

The Clarity-Brevity Trade-off in Generating Referring Expressions *

Imtiaz Hussain Khan and Graeme Ritchie and Kees van Deemter

Department of Computing Science

University of Aberdeen

Aberdeen AB24 3UE, U.K.

{ikhan, gritchie, kvdeemte}@csd.abdn.ac.uk

Abstract

Existing algorithms for the Generation of Referring Expressions (GRE) aim at generating descriptions that allow a hearer to identify its intended referent uniquely; the length of the expression is also considered, usually as a secondary issue. We explore the possibility of making the trade-off between these two factors more explicit, via a general cost function which scores these two aspects separately. We sketch some more complex phenomena which might be amenable to this treatment.

1 Introduction

Until recently, GRE algorithms have focussed on the generation of distinguishing descriptions that are either as short as possible (e.g. (Dale, 1992; Gardent, 2002)) or almost as short as possible (e.g. (Dale and Reiter, 1995)). Since reductions in ambiguity are achieved by increases in length, there is a tension between these factors, and algorithms usually resolve this in some fixed way. However, the need for a distinguishing description is usually assumed, and typically built in to GRE algorithms. We will suggest a way to make explicit this balance between *clarity* (i.e. lack of ambiguity) and *brevity*, and we indicate some phenomena which we believe may be illuminated by this approach. The ideas in this paper can be seen as a loosening of some of the many simplifying assumptions often made in GRE work.

*This work is supported by a University of Aberdeen Sixth Century Studentship, and the TUNA project (EPSRC, UK) under grant number GR/S13330/01. We thank Ielka van der Sluis and Albert Gatt for valuable comments.

2 Clarity, Brevity and Cost

We consider only simple GRE, where the aim is to construct a conjunction of unary properties which distinguish a single target object from a set of potential distractors. Our notation is as follows. A domain consists of a set \mathbf{D} of objects, and a set \mathbf{P} of properties applicable to objects in \mathbf{D} . A *description* is a subset of \mathbf{P} . The *denotation* of S , written $\llbracket S \rrbracket$, is $\{x \in \mathbf{D} \mid \forall p \in S : p(x)\}$.

(Krahmer et al., 2003) describe an approach to GRE in which a *cost function* guides search for a suitable description, and show that some existing GRE algorithms fit into this framework. However, they follow the practice of concentrating solely on distinguishing descriptions, treating cost as a matter of brevity. We suggest that decomposing cost into two components, for the clarity and brevity of descriptions, permits the examination of trade-offs. For now, we will take the cost of a description S to be the sum of two terms:

$$\text{cost}(S) = f_C(S) + f_B(S).$$

where f_C counts ambiguity (lack of clarity) and f_B counts size (lack of brevity). Even with this decomposition of cost, some existing algorithms can still be seen as cost-minimisation. For example, the cost functions:

$$\begin{aligned} f_C(S) &= |\mathbf{P}| \times |\llbracket S \rrbracket| \\ f_B(S) &= |S| \end{aligned}$$

allow the Full Brevity algorithm (Dale, 1992) to be viewed as minimising $\text{cost}(S)$, and the incremental algorithm (Dale and Reiter, 1995) as hill-climbing (strictly, hill-descending), guided by the property-ordering which that algorithm requires. Whereas Krahmer et al.'s cost functions are (brevity-based) heuristic guidance functions, our alternative here is a global quantity for optimisation. Hence their simulation of Full Brevity

relies on the details of their algorithm (rather than cost) to ensure clarity, while our own cost function ensures both brevity and clarity.

3 Exploring the Trade-off

3.1 Varying penalties for distractors

Imagine the following situation. You are preparing a meal in a friend's house, and you wish to obtain, from your own kitchen, a bottle of Italian extra virgin olive oil which you know is there. The only way open to you is to phone home and ask your young child to bring it round for you. You know that also in your kitchen cupboard are some distractors: one bottle each of Spanish extra virgin olive oil, Italian non-virgin olive oil, cheap vegetable oil, linseed oil (for varnishing) and camphorated oil (medicinal). It is imperative that you do not get the linseed or camphorated oil, and preferable that you receive olive oil. A full expression, *Italian extra virgin olive oil*, guarantees clarity, but may overload your helper's abilities. A very short expression, *oil*, is risky. You might well settle for the intermediate *olive oil*.

To model this situation, f_C could take a much higher value if $\llbracket S \rrbracket$ contains a distractor which must not be selected (e.g. varnish rather than cooking oil). That is, instead of a simple linear function of the size of $\llbracket S \rrbracket$, there is a curve where the cost drops more steeply as the more undesirable distractors are excluded. For example, each object could be assigned a numerical rating of how undesirable it is, with the target having a score of zero, and the f_C value for a set A could be the *maximum* rating of any element of A . (This would, of course, require a suitably rich domain model.)

The brevity cost function f_B could still be a relatively simple linear function, providing f_B values do not mask the effect of the shape of the f_C curve.

3.2 Fuzziness of target

Suppose Mrs X has dropped a piece of raw chicken meat on the kitchen table, and immediately removed the meat. She would now like Mr X to wipe the area clean. The meat leaves no visible stain, so she has to explain where it was. In this case, it appears that there is no such thing as a distinguishing description (i.e. a description that pins down the area precisely), although Mrs X can arbitrarily increase precision, by adding properties:

- *the edge of the table,*
- *the edge of the table, on the left (etc.)*

The ideal description would describe the dirty area and nothing more, but a larger area will also do, if not too large. Here, the domain \mathbf{D} is implicitly defined as all conceivable subareas of the table, the target is again one element of \mathbf{D} , but – unlike the traditional set-up with discrete elements – a description (fuzzily) defines *one* such area, not a disjoint collection of individual items. Our f_C operates on the description S , not just on the number of distractors, so it can assess the aptness of the denotation of any potential S . However, it has to ensure that this denotation (subarea of the surface) contains the target (contaminated area), and does not contain too much beyond that. Hence, we may need to augment our clarity cost function with another argument: the target itself. In general, more complex domains may need more complicated functions.

