Resolving Ellipsis in Clarification

Jonathan Ginzburg
Dept of Computer Science
King's College, London
The Strand, London WC2R 2LS
UK
ginzburg@dcs.kcl.ac.uk

Robin Cooper
Dept of Linguistics
Göteborg University
Box 200, 405 30 Göteborg,
Sweden
cooper@ling.gu.se

Abstract

We offer a computational analysis of the resolution of ellipsis in certain cases of dialogue clarification. We show that this goes beyond standard techniques used in anaphora and ellipsis resolution and requires operations on highly structured, linguistically heterogeneous representations. We characterize these operations and the representations on which they operate. We offer an analysis couched in a version of Head-Driven Phrase Structure Grammar combined with a theory of information states (IS) in dialogue. We sketch an algorithm for the process of utterance integration in ISs which leads to grounding or clarification.

1 Introduction

Clarification ellipsis (CE), nonsentential elliptical queries such as (1a(i),(ii)) are commonplace in human conversation. Two common readings/understandings of CE are exemplified in (1b,c): the *clausal* reading is commonly used simply to confirm the content of a particular subutterance. The main function of the *constituent* reading is to elicit an alternative description or ostension to the content (referent or predicate etc) intended by the original speaker of the reprised subutterance.

(1)   a. A: Did Bo finagle a raise?
        B: (i) Bo?/ (ii) finagle?

   b. **Clausal reading**: Are you asking if BO (of all people) finagled a raise/Bo FINAGLED a raise (of all actions)
   c. **Constituent reading**: Who is Bo?/What does it mean to finagle?

The issue of whether CE involves an ambiguity or is simply vague is an important one.[1,2] Clearly, pragmatic reasoning plays an important role in understanding CEs. Some considerations do, nonetheless, favour the existence of an ambiguity. First, the BNC provides numerous examples of misunderstandings concerning CE interpretation,[3] where a speaker intends one reading, is misunderstood, and clarifies his original interpretation:

(2)   a. A: ... you always had er er say every foot he had with a piece of spunyarn in the wire/B: Spunyarn?/A: Spunyarn, yes/B: What's spunyarn?

   b. A: Have a laugh and joke with Dick./ B: Dick?/A: Have a laugh and joke with Dick./B: Who's Dick?

---

[1] An anonymous ACL reviewer proposed to us that all CE could be analyzed in terms of a single reading along the lines of "I thought I heard you say Bo, and I don't know why you would do so?".

[2] Closely related to this issue is the issue of what other readings/understandings CE exhibits. We defer discussion of the latter issue to (Purver et al., 2001), which provides a detailed analysis of the frequency of CEs and their understandings among clarification utterances in the British National Corpus (BNC).

[3] This confirms our (non-instrumentally tested) impression that these understandings are *not* on the whole disambiguated intonationally. All our CE data from the BNC was found using SCoRE, Matt Purver's dialogue oriented BNC search engine (Purver, 2001).

More crucially, the clausal and constituent readings involve distinct syntactic and phonological parallelism conditions. The constituent reading seems to actually require phonological identity. With the resolution associated with clausal readings, there is no such requirement. However, partial syntactic parallelism does obtain: an XP used to clarify an antecedent sub-utterance u_1 must match u_1 categorially, though there is no requirement of *phonological* identity:

(3)   a. A: I phoned him. B: him? / #he?
         b. A: Did he adore the book. B: adore? / #adored?
         c. A: We're leaving? B: You?

We are used to systems that will confirm the user's utterances by repeating part of them. These presuppose no sophisticated linguistic analysis. However, it is not usual for a system to be able to process CEs produced by the user. It would be a great advantage in negotiative dialogues, where, for example, the system and the user might be discussing several options and the system may make alternative suggestions, for a system to be able to recognize and interpret a CE. Consider the following (constructed) dialogue in the route-planning domain:

(4) Sys: Would you like to make that trip via Malvern? User: Malvern?

At this point the system has to consider a number of possible intepretations for the user's utterance all of which involve recognizing that this is a clarification request concerning the system's last utterance.

Appropriate responses might be (5a-c); the system should definitely not say (5d), as it might if it does not recognize that the user is trying to clarify its previous utterance.

(5)   a. Yes, Malvern
         b. Malvern – M-A-L-V-E-R-N
         c. Going via Malvern is the quickest route
         d. So, you would like to make that trip via Malvern instead of Malvern?

