CLARIE: the Clarification Engine

Matthew Purver
Department of Computer Science
King's College London, Strand, London WC2R 2LS, UK
matthew.purver@kcl.ac.uk

Abstract

This paper describes the CLARIE system, a prototype information-state-based text dialogue system designed to deal with many types of *clarification requests* (CRs) by using a highly contextualised semantic representation together with a suitable grounding process. This allows it to interpret and respond to user CRs, and generate its own CRs in order to clarify unknown reference and learn new words, with both integrated within the standard dialogue update processes.

1 Introduction

CLARIE is a prototype information-state (IS)-based dialogue system designed to generate, interpret and respond to many types of *clarification requests* (CRs), allowing it to clarify problematic features of utterances – including unknown or surprising reference and meaning – and allowing users to do the same. This is achieved via a view of utterances as *contextual abstracts* requiring a *grounding* process to fully specify their content; a highly contextualised semantic representation including a view of ellipsis as abstraction; and a simple set of pragmatic contextual operations implemented as IS update rules. The system itself is implemented using the TrindiKit (Larsson et al., 2002), building upon the GoDiS dialogue system (Larsson et al., 2000) and SHARDS ellipsis reconstruction system (Ginzburg et al., 2001). Being a prototype, it is currently text-based and has only a small narrow-coverage grammar and a toy domain. This paper will concentrate on the novel semantic representation and the grounding process which enable its clarificational capabilities.

Motivation CRs (questions about a previous (sub-)utterance's meaning or form) are common in dialogue (3-4% of human-human dialogue turns according to a corpus study (Purver et al., 2003)) but are often not paid a great deal of theoretical or implementational attention. Dialogue systems generally have the capability of indicating inability to recognize or understand an entire user turn (or inability to do so to a reasonable degree of confidence), and will usually be able to produce outputs like *"I did not understand what you said. Please rephrase"* or *"You want to go to Paris, is that right?"* (from IBiS, (Larsson, 2002)). However, they are not usually able to clarify problems in a finer-grained way (e.g. at the word or phrase level, as argued for by (Gabsdil, 2003)), nor to understand and respond to CRs generated by the user. While recent advances have led to some systems that can highlight problematic words (Hockey et al., 2002), or ask about NP reference (Traum, 2003), these are so far restricted to particular phenomena, and tend to treat CRs and clarificational dialogue as governed by different rules from standard questions and standard dialogue.[1] As the following imagined example (Stone, 2003) shows,

[1] Hockey et al. (2002) use a separate module to highlight a problematic word and suggest reformulation; Traum (2003) sees CRs and their answers as pairs of dedicated dialogue moves *request-repair* and *repair*, and restricts them to wh- or alternative-questions about NP reference.

clarification need not be restricted to NPs, may involve extended sequences, and will ideally be seamlessly integrated within the dialogue.

(1)
Q: What do I do next?
A: Slide the sleeve onto the elbow.
Q: What do you mean sleeve?
A: That tube around the pipe at the joint.
Q: What do you mean slide?
A: Just push the sleeve gently over along the pipe.
Q: What do you mean onto?
A: The sleeve can hang there safely out of the way while you complete the repair.

Background While there has been extensive research into the possible levels of information which CRs can query, e.g. (Larsson, 2002; Gabsdil, 2003; Schlangen, 2004) but going back at least to (Clark, 1996), there has been little which examines the precise relation between their surface form and the question they ask.[2] A suitable analysis of CRs must provide two things: it must give a representation to normal utterances that explains how and why they can cause CRs; and it must allow CRs themselves (including their often elliptical forms) to be given a suitable representation.

Ginzburg and Cooper (2004) (hereafter G&C) provide a HPSG analysis of CRs that promises both. Utterances are represented as encoding *meaning* rather than *content*: functions from context to fully specified content. Contextually dependent parameters such as the reference of proper names (as well as speaker, hearer and utterance time) are abstracted to a set expressed in HPSG terms as a C-PARAMS feature, but shown here as the abstracted set in a simultaneous λ-abstract.[3]

An utterance *"I want to go to Paris"* would be given a representation such as (3), or simplifying by removing the parameters for speaker and addressee as will be done hereafter, as in example (4):[4]

