

Vagueness: Facts, Models, and Programs

Kees van Deemter
ITRI, University of Brighton
Kees.van.Deemter@itri.brighton.ac.uk
<http://www.itri.brighton.ac.uk/~Kees.van.Deemter/>

Preface

This course will explore some of the formal semantics literature on vagueness, juxtapose it with experimental results, and ask what it would take for an NLP program to behave in a natural way (e.g., to pass the Turing test) with respect to its treatment of vagueness. A number of language-generating programs will be discussed and, if time allows, some language-interpreting ones as well.

This *reader* contains a number of recent papers in this area. References to classic papers related to this course (e.g., by Dummett, Kamp, and Klein) can be found in these papers.

Contents

- Goyal and Shoham (1993). “Reasoning Precisely with Vague Concepts”. In Procs. of AAAI Conference, Washington 1993.¹
- Dominic Hyde (2002). “Sorites Paradox”.² The Stanford Encyclopedia of Philosophy (Fall 2002 Edition), Edward Zalta (Ed.) Also downloadable from <http://plato.stanford.edu/archives/fall2002/entries/sorites-paradox/>.
- Kyburg and Morreau (2000). M. Morreau and A. Kyburg, “Fitting Words: Vague Language in Context”.³ Linguistics and Philosophy 23:577-597.
- Stone and DeVault (2004). “Interpreting Vague Utterances in Context”.⁴ To appear in Procs. of COLING 2004, Geneva.
- van Deemter (1992). “The Role of Ambiguity in the Sorites Fallacy”. In: P.Dekker and M.Stokhoff (Eds.), Procs. of 9th Amsterdam Colloquium, pp.209-227, Amsterdam.
- van Deemter (2000). “Generating Vague Descriptions”, In Procs. of First International Conference on Natural Language Generation (INLG-2000), Mitzpe Ramon.
- van Deemter (2004). “Finetuning an NLG System through Experiments with Human Subjects: the Case of Vague Descriptions”. In procs. of 3rd International Conference on Natural Language Generation (INLG-04).

¹Printed with kind permission from the American Association for Artificial Intelligence.

²Printed with kind permission from the author.

³Printed with kind permission from Kluwer Academic Publishers.

⁴Printed with kind permission from the authors.

the AAAI Press / The MIT Press

distributed by The MIT Press
Massachusetts Institute of Technology
Cambridge, Massachusetts 02142

ISBN 0-262-51071-5

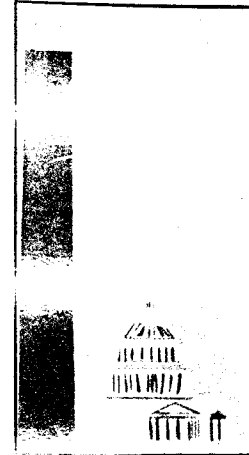
AAAI93



AAAI-93

Proceedings, Eleventh National Conference on Artificial Intelligence

AAAI-93



Proceedings of the Eleventh National Conference on Artificial Intelligence

Sponsored by the American Association for Artificial Intelligence

July 11 - 15, 1993

Washington, DC

Reasoning Precisely with Vague Concepts*

Nita Goyal and Yoav Shoham
Robotics Laboratory, Computer Science Department
Stanford University
Stanford, CA 94305
{nita,shoham}@cs.stanford.edu

Abstract

Many knowledge-based systems need to represent vague concepts. Although the practical approach of representing vague concepts as precise intervals over numbers is well-accepted in AI, there is no systematic method to delimit the boundaries of intervals, only *ad hoc* methods. We present a framework to reason precisely with vague concepts based on the observation that the vague concepts and their interval-boundaries are constrained by the underlying domain knowledge. The framework is comprised of a constraint language to represent logical constraints on vague concepts, as well as numerical constraints on the interval-boundaries; a query language to request information about the interval boundaries; and a computational mechanism to answer the queries. A key step in answering queries is preprocessing the constraints by extracting the numerical constraints from the logical constraints and combining them with the given numerical constraints.

1 Introduction

The input to an AI system embedded in a real-world environment is often numerical whereas the reasoning is done with abstract symbols. Many abstract symbols embody vague concepts over continuous numerical ranges. To quote Davis, "In some respects, the concepts of commonsense knowledge are *vague*,... Many categories of common sense have no well-marked boundary lines; there are clear examples and clear nonexamples, but in between lies an uncertain region that we cannot categorize, even in principle." [Davis 1990]. For example, there is no minimum precise body temperature that a doctor considers high and there is no maximum number of hairs that a person might have and still be considered bald.

The representation of such vagueness poses a problem. "From a theoretical point of view, this vagueness is extremely difficult to deal with, and no re-

*This research has been supported by grant AFOSR-89-0326.

ally satisfactory solutions have been proposed." [Davis 1990]. Some of the approaches that try to address this theoretical difficulty are fuzzy logic [Zadeh 1983] and vague predicates [Parikh 1983]. However, it is commonly accepted in AI that, though inadequate theoretically, in practice it is often adequate to assume that a vague concept is precise and that there is indeed a well-defined boundary. In fact, most system builders who encounter the vagueness problem [Hayes-Roth et al. 1989; Shahar, Tu & Musen 1992] adopt a similar approach of representing a vague concept as an interval over the range of numbers. This practical approach is illustrated by an example from [Davis 1990]: "Suppose that 'bald' did refer to some specific number of hairs on the head, only we do not know which number. We know that a man with twenty thousand hairs on his head is not bald, and that a man with three hairs on his head is bald, but somewhere in between we are doubtful." This precise representation of the vague concept bald is still useful for reasoning.

Despite the pervasiveness of the vagueness problem, and the pervasiveness of the practical approach of representing vague concepts as intervals, there has been no effort in AI to provide a systematic account of this practical approach. We propose a framework for representing and reasoning with vague concepts as intervals that has the advantages of (1) improving our understanding of the issues involved in the practical approach, and (2) replacing the *ad hoc* approach used by system designers to delimit the interval-boundaries.

The framework is based on the observation that vague concepts and their interval-boundaries (also referred to as *thresholds*) are constrained by the underlying domain knowledge that must be used to reason about the thresholds. We motivate the components of the framework by extending Davis' baldness example.

Example 1: "Anyone with 3 or fewer hairs is bald and anyone with 20000 or more hairs is not bald" "All old people are bald" (note that "old" itself is a vague concept that we will assume has a well-defined boundary) "Anyone who is 50 years or younger is not old whereas anyone over 80 is old"

"All presidents of companies are old"

"Tom's age is 70 years, he has 500 hairs and is the president of a company"

"Jim's age is 75 years and he has 800 hairs"

"Sam's age is 45 and he has 650 hairs"

Is Tom bald? Logical reasoning tells us that since Tom is president of a company, he is old and therefore bald. Note that here we used only the logical relations between the concepts *president*, *old* and *bald*, where old and bald are vague concepts but *president* is not.

Is Jim bald? We can reason that since Tom is old, the oldness threshold¹ can be at most 70. Since Jim's age is 75 which is over the oldness threshold, he must be old and therefore bald. Note that here we needed numerical reasoning with Tom and Jim's ages and oldness threshold, as well as logical reasoning that since Jim is old he must be bald.

We can ask if the baldness threshold is necessarily more than 800? Since Jim is bald and has 800 hairs, the baldness threshold must be at least 800. Therefore, the answer to the query is *yes* and hence anyone with less than 800 hairs is bald. Here we needed numerical reasoning about Jim's hairs and the baldness threshold.

Is Sam bald? Since anyone with less than 800 hairs is bald, and Sam has only 650 hairs, he must be bald. Here we needed numerical reasoning with number of hairs on Sam's head and the baldness threshold. ■

As illustrated by this example, we need to *represent* both logical relations between symbolic concepts and numerical relations on thresholds. Also, logical as well as numerical *reasoning* is required to answer the interesting queries. Hence, the proposed framework facilitates this representation and supports queries about the thresholds.

Framework: The framework is comprised of three main parts – a constraint language to express domain knowledge, a query language to query the domain knowledge and a computational mechanism to answer the queries.

The first part of the framework is a *constraint language* that captures the domain knowledge. The language enables the expression of logical constraints on the vague concepts as well as numerical constraints on the thresholds of these concepts. Explicit representation of the thresholds is important to represent the numerical constraints and as we shall see, to ask queries.

The second part of the framework is a *query language* that extracts relevant information about the thresholds implied by the domain knowledge. In particular, the queries enable us to delimit the thresholds based on the information provided in the domain knowledge².

¹By oldness threshold we mean that age such that everyone of higher age is old whereas everyone of lower age is not old. The baldness threshold is defined analogously.

²Note that it is not necessary to assign specific values to the thresholds to answer any queries, although this as-

This is exactly what a system designer needs to define intervals for a vague concept that are consistent with the domain knowledge. For example, the answer to the query "what is the minimum permissible value for the baldness threshold?" provides the designer with useful information to define the interval for bald.

The third part of the framework is a *computational mechanism* to answer the queries in the query language using the domain knowledge expressed in the constraint language.

In Sections 2 and 3, we introduce particular constraint and query languages. In Section 4, we describe a computational mechanism to answer queries for these languages. It includes a sound and complete algorithm, a discussion of the complexity, and experimental results illustrating the applicability of the framework. Our experiments were carried out in the domain of medical diagnosis where numerical measurements of parameters such as blood pressure and heart rate are abstracted to vague concepts such as high and low blood pressure and used for the diagnosis of the patient's condition. In this paper, for the sake of clarity and understanding, we stick to the more everyday example of bald people.

2 Constraint Language

To express the domain knowledge, the constraint language must have an explicit representation of thresholds, a language to express numerical constraints and a language to express the logical constraints. We present such a language here, chosen for its familiarity as well as to strike a tradeoff between expressivity and efficiency of answering queries.

The vague predicates in the logical language are distinguished from the other predicates. We refer to the vague predicates, which must all be unary, as *interval-predicates* and to all other predicates as *noninterval-predicates*. The set of *interval-predicates* is denoted by *IP*, and the set of *noninterval-predicates* by *NIP*. With every $P \in IP$ we associate two *threshold terms* P^- and P^+ , called the *lower* and *upper thresholds* of P , respectively. The set of all threshold terms is denoted by T ($T = \{P^-, P^+ \mid P \in IP\}$). The interval-predicates will be interpreted in a special way to reflect our intuition about the vague predicates: P will be interpreted as the interval $[P^-, P^+]$ over $\mathbb{R}$, the set of real numbers. We will refer to this interpretation as the *predicate-as-interval assumption*.

1. Numerical Constraints: The language of numerical constraints is that of linear arithmetic inequalities where the threshold terms in T are the variables of the inequalities. A numerical constraint must be reducible to the form $(a_1x_1 + \dots + a_nx_n) \text{ rel } b$, where $a_1, \dots, a_n, b \in \mathbb{R}$, $x_1, \dots, x_n \in T$, and $\text{rel} \in \{<, >, \leq, \geq, =\}$. We denote the set of numerical constraints by *NC*.

signment is made much easier in our framework.

2. **Logical Constraints:** These are definite Horn clauses without function symbols (also called Datalog sentences in the deductive database literature [Ullman 1988]). The predicates of these logical constraints are interval-predicates IP as well as noninterval-predicates NIP . We denote the set of logical constraints by LC .

The constraints in Example 1 are represented in the language as follows. We extend the example to include another constraint that all rich VPs become presidents of companies.

Example 2:

$IP = \{\text{bald}, \text{old}, \text{rich}\}$
 $NIP = \{\text{age}, \text{hairs}, \text{pres}, \text{money}, \text{was.VP}\}$
 $NC = \{\text{bald}^- = 0, 3 \leq \text{bald}^+ \leq 20000, \text{old}^+ = \infty, 50 \leq \text{old}^- \leq 80, 0.1 \leq \text{rich}^- \leq 1, \text{rich}^+ = \infty\}$
 $\cup \{P^- \leq P^+ \mid P \in IP\}$

The unit for bald is number of hairs, for old is age in years, and for rich is money in millions of dollars.

