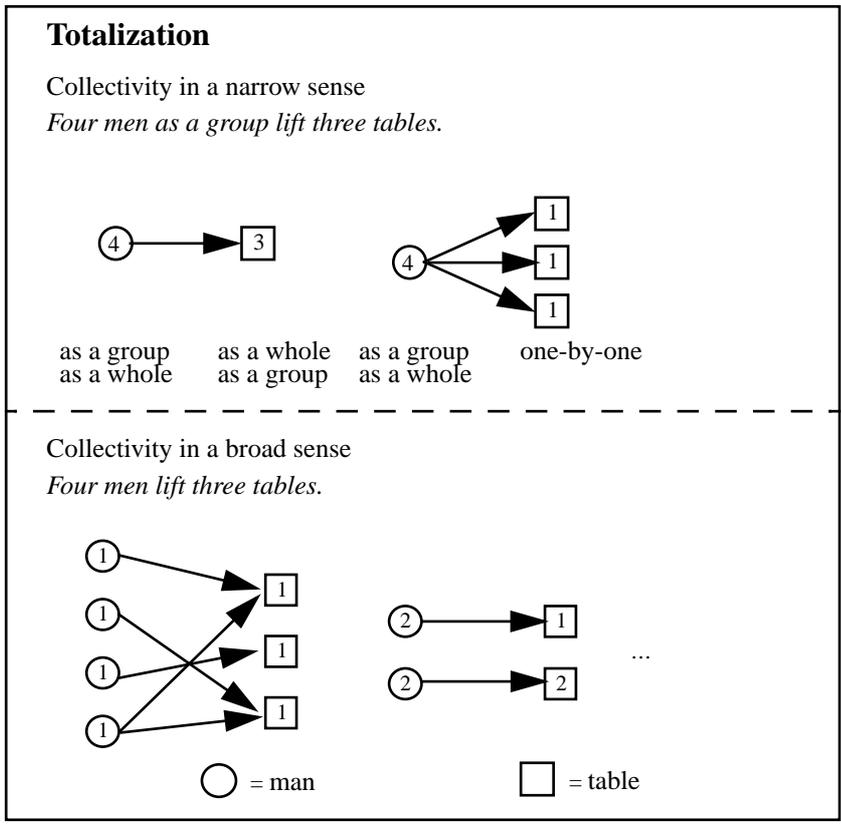


Figure 2 Two Readings of *Four men lift three tables* in ACE.

(34) The tables each are heavy.
 preserves the anaphoricity of *the tables*. However, (34) sounds rather unnatural and, what is more, the role of *each* does not correspond to the disambiguating use presented above where *each* was only introduced after the external argument. A partitive construction

(35) Each of the tables is heavy.

would do the job best. Since partitives like *each of the tables* are a powerful means to explicitly disambiguate or precisify plurals we plan to introduce them into the language.

There are other predicates that seem to cause problems for our approach, e.g. verbs requiring a plural argument or a group denoting noun (*group, herd, committee*):

(36) The boys gather.

(37) Three customers meet.

(38) John counted the cards.

It has been claimed that these examples enforce a narrowly collective reading since the predicate cannot distribute down to individual members. Our opinion, however, is that verbs like *gather* and *meet* do have distributive subentailments similar to *disappear*, however, the bottom elements of the distribution are not individuals but groups. To exclude constructions like *John gathers* from ACE, users would have to classify *gather* as a verb taking a plural argument. We currently check, whether this additional lexical information can be realistically given by users.

3.3 Extensions

We have explained the basic techniques to control plural ambiguities in ACE. These techniques constitute just the starting point. Modifications might become necessary as the coverage of ACE is extended.

ACE has to permit three place-predicates with plural arguments, e.g. *Five men show three cars to two customers*. ACE currently interprets these constructions as totalizing in each argument position when no markers are present. The disambiguation marker *each* can be added after the external argument *five men* and/or the direct object *three cars*. For instance, the sentence *Five men show three cars each to two customers* means that each of the three involved cars is shown to two possibly different customers (allowing for three cars and up to six customers). Since ACE does currently not allow to reverse the order of direct and indirect object the reading *Five men show two customers each three cards* cannot yet be easily expressed. Additionally, inference rules that control distributive subentailments concerning the direct object have to be formulated.