(3) $\lambda\Big\{a, b, x\Big\}[speaker(a),\ addressee(b),\ name(x, Paris)].assert(a, b, go\_to(a, x))$

(4) $\lambda\Big\{x\Big\}[name(x, Paris)].assert(a, b, go\_to(a, x))$

This abstract must then be *grounded* – the abstract applied to the context – in order to fully instantiate the content by finding a suitable referent x which satisfies the given restriction $name(x, Paris)$. If this cannot be done (the hearer may not know what/where Paris is, or perhaps instantiating x to Paris leads to this new assertion being inconsistent with previous beliefs), the utterance cannot be grounded and this can lead to a CR concerning the intended reference of the problematic parameter $[x : name(x, Paris)]$.

This CR may take many forms, one of the most common being an elliptical *reprise fragment*, an echo *"Paris?"* (although others are also possible including reprise sluices *"Where?"* and full reprise sentences *"You want to go to Paris?"* or *"You want to go where?"*). G&C analyse such reprises using a question-under-discussion (QUD)-based approach to ellipsis (Ginzburg et al., 2001): briefly, reprises and other elliptical fragments are given a content which depends on the maximal QUD and a salient utterance, encoded as MAX-QUD and SAL-UTT features which are taken to be provided by context. In the case of standard non-reprise fragments (such as a bare answer to a wh-question), values for these contextual features will be provided by standard dialogue mechanisms triggered by the prior asking of the question. In the case of CRs, failure of grounding for a particular parameter licenses one of a set of *coercion operations* which produce a context in which the values of MAX-QUD and SAL-UTT allow the fragment to be resolved as a question concerning the constituent associated with the problematic parameter.

They give two specific such operations, termed *parameter focussing* and *parameter identification*, which lead to different contexts and thus eventually lead to different reprise readings. The first

[2]Although see (Rodríguez and Schlangen, 2004) in this volume.

[3]More specifically, they are interpreted as *simultaneous abstracts with restriction* as shown in (2): $\{ABS\}$ is the set of abstracted indices, $[RESTR]$ a set of restrictions which must be satisfied during application, and $BODY$ the body of the abstract (in this case, the semantic content). For further formal details, see (Ginzburg and Sag, 2000).

(2) $\lambda\Big\{ABS\Big\}[RESTR].BODY$

[4]The representation of examples (3) and (4) is simplified for clarity; in particular the use of a go_to predicate ignores details of the representation of verbs and modification.

will be used in cases where grounding produces a surprising or inconsistent content: the new context makes the question *"For which X did you say you want to go to X?"* under discussion, resulting in an elliptical CR *"Paris?"* being resolved as asking the yes/no question *"Is it really **Paris** you are saying you want to go to?"*, what they call the *clausal* reading. The second will be used in cases where no referent for *Paris* can be found: it produces a context where the QUD, and the resolved content of the CR, is the *wh*-question *"What do you intend the word 'Paris' to refer to?"*, what they call the *constituent* reading. In this second case, the elliptical fragment must be given an *utterance-anaphoric* analysis, allowing it to refer to the previous utterance *Paris* and ask a question about its intended content.

2 Utterance Representation

The analysis of G&C applies (explicitly at least) only to proper names. The general approach has now been extended to cover a wide range of word and phrase types and a wide range of CR forms, together with an integrated account of ellipsis and reprises.

Contextual Abstraction Given the view of clarification as querying contextual parameters, a suitable semantic representation must require *all* those elements of an utterance with clarificational potential (i.e. that can function as *sources* of CRs) to be included in the abstracted set. This leads to a highly contextualised representation. As shown in (5) for a simple utterance *"The dog snores"*, the abstracted set must include not only the referents of proper names, but the referents of definite NPs, and the denotations of common nouns, verbs and even function words such as determiners, as any of these can be subsequently clarified:

(5) $$\lambda\Big\{w, Q, P, S\Big\}.[w = Q(P),\ Q = the',\ name(P, dog),\ name(S, snore)].\ assert(a, b, S(w))$$

As detailed in (Purver and Ginzburg, 2003; Purver and Ginzburg, 2004), nouns and verbs are taken to denote named predicates,[5] while determiners denote logical relations. NPs are given a lower-order representation, denoting sets of individuals (rather than generalised quantifiers) which for definites must be made part of the abstracted set, but for indefinites and other quantifiers are existentially quantified within the utterance.[6] The overall representation is built up compositionally by a HPSG grammar; space precludes details here, and the use of HPSG is not essential, but it is important to note that the output of the grammar (here, an HPSG sign) must associate each subconstituent with all and only the contextual parameters which it contributed, thus ensuring that clarifying a particular word or phrase can ask only about its contributions to the utterance.