$LC = \{\text{pres}(x) \leftarrow \text{was.VP}(x) \wedge \text{money}(x, y) \wedge \text{rich}(y) \mid \text{bald}(z) \leftarrow \text{old}(y) \wedge \text{age}(x, y) \wedge \text{hairs}(x, z) \mid \text{old}(y) \leftarrow \text{pres}(x) \wedge \text{age}(x, y) \mid \text{age}(\text{Tom}, 70), \text{hairs}(\text{Tom}, 500), \text{was.VP}(\text{Tom}), \text{money}(\text{Tom}, 6), \text{age}(\text{Jim}, 75), \text{hairs}(\text{Jim}, 800) \mid \text{age}(\text{Sam}, 45), \text{hairs}(\text{Sam}, 650)\}$

There are other languages that combine quantitative and qualitative constraints. For instance, Williams' qualitative algebra [Williams 1988] expresses operations on reals and signs of reals, but is not concerned with logical constraints. Similarly, [Meiri 1991] and [Kautz & Ladkin 1991] present frameworks for expressing and processing both quantitative and qualitative temporal constraints. Their language limits the constraints, whether numerical or logical, to be binary whereas our language does not. On the other hand, their language can express disjunctive relations between intervals which our language does not.

Most closely related to our language are languages for *constraint logic programming* (CLP) in the style of Lassez et al. [Jaffar & Lassez 1987]. CLP considers general Horn theories, as opposed to our limited Datalog theories. However, CLP does not allow numerical constraints in the head of a clause. In our language the interval-predicates can occur in the head which, if represented in CLP, would correspond to numerical constraints occurring in the head.

3 Query Language

The purpose of the query language is to enable a user to extract information about the thresholds that is implied by the domain constraints. It is a useful tool for a system designer to find the threshold values allowed by the constraints. The kind of queries supported are informally described below. Here $P_1^{th}, \dots, P_n^{th} \in \mathcal{T}$, $a_1, \dots, a_n \in \mathbb{R}$, $rel_1, \dots, rel_n \in \{\leq, \geq, <, >, =\}$, and $i \in \{1, \dots, n\}$.

1. Is it *necessarily* the case that $(P_1^{th} rel_1 a_1) \wedge \dots \wedge (P_n^{th} rel_n a_n)$?
2. Is it *possibly* the case that $(P_1^{th} rel_1 a_1) \wedge \dots \wedge (P_n^{th} rel_n a_n)$?
3. What is the *minimum* value that P_i^{th} can take?
4. What is the *maximum* value that P_i^{th} can take?

Many queries may be derived using the above primitives. For example, the query " $P(a)$?" can be cast as "Is it necessarily the case that $(P^- \leq a) \wedge (P^+ \geq a)$?" If the answer is *yes* then $P(a)$ is *true*, otherwise it is unknown. If the answer to "Is it possibly the case that $(P^- \leq a) \wedge (P^+ \geq a)$?" is *no* then $P(a)$ is *false*, otherwise it is unknown.

In addition, it is possible to request a specific assignment of values to the thresholds that satisfies the constraints. For Example 2, assigning the value 2000 to the baldness threshold is consistent with the given constraints. We will indicate briefly in Section 4.3 how this assignment is made. This procedure is particularly useful for a system designer who assigns specific numbers to the thresholds in the design stage of the system.

4 Computational Mechanism for Answering Queries

The final component of our framework is the computational mechanism responsible for answering queries on constraints. A key step in answering queries is preprocessing the constraints by extracting the numerical constraints from the logical constraints and combining them with the given numerical constraints. The preprocessing is a two-step procedure: first, using the predicate-as-interval assumption, the procedure extracts the numerical information on the interval-predicates from the logical constraints LC . This derived information is in the form of disjunctions of numerical constraints. Next, the procedure combines these disjunctive constraints with the given numerical constraints NC . We describe the procedure to extract numerical information from LC in Section 4.1. In Section 4.2 we prove that the procedure is sound and complete and also discuss the complexity issues. In Section 4.3 we describe how to combine the numerical information from LC with NC .

4.1 Extracting Numerical Information from Logical Constraints

The algorithm *Symb.to.Numeric* described in this section takes as input the logical constraints LC and the sets of interval and noninterval-predicates IP and NIP , and returns a set of numerical constraints $quant.LC$. This process of conversion from logical to numerical constraints preserves the information about the thresholds of interval-predicates but discards the information on the noninterval-predicates. A formal discussion is deferred to Section 4.2.

```
Function Symb.to.Numeric(LC, IP, NIP) : quant.LC
S ← ∅;
for every clause c ∈ LC such that head(c) ∈ IP do
  S_c ← Expand(c, LC, IP, NIP);
  S ← S ∪ S_c; /* S has no NIP predicates */
endfor
quant.LC ← ∅;
for every clause c ∈ S do
  quant.LC ← quant.LC ∪ Convert.LC.to.NC(c);
return(quant.LC)
endfunction
```

Figure 1: Numerical Information from Logical Constraints

The algorithm *Symb.to.Numeric* is described in Figure 1. Starting with all those clauses in LC that have interval-predicates at the head, we *expand* their bodies using other clauses in LC until all noninterval-predicates are eliminated from the body. *Expand* is very similar to SLD resolution [Lloyd 1987] but with two differences: (1) only noninterval-predicates are expanded (2) all possible expansions are computed. Thus, each clause in set S of Figure 1 has only interval-predicates. Using the predicate-as-interval assumption, we convert the resultant clauses to numerical constraints as described by function *Convert.LC.to.NC* in Figure 2. This function works by fragmenting each clause into subclauses such that each subclause has at most one variable and no two subclauses have the same variable³. For example, the clause $P(a) \leftarrow Q(x) \wedge R(x) \wedge S(b)$ is a disjunction of three subclauses: " $P(a)$ ", " $\leftarrow Q(x) \wedge R(x)$ " and " $\leftarrow S(b)$ ". In general, each subclause thus obtained will be one of the six basic types described in Figure 2. Each type of subclause is converted to numerical constraint by using the predicate-as-interval assumption, and by interpreting the connectives $\neg, \vee, \wedge$ as complement, union and intersection of intervals, respectively.

An application of the algorithm on Example 2 is illuminating:

Example 3: The first step in the procedure is to locate clauses with interval-predicates at the head in LC and expand them until all noninterval-predicates are eliminated. Here there are two such clauses with old and bald at the head. On expansion, we obtain set S :

$\text{bald}(500) \leftarrow \text{old}(70) \quad \text{old}(70) \leftarrow \text{rich}(6)$
 $\text{bald}(800) \leftarrow \text{old}(75) \quad \text{bald}(650) \leftarrow \text{old}(45)$

On applying the function *Convert.LC.to.NC*, each of these clauses fragments into subclauses of the first two types: " $P(a)$ " and " $\leftarrow P(a)$ ". On conversion we obtain the set $quant.LC$ ⁴:

$quant.LC = \{(\text{bald}^- \leq 500 \leq \text{bald}^+) \vee (70 < \text{old}^-) \vee (\text{old}^+ < 70)$

³Note that this is always possible because all interval-predicates are unary.

⁴Note that each of the 4 clauses obtained here will actually split into 2 clauses.

```
Function Convert.LC.to.NC(lc) : nc
nc ← ∅;
lc.subclauses ← Make.Subclauses(lc); /* Every
subclause of lc with a constant or the same variable */
for every subclause subcl ∈ lc.subclauses do
  Case subcl of: /* a is a constant */
    "P(a)": subcl' ← (P^- ≤ a ≤ P^+)
    "← P(a)": subcl' ← (a < P^-) ∨ (a > P^+)
    "P(x)": subcl' ← (P^- = -∞) ∧ (P^+ = +∞)
    "← P(x)": subcl' ← P^- > P^+
    "← P_1(x), ..., P_n(x)":
      subcl' ← ∨_{i=1}^n ∨_{j=1}^n (P_i^- > P_j^+)
    "P(x) ← Q_1(x), ..., Q_n(x)":
      subcl' ← (∨_{i=1}^n P_i^- ≤ Q_i^-) ∧ (∨_{i=1}^n P_i^+ ≥ Q_i^+);
nc ← nc ∨ subcl'
endfor
return(nc)
endfunction
```

Figure 2: Conversion from Logical to Linear Arithmetic Constraint

$(\text{old}^- \leq 70 \leq \text{old}^+) \vee (6 < \text{rich}^-) \vee (\text{rich}^+ < 6)$
 $(\text{bald}^- \leq 800 \leq \text{bald}^+) \vee (75 < \text{old}^-) \vee (\text{old}^+ < 75)$
 $(\text{bald}^- \leq 650 \leq \text{bald}^+) \vee (45 < \text{old}^-) \vee (\text{old}^+ < 45)$

4.2 Formal results on conservation of numerical information

We establish formally that no numerical information is lost in the conversion performed by algorithm *Symb.to.Numeric*. We begin by defining the models of LC that are faithful to the predicate-as-interval assumption; we call these the *standard models*. Specifically, in all standard models $M = (D, \mu)$ over a domain D , the interpretation function μ will have to map interval-predicates to intervals over the reals. In the following, $\mathbb{R}$ denotes the set of real numbers.

Definition 1: Given a set of logical constraints LC , the set of interval-predicates IP , and the set of threshold terms $\mathcal{T}$, a *standard model of LC w.r.t. IP* is a model $M = (D, \mu)$ such that $M \models LC$, and for every $P \in IP$ there exist $P^-, P^+ \in \mathcal{T}$ and it is the case that $\mu(P^-), \mu(P^+) \in \mathbb{R}$ and $\mu(P) = \{x \mid \mu(P^-) \leq x \leq \mu(P^+), x \in \mathbb{R}\}$. ■

Definition 2: Given LC , IP and $\mathcal{T}$ as above, a *numerical submodel of LC w.r.t. IP* is a model $M = (\mathbb{R}, \mu)$ such that there is some standard model $M' = (D, \mu')$ of LC w.r.t. IP , and μ is the restriction of μ' to terms in $\mathcal{T}$. ■

The following theorem establishes that the algorithm *symb.to.numeric* is sound and complete w.r.t. the numerical information (complete proof in [Goyal 1993]).

Theorem 4: (Soundness and Completeness) *The class of numerical submodels of LC w.r.t. IP is identical to the class of models of quant.LC.*

Proof: (Sketch for Completeness) An arbitrary model M of $quant.LC$ is extended to a standard model M' of

LC such that its numerical submodel is exactly M . M' is constructed by first building a dependency graph of predicates in LC and then by defining the interpretation of the predicates in NIP in the topological order of the graph. The intuition is that when a clause in LC is used to build the interpretation of the predicate in the head from the predicates in the body, the body predicates would have been already interpreted because of the order of interpretation. The equivalence of the numerical submodel of M' and the model M is proved through mathematical induction on the topological order of predicates. ■

In the worst case, the space and time complexity of computing $quant_LC$ is exponential in the size of LC . This is not surprising, since in the worst case, $quant_LC$ is of exponential size. However, we have identified syntactic restrictions on the constraint language for which we can avoid such exponential blowup. In practice, the performance of the algorithm has been found to be quite acceptable for the following reasons. First, we observe that the algorithm is exponential only in the size of the non-ground constraints. Typically, the number of non-ground constraints is small compared to the number of ground literals. Second, this algorithm is invoked only once for all the queries on a given set of constraints; hence, the cost is amortized over all the queries. Thus, the overall performance of the system is not severely affected despite the apparent intractability. A more detailed discussion of the complexity issues may be found in [Goyal 1993].

