

Everyday language is highly intensional

Allan Ramsay

School of Computer Science
University of Manchester
Manchester M60 1QD, United Kingdom

Debora Field

Informatics Collaboratory of the Social Sciences
University of Sheffield
Sheffield S1 4DP, United Kingdom

Abstract

There has recently been a great deal of work aimed at trying to extract information from substantial texts for tasks such as question answering. Much of this work has dealt with texts which are reasonably large, but which are known to contain reliable relevant information, e.g. FAQ lists, on-line encyclopaedias, rather than looking at huge unorganised resources such as the web.

We believe, however, that even this work underestimates the complexity and subtlety of language, and hence will inevitably be restricted in what it can cope with. In particular, everyday use of language involves considerable amounts of reasoning over intensional objects (properties and propositions). In order to respond appropriately to simple-seeming questions such as *'Is going for a walk good for me?'*, for instance, you have to be able to talk about event-types, which are intrinsically intensional. We discuss the issues involved in handling such items, and shows the kind of background knowledge that is required for drawing the appropriate conclusions about them.

1 Introduction

The work reported here aims to allow users to interact with a health information system via natural language. In this context, allowing a user to make simple statements about their condition and then ask questions about what they can or should do, as in (1), seems to be a minimal requirement.

- (1) My doctor says I am allergic to eggs. Is it safe for me to eat cake?

Understanding such utterances requires the use of a highly intensional representation language, and responding to them requires a surprising amount of background knowledge. We will consider below the problems that such everyday utterances bring for formal paraphrases of natural language, and we will look at the kind of background knowledge that is required for producing the right kinds of response. In order to produce a system that carries out the required inference we need access to an inference engine for carrying out proofs in a representation language with the required expressive power. The details of the engine we use are beyond the scope of this paper. (Ramsay, 2001; Ramsay and Field, 2007). For the purposes of the current paper we will simply show the results that can be obtained by using it.

The work reported here is complementary to work on corpus-based approaches such as textual entailment: approaches that ignore the intensionality of everyday language will inevitably fail to capture important inference patterns, but on the other hand the work reported here cannot deal with large amounts of information provided as free text. Ideally, the two approaches will be combined. The aim of the current paper is to provide a reminder of the prevalence of intensionality in everyday language, and to demonstrate that modern theorem proving techniques can cope with this kind of knowledge without introducing undue processing delays.

2 Background

The general idea behind the work reported here is that users will input statements about their health,

either spontaneously or in response to prompts from the system, and will ask questions about what they can and should do, and the system will provide them with appropriate guidance. The overall architecture is completely classical:

1. The user's input is translated into a meaning representation (logical form, LF) in some suitable representation language.
2. This LF contains a specification of the illocutionary force of the input (is it a statement, or a question, or a command, or ...?).
3. If the utterance is classified as a statement, its propositional content is added to the system's view of the user's beliefs, and if it is classified as a question, the system will attempt to use its background knowledge of the domain to answer it. We are not currently attempting to make the system do anything in response to a command from the user, since users do not generally issue commands in our chosen domain, but clearly if this did happen then we would want to make the system construct a plan to carry out the required action.

This part of the system's activity requires it be able to access and exploit relevant background knowledge. This is obvious in the case of questions, but in the given domain it is also important to be able to spot situations where the user's beliefs are incomplete or are in conflict with the system's beliefs, since most people's understanding about medical topics is flawed. The ability to reason about what has been said, then, is crucial to the construction of appropriate responses.

This architecture is entirely orthodox. What is unusual about the current work is the emphasis on intensionality, so the first thing to do is examine why we believe that this is such a significant problem.

1. Doctors and patients make extensive use of generic NPs and bare plurals: *'If you follow this diet you should manage to control them without drugs'*, *'Do you normally have snacks?'*, *'When I started chemotherapy, on*

the 2nd of August, glycaemia was still rather high' ...

Such NPs are not, in fact, all that much more prevalent in this domain than in general language. Across the BNC, for instance, it turns out that 27% of NPs have 'the' as their determiner, 19% are bare plurals, 29% are bare singulars, 11% have 'a' or 'an' as their determiner, and the remainder have a variety of other determiners¹.