In this paper we examine the interpretation of CEs. CE is a singularly complex ellipsis/anaphoric phenomenon which cannot be handled by standard techniques such as first order unification (as anaphora often is) or by higher order unification (HOU) on logical forms (see e.g. (Pulman, 1997)). For a start, in order to capture the syntactic and phonological parallelism exemplified in (3), logical forms are simply insufficient. Moreover, although an HOU account could, given a theory of dialogue that structures context appropriately, generate the clausal reading, the constituent reading cannot be so generated. Clark (e.g. (Clark, 1996)) initiated work on the grounding of an utterance (for computational and formal work see e.g. (Traum, 1994; Poesio and Traum, 1997)). However, existing work, while spelling out in great detail what updates arise in an IS as a result of grounding, do not offer a characterization of the clarification possibilities spawned by a given utterance. A sketch of such a characterization is provided in this paper. On the basis of this we offer an analysis of CE, integrated into a large existing grammar framework, Head-Driven Phrase Structure Grammar (HPSG) (specifically the version developed in (Ginzburg and Sag, 2000)). We start by informally describing the grounding/clarification processes and the representations on which they operate. We then provide the requisite background on HPSG and on the KOS framework (Ginzburg, 1996; Bohlin et al., 1999), in which our analysis of ISs is couched. We sketch an algorithm for the process of utterance integration which leads to grounding or clarification. Finally, we formalize the operations which underpin clarification and sketch a grammatical analysis of CE.

2 Utterance Representation: grounding and clarification

We start by offering an informal description of how an utterance u such as (6) can get grounded or spawn a clarification by an addressee B:

(6) A: Did Bo leave?

A is attempting to convey to B her question whether the property she has referred to with her utterance of *leave* holds of the person she has referred to with the name *Bo*. B is required to try and find values for these references. Finding values is, with an important caveat, a necessary condition for B to ground A's utterance, thereby signalling that its content has been integrated in

B's IS.[4] Modelling this condition for successful grounding provides one obvious constraint on the representation of utterance types: such a representation must involve a function from or λ-abstract over a set of certain parameters (the *contextual parameters*) to contents. This much is familiar already from early work on context dependence by (Montague, 1974) et seq. What happens when B cannot or is at least uncertain as to how he should instantiate in his IS a contextual parameter i? In such a case B needs to do at least the following: (1) perform a partial update of the existing context with the successfully processed components of the utterance (2) pose a clarification question that involves reference to the sub-utterance u_i from which i emanates. Since the original speaker, A, can coherently integrate a clarification question once she hears it, it follows that, for a given utterance, there is a predictable range of $<$ partial updates + consequent clarification questions $>$. These we take to be specified by a set of *coercion operations* on utterance representations.[5] Indeed we assume that a component of dialogue competence is knowledge of these *coercion operations*.

CE gives us some indication concerning both the input and required output of these operations. One such operation, which we will refer to as parameter identification, essentially involves as output a question paraphrasable as *what is the intended reference of sub-utterance u_i?*. The partially updated context in which such a clarification takes place is such that simply repeating the segmental phonology of u_i using rising intonation enables that question to be expressed. Another existent coercion operation is one which we will refer to as parameter focussing. This involves a (partially updated) context in which the issue under discussion is a question that arises by instantiating all contextual parameters except for i and abstracting over i. In such a context, one

can *confirm* that i gets the value B suspects it has by uttering with rising intonation any apparently co-referential phrase whose syntactic category is identical to u_1's.

From this discussion, it becomes clear that coercion operations (and by extension the grounding process) cannot be defined simply on meanings. Rather, given the syntactic and phonological parallelism encoded in clarification contexts, these operations need to be defined on representations that encode in parallel for each sub-utterance down to the word level phonological, syntactic, semantic, and contextual information. With some minor modifications, signs as conceived in HPSG are exactly such a representational format and, hence, we will use them to define coercion operations.[6] More precisely, given that an addressee might not be able to come up with a unique or a complete parse, due to lexical ignorance or a noisy environment, we need to utilize some 'underspecified' entity (see e.g. (Milward, 2000)). For simplicity we will use *descriptions* of signs. An example of the format for signs we employ is given in (7):[7]

---

[4]The caveat is, of course, that the necessity is goal driven. Relative to certain goals, one might decide simply to existentially quantify the problematic referent. For this operation on meanings see (Cooper, 1998). We cannot enter here into a discussion of how to integrate the view developed here in a plan based view of understanding, but see (Ginzburg, (forthcoming)) for this.

[5]The term *coercion operation* is inspired by work on utterance representation within a type theoretic framework reported in (Cooper, 1998).