4.3 Combining with Numerical Constraints

The constraints in the set $quant_LC$, obtained by converting the logical constraints to numerical constraints, are disjunctive. These constraints must be combined with the set of given numerical constraints NC to answer the queries. In principle, we can convert the set $quant_LC$ to disjunctive normal form (DNF) and add the constraints NC to each disjunct. The disjunction thus obtained is referred to as $output_C$. However, in practice, we leave $quant_LC$ in its conjunctive normal form to save space, and generate the disjuncts of $output_C$ one by one through backtracking. Furthermore, to make the process more efficient, we first reduce the size of the set $quant_LC$ using the constraints NC . We elaborate on these below and also discuss how existing methods are applicable to answer queries on a single disjunct of $output_C$.

Pruning $quant_LC$: We have developed a procedure that uses the constraints in NC to reduce the size of the set $quant_LC$ significantly. This procedure, called *reduce*, uses the upper and lower bounds of all thresholds implied by the constraints in NC to prune $quant_LC$ in two ways. First, if a disjunct of a constraint in $quant_LC$ is already satisfied by the bounds, then that constraint can be deleted from $quant_LC$. In

Example 3, the lower and upper bounds for old^- are 50 and 80 respectively, hence $(old^- > 45)$ is already satisfied. Second, if a disjunct is inconsistent with the bounds, then that disjunct can be deleted. For instance, $(old^- > 82)$ is inconsistent with the bounds for old^- .

The experimental results confirm that the procedure *reduce* reduces the size of $quant_LC$ significantly. In Example 3, $quant_LC$ has 8 constraints with 3 disjuncts each that should give rise to 3^8 disjuncts (in DNF) in the worst case. Applying procedure *reduce* eliminates all but 1 disjunct, that is:

$$output_C = NC \cup \{(old^- \leq 70), (bald^+ \geq 800)\}$$

When the procedure was applied to the medical diagnosis domain, in the first application $quant_LC$ had 12 constraints with 3 disjuncts each, giving rise to 3^{12} disjuncts in the worst case. Procedure *reduce* eliminated all but 2 disjuncts. In a second medical application, $quant_LC$ had 416 constraints with 2 or 3 disjuncts per constraint that would have given rise to at least 2^{416} disjuncts. Procedure *reduce* eliminated all but 2592 disjuncts.

Generating disjuncts of $output_C$: Once the set $quant_LC$ has been pruned, queries are answered by generating the remaining disjuncts of $output_C$ one at a time through backtracking. We avoid generating redundant disjuncts in $output_C$ by recognizing the presence of common disjuncts in the constraints of $quant_LC$. For instance, in the second medical application, only 184 disjuncts had to be generated out of the 2592 that were possible.

In practice, most queries do not require backtracking even over all possible distinct disjuncts that are generated. For instance, a query whether a constraint is *possibly* true or not, has to find any one disjunct over which the constraint is satisfied. Furthermore, even for queries where all disjuncts have to be checked, an approximate answer can be obtained by computing only on a few disjuncts. For instance, a query to find the minimum value of a threshold can return the minimum over only a few disjuncts. This approximate answer is still useful since it supplies a lower bound on the threshold, even though not the tightest lower bound. Thus, this procedure gives a useful approximate answer any time that an answer is required, and the approximation gets closer to the optimal as the allowed time increases. The experimental results on answering queries are available in [Goyal 1993].

We can even assign a specific value to the thresholds using heuristic criteria. For Example 3, the baldness threshold ($bald^+$) could be 10400 which is halfway between its bounds of 800 and 20000, and satisfies all the given constraints. When a large amount of ground data is available, clustering techniques are utilized to assign a specific value.

Answering for each disjunct: We have discussed previously how the set $quant_LC$ is pruned *a priori* to eliminate redundant disjuncts and how the disjuncts

of $output_C$ are generated. Each disjunct thus generated is a set of linear arithmetic constraints. We now discuss how any query is answered on a single disjunct. The queries for maximum and minimum values of thresholds (queries 3 and 4 in Section 3) require the computation of lower and upper bounds of thresholds. Queries for checking a constraint for consistency (queries 1 and 2 in Section 3) require a consistency check on a set of constraints. Thus, any existing method for computing bounds and checking consistency of linear arithmetic constraints can be used. If NC has only simple order relations or bounded differences, we can use an efficient $O(n^3)$ procedure (where n is the number of variables) from [Davis 1987] or [Meiri 1991]. Sacks' *bounder* [Sacks 1990] is applicable but more useful for nonlinear constraints. For more general linear constraints, we have to use a linear programming method that is still tractable $O(n^{3.5}L)$ (L is size of input) [Karmarkar 1984]. Lassez's work on canonical form of generalized linear constraints [Huynh et al. 1990] has potential applications, though the advantage of a canonical form would be offset by the cost of maintaining the canonical form because we backtrack on disjunctive constraints.

5 Conclusions

We have provided a systematic account of the practical approach of representing vague concepts as precise intervals over numbers. Based on the observation that the vague concepts and their interval-boundaries are constrained by the underlying domain knowledge, we motivated and proposed a framework to reason precisely with vague concepts. The framework is comprised of a constraint language to represent the domain knowledge; a query language to request information about the interval boundaries; and a computational mechanism to answer the queries.

We described the constraint and query languages and a computational mechanism to answer queries. A key step in answering queries is preprocessing the constraints by extracting the numerical constraints from the logical constraints and combining them with the given numerical constraints. We proved this algorithm to be sound and complete and also discussed the complexity issues. Some experimental results of applying this framework to a medical domain were discussed.

The main contribution of our work is in providing a systematic framework to understand the common though *ad hoc* approach of representing vague predicates as intervals. This work is particularly applicable to a knowledge base during its development stage where the vague concepts over numbers need to be defined precisely.

Acknowledgements We would like to thank Surajit Chaudhuri, Ashish Gupta, Alon Levy, Pandu Nayak, Moshe Tennenholtz, Becky Thomas and the anonymous reviewers.

References

- Davis, E. 1987. Constraint Propagation with Interval Labels. *Artificial Intelligence* 32(3):281-331.
- Davis, E. 1990. *Representations of Commonsense Knowledge*. Morgan Kaufmann Publishers, 19-20.
- Goyal, N. 1993. A Framework for Reasoning Precisely with Vague Concepts. Ph.D. diss. (in preparation), Dept. of Computer Science, Stanford University.
- Hayes-Roth, B.; Washington, R.; Hewett, R.; Hewett, M.; and Seiver, A. 1989. Intelligent Monitoring and Control. In *Proc. of Eleventh International Joint Conference on Artificial Intelligence*, 43-249.
- Huynh, T.; Joskowicz, L.; Lassez, C.; and Lassez, J.-L. 1990. Reasoning about Linear Constraints using Parametric Queries. In *Proc. of Tenth FST TCS*, Bangalore, India.
- Jaffar, J.; and Lassez, J.-L. 1987. Constraint Logic Programming. In *Proc. of 14th ACM Symposium on Principles of Programming languages*, 111-119.
- Karmarkar, N. 1984. A New Polynomial-Time Algorithm for Linear Programming. *Combinatorica* 4:373-395.
- Kautz, H.A.; and Ladkin, P.B. 1991. Integrating Metric and Qualitative Temporal Reasoning. In *Proc. of Ninth National Conference on Artificial Intelligence*, 241-246.
- Lloyd, J.W. 1987. *Foundations of Logic Programming*, 2nd. ed.. Springer-Verlag.
- Meiri, I. 1991. Combining Qualitative and Quantitative Constraints in Temporal Reasoning. In *Proc. of Ninth National Conference on Artificial Intelligence*, 260-267.
- Parikh, R. 1983. The problem of Vague Predicates. In Cohen and Wartofsky (eds.), *Language, Logic, and Method*. Reidel Publishers, 241-261.
- Sacks, E. 1990. Hierarchical Reasoning about Inequalities. In *Readings in Qualitative Reasoning about Physical Systems*, eds. D.S. Weld and J. de Kleer. Morgan Kaufmann Publishers, 344-350.
- Shahar, Y.; Tu, S.W.; and Musen, M.A. 1992. Knowledge Acquisition for Temporal-Abstraction Mechanisms. *Knowledge Acquisition* 4:217-236.
- Ullman, J.D. 1988. *Principles of Database and Knowledge-Base Systems*, Vol. 1. Computer Science Press.
- Williams, B.C. 1988. MINIMA: A Symbolic Approach to Qualitative Algebraic Reasoning. In *Proc. of Seventh National Conference on Artificial Intelligence*, 264-269.
- Zadeh, L.A. 1983. Commonsense and Fuzzy Logic. In *The Knowledge Frontier: Essays in the Representation of Knowledge*, eds. N. Cercone and G. McCalla. New York: Springer-Verlag, 103-136.

Sorites Paradox

The sorites paradox is the name given to a class of paradoxical arguments, also known as little-by-little arguments, which arise as a result of the indeterminacy surrounding limits of application of the predicates involved. For example the concept of a heap appears to lack sharp boundaries and, as a consequence of the subsequent indeterminacy surrounding the extension of the predicate ‘is a heap’, no one grain of wheat can be identified as making the difference between being a heap and not being a heap. Given then that one grain of wheat does not make a heap, it would seem to follow that two do not, thus three do not, and so on. In the end it would appear that no amount of wheat can make a heap. We are faced with paradox since from apparently true premises by seemingly uncontroversial reasoning we arrive at an apparently false conclusion.

This phenomenon at the heart of the paradoxes is now recognised as the phenomenon of vagueness. Once identified, vagueness can be seen to be a feature of syntactic categories other than predicates, nonetheless one speaks primarily of the soriticality of predicates. Names, adjectives, adverbs and so on are only susceptible to paradoxical sorites reasoning in a derivative sense.

Sorites arguments of the paradoxical form are to be distinguished from multi-premise syllogisms (polysyllogisms) which are sometimes also referred to as sorites arguments. Whilst both polysyllogisms and sorites paradoxes are chain-arguments, the former need not be paradoxical in nature and the latter need not be syllogistic in form.

- 1. The Sorites In History
- 2. Its Paradoxical Forms
- 3. Responses
- Bibliography
- Other Internet Resources
- Related Entries

1. The Sorites In History

The name ‘sorites’ derives from the Greek word *soros* (meaning ‘heap’) and originally referred, not to a paradox, but rather to a puzzle known as *The Heap*: Would you describe a single grain of wheat as a heap? No. Would you describe two grains of wheat as a heap? No. ... You must admit the presence of a heap sooner or later, so where do you draw the line?

It was one of a series of puzzles attributed to the Megarian logician Eubulides of Miletus. Also included were:

The Liar: A man says that he is lying. Is what he says true or false?

The Hooded Man: You say that you know your brother. Yet that man who just came in with his head covered is your brother and you did not know him.

The Bald Man: Would you describe a man with one hair on his head as bald? Yes. Would you describe a man with two hairs on his head as bald? Yes. ... You must refrain from describing a man with ten thousand hairs on his head as bald, so where do you draw the line?

This last puzzle, presented as a series of questions about the application of the predicate 'is bald', was originally known as the *falakros* puzzle. It was seen to have the same form as the Heap and all such puzzles became collectively known as sorites puzzles.

It is not known whether Eubulides actually invented the sorites puzzles. Some scholars have attempted to trace its origins back to Zeno of Elea however the evidence seems to point to Eubulides as the first to employ the sorites. Nor is it known just what motives Eubulides may have had for presenting this puzzle. It was, however, employed by later Greek philosophers to attack various positions. Most notably by the Sceptics against the Stoics' claims to knowledge.

These puzzles of antiquity are now more usually described as paradoxes. Though the conundrum can be presented informally as a series of questions whose puzzling nature gives it dialectical force it can be, and was, presented as a formal argument having logical structure. The following argument form of the sorites was common:

1 grain of wheat does not make a heap.
If 1 grain of wheat does not make a heap then 2 grains of wheat do not.
If 2 grains of wheat do not make a heap then 3 grains do not.
...
If 9,999 grains of wheat do not make a heap then 10,000 do not.