Ellipsis & Reprises via Abstraction The treatment of reprises and other elliptical fragments is based on that of SHARDS (as assumed by G&C): their content is specified by the grammar as being identified with features of the context, specifically the features MAX-QUD and SAL-UTT. However, these features are now also taken to be members of the utterance's abstracted set. A fragment *"Paris"* is therefore given an abstracted representation such as that in (6): its content will be an assertion of a proposition concerning some object x named Paris, but first not only x but a maximal QUD question Q and a salient utterance S must be found in context to fully specify that proposition and the role of x in it:

(6) $$\lambda\Big\{x, Q, S\Big\}.[name(x, paris),\ max\_qud(Q) \wedge Q = \lambda\{\ldots\}.P,\ sal\_utt(S) \wedge content(S, x)].\ assert(a, b, P(\ldots x \ldots))$$

This has several advantages. Firstly, it avoids some potential problems with the SHARDS approach (see (Schlangen, 2003)): the representation of the fragment is now a well-defined object (a simultaneous abstract) rather than being underspecified (with the potential problems that can lead to when implementing within a standard grammar and/or parser), and is derived entirely compositionally, with the non-abstracted parts derived entirely from the constituent words and the ab-

[5] Mass nouns and bare plurals are more complex, seen as ambiguous between predicates (or kinds) and existentially quantified individuals.

[6] Quantifier scope is treated via a functional analysis, and monotone decreasing quantifiers via a representation as pairs of sets – see (Purver and Ginzburg, 2004) for more details.

stracted set expressing only its contextual dependence, specifying the type of context that the abstract can be applied to. Secondly, resolution no longer has to be performed by a separate module, as with SHARDS: as all contextual dependence is now expressed together, resolution can be part of the grounding process, instantiating all parameters together to obtain the fully specified content.[7]

All fragments show this kind of abstraction; so do all CRs (which depend on their source SAL-UTT utterance, and on MAX-QUD if elliptical or reprise). The representation in example (6) above is for a standard declarative fragment, where the word *Paris* is taken as denoting an object named Paris, and the overall content of the fragment is an assertion. Other versions are also possible (and required for certain CR types): firstly equivalent interrogative fragments; secondly utterance-anaphoric fragments, where the word *Paris* is taken to denote a previous salient utterance *'Paris'*, as in (7):

(7) $$\lambda\Big\{Q, S\Big\}.[max\_qud(Q),\ sal\_utt(S) \wedge phon(S, \texttt{paris})].\ ask(a, b, Q(\ldots S \ldots))$$

As (7) shows, CRs are treated as standard interrogative *ask* moves (rather than special e.g. *request-repair* moves). Their CR nature comes only from the question asked (concerning some feature of the source utterance). They also have contextual parameters which must be grounded, and CRs-of-CRs are therefore possible (and do occur in corpora – see (Purver, 2004)).

3 Utterance Processing & Grounding

The system's ability to handle clarificational dialogue centres around the grounding process: application of the abstract to the current context (the IS), finding suitable referents for each of the abstracted parameters (including values for the MAX-QUD and SAL-UTT features) such that the utterance then receives a fully specified content.[8] It is the inability to ground a particular parameter in context (or to ground it in a way that is consistent with what is already known in context) that gives rise to system CRs; it is the grounding of parameters in a suitable way that allows user CRs (particularly elliptical forms) to be interpreted correctly.

In CLARIE this process is implemented in as simple a way as possible. Rather than using general reasoning or inference, a set of logical constraints and preferences that govern the process are defined as TrindiKit IS update rules. Prolog backtracking is then used to find an assignment for the abstracted set such that all constraints are satisfied. The constraints are expressed as preconditions on particular rules, and express general requirements on the way parameters are instantiated: for example, to ensure that utterances are interpreted in such a way that their content is internally consistent and consistent with what is already known (where possible). The preferences are expressed in the ordering of the update rules, and ensure that utterances are grounded in a maximally relevant way: e.g. that an ambiguous utterance be instantiated as an answer to a question currently under discussion if possible, and only as a CR if not.