[6]We make two minor modifications to the version of HPSG described in (Ginzburg and Sag, 2000). First, we revamp the existing treatment of the feature C-INDICES. This will now encode the entire inventory of contextual parameters of an utterance (proper names, deictic pronouns, indexicals) not merely information about speaker/hearer/utterance-time, as standardly. Indeed, in principle, relation names should also be included, since they vary with context and are subject to clarification as well. Such a step involves a significant change to how argument roles are handled in existing HPSG. Hence, we do not make such a move here. This modification of C-INDICES will allow signs to play a role akin to the role associated with 'meanings', i.e. to function as abstracts with roles that need to be instantiated. The second modification we make concerns the encoding of phrasal constituency. Standardly, the feature DTRS is used to encode immediate phrasal constituency. To facilitate statement of coercion operations, we need access to all phrasal constituents—given that a contextual parameter emanating from deeply embedding constituents are as clarifiable as immediate constituents. We posit a set valued feature CONSTIT(UENT)S whose value is the set of all constituents immediate or otherwise of a given sign (Cf. the mother-daughter predicates used in (Gregory and Lappin, 1999).) In fact, having posited CONSTITS one could eliminate DTRS: this by making the value of CONSTITS be a set of sets whose first level elements are the immediate constituents. For current purposes, we stick with tradition and tolerate the redundancy of both DTRS and CONSTITS.

[7]Within the phrasal type system of (Ginzburg and Sag, 2000) *root-cl* constitutes the 'start' symbol of the grammar. In particular, phrases of this type have as their content an illocutionary operator embedding the appropriate semantic

(7)
$$
\begin{bmatrix}
\textit{root-cl} \\
\text{PHON did bo leave} \\
\text{CAT V[+fin]} \\
\text{C-INDICES } \{\boxed{1},\boxed{2},\boxed{3}, i,j\} \\
\text{CONT }
\begin{bmatrix}
\text{ASK-REL} \\
\text{ASKER } i \\
\text{ASKED } j \\
\text{MSG-ARG }
\begin{bmatrix}
\textit{question} \\
\text{PARAMS } \{\,\} \\
\text{PROP}|\text{SOA}
\begin{bmatrix}
\textit{leave-rel} \\
\text{AGT} \quad \boxed{1} \\
\text{TIME} \quad \boxed{2}
\end{bmatrix}
\end{bmatrix}
\end{bmatrix} \\
\text{CTXT}|\text{BCKGRD }
\left\{
\begin{array}{l}
\textit{utt-time}(\boxed{3}), \\
\textit{precede}(\boxed{2},\boxed{3}), \textit{named}(\text{bo})(\boxed{1})
\end{array}
\right\} \\
\text{CONSTITS }
\left\{
\begin{array}{l}
\boxed{4}[\text{PHON Did}], \boxed{5}[\text{PHON Bo}], \\
\boxed{6}[\text{PHON leave}], \boxed{7}[\text{PHON Did Bo leave}]
\end{array}
\right\}
\end{bmatrix}
$$

Before we can explain how these representations can feature in dialogue reasoning and the resolution of CE, we need to sketch briefly the approach to dialogue ellipsis that we assume.

3 Contextual evolution and ellipsis

We adopt the situation semantics based theory of dialogue context developed in the KOS framework (Ginzburg, 1996; Ginzburg, (forthcoming); Bohlin et al., 1999). The common ground component of ISs is assumed to be structured as follows:[8]

(8)
$$
\begin{bmatrix}
\text{FACTS} & \textit{set of facts} \\
\text{LATEST-MOVE} & \textit{(illocutionary) fact} \\
\text{QUD} & \textit{p.o. set of questions}
\end{bmatrix}
$$

In (Ginzburg and Sag, 2000) this framework is integrated into HPSG (Pollard and Sag, 1994); (Ginzburg and Sag, 2000) define two new attributes within the CONTEXT (CTXT) feature structure: Maximal Question Under Discussion (MAX-QUD), whose value is of sort *question*;[9]

and Salient Utterance (SAL-UTT), whose value is a set (singleton or empty) of elements of type *sign*. In information structure terms, SAL-UTT can be thought of as a means of underspecifying the subsequent focal (sub)utterance or as a potential *parallel element*. MAX-QUD corresponds to the ground of the dialogue at a given point. Since SAL-UTT is a *sign*, it enables one to encode syntactic *categorial* parallelism and, as we will see below, also phonological parallelism. SAL-UTT is computed as the (sub)utterance associated with the role bearing widest scope within MAX-QUD.[10] Below, we will show how to extend this account of parallelism to clarification queries.