10,000 grains of wheat do not make a heap.

The argument certainly seems to be valid, employing only *modus ponens* and *cut* (enabling the chaining together of each sub-argument which results from a single application of *modus ponens*). These rules of inference are endorsed by both Stoic logic and modern classical logic.

Moreover its premises appear true. Some Stoic presentations of the argument recast it in a form which replaced all the conditionals, 'If A then B', with 'Not(A and not-B)' to stress that the conditional should not be thought of as being a strong one, but rather the weak Philonian conditional (the modern material conditional) according to which 'If A then B' was equivalent to 'Not(A and not-B)'. Such emphasis was deemed necessary since there was a great deal of debate in Stoic logic regarding the correct analysis for the conditional. In thus judging that a connective as weak as the Philonian conditional underpinned this form of the paradox they were forestalling resolutions of the paradox that denied the truth of the conditionals based on a strong reading of them. This interpretation then presents the argument in its strongest form since the validity of *modus ponens* seems assured whilst the premises are construed so weakly as to be difficult to deny. The difference of one grain would seem to be too small to make any difference in the application of the predicate; it is a difference so negligible as to make no apparent difference in the truth values of the respective antecedents and consequents.

Yet the conclusion seems false. Thus paradox confronted the Stoics just as it does the modern classical logician. Nor are such paradoxes isolated conundrums. Innumerable sorites paradoxes can be expressed in this way. For example, one can present the puzzle of the Bald Man in this manner. Since a man with one hair on his head is bald and if a man with one is then a man with two is, so a man with two hairs on his head is bald. Again, if a man with two is then a man with three is, so a man with three hairs on his head is bald, and so on. So a man with ten thousand hairs on his head is bald yet we rightly feel that such men are hirsute, i.e. not bald. Indeed, it seems that almost any vague predicate admits of such a sorites paradox and vague predicates are ubiquitous.

As presented, the paradox of the Heap and the Bald Man proceed by addition (of grains of wheat and hairs on the head respectively). Alternatively though, one might proceed in reverse, by subtraction. If one is prepared to admit that ten thousand grains of sand do make a heap then would can argue that one grain of sand does since the removal of any one grain of sand cannot make the difference. Similarly, if one is prepared to admit a man with ten thousand hairs on his head is not bald, then one could prove that even with one hair on his head he is not bald since the removal of any one hair from the originally hirsute scalp cannot make the relevant difference. It was thus recognised even in antiquity that sorites arguments come in pairs: non-heap and heap; bald and hirsute; poor and rich; few and many; small and large; and so on. For every argument which proceeds by addition there is another reverse argument which proceeds by subtraction.

The paradox attracted little subsequent interest until the late nineteenth century when formal logic once again assumed a central role in philosophy. With the demise of ideal language doctrines in the latter half of the twentieth century interest in the vagaries of natural language and the sorites paradox in particular has greatly increased.

2. Its Paradoxical Forms

The common form of the sorites paradox presented for discussion in the literature is the form discussed above. Let ' F ' represent the soritical predicate (e.g., 'is bald', or 'does not make a heap') and let the expression ' a_i ' (where i is a natural number) represent a subject expression in the series with regard to which ' F ' is soritical (e.g., 'a man with i hair(s) on his head' or ' i grain(s) of wheat', depending on F). Then the sorites proceeds by way of a series of conditionals and can be schematically represented as follows:

Conditional Sorites

$$\begin{array}{l}
 Fa_1 \\
 \text{If } Fa_1 \text{ then } Fa_2 \\
 \text{If } Fa_2 \text{ then } Fa_3 \\
 \dots \\
 \text{If } Fa_{i-1} \text{ then } Fa_i \\
 \hline
 Fa_i \quad \quad (\text{where } i \text{ can be arbitrarily large})
 \end{array}$$

Whether the argument is taken to proceed by addition or subtraction will depend on how one views the series.

Barnes (1982) states conditions under which any argument of this form is soritical. Initially, the series $\langle a_1, \dots, a_i \rangle$ must be ordered; for example, scalps ordered according to number of hairs, heaps ordered according to number of grains of wheat, and so on. Secondly, the predicate ' F ' must satisfy the

following three constraints: (i) it must appear true of a_1 , the first item in the series; (ii) it must appear false of a_i , the last item in the series; and (iii) each adjacent pair in the series, a_n and a_{n+1} , must be sufficiently similar as to appear indiscriminable in respect of ' F ' -- that is, both a_n and a_{n+1} appear to satisfy ' F ' or neither do. Under these conditions ' F ' will be soritical relative to the series $\langle a_1, \dots, a_i \rangle$ and any argument of the above form using ' F ' and $\langle a_1, \dots, a_i \rangle$ will be soritical.

In recent times the explanation of the fact that sorites arguments come in pairs has shifted from consideration of the sorites series itself to the predicate involved. It is now common to focus on the presence or absence of a negation in the predicate, noting the existence of both a positive form which bloats the predicate's extension and negative form which shrinks the predicate's extension. With the foregoing analysis of the conditions for sorites susceptibility it is easy to verify that ' F ' will be soritical relative to $\langle a_1, \dots, a_i \rangle$ if and only if ' $\text{not-}F$ ' is soritical relative to $\langle a_i, \dots, a_1 \rangle$. Thus verifying that for every positive sorites there is an analogous negative variant.

The key feature of soritical predicates which drives the paradoxes, constraint (iii), is described in Wright (1975) as "tolerance" and is thought to arise as a result of the vagueness of the predicate involved. Predicates such as 'is a heap' or 'is bald' appear tolerant of sufficiently small changes in the relevant respects -- namely number of grains or number of hairs. The degree of change between adjacent members of that series relative to which ' F ' is soritical would seem too small to make any difference to the application of the predicate ' F '. Yet large changes in the relevant respects will make a difference, even though large changes are the accumulation of small ones which don't seem to make a difference. This is the very heart of the conundrum which has delighted and perplexed so many for so long.

Any resolution of the paradoxes is further complicated by the fact that they can be presented in a variety of forms and the problem they present can only be considered solved when all forms have been defused.

One variant replaces the set of conditional premises with a universally quantified premises. Let ' n ' be a variable ranging over the natural numbers and let ' $\text{forall } n(\dots n \dots)$ ' assert that every number n satisfies the condition $\dots n \dots$. Further, let us represent the claim of the form "for all n , if Fa_n then Fa_{n+1} " as follows:

$$\text{forall } n(Fa_n \rightarrow Fa_{n+1})$$

Then the sorites is now seen as proceeding by the inference pattern known as mathematical induction:

Mathematical Induction Sorites

$$Fa_1$$

$$\text{forall } n(Fa_n \rightarrow Fa_{n+1})$$

$$\text{forall } nFa_n$$

So, for example, it is argued that since a man with 1 hair on his head is bald and since the addition of one hair cannot make the difference between being bald and not bald (for any number n , if a man with n hairs is bald then so is a man with $n+1$ hairs), then no matter what number n you choose, a man with n hairs on his head is bald.

Yet another form is a variant of this inductive form. Assume that it is not the case that for every n , a man with n hairs on his head is not bald, i.e., that for some number n , it is not the case that a man with n hairs on his head is bald. Then by the least number principle (equivalent to the principle of mathematical induction) there must be a least such number, say $i+1$, such that it is not the case that a man with $i+1$ hairs on his head is bald. Since a man with 1 hair on his head is bald it follows that $i+1$ must be greater than 1. So, there must be some number $n (= i)$ such that a man with n hairs counts as bald whilst a man with $n+1$ does not. Thus it is argued that though a_1 is bald, not every number n is such that a_n is bald, so there must be some point at which baldness ceases. Let 'exists $n(\dots n \dots)$ ' assert that some number n satisfies the condition $\dots n \dots$. Then we can represent the chain of reasoning just described as follows:

Line-drawing Sorites

Fa_1

$\sim \text{forall } n Fa_n$

exists $n \geq 1 (Fa_n \ \& \ \sim Fa_{n+1})$

Now obviously, given that sorites arguments have been presented in these three forms, "the sorites paradox" will not be solved by merely claiming, say, mathematical induction to be invalid for soritical predicates. All forms need to be addressed one way or another. One would hope to solve it, if at all, by revealing some general underlying fault common to all forms of the paradox. No such general solution could depend on the diagnosis of a fault peculiar to any one form. On the other hand, were no general solution available then "the sorites paradox" will only be solved when each of its forms separately have been rendered toothless. This piecemeal approach holds little attraction though. It is less economical than a unified approach, arguably less elegant, and would fail to come to grips with the underlying unifying phenomenon which is considered to give rise to the paradoxes, namely vagueness. A logic of vagueness, be it classical or otherwise, ought to be able to defuse all those paradoxes that have their source in this phenomenon.

3. Responses

- Ideal Language Approaches
- The Epistemic Response
- Supervaluationism
- Many-Valued Logic
- Embracing the Paradox

The various responses to soritical reasoning can be most easily catalogued by focussing on that form most commonly discussed in the literature -- the conditional form. As with any paradox, four responses appear to be available. One might:

(1) deny that logic applies to soritical expressions.

According to this response the problem cannot legitimately be set up in the first place. On the other hand one might accept that the sorites paradox constitutes a legitimate argument to which logic applies and deny its soundness by:

(2) denying some premise(s),

or

(3) denying its validity.

Finally, seemingly as a last resort, one might embrace the paradox and

(4) accept it as sound.

3.1 Ideal Language Approaches

Committed as Frege and Russell were to ideal language doctrines, it is not surprising to find them pursuing response (1). A key attribute of the ideal language is its precision; the vagueness of natural language is a defect to be eliminated. Since soritical terms are vague, the elimination of vagueness will entail the elimination of soritical terms. They cannot then, as some theorists propose, be marshalled as a challenge to classical logic.

A modern variation on this response, promoted most notably by Quine, sees vagueness as an eliminable feature of natural language. The class of vague terms, including soritical predicates, can as a matter of fact be dispensed with. There is, perhaps, some cost to ordinary ways of talking, but a cost that is nonetheless worth paying for the simplicity it affords -- namely, our thereby being able to defend classical logic with what Quine describes as its "sweet simplicity".

However, with the demise of ideal language doctrines and subsequent restoration of respect for ordinary language, vagueness is increasingly considered less superficial than response (1) suggests. If logic is to have teeth it must be applicable to natural language as it stands. Soritical expressions are unavoidable and the paradox must be squarely faced.

Responses of type (2) do just this. Logic is seen as applicable to natural language, in particular the sorites paradox, but the conditional form of the argument is seen as proceeding from a faulty premise.

3.2 The Epistemic Theory

According to Williamson (1994) Stoic logicians pursued just such a route. Given their acceptance of the principle of bivalence and their presentation of the argument as invoking a material conditional, they blocked the sorites by claiming some one conditional to be false (since not true) and that there comes a point in any sorites series where the relevant predicate ceases to apply and its negation does. For example vague terms like 'heap' or 'knowledge', though soritical relative to an appropriately chosen series, are semantically determinate so, in spite of appearances to the contrary, there is a sharp cut-off point to their application. The inclination to validate all the premises of a sorites argument (along with the inference pattern employed, which the Stoics accepted) was to be explained via ignorance -- more exactly, the unknowable nature of the relevant sharp semantic boundary.

In this way the threat of wholesale scepticism urged by the Sceptics was met by the limited scepticism arising from our inability to know the precise boundaries to knowledge. "Nothing can be known" was rejected in favour of "The precise boundaries to knowledge itself cannot be known". This epistemological response has been elaborated on in Sorensen (1988) and Williamson (1994). Though soritical predicates are admittedly indeterminate in their extension the indeterminacy is not semantic. The conundrum presented by the sorites paradox is an epistemological one which in no way challenges classical semantics or logic.