Thus bare plural and generic singular NPs occur about as frequently as 'the' and 'a', and substantially more frequently than 'some', 'all' and 'every' (less than 1% each). They have, however, been much less widely discussed by formal semanticists, and there are a number of serious problems with the analyses that have been proposed (Carlson, 1989; Ramsay, 1992; Cohen, 1994).

2. Everyday language is littered with words that can be used either as nouns or verbs, and many of the apparently verbal uses of such words occur in essentially nominal contexts. Fig. 1 shows the pattern of usage for three common words², but it should be noted that about 25% of the instances that are classified as verbs are present participle forms, many of which are actually nominal or verbal gerunds and hence should be regarded as nouns.

	verb	noun	other
walk	75%	22%	3%
run	70%	24%	6%
kick	63%	35%	2%

Figure 1: Uses of common words in the BNC

Axiomatisation of the semantics of such words requires considerable care, since we need to ensure that all the examples in (2) have very similar consequences.

¹The count of bare singulars is in fact a slight overestimate, since it includes some uses of singular nouns as modifiers.

²The classification is taken directly from the BNC tags.

- (2)
 - a. Swimming is good for you.
 - b. Going for a swim is good for you.
 - c. It is good for you to go swimming.
3. The goal of the project is to produce appropriate responses to simple statements and queries about a patient's health. To do this, we need to be able to specify a body of background knowledge in this area. We believe that for applications such as medical information provision it is important that the information provided be as accurate as possible, and hence that it may be necessary to provide the required background knowledge from scratch. This is, of course, a very time-consuming and challenging activity, and it would be nice to be able to side-step it by extracting the required information from existing texts. Unfortunately, it seems likely that any such existing text will contain gaps which will lead to the generation of partial, or wrong, answers. As noted above, ideally we would want to link special purpose knowledge of the kind outlined here with information extracted from existing texts, but for the current paper we are just looking at what is involved in providing the required knowledge from scratch.

It turns out, as will be seen below, that much of this knowledge involves quantification over situation types (of roughly the kind discussed by (Barwise and Perry, 1983)), and in particular it involves statements about whether one situation type is a subset of another, or is incompatible with it. This kind of knowledge is intrinsically intensional, but it is hard to see how it can be avoided in this domain.

3 Logical forms

The logical forms that we use are fairly orthodox.

- We assume that events are first-class objects, as suggested by Davidson Davidson (1967, 1980).
- We allow other entities to play named roles with respect to these events, where we de-

note that some item X is, for instance, the agent of some event E by writing $\theta(E, agent, X)$: using this notation, rather than writing $agent(E, X)$, allows us to quantify over thematic roles, which in turn allows us to state generalisations that would otherwise be awkward.

- We treat tense as a relation between speech time and 'reference time', and aspect as a relation between reference time and event time, as suggested by Reichenbach Reichenbach (1947, 1956).
- We use 'reference terms' to denote referring expressions, so that $ref(\lambda X man(X))$ is used to denote 'the man'. Reference terms are similar to 'anchors' from (Barwise and Perry, 1983), though the treatment is essentially proof-theoretic (similar to the discussion of presupposition in (Gazdar, 1979; van der Sandt, 1992)) rather than model theoretic.
- Given that we are particularly concerned with the intensional nature of natural language, we need to use a formal language that supports intensionality. The language we choose is a constructive version of property theory (Turner, 1987; Ramsay, 2001). We have extended the theorem prover described in (Ramsay, 2001) to cope with reasoning about knowledge and belief, and we have shown how this can be used to carry out interesting inferences in cooperative and non-cooperative situations (Ramsay and Field, 2007).

We also include the surface illocutionary force in the LF, since this is part of the meaning of the utterance and hence it seems sensible to include it in the LF. In particular, there are interactions between surface illocutionary force and other aspects of the meaning which are hard to capture if you treat them independently. This is slightly less standard than the other aspects of our LFs, but it does have the advantage that these LFs keep all the information that we can obtain by inspecting the form of the utterance in one place.