To account for elliptical constructions such as short answers and sluicing, Ginzburg and Sag posit a phrasal type *headed-fragment-phrase* (*hd-frag-ph*)—a subtype of *hd-only-ph*—governed by the constraint in (9). The various fragments analyzed here will be subtypes of *hd-frag-ph* or else will contain such a phrase as a head daughter.[11]
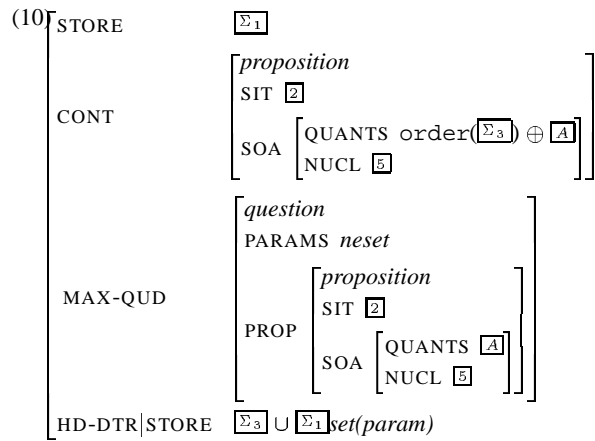
(9)
$$
\begin{bmatrix}
\text{HEAD } v \\
\text{CTXT}|\text{SAL-UTT}
\begin{bmatrix}
\text{CAT} & \boxed{1} \\
\text{CONT}|\text{INDEX} & \boxed{2}
\end{bmatrix} \\
\text{HD-DTR}
\begin{bmatrix}
\text{CAT } \boxed{1}[\text{HEAD } \textit{nominal}] \\
\text{CONT}|\text{INDEX} \ \boxed{2}
\end{bmatrix}
\end{bmatrix}
$$

This constraint coindexes the head daughter with the SAL-UTT. This will have the effect of 'unifying in' the content of the former into a contextually provided content. A subtype of *hd-frag-ph* relevant to the current paper is (*decl-frag-cl*)—also a subtype of *decl-cl*—used to analyze short answers:

---

object (an assertion embedding a proposition, a query embedding a question etc.). Here and throughout we omit various features (e.g. STORE, SLASH etc that have no bearing on current issues wherever possible.

[8]Here FACTS corresponds to the set of commonly accepted assumptions; QUD('questions under discussion') is a set consisting of the currently discussable questions, partially ordered by $\prec$ ('takes conversational precedence'); LATEST-MOVE represents information about the content and structure of the most recent accepted illocutionary move.

[9]Questions are represented as semantic objects comprising a set of parameters—empty for a polar question—and a

proposition. This is the feature structure counterpart of the λ-abstract $\lambda\pi(\ldots\pi\ldots)$.

[10]For *Wh*-questions, SAL-UTT is the *wh*-phrase associated with the PARAMS set of the question; otherwise, its possible values are either the empty set or the utterance associated with the widest scoping quantifier in MAX-QUD.

[11]In the (Ginzburg and Sag, 2000) version of HPSG information about phrases is encoded by cross-classifying them in a multi-dimensional type hierarchy. Phrases are classified not only in terms of their phrase structure schema or X-bar type, but also with respect to a further informational dimension of CLAUSALITY. Clauses are divided into *inter alia* declarative clauses (*decl-cl*), which denote propositions, and interrogative clauses (*inter-cl*) denoting questions. Each maximal phrasal type inherits from both these dimensions. This classification allows specification of systematic correlations between clausal construction types and types of semantic content.

(10)
$$
\begin{bmatrix}
\text{STORE} & \boxed{\Sigma_1} \\[4pt]
\text{CONT} & \begin{bmatrix}
\textit{proposition} \\
\text{SIT } \boxed{2} \\
\text{SOA } \begin{bmatrix} \text{QUANTS order}(\boxed{\Sigma_3}) \oplus \boxed{A} \\ \text{NUCL } \boxed{5} \end{bmatrix}
\end{bmatrix} \\[20pt]
\text{MAX-QUD} & \begin{bmatrix}
\textit{question} \\
\text{PARAMS } \textit{neset} \\
\text{PROP } \begin{bmatrix}
\textit{proposition} \\
\text{SIT } \boxed{2} \\
\text{SOA } \begin{bmatrix} \text{QUANTS } \boxed{A} \\ \text{NUCL } \boxed{5} \end{bmatrix}
\end{bmatrix}
\end{bmatrix} \\[20pt]
\text{HD-DTR}|\text{STORE} & \boxed{\Sigma_3} \cup \boxed{\Sigma_1} \textit{set(param)}
\end{bmatrix}
$$

The content of this phrasal type is a proposition: whereas in most headed clauses the content is entirely (or primarily) derived from the head daughter, here it is constructed for the most part from the contextually salient question. This provides the concerned situation and the nucleus, whereas if the fragment is (or contains) a quantifier, that quantifier must outscope any quantifiers already present in the contextually salient question.