Until recently such a solution was ruled out by definition. Vagueness was characterised as a semantic phenomenon whereby the apparent semantic indeterminacy surrounding a soritical term's extension was considered real. In the absence of any apparent barrier to knowledge of a soritical predicate's precise extension it was assumed that there was simply no precise extension to be known. The philosophical landscape has now changed. Williamson (1994) contains an impressive array of arguments defending an epistemological account of vagueness which, if successful, would make possible an epistemological solution to the sorites.

The key concern with the epistemological approach however is its counter intuitive nature. Even if such an analysis is possible, the indeterminacy surrounding the application of soritical terms is generally considered to be a semantic phenomenon. Once seen in this way, classical semantics appears in need of revision, and with it classical logic. In the second half of this century there have been a number of attempts to develop non-classical logics of vagueness, a major constraint being the provision of a solution to the sorites paradox. The extent of the proposed logical innovation varies.

3.3 Supervaluationism

In accord with a principle of least mutilation, Dummett (1975) and Fine (1975) adapt Van Fraassen's supervaluation semantics to the sorites paradox, and vagueness more generally, resulting in a non-bivalent logic which, initially at least, retains the classical consequence relation and classical laws whilst admitting truth value gaps. The challenge posed by the sorites paradox can, on this view, to be met by logical revision in the metatheory alone.

According to the semantic conception of vagueness, a vague predicate is characterised by the existence of border cases; that is, cases to which the predicate neither definitely applies nor definitely doesn't apply. If we define the positive extension as those objects to which the predicate definitely applies, the negative extension as those to which the predicate definitely does not apply, and any remaining border cases as the penumbra, then vague predicates are characterised by their having a penumbra.

Given a vague predicate, for example 'heap', we can then stipulate a sharpening thereof, 'heap*' which resolves any border cases by placing them either in the positive or negative extension of 'heap*'. Intuitively, for a sharpening to be admissible as a sharpening of the original vague predicate it ought to also be constrained by its only resolving vague semantic aspects of the predicate's meaning. For example if a pile of i grains of wheat definitely counts as a heap then it ought to be definitely counted a heap*; positive cases ought remain unchanged, as should negative cases. Additionally it should not draw a line between positive and negative cases in such a way as to alter definite ordering relations. Since it is definitely the case that $i+1$ grains counts as a heap if i does, $i+1$ will count as a heap* if i does. An admissible sharpening of any soritical predicate will simply extend the positive and negative extensions to form a sharp boundary somewhere in the predicate's penumbra.

It is easy to see that vague predicates in general have no unique admissible sharpening; there will be a number of possibilities depending on where the cut-off point between the positive and negative extension of the now precise predicate is drawn. Nonetheless, to predicate 'heap' of something in the positive extension will be true for all admissible sharpenings of the predicate and to predicate 'heap' of something in the negative extension will be false for all admissible sharpenings. Defining "truth" *simpliciter* (or, as it is sometimes called, "supertruth") as "truth on all admissible sharpenings", the former predications will be, quite simply, true and the latter false. To predicate 'heap' of a border case will be true on some admissible sharpenings and false on others and so is considered neither true nor false but indeterminate in truth value.

Validity can now be defined so that an argument is supervalueationally valid just in case every model in which the premises are true is one in which the conclusion is also true. Validity thus defined coincides with classical validity, reflecting the fact that the basic underlying notion of ‘truth on an admissible sharpening’ is a classical two-valued valuation function. In particular, treating laws as zero-premise arguments, such supervaluationism preserves all classical laws. In spite of its being non-bivalent then it validates the law of excluded middle. For example, it is true in every model that i grains of wheat either does or does not make a heap since in any model the corresponding disjunction with ‘heap*’ in place of ‘heap’ will be true regardless of which admissible sharpening ‘heap*’ is considered.

Supervaluation semantics then is no longer truth-functional. Consider disjunction for example. Some disjunctions with indeterminate disjuncts will count as indeterminate, for example ‘(Border case) a_i makes a heap or it makes a heap’; whereas some will count as true, for example ‘(Border case) a_i makes a heap or it does not make a heap’. Moreover the semantics countenances instances of true disjunctions neither of whose disjuncts is true. Conjunction and the conditional exhibit analogous non-classical features.

Since all the forms taken by the sorites are classically valid, they are also supervalueationally valid. The conclusion of the conditional form is resisted by noticing that some conditional premise fails to be true. For example, the premise ‘If (border case) a_i makes a heap then (border case) a_{i+1} makes a heap’ will admit of a sharpening which draws the line between a_i and a_{i+1} thus making the conditional false on that sharpening, and one which does not so draw the line, making the conditional true on that sharpening. It is therefore neither true nor false but indeterminate. The conditional sorites is valid but unsound.

More revealing is the diagnosis with regard to the mathematical induction form. It is also deemed unsound due to the failure of one of the premises -- the universal premise. The universally quantified conditional is not true; in fact it’s false. While there is no one conditional premise of the conditional form which is false, it is nonetheless true according to supervaluation theory that some conditional is. That is to say, it is false that for all n , if Fa_n then Fa_{n+1} (where ‘ F ’ is soritical relative to the subjects of the form a_i). For any sharpening of the vague predicate ‘ F ’ there will always be some a_i which counts as the cut-off point relative to that sharpening and thus falsifies the relevant conditional ‘if Fa_i then Fa_{i+1} ’. Every sharpening produces a cut-off point even though no single cut-off point is produced by every sharpening. Hence “For all n , if Fa_n then Fa_{n+1} ” is false on every sharpening and so false *simpliciter*.

Given that supervaluation semantics admits that the falsity of “For all n , if Fa_n then Fa_{n+1} ” is logically equivalent to the truth of “For some n , Fa_n & $\sim Fa_{n+1}$ ”, the line-drawing form of the sorites is also solved. The argument is supervalueationally valid since classically valid and its premises are uncontestably true. What supervaluation semantics provides is an account of how it is that such a conclusion could be true; it is true since true no matter how one admissibly sharpens the soritical predicate involved.

In this way then the sorites paradoxes are said to be defused. Classical logic is no longer appropriate to reasoning in vague contexts and supervaluation semantics is proposed in its place. One immediate concern facing this solution however is the fact that it ultimately treats the mathematical induction and line-drawing forms of the sorites in just the same way as the logically conservative epistemic theory. We are forced to accept the avowedly counter intuitive truth of “For some n , Fa_n & $\sim Fa_{n+1}$ ” which seems to postulate the existence of a sharp boundary yet the existence of just such a boundary is what the semantic theory of vagueness is supposed to deny.

Supervaluationists respond by denying that the conclusion of the line-drawing sorites expresses the existence of a sharp boundary. Though committed to the claim

(a) $\text{T 'exists } n (Fa_n \ \& \ \sim Fa_{n+1})'$,

semantic precision is only properly captured by the claim that

(b) $\text{exists } n \text{ T '}(Fa_n \ \& \ \sim Fa_{n+1})'$

and this is clearly denied by supervaluation theory. Whilst it is true that there is some cut-off point, there is no particular point of which it is true that it is the cut-off point. Since it is only this latter claim which is taken to commit one to the existence of a sharp boundary there is no commitment to there being such a boundary of which we are ignorant (contra the epistemic theorist).

With this explanation however, doubts arise as to the adequacy of the logic. Not only must (b) be properly taken to represent the semantic precision of '*F*' but we must also be prepared to admit that some existential statements can be true without having any true instance, thus blocking any inference from (a) to (b). Just as the failure of the metatheoretic principle of bivalence in conjunction with the retention of the law of excluded middle leads to the presence of true disjunctions lacking true disjuncts, so too must we countenance analogous non-standard behaviour in the logic's quantification theory. In effect, the counter intuitive aspects of the epistemic theory are avoided only at a cost to other intuitions.

At this point the supervaluationist might seek to explain the non-standard behaviour of the quantifiers, and for that matter the non-truth-functional two-place connectives, by showing how such behaviour follows from a proper understanding of the underlying phenomenon of vagueness. More exactly, the suggestion is that a view of vagueness as merely semantic and in no way a reflection of any underlying phenomenon of ontological vagueness might underpin a supervaluationist approach to vagueness. Fine (1975) promotes this representational view of vagueness when defending the law of excluded middle for example. The suggestion would not only prescribe the counter intuitive aspects of supervaluation semantics but would also provide a principled justification of the common *de facto* linkage of supervaluation theory and a representational view of vagueness.

If this explanation is to be pursued then the formal machinery of supervaluation semantics dissolves the paradox only in conjunction with the metaphysical assumption of the impossibility of ontological vagueness. The metaphysical debate is ongoing.

The supervaluation approach has also come under fire for its semantic ascent into the metatheory when defusing the sorites. The problem with accepting the major premise of the mathematical induction sorites as false is simply that it runs counter to our conviction that a grain of wheat can make the difference between a heap and a non-heap; but this conviction can be expressed in the object-language, so why should the elaborate metalinguistic theory be relevant here?

As it happens, such assent is not essential to the account. The object language can be extended to include an operator 'It is determinate (or definite) that ...' ('Det') appropriate to the expression of vagueness in the object language. The vagueness of expressions like 'heap' is characterised by their possessing border cases; this can now be expressed as the existence of cases which are neither determinately heaps nor determinately non-heaps.

By means of the extended language, (a) and (b) are now recast as claims within the object language:

(a') Det exists n ($Fa_n \ \& \ \sim Fa_{n+1}$)

(b') exists n Det ($Fa_n \ \& \ \sim Fa_{n+1}$).

The first is again affirmed and the latter denied. Any inference from (a') to (b') is analogous to the modal inference from 'Nec exists x Fx ' to 'exists x Nec Fx ' and is seen as fallacious, just as the modal inference is commonly said to be. Vagueness, like modality, is viewed as *de dicto*, and, as in many modal logics, the resulting quantification theory reflects important scope distinctions.

Williamson (1994) points to two further problems which now beset the account. If the definition of validity as necessary truth-preservation is retained then classical inferences like *conditional proof*, *dilemma* and *reductio ad absurdum* are no longer valid. Moreover problems arise with regard to the phenomenon of higher order vagueness.

3.4 Many-Valued Logic

As alternatives to the non-truth-functional supervaluation semantics, non-classical logics have been proposed, and in particular, 'many-valued logics'. Again vagueness is seen as grounds for rejecting the principle of bivalence, however truth-functionality is preserved. The approaches vary as regards the number of non-classical truth values deemed appropriate to model vagueness and defuse the sorites paradox.

An initial proposal, first developed in Halldén (1949) and Körner (1960) and recently revamped in Tye (1994), uses a three-valued logic. The motivation for such a logic is similar to the supervaluationist's. Just as a vague predicate divides objects into the positive extension, negative extension and the penumbra, vague sentences can be divided into the true, the false and the indeterminate. The truth-set is then $\{1$ (true), $1/2$ (indeterminate), 0 (false) $\}$. Unlike supervaluation semantics, however, the connectives can then be defined by means of truth tables. There are a range of proposals with Kleene's strong three-valued tables as the preferred choice. (See Haack (1974), Appendix.)

The particular response to the sorites paradox then depends on the definition of validity adopted. A common generalisation of the concept of validity to many-valued logic involves the designation of certain values. A sentence holds (or is assertible) in a many-valued interpretation just if it takes a designated value. Validity is then defined as the necessary preservation of designated value. (In classical logic, of course, only truth is designated and thus the generalised concept reduces to the classical concept of necessary truth preservation.) There are then two non-trivial choices: let the set of designated values be $\{1\}$ or $\{1, 1/2\}$. The former proposal results in a type (2) response, the latter a type (3) response.