Disambiguation It is therefore the grounding process which performs disambiguation between all the possible moves that the utterance can make. The abstracted representation means that lexical ambiguity, as well as ambiguity of reference and elliptical resolution, is now represented as contextual dependence (to be fixed by grounding). Any other ambiguity (i.e. multiple possible parses returned by the grammar) means that more than one possible abstract will be available,[9] and again it is the grounding process that must choose between them based on their consistency and relevance. This therefore allows all IS information to be used in disambiguation: not only the possible referents for parameters, but the current state of the dialogue (QUDs, beliefs etc.).

[7] Note that this use of abstraction in ellipsis is not the same as the higher-order abstraction approach of (Dalrymple et al., 1991) for VP ellipsis, in which abstracts are formed from the antecedent and used in resolving the ellipsis. Here, the elliptical fragment is the abstract, to be applied to the context.

[8] The term *grounding* is often used in a wider sense to incorporate the general process of understanding and addition to the common ground, including acknowledgement and acceptance. Here it is used narrowly to refer to the fixing of an utterance's content in context.

[9] The grammar will in fact *always* assign more than one possible parse – a typical fragment will be given at least four representations, two in each of the dimensions declarative/interrogative and standard/utterance-anaphoric – see above.

Utterance Processing The CLARIE IS is shown in AVM (8). Like GoDiS, it is divided into two parts, with PRIVATE for system plans and beliefs that have not been explicitly introduced into the dialogue, and SHARED representing the system's view of the common ground:

$$(8)\quad \left[\begin{array}{ll} \textrm{PRIVATE} & \left[\begin{array}{ll} \textrm{AGENDA} & \left[\textit{stack(action)}\right] \\ \textrm{PLAN} & \left[\textit{stackset(action)}\right] \\ \textrm{BEL} & \left[\textit{set(proposition)}\right] \\ \textrm{BG} & \left[\textit{set(parameter)}\right] \end{array}\right] \\ \textrm{SHARED} & \left[\begin{array}{ll} \textrm{COM} & \left[\textit{set(proposition)}\right] \\ \textrm{BG} & \left[\textit{set(parameter)}\right] \\ \textrm{QUD} & \left[\textit{stack(question)}\right] \\ \textrm{SAL-UTT} & \left[\textit{stack(sign)}\right] \\ \textrm{UTT} & \left[\textit{nstackset(4,sign)}\right] \\ \textrm{PENDING} & \left[\textit{stack(set(sign))}\right] \end{array}\right] \end{array}\right]$$

In the shared part, COM is a set of commitments and BG a set of descriptions of referents that have been explicitly introduced in the dialogue. QUD and SAL-UTT are stacks of QUDs and salient utterances respectively, used for ellipsis resolution and answerhood. UTT is an utterance record, a stack of utterances in linear dialogue order which is used to find CR sources, allowing the questions asked by user CRs to be fully interpreted, and their answers to be determined. It is therefore important that its members are *signs*, including all attendant phonological, syntactic and semantic information which may be clarified, rather than just semantic representations such as *moves*. It has a limited length, currently 4 utterances, as Purver et al. (2003) found that CRs beyond this distance are rare. PENDING holds ungrounded utterance abstracts during the grounding process.

Utterance processing is based on the protocol proposed by G&C and proceeds as follows: the utterance abstracts produced by the parser are pushed onto the PENDING and UTT stacks while grounding is attempted:

$$(9)\quad \left[\begin{array}{ll} \textrm{AGENDA} & \langle \ldots \rangle \\ \textrm{QUD} & \langle \ldots \rangle \\ \textrm{UTT} & \langle U, \ldots \rangle \\ \textrm{PENDING} & \langle U \rangle \end{array}\right]$$