4   Integrating Utterances in Information States

Before we turn to formalizing the coercion operations and describing CE, we need to explain how on our view utterances get integrated in an agent's IS. The basic protocol we assume is given in (11) below.[12]

(11)   **Utterance processing protocol**

For an agent B with IS I: if an utterance u is Maximal in PENDING:
(a) Try to:
(1) find an assignment f in I for σ, where σ is the (maximal description available for) the sign associated with u
(2) update LATEST-MOVE with u:

1. If LATEST-MOVE is grounded, then FACTS:= FACTS + LATEST-MOVE;

2. LATEST-MOVE := $\langle \sigma, f \rangle$

(3) React to content(u) according to querying/assertion protocols.
(4) If successful, u is removed from PENDING
(b) Else: Repeat from stage (a) with MAX-QUD and SAL-UTT obtaining the various values of $\text{coe}^i(\tau)|max - qud/sal - utt$, where τ is the sign associated with LATEST-MOVE and coe^i is one of the available coercion operations;

(c) Else: make an utterance appropriate for a context such that MAX-QUD and SAL-UTT get values according to the specification in $\text{coe}^i(u, \sigma)$, where coe^i is one of the available coercion operations.
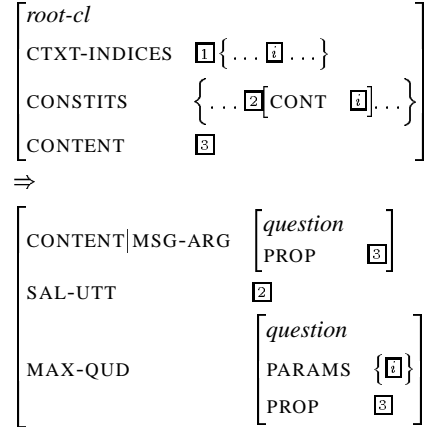
The protocol involves the assumption that an agent always initially tries to integrate an utterance by assuming it constitutes an adjacency pair with the existing LATEST-MOVE. If this route is blocked somehow, because the current utterance cannot be grounded or the putative resolution leads to incoherence, only then does she try to repair by assuming the previous utterance is a clarification generated in accordance with the existing coercion operations. If that too fails, then, she herself generates a clarification. Thus, the prediction made by this protocol is that A will tend to initially interpret (12(2)) as a response to her question, not as a clarification:

(12)   A(1): Who do you think is the only person that admires Mary? B(2): Mary?

5   Sign Coercion and an Analysis of CE

We now turn to formalizing the coercion operations we specified informally in section 2. The first operation we define is parameter focussing:

(13)   parameter focussing$_i$:
$$
\begin{bmatrix}
\textit{root-cl} \\
\text{CTXT-INDICES} & \boxed{1}\{\ldots\boxed{i}\ldots\} \\
\text{CONSTITS} & \{\ldots\boxed{2}[\text{CONT } \boxed{i}]\ldots\} \\
\text{CONTENT} & \boxed{3}
\end{bmatrix}
$$
$\Rightarrow$
$$
\begin{bmatrix}
\text{CONTENT}|\text{MSG-ARG} & \begin{bmatrix} \textit{question} \\ \text{PROP } \boxed{3} \end{bmatrix} \\[10pt]
\text{SAL-UTT} & \boxed{2} \\[6pt]
\text{MAX-QUD} & \begin{bmatrix} \textit{question} \\ \text{PARAMS } \{\boxed{i}\} \\ \text{PROP } \boxed{3} \end{bmatrix}
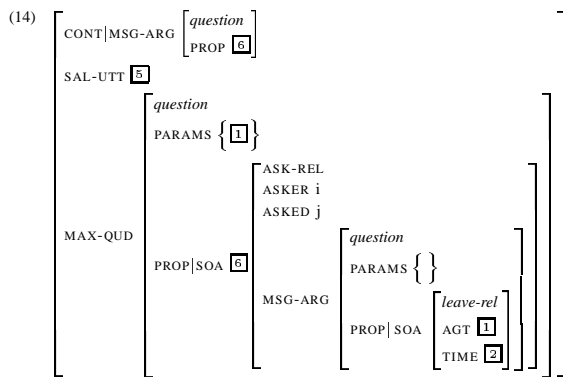\end{bmatrix}
$$

This is to be understood as follows: given an utterance (whose associated sign is one) which satisfies the specification in the LHS of the rule, a CP may respond with any utterance which satisfies the specification in the RHS of the rule.[13] More specifically, the input of the rules singles out a

---

contextual parameter i, which is the content of an element of the daughter set of the utterance $\boxed{2}$. Intuitively, i is a parameter whose value is problematic or lacking. The sub-utterance $\boxed{2}$ is specified to constitute the value of the feature SAL-UTT associated with the context of the clarification utterance cu_0. The descriptive content of cu_0 is a question, any question whose open proposition $\boxed{3}$ (given in terms of the feature PROP) is identical to the (uninstantiated) content of the clarified utterance. MAX-QUD associated with the clarification is fully specified as a question whose open proposition is $\boxed{3}$ and whose PARAMS set consists of the 'problematic' parameter i.