If validity is defined as necessary preservation of truth-only then *modus ponens* and *cut* are valid and so, subsequently, is the conditional sorites. Yet not all its premises hold since they are not all true. Some conditional premise is considered to have a true antecedent and indeterminate consequent and so counts as indeterminate. Like supervaluationism such a logic may be said to be paracomplete, admitting non-trivial incomplete theories. (That is, a vague sentence ' A ' and its negation ' $\sim A$ ' can both fail to be designated in an interpretation which nonetheless designates some sentences.)

If, on the other hand, validity is defined so that the conclusion of an argument is either true or indeterminate whenever the premises are either true or indeterminate then a paraconsistent logic results, admitting non-trivial inconsistent theories. (A vague sentence ' A ' and its negation ' $\sim A$ ' might both be designated in an interpretation which nonetheless fails to designate every sentence.) Such an

approach solves the sorites paradox by declaring it invalid. The premises hold since they are at worst indeterminate in truth-value, yet *modus ponens* is no longer valid. In an interpretation where ' Fa_i ' is indeterminate and ' Fa_{i+1} ' is false both ' Fa_i ' and ' $\text{If } Fa_i \text{ then } Fa_{i+1}$ ' are designated yet ' Fa_{i+1} ' is not.

A general concern with three-valued approaches is that their tripartite division of sentences faces similar objections to those which led to the abandonment of the bipartite division effected by two-valued classical logic. Due to the phenomenon of higher order vagueness (in particular second order vagueness) there would seem to be no more grounds for supposing there to exist a sharp boundary between the true sentences and indeterminate ones or the indeterminate sentences and false sentences than there was for supposing a sharp boundary to exist between the true sentences and the false ones. The phenomenon of vagueness which drives the sorites paradox no more suggests two sharp boundaries than it did one. Vague concepts appear to be concepts without boundaries at all. No finite number of divisions seems adequate. Tye (1994) seeks to avoid such difficulties by employing a vague metalanguage.

Goguen (1969) and Zadeh (1975) suggest replacing classical two-valued logic with an infinite-valued one. Infinite-valued or fuzzy logics have also been promoted for their recognition of degrees of truth. Just as baldness comes in degrees so too it is argued does the truth of sentences predicating baldness of things. The fact that John is more bald than Jo is reflected in the sentence 'John is bald' having a higher degree of truth than 'Jo is bald'. With this logical innovation infinite-valued logics are then offered as a means to solve the sorites paradox.

The classical two-valued truth set $\{0, 1\}$ is replaced by set of real numbers in the interval $[0, 1]$. Sentences which are neither definitely true nor definitely false take values other than 0 or 1, but some number in between. As with all many-valued logics, the connectives can be defined in a number of ways, giving rise to a number of distinct logics. A standard proposal proceeds by way of the continuum-valued truth-functional semantics devised by Łukasiewicz. (See Haack (1974), Appendix.) As with the three-valued case, the type of response offered to the paradox depends on the definition of validity. If only the maximum value 1 is designated then the conditional form of the argument is valid, however the conditional premises are not completely true (that is, they do not take the value 1) and do not hold. The diagnosis of the paradox is that we mistake nearly true sentences for completely true ones and the error compounds each time a new conditional premise is invoked in the chain of reasoning. Beginning with a completely true categorical premise we thus proceed to heap neglected difference on neglected difference until finally complete falsity is reached. This type (2) response contrasts with another approach sometimes advocated.

In the foregoing account the laws of excluded middle and non-contradiction fail. Consider some sentence ' A ' with truth-value $1/2$. By the definition of negation ' $\sim A$ ' has value $1/2$ and so their disjunction also has value $1/2$, which is not designated. Thus the law of excluded middle fails. Similarly for the law of non-contradiction. To reinstate these classical laws one can consider all values in the interval $[1, 1/2]$ as designated. In this case all premises of the paradox hold yet *modus ponens* fails and a type (3) response results. In an interpretation where ' Fa_i ' takes the value $1/2$ and ' Fa_{i+1} ' takes the value 0 both ' Fa_i ' and ' $\text{If } Fa_i \text{ then } Fa_{i+1}$ ' take the value $1/2$ and so are designated yet ' Fa_{i+1} ' is not. Such a proposal is paraconsistent, admitting contradictions as sometimes taking a designated value; namely when a sentence and its negation both take the value $1/2$. Other type (3) responses can be developed by taking the set of designated values to be the interval $[1, n]$ where $1/2 < n < 1$.

There are, however, a number of problems which beset any infinite-valued approach to vagueness. Firstly, the very idea of a degree of truth needs explanation. Secondly, if numerical truth-values are to be used some justification seems required for the particular truth-value assignments. Thirdly, the full implications of abandoning the well-understood classical theory in favour of degree theory need spelling out before a proper evaluation of its worth can be made. (On these points see Sainsbury (1995) Ch 2, sec. 6.) Furthermore, it is far from clear whether such an approach successfully avoids problems of higher order vagueness. And the assumption of a totally ordered truth-set is overly simple. Not all natural language sentences are comparable as regards their truth. Due to the multi-dimensional nature of a concept such as redness we may be unable to say of two reddish patches which differ in hue or brightness or colour-saturation whether one is redder than the other. (On these latter points see Williamson (1994) Ch 4, sec. 12-13.)

3.5 Embracing the Paradox

A final option is to simply embrace the paradox. (See Dummett (1975), Wright (1975).) Conditional sorites paradoxes are, contrary to appearances, sound. For example, no amount of grains of wheat makes a heap. This initial claim in favour of a universal type (4) response immediately runs into difficulty however with the realisation that, as noted above, such paradoxes come in pairs. There are negative and positive versions depending on whether the soritical predicate is negated or not. To accept all sorites as sound requires assent to the additional claim that, since one grain of wheat makes a heap, any number do. A radical incoherence follows since there is a commitment to all and any number both making a heap and not making a heap. Similarly, everyone is bald and no-one is; everyone is rich and no-one is, and so on.

The problem is that the soundness of any positive conditional sorites undercuts the truth of the unconditional premise of the corresponding negative version, and vice versa. Unless one is prepared to countenance the almost total pervasiveness of contradictions in natural language, it seems that not all sorites can be sound. Unger (1979) and Wheeler (1979) propose a more restricted embrace. Following dissatisfaction with responses of type (1) and (3) one accepts the applicability and validity of classical norms of reasoning. Nonetheless, dissatisfaction with responses of type (2) considered so far -- rejecting some conditional premise -- leaves open the possibility of either rejecting the unconditional premise or accepting it and, with it, the soundness of the paradox. What is advocated is the soundness of those sorites which deny heapness, baldness, hirsuteness, richness, poverty, etc. of everything -- a type (4) response -- and the corresponding falsity of the unconditional premise of all respective positive variants of the argument -- a type (2) response. Terms like 'heap', 'bald', 'hirsute', 'rich' and 'poor' apply to nothing. It is admitted that they apply to everything if they apply to anything, but the all-or-nothing choice is resolved in favour of the latter option. (See Williamson (1994) Ch 6.)

Bibliography

General

- Sainsbury, M. 1995 (2nd edn). *Paradoxes*. Cambridge: Cambridge University Press. Chapter 2.
- Williamson, T. 1994. *Vagueness*. London: Routledge. Includes detailed exposition and criticism of all the major approaches to the paradox.

Sorites in History

- Barnes, J. 1982. 'Medicine, experience and logic', in J. Barnes, J. Brunschwig, M.F. Burnyeat and M. Schofield (eds), *Science and Speculation*. Cambridge: Cambridge University Press.
- Burnyeat, M.F. 1982. 'Gods and heaps', in M. Schofield and M.C. Nussbaum (eds), *Language and Logos*. Cambridge: Cambridge University Press.
- Williamson, T. 1994. *Vagueness*. London: Routledge. Chapter 1.

Its Paradoxical Forms

- Sainsbury, M. and Williamson, T. 1995. 'Sorites' in B. Hale and C. Wright (eds), *Blackwell Companion to the Philosophy of Language*. Oxford: Blackwell.

Responses

- Quine, W.V. 1981. 'What price bivalence?', *Journal of Philosophy* 77: 90-5.
- Russell, B. 1923. 'Vagueness', *The Australian Journal of Philosophy and Psychology* 1: 84-92.

The Epistemic Response

- Sorensen, R. 1988. *Blindspots*. Oxford: Clarendon Press.
- Williamson, T. 1994. *Vagueness*. London: Routledge. Chapters 7-8.

Supervaluationism

- Dummett, M. 1975. 'Wang's paradox', *Synthese* 30: 301-24; reprinted in his *Truth and Other Enigmas*.
- Fine, K. 1975. 'Vagueness, truth and logic', *Synthese* 30: 265-300.

Many-Valued Approaches

- Goguen, J.A. 1969. 'The logic of inexact concepts', *Synthese* 19: 325-373.
- Haack, S. 1974. *Deviant Logic*. Cambridge: Cambridge University Press.
- Halldén, S. 1949. *The Logic of Nonsense*. Uppsala: Uppsala Universitets Arsskrift.
- Körner, S. 1960. *The Philosophy of Mathematics*. London: Hutchinson.
- Tye, M. 1994. 'Sorites paradoxes and the semantics of vagueness', in J. Tomberlin (ed.), *Philosophical Perspectives: Logic and Language*. Atascadero, California: Ridgeview.
- Zadeh, L. 1975. 'Fuzzy logic and approximate reasoning', *Synthese* 30: 407-428.

Embracing the Paradox

- Dummett, M. 1975. 'Wang's paradox', *Synthese* 30: 301-24; reprinted in his *Truth and Other Enigmas*.
- Unger, P. 1979. 'There are no ordinary things', *Synthese* 41: 117-54.
- Wheeler, S.C. 1979. 'On that which is not', *Synthese* 41: 155-94.
- Wright, C. 1975. 'On the coherence of vague predicates', *Synthese* 30: 325-65.

Other Internet Resources

- Bibliography of literature on vagueness and the Sorites Paradox, maintained by Justin Needle (Ph.D./Philosophy, U. Dundee, 1995)

Related Entries

logic: non-classical | logic: paraconsistent | vagueness

Copyright © 1997, 2002

Dominic Hyde

D.Hyde@mailbox.uq.edu.au

A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z

Stanford Encyclopedia of Philosophy

FITTING WORDS: VAGUE LANGUAGE IN CONTEXT

1. INTRODUCTION

Consider the possibilities for reference using vague definite descriptions. There are two pigs in a pen: the first a runt, the second quite round but not really fat. Now the farmer says, “The fat pig won a prize”. His neighbor could protest that neither pig is really fat. But he does not protest. He understands that the rounder pig is meant, and both of them know it.

We can chalk this up to negotiation about how to use vague words. The farmer, by using the definite description, implicitly proposes to stretch the sense of ‘fat’ to cover one of the two pigs. This looser sense is needed in order to satisfy the presupposition that there is a single fat pig. Now, given the pigs’ sizes, there is really just one way to stretch. The rounder of the two will be the one that is covered, and so the definite description will come to pick out this pig. In this way, using ‘fat’ more loosely can help to distinguish between the pigs. That is why the farmer makes his proposal. When the neighbor does not protest he tacitly accepts the proposal, and lets the farmer know that he has understood.

Of course, such agreements about how to use vague words are bound to the particular contexts in which they are reached. If the rounder pig were instead among some truly fat pigs and the farmer said, “The pig that is not fat won a prize”, then the neighbor would understand that he meant the same pig. In such a context they would settle on a stricter sense of ‘fat’, not a looser one.

This contextuality and the *accommodation* that accompanies it are in our view characteristic of vague language.¹ Just as a home handyman can fit an adjustable wrench to a nut, we think, a speaker can adjust the extension of a vague expression to suit his needs, relying on the hearer to recognize his intentions and to accommodate him.²

¹ We borrow the term “accommodation” from David Lewis (1979).

² The contextuality of vague language has been pointed out by others, e.g., Manfred Pinkal (1983, 1995) and Ruth Manor (1995).