Grounding can be achieved via three sets of rules, tested in order. Firstly, standard *integration* rules attempt to ground the utterance given the current IS (and in particular the current top members of QUD and SAL-UTT). Secondly, *accommodation* rules can be used to achieve the same effect using a new QUD determined from a relevant but as yet not explicitly asked question from the plan (see (Larsson et al., 2000)). If neither of these succeed, a third set of *coercion* rules attempt to ground the utterance as a user CR, by using contextual coercion operations which produce new CR-related values of QUD and SAL-UTT. In all cases, the newly grounded utterance (with all parameters now fully instantiated) is removed from PENDING and its update effects applied to the IS (e.g. for a question, raising a new QUD Q and an action to respond):

$$(10)\quad \left[\begin{array}{ll} \textrm{AGENDA} & \langle \texttt{respond}(Q), \ldots \rangle \\ \textrm{QUD} & \langle Q, \ldots \rangle \\ \textrm{UTT} & \langle U, \ldots \rangle \\ \textrm{PENDING} & \langle \rangle \end{array}\right]$$

If all grounding rules fail, the utterance cannot be grounded in the current IS, and a set of *clarification* rules use its ungrounded parameters to form a suitable CR and add an agenda action to ask it. Importantly, the ungrounded utterance is left in PENDING so that grounding can be re-attempted once a user answer to the CR has provided sufficient new information:

$$(11)\quad \left[\begin{array}{ll} \textrm{AGENDA} & \langle \texttt{clarify}(U), \ldots \rangle \\ \textrm{QUD} & \langle \ldots \rangle \\ \textrm{UTT} & \langle U, \ldots \rangle \\ \textrm{PENDING} & \langle U \rangle \end{array}\right]$$

Grounding Non-CRs The *integration* rules use the current IS unchanged and only general constraints apply: the parameters corresponding to MAX-QUD and SAL-UTT must be identified with the current heads of the QUD and SAL-UTT IS stacks. Parameters concerning individual referents (e.g. of names and definites) must be identified with unique antecedents in the BG set or available from the domain model. Parameters concerning predicates (from nouns and verbs) must again refer to explicit antecedents in BG or be available from the lexicon.[10] Logical relations (contributed by closed-class function words) are always available from the lexicon. A simple consistency check is also applied, preventing grounding of e.g. assertions that conflict with previously held beliefs (conflicting assertions will lead to clarification – see below). The protocol is as follows:

1. Interpret as answering a question which is under discussion (currently maximal in QUD).
2. Interpret as asking a question which is relevant to the current IS.
3. Interpret as a greeting, closing or thanking move.

Accommodation rules follow similar lines, but use a new accommodated QUD to ground the MAX-QUD parameter. Coercion rules are described in the next section.

4 Processing CRs

A Sample Dialogue The dialogue below shows some of CLARIE's current capabilities:

```
S1> Where do you want to go to?
U1> pardon?
S2> Where do you want to go to?
U2> i want to go to that city
S3> Which city?
U3> paris
S4> How do you want to travel?
U4> no i want to go to london
S5> London?
U5> yes london
S6> How do you want to travel?
U6> by pullman
S7> 'Pullman'?
U7> train
S8> The trip will cost fifty pounds.
U8> sixty?
S9> No, fifty.
U9> fifty?
S10> Yes.
```

Note that both system CRs (S3, S5, S7) and user CRs (U1, U8, U9) are included, as are questions about utterance form (U2), NP reference (S3) and noun predicate reference (S7), and both clausal (S5, U8, U9) and constituent (S7) reprises.[11]

User CRs The *coercion* rules allow user CRs to be grounded (and thus fully interpreted) by modifying the context: using coercion operations in order to ground the MAX-QUD and SAL-UTT parameters. Firstly, for all CRs, the *source* utterance (the utterance being asked about) must be identified by examining constituents from the UTT record until a suitable one is found which meets the constraints associated with the CR. This corresponds to grounding the SAL-UTT parameter. Some CRs only require this step – those which have their propositional content specified directly by the grammar, thus not requiring a MAX-QUD question to fill in, merely requiring utterance reference to be established, e.g. non-reprise CRs such as *"Did you say 'Paris'?"*, *"What do you mean by 'pullman'?"* or conventional expressions such as *"What?"*, *"Pardon?"*. For these, then, a simple coercion operation which provides possible SAL-UTT values from the UTT record suffices.