We can exemplify the effect of **parameter focussing** with respect to clarifying an utterance of (7). The output this yields, when applied to *Bo*'s index $\boxed{1}$, is the partial specification in (14). Such an utterance will have as its MAX-QUD a question cq_0 paraphrasable as *who$_t$, named Bo, are you asking if t left*, whereas its SAL-UTT is the sub-utterance of *Bo*. The content is underspecified:

(14)
$$
\begin{bmatrix}
\text{CONT}|\text{MSG-ARG} \begin{bmatrix} \textit{question} \\ \text{PROP} \boxed{6} \end{bmatrix} \\
\text{SAL-UTT } \boxed{5} \\
\text{MAX-QUD} \begin{bmatrix} \textit{question} \\ \text{PARAMS} \left\{ \boxed{1} \right\} \\ \text{PROP}|\text{SOA } \boxed{6} \begin{bmatrix} \text{ASK-REL} \\ \text{ASKER i} \\ \text{ASKED j} \\ \text{MSG-ARG} \begin{bmatrix} \textit{question} \\ \text{PARAMS} \{\,\} \\ \text{PROP}|\text{SOA} \begin{bmatrix} \textit{leave-rel} \\ \text{AGT } \boxed{1} \\ \text{TIME } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}
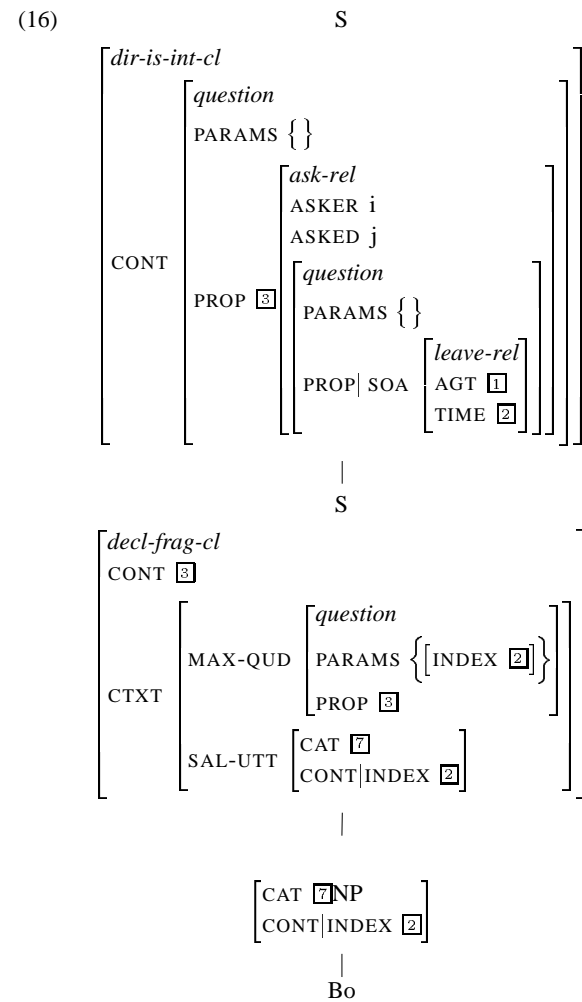\end{bmatrix}
$$

This (partial) specification allows for clarification questions such as the following:

(15)   a. Did WHO leave?
       b. WHO?
       c. BO? (= Are you asking if BO left?)

Given space constraints, we restrict ourselves to explaining how the clausal CE, (15c), gets analyzed. This involves direct application of the type *decl-frag-cl* discussed above for short answers. The QUD-maximality of cq_0 allows us to analyze the fragment as a 'short answer' to cq_0, using the type *bare-decl-cl*. And out of the proposition which emerges courtesy of *bare-decl-cl* a (polar)

question is constructed using the type *dir-is-int-cl*.[14]

(16)
$$
S
$$
$$
\begin{bmatrix}
\textit{dir-is-int-cl} \\
\text{CONT} \begin{bmatrix} \textit{question} \\ \text{PARAMS} \{\,\} \\ \text{PROP } \boxed{3} \begin{bmatrix} \textit{ask-rel} \\ \text{ASKER i} \\ \text{ASKED j} \\ \text{PROP}|\text{SOA} \begin{bmatrix} \textit{question} \\ \text{PARAMS} \{\,\} \\ \text{PROP}|\text{SOA} \begin{bmatrix} \textit{leave-rel} \\ \text{AGT } \boxed{1} \\ \text{TIME } \boxed{2} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$
$$
|
$$
$$
S
$$
$$
\begin{bmatrix}
\textit{decl-frag-cl} \\
\text{CONT } \boxed{3} \\
\text{CTXT} \begin{bmatrix} \text{MAX-QUD} \begin{bmatrix} \textit{question} \\ \text{PARAMS} \left\{ \begin{bmatrix} \text{INDEX } \boxed{2} \end{bmatrix} \right\} \\ \text{PROP } \boxed{3} \end{bmatrix} \\ \text{SAL-UTT} \begin{bmatrix} \text{CAT } \boxed{7} \\ \text{CONT}|\text{INDEX } \boxed{2} \end{bmatrix} \end{bmatrix}
\end{bmatrix}
$$
$$
|
$$
$$
\begin{bmatrix}
\text{CAT } \boxed{7}\,\text{NP} \\
\text{CONT}|\text{INDEX } \boxed{2}
\end{bmatrix}
$$
$$
|
$$
$$
Bo
$$