Maybe it is possible to account for examples like these without supposing a change in the senses of vague expressions. Keith Donnellan describes the case of a usurper to the throne: all refer to him somehow insincerely, but understandably enough, as “the king”, although it is common knowledge that he is not a king.³ Perhaps it is the same with the pigs. The farmer *calls* the rounder pig “fat”, though really it is borderline.

No doubt ‘fat pig’ can be used in this somehow insincere way. But the cases we have in mind differ from Donnellan’s example in two important ways. First, there are other individuals who are as entitled to be called “king”, or as lacking in entitlement, as our usurper. But whereas calling the usurper “king” does not bring with it any inclination, or obligation, also to call these other non-kings “king”, things are different with the pigs. Having called one a “fat pig”, it seems odd, even inconsistent, to withhold this description from comparable pigs.

The case of the usurper differs in another way too. In referring to the usurper as “the king”, a speaker has in mind some particular person. The description is used referentially, according to Donnellan. In contrast, in some of our examples the speaker need not have a particular referent in mind. A farmer might say, “The fat pigs are going to market”, even though he has not yet decided which are to be sent off. A hearer who understands this can nonetheless accommodate by agreeing to call pigs as round as the roundest ones “fat pigs”.

It has been suggested that the mechanism of the accommodation is in the compositional analysis of adjective-noun combinations. The import of an adjective such as ‘fat’ depends on a contextually determinate comparison class. That is how it can be that one and the same lady is a fat fashion model, but a skinny belly dancer. We think of accommodation as stretching the senses of vague expressions while keeping comparison classes fixed. But an obvious alternative is to think of it as switching comparison classes. Besides the noun phrases with which adjectives appear, comparison classes can depend on such contextual factors as salient groups of individuals. In our example, the rounder pig is not fat relative to pigs at large, but it is fat relative to the pigs in the pen. Accommodation might be switching from the one comparison class to the other.

We claim that the accommodation we have in mind can be accounted for without bringing in comparison classes. We do not have a conclusive empirical reason to leave them out, but we do have a theoretical reason. We expect that modeling accommodation as comparison-class change will require multiplying comparison classes in an *ad-hoc* manner. For most of this paper we shall hide comparison classes from view, to see how far we

³ (Donnellan, 1966).

can get without them. But we return briefly to this question in Section 6, with an example suggesting that they are needed anyway.

We shall proceed on the assumption that communication proceeds relative to a context, which includes shared presuppositions. Among these are items of general knowledge, along with whatever the situation makes obvious and whatever has been established in the course of the conversation, whether it is true or false. Often, perhaps normally, participants will *believe* these presuppositions, but they need not. Assumptions made for the sake of the argument are presuppositions in the relevant sense. So, perhaps, are propositions that children “make believe” in games. Following Robert Stalnaker, we shall call the body of these presuppositions the *common ground*.⁴ In keeping with a computational view of mind, we suppose that the common ground is represented, in the minds of the participants in a conversation, by sentences of some language. Since these presuppositions are to be shared, though, we think of the common ground metaphorically as a blackboard, visible to all participants, on which the presuppositions are written.

Utterances bring about changes in the common ground. Our purpose here is to propose a model of changes made in order to accommodate vague utterances. Our model is, in the sense of David Marr (1982), “computational”: we shall present a function for updating the common ground, but our object is not in the first place to present an algorithm that a language-understanding device might use. Rather, our object is to present a function we think such an algorithm ought to compute. With this goal in mind, we put to one side questions concerning the computational complexity of the accommodation function.

The model combines ideas from three domains: supervaluational accounts of vagueness, theories of conveyed meanings, and theories of belief revision. Supervaluational accounts of vagueness were suggested by, among others, Kit Fine (1975) and Hans Kamp (1975). From them we borrow an interpretation of vague language. By a theory of conveyed meanings we mean a specification of the presuppositions, conversational implicatures, and other meanings that are conveyed by utterances. Pieces of such a theory can be found in the work of Paul Grice (1967), and in subsequent work by Lauri Karttunen and Stanley Peters (1975, 1979) and Gerald Gazdar (1979a, 1979b), among others.⁵ We shall not present such an account, but our discussion shall take place against the background of some such theory, the existence of which we presuppose. Finally, one well-known account of belief revision is that of Carlos Alchourrón et al. (1985).

⁴ (Stalnaker, 1974).

⁵ For a survey see (Levinson, 1983) and, for more recent work, (Beaver, 1997).

The function that we shall propose for updating common ground derives directly from it.

In the next section, we shall describe a language for representing expressions of English. In Section 3, we shall briefly discuss some assumptions we make about a theory of conveyed meanings. With this background in place, in Section 4, we shall propose a model of accommodation, on which uttering a sentence with a vague description leads to changes in the common ground. We shall illustrate this accommodation in Section 5. Finally, in Section 6, we shall discuss several examples that point toward elaborations of our model.

2. VAGUE LANGUAGE

It will help to have a formal language in which to represent expressions of English. The language we have chosen is very limited in its expressive power; but it has the advantages for our purpose of its simplicity and familiarity: $\mathcal{L}$ is a language of predicate logic augmented with the sentential operator D , with the intended interpretation “It is definitely the case that”. $\mathcal{L}$ has individual constant symbols including *Arnold* and *Babe*, individual variable symbols $x, y, \dots$, and set variable symbols $X, Y, \dots$; $\mathcal{L}$ has monadic predicates including *pig*, *fat-pig*, *skinny-pig*, *winner*, and *curly-tail*; and dyadic predicates including *tail*, *at-least-as-fat-as*, and *fatter*. Some of these expressions – they include *pig*, *winner*, *tail*, *at-least-as-fat-as*, and *fatter* – are designated non-vague. The others – including *fat-pig*, *skinny-pig*, and *curly-tail* – are vague.⁶ The formulas of $\mathcal{L}$ are defined with the usual recursion over the logical symbols.

To illustrate, the intended interpretation of *fat-pig*(*Arnold*) is that Arnold is a fat pig. The intended interpretations of *at-least-as-fat-as*(*Arnold*,*Babe*) and *fatter*(*Arnold*,*Babe*) are that Arnold is at least as fat as Babe, and that Arnold is fatter than Babe. The intended interpretation of $D(\textit{skinny-pig}(\textit{Arnold}))$ is that it is definitely the case that Arnold is a skinny pig, and the intended interpretation of $\neg D(\textit{fat-pig}(\textit{Arnold})) \ \& \neg D(\neg \textit{fat-pig}(\textit{Arnold}))$ is that there is no matter of fact about whether Arnold is a fat pig, which is to say that Arnold is a borderline fat pig.

Notice that we have not included in $\mathcal{L}$ representations of vague adjectives like ‘fat’, which can be used to form vague composite expressions

⁶ We simplify in two ways. First, we consider only the vagueness of predicates, putting to one side questions relating to the vagueness of names and quantifiers. Second, we consider only some predicates to be vague, though just about all of them seem to be, to a greater or lesser extent.

like ‘fat pig’. We will treat such composite expressions as if they were semantically simple. A deeper treatment would interpret adjectives as functions mapping the semantic values of the nouns they combine with onto the semantic values of the resulting adjective-noun combinations. The semantic value of the noun could be thought of as picking out a “comparison class” relative to which the adjective is interpreted. An account of logical relations that hold between sentences in virtue of the adjective noun pairs that compose them would also be a part of this deeper treatment. For now, though, we shall set these matters to one side.

We turn now to the interpretation of $\mathcal{L}$. Imagine a pig that is on the threshold of becoming fat. Imagine also that you are well informed about the size of the pig – say, because you are looking straight at it. Now someone says that the pig is fat and someone else disagrees. In a case like this you may find yourself unable to agree with either of them. One possible explanation for this inability is that although there is a fact of the matter about whether or not the pig is fat, you are insufficiently informed about the extension of the word ‘fat’, and so you cannot tell. Either the pig in question is fat or it is not, but you just do not, and perhaps cannot, know which.⁷ Another explanation for this inability, which coheres well with the approach we take, is that there are some pigs that are neither definitely fat, nor definitely not fat, and that this pig is one of them.

This second explanation suggests a *partial* interpretation for vague predicates. A partial interpretation assigns to each predicate both a positive and a negative extension, which need not exhaust the entire domain of quantification. A *borderline* individual is one that is neither in the positive nor in the negative extension of the predicate in question.

The notion that vagueness is semantic partiality is one of the main ideas underlying supervaluational treatments of vagueness. The other is the idea that vagueness can be reduced by resolving borderline cases one way or the other, while leaving determinate cases be. The different ways of making vague expressions more precise are *precisifications*.

We shall use precisifications to explore the idea that accommodation involves the adoption of new senses of vague expressions. These senses might incorporate formerly borderline individuals either into the positive extensions of the expressions in question, or into their negative extensions. Or, they might exclude individuals from the extensions in question. The newly adopted senses correspond to new precisifications.

Precisifications are constrained by the intuitively understood meanings of expressions. Some of these constraints, following Fine (1975), we shall call *penumbral* constraints. For instance, the intuitive meanings of ‘fat pig’

⁷ Epistemic treatments are recommended by (Sorenson, 1988) and (Williamson, 1994).

and ‘skinny pig’ require that, on any *admissible* precisification, if a pig is at least as fat as a fat pig, then it too is fat, and that no skinny pig is fat. In addition, precisifications satisfy constraints placed on them by the meanings of non-vague expressions. The intuitive meaning of ‘at least as fat as’ requires that, on any precisification, given any two pigs, one is at least as fat as the other. Furthermore, on any precisification, one pig is fatter than another just in case it is at least as fat as the other, but not the other way around.

We evaluate a sentence at a precisification that is *appropriate* insofar as it retains a degree of indeterminacy that reflects the intuitively understood meanings of vague expressions. But the truth value of a sentence at this precisification depends, according to a supervaluational treatment, on the truth values it will obtain on the different ways of making it completely precise, i.e., on its *complete* precisifications. A sentence is counted true if it is true on each of its complete precisifications, and false if it is false on each of them. A sentence that is true on some but false on others has no truth value.

We shall now make these ideas precise. A *model* $\mathcal{M}$ for $\mathcal{L}$ is a quadruple $\langle \mathcal{U}, \mathcal{P}, \leq, \mathcal{I} \rangle$. Here $\mathcal{U}$ is a non-empty set, the domain of quantification; $\mathcal{P}$ is a set of evaluation points; $\leq$ is a partial ordering on $\mathcal{P}$; and $\mathcal{I}$ is an interpretation function. We require that $\leq$ orders $\mathcal{P}$ as a tree, with a unique minimal element $@$, the *appropriate point*, at the base of the trunk. $\mathcal{I}$ assigns to each n -place predicate or relation symbol R and evaluation point p two disjoint sets of n -tuples of elements of $\mathcal{U}$: $\mathcal{I}^+(R, p)$, the *positive extension* of R at p ; and $\mathcal{I}^-(R, p)$, the *negative extension* of R at p . $\mathcal{I}$ must respect $\leq$: for each n -place relation R and all points $p, q \in \mathcal{P}$, if $p \leq q$ then $\mathcal{I}^+(R, p) \subseteq \mathcal{I}^+(R, q)$ and $\mathcal{I}^-(R, p) \subseteq \mathcal{I}^-(R, q)$. A *complete point* is a point $p \in \mathcal{P}$ such that for each n -place relation R , $\mathcal{I}^+(R, p) \cup \mathcal{I}^-(R, p)$ is the set of all n -tuples of elements of $\mathcal{U}$. The appropriate precisification is that which is induced by $\mathcal{I}$ at the appropriate point, and complete precisifications are those induced by $\mathcal{I}$ at complete points. $\mathcal{I}$ assigns to each individual constant of $\mathcal{L}$ some element of $\mathcal{U}$.