3.3 Underspecification in dialogue

Standard GRE algorithms assume that the speaker knows what the hearer knows (Dale and Reiter, 1995). In practice, speakers can often only guess. It has been observed that speakers sometimes produce referring expressions that are only disambiguated through negotiation with the hearer, as exemplified in the following excerpt (quoted in (Hirst, 2002)).

1. A: What's that weird creature over there?
2. B: In the corner?
3. A: [*affirmative noise*]
4. B: It's just a fern plant.
5. A: No, the one to the left of it.
6. B: That's the television aerial. It pulls out.

A and B are in the same room, in an informal setting, so A can be relatively interactive in conveying information. Also, the situation does not appear to be highly critical, in comparison to a military officer directing gunfire, or a surgeon guiding an incision. Initially, A produces an expression which is not very detailed. It may be that he thinks this is adequate (the object is sufficiently salient that B will uniquely determine the referent), or he doesn't really know, but is willing to make an opening bid in a negotiation to reach the goal of reference. In the former case, a GRE algorithm which took account of salience (e.g. (Krahmer and Theune, 1999)), operating with A 's model of B 's knowledge, should produce this sort of effect. (A dialogue model might also be needed.) In the latter case, we need an algorithm which can

relax the need for complete clarity. This could be arranged by having f_C give similar scores to denotations where there are no distractors and to denotations where there are just a few distractors, with f_B making a large contribution to the cost.

3.4 Over-specification

Recently, interest has been growing in ‘overspecified’ referring expressions, which contain more information than is required to identify their intended referent. Some of this work is mainly or exclusively experimental (Jordan and Walker, 2000; Arts, 2004), but algorithmic consequences are also being explored (Horacek, 2005; Paraboni and van Deemter, 2002; van der Sluis and Kraemer, 2005). Over-specification could also arise in a dialogue situation (comparable to that in Section 3.3) if a speaker is unclear about the hearer’s knowledge, and so over-specifies (relative to his own knowledge) to increase the chances of success.

This goes beyond the classical algorithms, where the main goal is total clarity, with no reason for the algorithm to add further properties to an already unambiguous expression. That is, such algorithms assume that every description S for which $|\llbracket S \rrbracket| = 1$ has the same level of clarity (f_C value). This assumption could be relaxed. For example, the approach of (Horacek, 2005) to GRE allows degrees of uncertainty about the effectiveness of properties to affect their selection. Within such a framework, one could separately compute costs for clarity (e.g. likelihood of being understood) and brevity (which might include the complexity of expressing the properties).

4 Conclusion and Future Work

We have argued that the GRE task becomes very different when some commonly-made assumptions are abandoned: some distractors might be worse than others (section 3.1); the target may be impossible to distinguish precisely (section 3.2); the speaker may be unsure what the hearer knows (section 3.3); or there may be a need for over-specification (section 3.4)). As a result, it may be necessary to consider other aspects of the descriptions and their denotations, not simply counting distractors or numbers of properties. Some effects could perhaps be modelled using costs which are not simple linear functions, but which give varying importance to particular aspects of the denotation of a description, or of its content. We hope that

this approach will ultimately shed light not only on the effect of the discourse situation, but also some aspects of generating *indefinite* descriptions.

References

- Anja Arts. 2004. *Overspecification in Instructive Text*. Ph.D. thesis, Tilburg University, The Netherlands.
- Robert Dale and Ehud Reiter. 1995. Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 18:233–263.
- Robert Dale. 1992. *Generating Referring Expressions: Building Descriptions in a Domain of Objects and Processes*. MIT Press.
- Claire Gardent. 2002. Generating minimal distinguishing descriptions. In *Proceedings of the 40th Annual Meeting of the ACL (ACL’02)*, Philadelphia, USA.
- Graeme Hirst. 2002. Negotiation, compromise, and collaboration in interpersonal and human–computer conversations. In *Proceedings of Workshop on Meaning Negotiation, 18th National Conference on Artificial Intelligence*, pages 1–4, Edmonton, Canada.
- Helmut Horacek. 2005. Generating referential descriptions under conditions of uncertainty. In Graham Wilcock, Kristiina Jokinen, Chris Mellish, and Ehud Reiter, editors, *Proceedings of the 10th European Workshop on Natural Language Generation (ENLG-05)*, pages 58–67.
- Pamela Jordan and Marilyn Walker. 2000. Learning attribute selections for non-pronominal expressions. In *Proceedings of the 38th Annual Meeting of the ACL (ACL-00)*, pages 181–190.
- Emiel Kraemer and Mariët Theune. 1999. Efficient generation of descriptions in context. In *Proceedings of the ESSLLI workshop on the generation of nominals*, Utrecht, The Netherlands.
- Emiel Kraemer, Sebastiaan van Erk, and André Verleg. 2003. Graph-based generation of referring expressions. *Computational Linguistics*, 29(1):53–72.
- Ivandr  Paraboni and Kees van Deemter. 2002. Generating easy references: the case of document deixis. In *Proceedings of the Second International Conference on Natural Language Generation*, New York, USA.
- Ielka van der Sluis and Emiel Kraemer. 2005. Towards the generation of overspecified multimodal referring expressions. In *Proceedings of the Symposium on Dialogue Modelling and Generation at the 15th Annual Meeting of the ST & D (STD-05)*, Amsterdam, The Netherlands.