The second coercion operation we discussed previously is **parameter identification**: for a given problematic contextual parameter its output is a partial specification for a sign whose content and MAX-QUD involve a question querying the content of that utterance parameter:

---

[14]The phrasal type *dir-is-int-cl* which constitutes the type of the mother node in (16) is a type that *inter alia* enables a polar question to be built from a head daughter whose content is propositional. See (Ginzburg and Sag, 2000) for details.

(17)   parameter identification$_i$:

$$\begin{bmatrix} \textit{root-cl} \\ \text{CTXT-INDICES} \ \{\dots \boxed{1} \dots\} \\ \text{CONSTITS} \ \{\dots \boxed{2}[\text{CONT} \quad \boxed{4}]\dots\} \\ \dots \end{bmatrix}$$

$$\Rightarrow$$

$$\begin{bmatrix} \text{CONTENT}|\text{MSG-ARG} \ \begin{bmatrix} \textit{question} \\ \text{PROP} \quad \boxed{3} \end{bmatrix} \\ \text{SAL-UTT} \ \boxed{2} \\ \text{MAX-QUD} \ \begin{bmatrix} \textit{question} \\ \text{PARAMS} \ \{[\text{INDEX} \quad \boxed{4}]\} \\ \text{PROP} \ \boxed{3} \begin{bmatrix} \text{SOA} \begin{bmatrix} \textit{content-rel} \\ \text{SIGN} \quad \boxed{2} \\ \text{CONT} \quad \boxed{4} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

To exemplify: when this operation is applied to (7), it will yield as output the partial specification in (18):

(18)
$$\begin{bmatrix} \text{CONT}|\ \text{MSG-ARG} \ \begin{bmatrix} \textit{question} \\ \text{PROP} \ \boxed{6} \end{bmatrix} \\ \text{SAL-UTT} \ \boxed{5} \begin{bmatrix} \text{PHON bo} \\ \text{CAT NP} \\ \text{CONT}|\text{INDEX} \ \boxed{i} \\ \text{CTXT}|\text{BCKGRD} \ \{\textit{named(Bo)}(\boxed{i})\} \end{bmatrix} \\ \text{MAX-QUD} \ \begin{bmatrix} \textit{question} \\ \text{PARAMS} \ \{[\text{INDEX} \ \boxed{8}]\} \\ \text{PROP} \ \boxed{6} \begin{bmatrix} \text{SOA} \begin{bmatrix} \textit{content-rel} \\ \text{SIGN} \ \boxed{5} \\ \text{CONT} \ \boxed{8} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

This specification will allows for clarification questions such as the following:

(19)   a. Who do you mean BO?
        b. WHO? (= who is Bo)
        c. Bo? (= who is Bo)

We restrict attention to (19c), which is the most interesting but also tricky example. The tricky part arises from the fact that in a case such as this, in contrast to all previous examples, *the fragment does not contribute its conventional content to the clausal content.* Rather, as we suggested earlier, the semantic function of the fragment is merely to serve as an anaphoric element to the phonologically identical to–be–clarified sub-utterance. The content derives entirely from MAX-QUD. Such utterances can still be analyzed as subtypes of *head-frag-ph*, though not as *decl-frag-cl*, the short-answer/reprise sluice phrasal type we have been appealing to extensively. Thus, we posit *constit(uent)-clar(ification)-int-cl*, a new phrasal subtype of *head-frag-ph* and of *inter-cl* which encapsulates the two idiosyncratic facets of such utterances, namely the phonological parallelism and the max-qud/content identity:

(20)
$$\begin{bmatrix} \text{CONT} \quad \boxed{1} \\ \text{CTXT} \begin{bmatrix} \text{MAX-QUD} \quad \boxed{1} \\ \text{SAL-UTT}|\text{PHON} \quad \boxed{2} \end{bmatrix} \end{bmatrix} \rightarrow \mathbf{H}[\text{PHON} \ \boxed{2}]$$