The interpretation of prepositional phrases with plural noun phrases like *The customers enter two cards into three slots* or *Two customers of three countries have seven cards* yet have to be worked out.

Cardinality quantifiers like *at least three, at most three, exactly two, three or more* can also lead to a collective/distributive ambiguity. In ACE they are currently treated analogous to indefinite determiners and require disambiguation markers if they occur in external argument position: *At least three customers (each) own a house*.

Measurement phrases like *five kilos of apples* are partially included in ACE. In ACE measurement phrases do not introduce collective/distributive ambiguities. They are analysed via measure relations. For instance, the predefined measure word *kilo* introduces a relation into the logical form that expresses the measure unit being used to count the objects denoted by the head noun (*apples* in our example). The numeral *five* indicates the amount. As a further extension, axioms for measure relations have to be formulated so that measures can be added, subtracted, multiplied or otherwise compared.

Bare plurals like *cars, students, books* etc. are a frequent source of ambiguities in full natural lan-

guage. Roughly speaking, bare plurals can be interpreted existentially or generically. In the sentence *John owns old cars* the bare plural *cars* expresses an existential reading, meaning that there are some old cars that John owns. In the sentence *Cars have four wheels* the bare plural *cars* is interpreted generically, meaning that all typical cars have four wheels. The generic interpretation very often occurs when bare plurals are subjects of stative verbs. To prevent ambiguities in ACE, users are strongly discouraged to use bare plurals at all. In many cases, bare plurals can be avoided by either using a singular indefinite article or a universal quantifier, e.g. *Visacards are credit cards* has to be expressed in ACE as *Every Visacard is a credit card*. There are, however, cases where bare plurals cannot be neatly avoided, especially in predicative position with a non-quantifying subject, e.g. *The cards are Visacards*. We therefore decided to allow bare plurals in ACE but always interpret them as indefinite noun phrases with an existential reading. A generic interpretation of bare plurals is not available.

ACE allows for a limited form of plural anaphora. In the sequence *John enters two cards. They are valid.* the second sentence is paraphrased as *[The two cards] are valid*. The anaphora resolution is basically analogous to singular anaphora resolution: A plural anaphor refers to the closest preceding accessible plural noun phrase. Anaphoric reference to split antecedents, e.g. *John likes Mary. They are in London.*, is not allowed in ACE to avoid possible ambiguities.

Further issues are to predefine the meaning of modifiers that often occur in connection with plurals. Among these are temporal modifiers like *simultaneously* which modify the relative temporal order of the subevents of a complex event, e.g. *John enters two cards simultaneously*. Temporal modifiers have to be distinguished from precisification markers like *as a whole*. Things can be done simultaneously, yet not necessarily involving the objects to be treated as a whole, e.g. *John wraps two presents simultaneously*. The meaning of other modifiers like *separately, individually, together* is interesting in this respect as well.

Furthermore, coordination of noun phrases and enumerations are an important topic for the treatment of plurals in ACE.

4 Conclusion

The best solution to the problem of plural ambiguity is of course to avoid plurals wherever possible and use singular sentences instead. Unfortunately, there are phenomena that cannot be simplified this way because there are properties that only hold of groups and cannot be reduced to the properties of the members of this group.

We observed that the problem of plural ambiguity is particularly tricky in complicated cases, whereas the majority of cases tends to be simple and does not pose problems for humans. Our approach therefore tries to give a simple theory of plurals that a user of ACE can realistically learn and use. The approach is based on the avoidance of plurals wherever possible, on a small set of interpretation rules, on the permission of indeterminacy, and, as a new and important element of ACE, on the addition of explicit predefined markers. These simple elements combined with the constructive disambiguation strategy of ACE allow for an automatic processing of plural sentences that does not face the problem of a combinatorial explosion of readings. We are aware that the plural fragment of ACE has to be extended before a realistic assessment of the approach can be made – particularly as for the question whether the proposed disambiguation triggers suffice.

Any elements added to the plural disambiguation algorithm must, however, conform to our constructive disambiguation approach which is a concession to the observation that human disambiguation knowledge can (currently) not be reliably formalized.