A typical example of an LF for a simple sentence is given in Fig. 2³.

- (3) The man loves a woman.

$$\begin{aligned} \text{claim}(\exists B : \{ \text{woman}(B) \} \\ \exists C : \{ \text{past}(\text{now}, C) \} \\ \exists D : \{ \text{aspect}(C, \text{simplePast}, D) \} \\ \theta(D, \text{agent}, \text{ref}(\lambda E(\text{man}(E)))) \\ \&\theta(D, \text{object}, B) \\ \&\text{event}(D, \text{love})) \end{aligned}$$

Figure 2: Logical form for (3)

If you want to reason about utterances in natural language, e.g. in order to answer questions on the basis of things you have been told, then there seems to be no alternative to constructing LFs of the kind in Fig. 2, axiomatising the relevant background knowledge, and then invoking your favourite theorem prover. Shallow semantic analysis simply does not provide the necessary detail, and it is very hard to link textual entailment algorithms (Dagan *et al.*, 2005) to complex domain knowledge. The critical issue in connecting NLP systems to rich axiomatisations of domain knowledge seems likely to be that existing frameworks for constructing meaning representations are not rich enough, not that they are too rich. In the remainder of this paper we will explore three specific issues that have arisen in our attempt to use natural language as a means for accessing medical knowledge. We have become sensitised to these issues because of their importance for our application, but we believe that they are actually widespread, and they will need to be solved for any system which links natural language to complex domain knowledge.

4 Bare NPs

Consider (4):

- (4) a. I am eating eggs.
b. I eat eggs.
c. I am allergic to eggs.

What is the status of ‘eggs’ in these sentences?

It is clear that in (4a) there are some eggs that I am eating, so that (4a) means something quite like

³All the formal paraphrases in this paper are obtained from the target sentences by parsing the text and using the standard techniques of compositional semantics.

‘I am eating some eggs.’ (4b), on the other hand, means something fairly different from ‘There are some eggs that I eat’, since it does not seem to commit the speaker to the existence of any specific set of eggs. The use of the simple aspect with a non-stative verb gives (4b) a habitual/repeated interpretation, saying that there are numerous eating events, each of which involves at least one egg.

It seems, then, that it is possible to treat ‘eggs’ in (4a) and (4b) as a narrow scope existential, with the simple aspect introducing a set of eating events of the required kind.

You would not, however, want to paraphrase (4c) by saying that there are some eggs to which I am allergic. (4b) says that there is a relationship between me and situations where there is an egg present, namely that if I eat something which has been made out of some part of an egg then I am likely to have an allergic reaction. The bare plural ‘eggs’ in (4c) seems to have some of the force of a universal quantifier. This is problematic: does the bare plural ‘eggs’ induce an existential or a universal reading, or something entirely different?

Note that the word ‘eggs’ can appear as a free-standing NP (as in (4a)) or as the head noun of an NP with an explicit determiner (as in ‘He was cooking some eggs.’). In the latter context, the meaning of ‘eggs’ is normally taken to be the property $\lambda X(\text{egg}(X))$, to be combined with the determiner ‘some’ to produce an existentially quantified expression which can be used as part of the interpretation of the entire sentence.

It is clear that there are constructions that involve allowing prepositions to take nouns rather than NPs as their complements, in examples like ‘For example, cockerels generally have more decorative plumage than hens’, where ‘example’ is evidently a noun rather than an NP. If we allow the adjective ‘allergic’ to select for a PP with a noun complement rather than an NP complement, we can obtain an interpretation of (4c) which says that my allergy is a relation between me and the property of being an egg (= the set of eggs) (Fig. 3).