Given this, (19c) receives the following analysis:

(21)
$$\begin{bmatrix} \textit{constit-repr-int-cl} \\ \text{CONT} \quad \boxed{1} \begin{bmatrix} \textit{question} \\ \text{PARAMS} \ \{\boxed{2}\} \\ \text{PROP} \ \textit{content}(\boxed{3},\boxed{2}) \end{bmatrix} \\ \text{CTXT} \begin{bmatrix} \text{MAX-QUD} \ \boxed{1} \\ \text{SAL-UTT} \ \boxed{3} \begin{bmatrix} \text{PHON} \ \boxed{5} \\ \text{CAT} \quad \boxed{4}\text{NP} \end{bmatrix} \end{bmatrix} \\ \text{HD-DTR} \begin{bmatrix} \text{PHON} \ \boxed{5} \\ \text{CAT} \quad \boxed{4} \end{bmatrix} \end{bmatrix}$$

6 Summary and Future Work

In this paper we offered an analysis of the types of representations needed to analyze CE, the requisite operations thereon, and how these update ISs during grounding and clarification.

Systems which respond appropriately to CEs in general will need a great deal of background knowledge. Even choosing among the responses in (5) might be a pretty knowledge intensive business. However, there are some clear strategies that might be pursued. For example, if Malvern has been discussed previously in the dialogue and understood then (5a,b) would not be appropriate responses. In order to be able to build dialogue systems that can handle even some restricted aspects of CEs we need to understand more about what the possible interpretations are and this is what we have attempted to do in this paper. We are currently working on a system which integrates SHARDS (see (Ginzburg et al., 2001), a system which processes dialogue ellipses) with GoDiS (see (Bohlin et al., 1999), a dialogue system developed using TRINDIKIT, which makes use of ISs modelled on those suggested in the KOS

framework. Our aim in the near future is to incorporate simple aspects of negotiative dialogue including CEs in a GoDiS-like system employing SHARDS.

Acknowledgements

References

Peter Bohlin, Robin Cooper, Elisabet Engdahl, and Staffan Larsson. 1999. Information states and dialogue move engines. *Gothenburg Papers in Computational Linguistics*.

Herbert Clark. 1996. *Using Language*. Cambridge University Press, Cambridge.

Robin Cooper. 1998. Mixing situation theory and type theory to formalize information states in dialogue exchanges. In J. Hulstijn and A. Nijholt, editors, *Proceedings of TwenDial 98, 13th Twente workshop on Language Technology*. Twente University, Twente.

Jonathan Ginzburg and Ivan Sag. 2000. *Interrogative Investigations: the form, meaning and use of English Interrogatives*. Number 123 in CSLI Lecture Notes. CSLI Publications, Stanford: California.

Jonathan Ginzburg, Howard Gregory, and Shalom Lappin. 2001. Shards: Fragment resolution in dialogue. In H. Bunt, editor, *Proceedings of the 1st International Workshop on Computational Semantics*. ITK, Tilburg University, Tilburg.

Jonathan Ginzburg. 1996. Interrogatives: Questions, facts, and dialogue. In Shalom Lappin, editor, *Handbook of Contemporary Semantic Theory*. Blackwell, Oxford.

Jonathan Ginzburg. forthcoming. *Semantics and Interaction in Dialogue*. CSLI Publications and Cambridge University Press, Stanford: California. Draft chapters available from http://www.dcs.kcl.ac.uk/staff/ginzburg.

Howard Gregory and Shalom Lappin. 1999. Antecedent contained ellipsis in HPSG. In Gert Webelhuth, Jean Pierre Koenig, and Andreas Kathol, editors, *Lexical and Constructional Aspects of Linguistic Explanation*, pages 331–356. CSLI Publications, Stanford.

David Milward. 2000. Distributing representation for robust interpretation of dialogue utterances. *ACL*.

Richard Montague. 1974. Pragmatics. In Richmond Thomason, editor, *Formal Philosophy*. Yale UP, New Haven.

Massimo Poesio and David Traum. 1997. Conversational actions and discourse situations. *Computational Intelligence*, 13:309–347.

Carl Pollard and Ivan Sag. 1994. *Head Driven Phrase Structure Grammar*. University of Chicago Press and CSLI, Chicago.

Stephen Pulman. 1997. Focus and higher order unification. *Linguistics and Philosophy*, 20.

Matthew Purver, Jonathan Ginzburg, and Patrick Healey. 2001. On the means for clarification in dialogue. Technical report, King's College, London.

Matthew Purver. 2001. Score: Searching a corpus for regular expressions. Technical report, King's College, London.

David Traum. 1994. *A Computational Theory of Grounding in Natural Language Conversations*. Ph.D. thesis, University of Rochester.