However, the most common forms of CR are of a reprise and/or elliptical nature and therefore also require a MAX-QUD parameter to be grounded to fully specify their propositional content. For these, there are currently four different possible coercion operations which not only take a possible source constituent from UTT but also use it to produce particular new CR-related MAX-QUDs: G&C's clausal and constituent versions, plus two further questions about lexical form (one querying the identity of an echoed word, one a *gap* question querying the identity of the word following an echoed word). These operations and the order and constraints of the rules which apply them are determined by corpus and experimental studies (see (Purver, 2004)). Constraints include factors such as source word/phrase category (e.g. no function words can lead to constituent readings as their meaning is mutual knowledge; only some function words such as number determiners seem likely

[10] This allows a degree of *alignment* behaviour (Pickering and Garrod, 2004) – once words and phrases have been grounded to particular referents, these become present in BG and will be used again for grounding before looking in the lexicon.

[11] The current behaviour is designed to demonstrate the elliptical CR interpretation and generation capabilities, hence the highly elliptical forms used. Many system CRs may benefit from less elliptical form in practice (e.g. S7 might be less ambiguously realised as *"What do you mean by 'pullman'?"* or *"What is a 'pullman'?"*) – this behaviour can be controlled by a user-settable flag.

to be given clausal readings), source parallelism (constituent reprise fragments require phonological parallelism) and common ground (words previously grounded in the dialogue do not cause constituent readings).[12]

1. Coerce SAL-UTT only and interpret as a conventional or non-reprise CR – see `U1`.
2. Perform *parameter identification* and interpret as a constituent fragment reprise *if* the source is the first mention of a content phrase fragment.
3. Perform *parameter focussing* and interpret as a clausal reprise (sentence, fragment or sluice) *if* the source is a content phrase or number determiner – see `U8`, `U9`.
4. Perform *gap identification* and interpret as a lexical reprise gap.
5. Perform *lexical identification* and interpret as a lexical reprise.

As CRs are *ask* moves, these coercion rules will all include in their effects the addition of a new question to QUD, and an agenda action to answer it. Answers can now be established from the relevant features of the antecedent utterance in the UTT record: in the case of clausal and constituent CRs, directly from the semantic content of the source utterance; in the case of lexical or gap questions, from the identity of the source utterance itself. This process is specified as a set of selection rules which produce corresponding *assert* moves – these are then passed to a generation module which uses the grammar to generate in the same way as answers to normal questions.[13]

System CRs If a user utterance cannot be grounded in any way, the *clarification* rules produce a system CR. Particular grounding problems lead to particular questions being asked (again derived from empirical findings – out-of-vocabulary nouns and verbs lead to constituent wh-questions, parameters which cause inconsistency lead to clausal yes/no "check" questions), and particular source types will lead to particular forms being used (definites and demonstratives are clarified using sluices, nouns using fragments). The protocol is as follows:

1. Parse failure (no move to ground): constituent CR about whole utterance meaning.
2. Unknown parameter (no unique referent can be found): clausal wh-question about source constituent *if* a definite or pronoun (leading to a sluice – see `S3`); constituent wh-question otherwise (leading to a fragment – see `S7`).
3. Inconsistent parameter (can only be grounded in a way which causes inconsistency with previous beliefs): clausal yn-question leading to a reprise fragment – see `S5`.
4. Inconsistent moves (can only be grounded inconsistently) and irrelevant moves (ungroundable MAX-QUD or SAL-UTT parameter): constituent question about whole intended utterance meaning.

Note that there is significant correspondence between these grounding problem types and the levels identified by (Larsson, 2002; Schlangen, 2004), but that the association of problematic parameters with their source words/phrases allows specific CR forms which target those phrases.[14]

As system CRs (being *ask* moves) introduce a new question to QUD, subsequent user answers (elliptical or not) can be interpreted according to standard answerhood rules – no special treatment is required – and that as long as such answers provide the required information, the problematic PENDING utterance will now be groundable without requiring repetition.

5 Summary

This paper shows how a basic dialogue system can be implemented which can handle many forms of CR, using them to clarify unknown reference and meaning, and allowing users to do the same. The grammar can parse and generate a wide range of

[12] While these coercion operations can perhaps be seen as a form of reasoning about context, they are highly constrained and far from unrestricted inference.

[13] There are two special cases: when answering CRs which ask about word meaning, alternative descriptions are used wherever possible rather than the original problematic form; when answering yn-questions negatively, an over-answer is produced by answering the coerced MAX-QUD as well as the explicit CR question – see `S9`.