Thus we can distinguish between cases where ‘eggs’ is being used as an NP, where it introduces a narrow scope existential quantifier, and ones where it is being used as an NN, where it de-

$$\text{utt}(\text{claim}, \\ \exists B \text{state}(B, \text{allergic}(\text{to}, \\ \lambda C(\text{egg}(C)), \\ \text{ref}(\lambda D(\text{speaker}(D))))!0) \\ \&\text{aspect}(\text{now}, \text{simple}, B))$$

Figure 3: Logical form for (4c)

notes, as usual, the property $\lambda(X, \text{egg}(X))$. We still have to work out saying that the relationship ‘*allergic*’ holds between me and the property of being an egg, but at least we have escaped the trap of saying that it holds between me and some eggs (or indeed all eggs). We will return to this in §6

5 Nominalisations and paraphrases

As noted above, there are often numerous ways of saying very much the same thing, and these often involve using combinations of nominal and verbal forms of the same root. To cope with these, we have to do two things: we have to construct appropriate logical forms, and we have to spot cases where we believe that there is no significant difference between the various natural language forms and introduce appropriate rules for treating one as canonical.

Gerunds and gerundives occur in very much the same places as bare NPs, and have very much the same feeling of being about types of entity.

- (5) a. Exercise is good for you.
b. Swimming is good for you.
- (6) a. I like watching old movies.
b. I like old movies.

It therefore seems natural to treat them in much the same way, as descriptions of event *types*, as in Fig. 4

$$\text{utt}(\text{claim}, \\ \exists B \text{state}(B, \\ \lambda C(\exists D \text{event}(D, \text{swim}) \& \theta(D, \text{agent}, C)), \\ \lambda E(\text{good}(E))) \\ \&\text{for}(B, \text{ref}(\lambda F(\text{hearer}(F))))!4) \\ \&\text{aspect}(\text{now}, \text{simple}, B))$$

Figure 4: Logical form for (5b)

The logical form in Fig. 4 says that there is a state of affairs relating events where someone does some swimming and the property of being good, and that this state of affairs concerns the

speaker. This does at least have the benefit of exposing the key concepts mentioned in (5b), and of doing so in such a way that it is possible to write rules that support appropriate chains of inference.

The kind of inference we are interested in concerns patterns like the ones in Fig. 5

Exercise is good for you if you are overweight
Swimming is a form of exercise
I am obese
Should I go swimming?

Figure 5: A simple(!) pattern of natural reasoning

We will discuss the rules and inference engine that are required in order to support this kind of reasoning in §6 and §7. For now we are concerned with the fact that the last line in Fig. 5 could have been replaced by a number of alternative forms such as ‘*Is swimming good for me?*’ or ‘*Is it good for me to go swimming?*’ without any substantial change of meaning.

In general, we believe that determining the relationships between sentences requires inference based on background rules which describe the relationships between terms. However, when we have forms which are essentially paraphrases of one another, these rules will tend to be bi-equivalences—rules of the form $P \leftrightarrow Q$. Such rules are awkward for any theorem prover, since they potentially introduce infinite loops: in order to prove P you can try proving Q , where one of the possible ways of proving Q is by proving P , ... It is possible to catch such loops, and our inference engine does monitor for various straightforward loops of this kind, but they do introduce an extra overhead. Equivalences of this kind are, in any case, not really facts about the world so much as facts about the way natural language describes the world. It seems therefore more sensible to capture them at the point when we construct our logical forms, when they can be dealt with by straightforward pattern matching and substitution on logical forms, rather than by embodying them as bi-directional rules to be used as required by the inference engine. We use rules of the kind given in Fig. 6 to canonical versions of logical forms for sentences which we regard as mutual paraphrases. These rules are matched against elements of the logical form, and the required sub-

7 Inference

Consider (7):

- (7) a. Eating eggs will make you ill if you are allergic to eggs.
- b. I am allergic to eggs.
- c. Will eating fried-egg sandwiches make me ill?

It is pretty obvious that the answer to (7c), given (7a) and (7b), must be ‘Yes’. The reasoning that is required to arrive at this answer turns out to be surprisingly complex.