Acknowledgements

The basis of the singular fragment of ACE has been developed and implemented by Norbert Fuchs and Rolf Schwitter before I joined the Attempto project. For extensions concerning the plural fragment I profited a lot from discussions with Norbert Fuchs. I would also like to thank three anonymous reviewers for their helpful comments.

References

- Alshawi, H. (1992). *The Core Language Engine*. MIT press, Cambridge Mass.
- CLAW (1998). *Proceedings of the 2nd International Workshop on Controlled Language Applications*. Pittsburgh, Pennsylvania.
- FRACAS Consortium (1994). *The State of the Art in Computational Semantics: Evaluating the Descriptive Capabilities of Semantic Theories*. FRACAS Deliverable D9, Centre for Cognitive Science, Edinburgh.
- Fuchs, N. E. (1992). Specifications are (Preferably) Executable. *Software Engineering Journal*, **7** (5), 323-334.
- Fuchs, N. E., Schwertel, U. and Schwitter, R. (1999a). Attempto Controlled English – Not Just Another Logic Specification Language. In P. Flener, ed., *Logic-Based Program Synthesis and Transformation*, 8th International Workshop LOPST98, LNCS 1559, 1-20, Springer-Verlag, Berlin, Heidelberg.
- Fuchs, N. E., Schwertel, U. and Schwitter, R. (1999b). *Attempto Controlled English (ACE), Language Manual Version 3.0*. Technical Report 99.03, Department of Computer Science, University of Zurich. Available via http://www.ifi.unizh.ch/techreports/TR_1999.html.
- Hamm, F. and Hinrichs, E. (1998). *Plurality and Quantification*. Studies in Linguistics and Philosophy 69. Kluwer, Dordrecht.
- Hindle, D. and Rooth, M. (1993). Structural Ambiguity and Lexical Relations. *Computational Linguistics*, **19** (1), 103-120.
- Kamp, H. and Reyle, U. (1993). *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Studies in Linguistics and Philosophy 42. Kluwer, Dordrecht.
- Link, G. (1991). Plural. In Wunderlich and von Stechow (1991).
- Link, G. (1998). Ten Years of Research on Plurals – Where Do We Stand? In Hamm and Hinrichs (1998).

- Lønning, J. T. (1987). Collective Readings of Definite and Indefinite Noun Phrases. In P. Gärdenfors, ed., *Generalized Quantifiers. Linguistic and Logical Approaches*. Dordrecht, Reidel.
- Macias, B. and Pulman, S. G. (1995). A Method for Controlling the Production of Specifications in Natural Language. *The Computer Journal*, **38** (4), 310-318.
- Reyle, U. (1993). Dealing with Ambiguities by Underspecification: Construction, Representation and Deduction. *Journal of Semantics*, **10**, 123-178.
- Reyle, U. (1994). Monotonic Disambiguation and Plural Pronoun Resolution. In H. Kamp, ed., *Ellipsis, Tense and Questions*, DYANA-2 Deliverable R2.2.B. University of Stuttgart.
- Roberts, C. (1987). *Modal Subordination, Anaphora and Distributivity*. Ph.D. Thesis, University of Massachusetts, Amherst.
- Scha, R. (1981). Distributive, Collective and Cumulative Quantification. In Groenendijk, J., Janssen, T. and Stokhof, M., eds., *Formal Methods in the Study of Language*. Amsterdam, Mathematical Centre, 483-512.
- Schwitler, R. (1998). *Kontrolliertes Englisch für Anforderungsspezifikationen*. Ph.D. Thesis, Department of Computer Science, University of Zurich.
- van den Berg (1994). Plurality. In *The Encyclopedia of Language and Linguistics*, 6, 3198-3200. Pergamon, Oxford.
- Verkuyl, H. (1999). *Aspectual Issues. Studies on Time and Quantity*. CSLI Publications, Stanford, California.
- Wunderlich, D. and von Stechow, A., eds. (1991). *Semantics. Ein internationales Handbuch der zeitgenössischen Forschung*. Handbücher zur Sprach- und Kommunikationswissenschaft 6.W. de Gruyter, Berlin, New York.