[14] Being text-based, the level of *perception* is currently ignored, but must be taken into account if a speech interface is to be added – see (Gabsdil, 2003).

CR forms from a wide range of source types. Importantly, this is achieved without having to use heavyweight inference about utterances or their relation to each other, or modelling the user's beliefs or context. User CRs are assigned straightforward (although heavily contextually dependent, and often ambiguous) representations; and the grounding process then gives them a full interpretation by instantiating their abstracted parameters in context. Problems with the grounding process, and with particular abstracted parameters, lead to system CRs. CRs are not treated in a significantly different way from other utterances: they are parsed by the same grammar and given a standard interrogative interpretation as *ask* moves which raise new questions for discussion – it is just that these questions concern other utterances.

Acknowledgements

The author is currently supported by the ESRC (RES-000-22-0355), although most of this work was carried out as part of the ROSSINI project (EPSRC GR/R04942/01). Thanks are due to the Catalog reviewers, and to Jonathan Ginzburg for enormous help and support as the supervisor of my thesis (Purver, 2004), of which this paper summarizes chapter 6.

References

H. H. Clark. 1996. *Using Language*. Cambridge University Press.

M. Dalrymple, S. Shieber, and F. Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy*, 14(4):399–452.

M. Gabsdil. 2003. Clarification in spoken dialogue systems. In *Proceedings of the AAAI Spring Symposium on Natural Language Generation in Spoken and Written Dialogue*.

J. Ginzburg and R. Cooper. 2004. Clarification, ellipsis, and the nature of contextual updates in dialogue. *Linguistics and Philosophy*, 27(3):297–365.

J. Ginzburg and I. A. Sag. 2000. *Interrogative Investigations: the Form, Meaning and Use of English Interrogatives*. CSLI Publications.

J. Ginzburg, H. Gregory, and S. Lappin. 2001. SHARDS: Fragment resolution in dialogue. In *Proceedings of the 4th International Workshop on Computational Semantics (IWCS-4)*, pages 156–172.

B. Hockey, J. Dowding, G. Aist, and J. Hieronymus. 2002. Targeted help and dialogue about plans. In *ACL-02 Companion Volume to the Proceedings of the Conference*.

S. Larsson, P. Ljunglöf, R. Cooper, E. Engdahl, and S. Ericsson. 2000. GoDiS - an accommodating dialogue system. In *Proceedings of ANLP/NAACL-2000 Workshop on Conversational Systems*.

S. Larsson, A. Berman, L. Grönqvist, and F. Kronlid. 2002. TrindiKit 3.0 manual. In *SIRIDUS Deliverable 6.4*.

S. Larsson. 2002. *Issue-based Dialogue Management*. Ph.D. thesis, Göteborg University.

M. Pickering and S. Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, forthcoming.

M. Purver and J. Ginzburg. 2003. Clarifying noun phrase semantics in HPSG. In *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar (HPSG-03)*.

M. Purver and J. Ginzburg. 2004. Clarifying noun phrase semantics. To appear in *Journal of Semantics*, 21(3).

M. Purver, J. Ginzburg, and P. Healey. 2003. On the means for clarification in dialogue. In R. Smith and J. van Kuppevelt, editors, *Current and New Directions in Discourse & Dialogue*. Kluwer Academic Publishers.

M. Purver. 2004. *The Theory and Use of Clarification Requests in Dialogue*. Ph.D. thesis, King's College, London.

K. Rodríguez and D. Schlangen. 2004. Form, intonation and function of clarification requests in German task-oriented spoken dialogues. In *Proceedings of the 8th Workshop on the Semantics and Pragmatics of Dialogue (Catalog)*.

D. Schlangen. 2003. *A Coherence-Based Approach to the Interpretation of Non-Sentential Utterances in Dialogue*. Ph.D. thesis, University of Edinburgh.

D. Schlangen. 2004. Causes and strategies for requesting clarification in dialogue. In *Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue*.

M. Stone. 2003. Specifying generation of referring expressions by example. In *Proceedings of the AAAI Spring Symposium on Natural Language Generation in Spoken and Written Dialogue*.

D. Traum. 2003. Semantics and pragmatics of questions and answers for dialogue agents. In *Proceedings of the International Workshop on Computational Semantics*.