The problem is, as noted above, that we need to reason about relationships between event types. We need to be able to spot that events where someone eats a fried-egg sandwich involve situations where they eat an egg. It is clearly quite easy, if tedious, to write rules that say that if someone eats something which contains an egg then they must eat an egg, and that fried-egg sandwiches contain eggs. The trouble is that we have to be able to invoke this rule in order to determine whether the *arguments* of ‘make’ are of the right kind. Because we are (correctly) allowing event types as arguments in intensional predicates, we have to be able to invoke arbitrary and unpredictable amounts of inference even to determine whether the arguments of a predicate are admissible. Roughly speaking, we have to be prepared to carry out arbitrary amounts of inference at the point where first-order theorem provers invoke unification.

There is nothing to stop us doing this. Sorted logics, for instance, use an extended notion of unification to try to ensure that items that are being considered as arguments have specific properties (Cohn, 1987). We can, indeed, do any computation we like in order to verify the suitability of arguments. The more complex the computations we perform, of course, the longer it may take to come to a decision. The key is thus to try to bound the potential costs without compromising what we can do too much. We exploit a notion of ‘guarded’ axioms, where we allow arbitrary amounts of reasoning to be performed to verify that some item fits a fully specified description, but we do not allow such reasoning to be used for generating candidates. We do, of course, have to

put a bound on the amount of work that will be done at any point, as indeed any inference engine for a language as expressive as first-order logic must do. In general, however, using guarded intensionality in this way allows us to cover a wide range of cases which are simply inexpressible using first-order logic (or any fragment of first-order logic, such as description logic) comparatively inexpensively.

8 Conclusions

We have argued that in order to cope properly with even quite straightforward uses of language, you need large amounts of background knowledge, much of which has to be couched in some highly intensional framework, and you need inference engines which can manipulate this knowledge. In the body of the paper we have shown a number of examples which we believe illustrate this argument, and have looked at the representations and rules that we employ for dealing with these cases. The natural question that arises at this point is: that’s all very well, but can the approach outlined here be extended to cover a more substantial set of cases?

There are two key issues here. How difficult is it to capture a reasonably substantial body of knowledge within the framework we have outlined, and what will happen to the inference engine when we do?

Writing rules in property theory is very hard work. Writing rules in property theory which will mesh nicely with logical forms obtained from natural language sentences is extremely hard work. If we had to hand-code the rules we want directly in property theory (or indeed in any formal language) then the approach discussed here would, clearly, be impossible to extend to cover more than a handful of cases. Fortunately, however, we have a much easier way of constructing rules. We have, after all, a mechanism for converting natural language sentences into logical forms. So if we state the rules we want in natural language we will obtain logical forms of those rules, and furthermore those paraphrases will automatically be couched in terms which mesh nicely with logical forms obtained from other natural language sentences. Thus (8) produces the rule in Fig. 8

- (8) Eating Y will make X ill if X is allergic to Y.

$$\begin{aligned}
& \forall C \forall D \exists E \text{state}(E, C, \lambda F(\text{allergic}(F))) \ \& \ \text{to}(E, D) \\
& \quad \& \ \text{aspect}(\text{now}, \text{simple}, E) \\
& \rightarrow \exists G : \{ \text{future}(\text{now}, G) \} \\
& \quad \exists \text{Event}(B, \text{make}) \\
& \quad \quad \& \ \theta(B, \text{object}, C) \\
& \quad \quad \& \ \theta(B, \text{object}1, \lambda H(\text{ill}(H))) \\
& \quad \quad \& \ \theta(B, \\
& \quad \quad \quad \text{agent}, \\
& \quad \quad \quad \lambda I \exists \text{Event}(J, \text{eat}) \\
& \quad \quad \quad \quad \& \ \theta(J, \text{object}, D) \\
& \quad \quad \quad \quad \& \ \theta(J, \text{agent}, I)) \\
& \quad \& \ \text{aspect}(G, \text{simple}, B)
\end{aligned}$$

Figure 8: Logical form for (8)

Writing rules like (8) is clearly easier than producing formulae like Fig. 8 by hand. Writing down all the knowledge you need in order to cope with a non-trivial domain is still a very substantial task, but doing it in English is at least feasible in a way that doing it directly in a formal language is not.

How will the inference engine cope when confronted with thousands of rules? Very large parts of everyday knowledge can, in fact, be expressed pretty much as Horn clauses. Our inference engine converts Horn clauses into (almost) pure Prolog, and there is certainly no problem in using very large sets of Horn clauses converted to this form (a modern Prolog system will cope comfortably with sets of several hundred thousand Horn clauses, and will carry out substantial inference chains involving such sets in small fractions of a second). The only concern here relates to non-Horn clauses (which do not tend to occur all that frequently in rules explaining the relationships between natural language terms) and intensional rules. The fact that most intensional rules are guarded has certainly meant that so far we have not encountered any problems when using them, and we are hopeful that this will remain the case.

In any case, there is an alternative question to be answered: what will happen if you *don't* take the approach outlined here? All the phenomena we have discussed are widespread—bare plurals, mutual paraphrases, intensional attitudes all occur all over the place. It is extremely hard to see that systems that rely on surface patterns (either directly, as in textual entailment, or indirectly through shallow parsing/information extraction)

can support the kind of reasoning required for getting from ‘*I have an allergy to eggs.*’ to ‘*It is dangerous for me to eat pancakes.*’, so at some point inference based on background knowledge will have to be invoked. There seems little alternative to constructing formal paraphrases that capture the subtleties of natural language in all its glory. If you don’t, then you will by definition lose some of the information that was expressed in the text, and that will inevitably mean that you get things wrong. There is no way round it: either you bite the bullet, construct formal paraphrases that capture the content of the input and use them to carry out inference, or you will get some things wrong.

References

- Aczel, P. (1988). *Non-Well-Founded-Sets*. CSLI Publications, Stanford.
- Barwise, J. and Perry, J. (1983). *Situations and Attitudes*. Bradford Books, Cambridge, MA.
- Carlson, G. (1989). On the semantic composition of English generic sentences. In G. Chierchia, B. H. Partee, and R. Turner, editors, *Properties, Types and Meaning II: Semantic Issues*, pages 167–192, Dordrecht. Kluwer Academic Press.
- Cohen, A. (1994). Reasoning with generics. In H. C. Bunt, editor, *1st International Workshop on Computational Semantics*, pages 263–270, University of Tilburg.
- Cohn, A. G. (1987). A more expressive formulation of many sorted logic. *Journal of Automated Reasoning*, **3**, 113–200.
- Dagan, I., Magnini, B., and Glickman, O. (2005). The PASCAL recognising textual entailment challenge. In *Proceedings of Pascal Challenge Workshop on Recognizing Textual Entailment*.
- Davidson, D. (1967). The logical form of action sentences. In N. Rescher, editor, *The Logic of Decision and Action*, Pittsburgh. University of Pittsburgh Press.
- Davidson, D. (1980). *Essays on actions and events*. Clarendon Press, Oxford.
- Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition and Logical Form*. Academic Press, New York.
- Jech, T. J. (1971). *Lectures in Set Theory, with Particular Emphasis on the Method of Forcing*. Springer Verlag (Lecture Notes in Mathematics 217), Berlin.
- Ramsay, A. M. (1992). Bare plural NPs and habitual VPs. In *Proceedings of the 14th International Conference on Computational Linguistics (COLING-92)*, pages 226–231, Nantes.
- Ramsay, A. M. (2001). Theorem proving for untyped constructive λ -calculus: implementation and application. *Logic Journal of the Interest Group in Pure and Applied Logics*, **9**(1), 89–106.
- Ramsay, A. M. and Field, D. G. (2007). Speech acts, epistemic planning and Grice’s maxims. *Logic and Computation*.
- Reichenbach, H. (1947). *Elements of Symbolic Logic*. The Free Press, New York.
- Reichenbach, H. (1956). *The Direction of Time*. University of California Press, Berkeley.
- Turner, R. (1987). A theory of properties. *Journal of Symbolic Logic*, **52**(2), 455–472.
- van der Sandt, R. (1992). Presupposition projection as anaphora resolution. *Journal of Semantics*, **9**, 333–377.
- Whitehead, A. N. and Russell, B. (1925). *Principia Mathematica*. Cambridge University Press, Cambridge.