Evaluation at a complete point is classical. For any assignment a of the individual variables of $\mathcal{L}$ into the domain of $\mathcal{M}$, and for any individual term t , let $[t]_{\mathcal{M},a}$ be $\mathcal{I}(t)$ if t is a constant, and $a(t)$ if t is a variable. If ϕ is an atomic formula $Rt_1, \dots, t_n$ and p is a complete point, then $[\phi]_{\mathcal{M},p,a} = \text{true}$ if and only if $\langle [t_1]_{\mathcal{M},a}, \dots, [t_n]_{\mathcal{M},a} \rangle \in \mathcal{I}^+(R, p)$. Otherwise, $[\phi]_{\mathcal{M},p,a} = \text{false}$. Formulas composed with the classical connectives and quantifiers obtain truth values in accordance with the classical truth functions.

At any point p that is not complete, the satisfaction value of a formula is supervaluational: $[\phi]_{\mathcal{M},p,a} = \text{true}$ [*false*] if and only if for all complete

points $q \in \mathcal{P}$ such that $p \leq q$, $[\phi]_{M,q,a} = \text{true}$ [*false*]. We say a sentence ϕ is *true in M at p , relative to a* , just in case $[\phi]_{M,p,a} = \text{true}$; ϕ is *false in M at p* just in case $[\phi]_{M,p,a} = \text{false}$; otherwise, ϕ is *undefined*. Finally, truth in a model goes by the appropriate precisification: a sentence ϕ is *true* [*false/undefined*] in $\mathcal{M}$ just in case ϕ is *true* [*false/undefined*] in $\mathcal{M}$ at $@$.

To complete the truth definition, we now consider expressions of the form $D(\phi)$. Informally speaking, it is definitely the case that ϕ if on all ways of making ϕ precise this sentence is true. Accordingly, we let the value of $D(\phi)$ at any point p depend on the value of ϕ at the appropriate point: $[D(\phi)]_{M,p,a} = \text{true}$ if $[\phi]_{M,@,a} = \text{true}$; otherwise, $[D(\phi)]_{M,p,a} = \text{false}$. Writing ' $I(\phi)$ ' as an abbreviation for ' $\neg D(\phi) \& \neg D(\neg\phi)$ ', it follows that $[I(\phi)]_{M,p,a} = \text{true}$ if there are complete points q and r such that $[\phi]_{M,q,a} = \text{true}$ and $[\phi]_{M,r,a} = \text{false}$; otherwise $[I(\phi)]_{M,p,a} = \text{false}$.

Note that the truth of $I(\phi)$ at a complete point p is compatible with the truth of ϕ there. This might seem odd at first, but it is as it should be. Where ϕ does not involve the operator D , the truth of ϕ at a complete point is a local matter, dependent only on the extensions of predicates at that point. But the truth of $I(\phi)$ is a global matter. It means, informally speaking, that ϕ is a sentence whose truth can go either way; on some ways of making its predicates precise it is true, on others it is false.⁸

We will be interested only in *acceptable* models. A model $\mathcal{M}$ is acceptable if for all points p and q : (1) for each non-vague predicate or relation symbol R , $\mathcal{I}(R, p) = \mathcal{I}(R, q)$; and (2) $\mathcal{I}$ induces an admissible precisification at p . The first constraint ensures that the interpretation of non-vague language does not vary from point to point within a model. From the second follows the logical validity, or truth in every model, of such *lexical truths* as:

$$\forall x, y (\text{at-least-as-fat-as}(x, y) \rightarrow (\text{fat-pig}(y) \rightarrow \text{fat-pig}(x))),$$

and

$$\forall x (\text{skinny-pig}(x) \rightarrow \neg \text{fat-pig}(x)).$$

⁸ For more discussion of the interpretation of ' $I(\phi)$ ' see (Fine, 1975), pp. 287–289; (Williamson, 1994), pp. 149–153; and (Morreau, 1999). This truth definition ignores the *higher order* vagueness that exists where borderlines themselves have borderlines. This simplification does not affect any of our main points.

Our models give rise to the following notion of entailment: a set of sentences Γ entails a sentence ϕ , written $\Gamma \models \phi$, if and only if ϕ is true at every point at which each member of Γ is true.⁹

We can now extend the notion of vagueness from simple expressions to sentences. A sentence ϕ is *vague* if and only if for some acceptable model $\mathcal{M}$ and points p and q in $\mathcal{P}_{\mathcal{M}}$, it is not the case that $[\phi]_{\mathcal{M},p} = [\phi]_{\mathcal{M},q}$. Not all sentences with vague symbols are vague; no lexical truth is, for instance. But it can easily be shown that containing vague symbols is a necessary condition for vagueness. It also follows that if we exclude higher order vagueness then the result of placing a vague sentence within the scope of D or I is a non-vague sentence.¹⁰

3. CONVEYED MEANINGS

A speaker often conveys more than he has literally said. We use ‘*conveyed meanings*’ as a general term for the contribution made by literal meanings, presuppositions, and conversational implicatures to what an utterance communicates.

Consider definite descriptions. When someone says a sentence with a definite description, his utterance will in many cases presuppose that there is exactly one (salient) object that fits that description. That is why, if he is not sure whether there is a pig but is sure there is none in the pen, it is misleading for him to say “The pig is not in the pen”. Instead he ought to say “The pig is not in the pen, if there is a pig”. By saying the former sentence, but not by saying the latter, he conveys something he does not believe to be true: that there is a pig.

Our account of accommodation rests on a theory of conveyed meanings, but we shall not assume any particular theory. For some time philosophers of language and linguists have worked to provide such an account, and we shall proceed on the assumption that one can be given. Partial accounts that fill the bill can be found in (Gazdar, 1979a, 1979b), (Karttunen & Peters, 1975, 1979), (Levinson, 1983) and, more recently, (Beaver, 1997). Allowing others to fill in the details, we shall write ‘ $Cnv(s,c)$ ’ to indicate the set of conveyed meanings of an utterance of the English sentence s in a context c ; here, $Cnv(s, c)$ is understood to be a set of sentences of $\mathcal{L}$.

⁹ For a discussion of this notion of entailment see (Williamson, 1994), pp. 147–149. Williamson favors an alternative “global” notion: Γ entails ϕ if ϕ is true in every model in which Γ is true.

¹⁰ For a discussion of higher-order vagueness as it pertains to a supervaluational treatment of vagueness see (Fine, 1975), pp. 287–298, and (Williamson, 1994), pp. 156–161.

When we come to worked examples, in Section 5, we shall make assumptions about the function Cnv , making claims that hold for all accounts of conveyed meanings satisfying these assumptions.

The above example illustrates one of these assumptions, namely, that while an utterance with a definite description normally conveys that there is exactly one thing satisfying the description, sometimes it does not. Saying “The pig is not in the pen” is in many contexts enough to convey that there is just one pig. But things are different if you say instead “The pig is not in the pen, if there is a pig”. One conversational implicature of this utterance is that you believe neither that there is a pig nor that there is not one. In light of this implicature, the meaning that would normally be conveyed in virtue of the definite description is not conveyed.

In the next section we shall extend this assumption to sentences that contain *vague* definite descriptions. We shall assume, for instance, that the farmer’s utterance, “The fat pig won a prize”, conveys that there exists just one fat pig and that this pig won a prize. Adapting slightly the notation of Whitehead and Russell, we write ‘ $Q(\iota x Px)$ ’ to abbreviate:

$$\exists x (\forall y (P(y) \leftrightarrow y = x) \& Q(x)).^{11}$$

Our assumption is that for certain contexts c :

$$\begin{aligned} &winner(\iota x \text{ fat-pig}(x)) \\ &\in Cnv(\text{“The fat pig won a prize”, } c). \end{aligned}$$

Similarly, we shall assume that, in certain contexts, saying “The fat pigs won prizes” conveys that each pig in the most inclusive set of fat pigs won a prize. We write ‘ $Ps(X)$ ’ for:

$$\exists x, y (x \in X \& y \in X \& x \neq y \& \forall z (z \in X \rightarrow P(z))),$$

and we write: ‘ $Qs(\iota_{\max} X Ps(X))$ ’ for:

$$\exists X (Ps(X) \& \forall Y (Ps(Y) \rightarrow Y \subseteq X) \& Qs(X)).$$

Our assumption is that for certain contexts c :

$$\begin{aligned} &winners(\iota_{\max} X \text{ fat-pigs}(X)) \\ &\in Cnv(\text{“The fat pigs won prizes”, } c). \end{aligned}$$

¹¹ See (Whitehead & Russell, 1927), p. 173.

4. A MODEL OF ACCOMMODATION

What is conveyed by an utterance will often become a part of the common ground. Unless there is reason to think a speaker is mistaken (or lying or sarcastic or something of the sort), the conveyed meanings of his utterance will often simply be added to the common ground. For instance, when a speaker says, “The pig is not in the pen”, the common ground will ordinarily grow to include conveyed meanings, such as the presupposition that there is a unique pig. Let $c + s$ be the context brought about by uttering sentence s in a context c and let $g(c)$ be the common ground of c . A simple hypothesis about the effect of an utterance on the common ground is that:

$$g(c + s) = g(c) \cup Cnv(s, c).$$

This simple model accounts for many cases of accommodation. Consider again the case of the farmer who first shows his neighbor the not-quite-fat pig Arnold. Prior to the farmer’s utterance the two share the belief that Arnold is a borderline fat pig. This piece of the common ground is represented by the sentence $I(fat-pig(Arnold))$. Suppose the pig in the pen with Arnold is Babe, and that the common ground includes sentences expressing that Arnold and Babe are the only two pigs, that Arnold is the fatter of the two, and that Babe is a skinny pig. Then $g(c)$ includes the following sentences:

$$\begin{aligned} &I(fat-pig(Arnold)), \\ &\forall x(pig(x) \rightarrow x = Arnold \vee x = Babe), \\ &fatter(Arnold, Babe), \\ &skinny-pig(Babe). \end{aligned}$$

In any satisfactory account of conveyed meaning, the farmer’s utterance, “The fat pig won a prize”, will convey that there is a unique fat pig, rendered $\exists x \forall y (fat-pig(y) \leftrightarrow x = y)$ or, for short, $\exists!x fat-pig(x)$. Now, the simple model has it that $g(c + \text{“The fat pig won a prize”}) = g(c) \cup Cnv(\text{“The fat pig won a prize”}, c)$. Since the presupposition $\exists!x fat-pig(x)$ and $g(c)$ together entail:

$$\begin{aligned} &fat-pig(Arnold) \vee fat-pig(Babe), \\ &skinny-pig(Babe), \\ &I(fat-pig(Arnold)), \end{aligned}$$

the updated common ground would then entail these sentences. In light of the constraint on models that skinny pigs are not fat pigs, it would also

entail *fat-pig(Arnold)*. This is just as it should be. We expect that Arnold is the pig picked out by the definite description. Indeed, the effect of simple growth of the common ground, in this case, would be that Arnold comes to satisfy the description.

In this case it is plausible that the common ground should simply grow. Notice that this is a case where the conveyed meanings of an utterance are consistent with the prior common ground. The sentence *fat-pig(Arnold)* reflects the presupposition introduced by the utterance that Arnold and other pigs his size are for the meantime to be called “*fat pigs*”. The continued presence of *I(fat-pig(Arnold))* in the common ground reflects the participants’ mutual belief that pigs this size *could* nonetheless be denied this description.

In other cases, though, it is not plausible that the common ground should simply grow. We shall consider cases where the meanings conveyed by an utterance are *not* consistent with the prior common ground. Suppose, for example, that earlier accommodation has already brought it about that Arnold is for the meantime to be called “*fat*”, although it is a matter of common knowledge that he is a borderline case. Now the farmer moves Arnold into a pen with two very round pigs, Babe and Cornelius, so that the common ground is:

I(fat-pig(Arnold)),

fat-pig(Arnold),

$\forall x(\text{pig}(x) \rightarrow x = \text{Arnold} \vee x = \text{Babe} \vee x = \text{Cornelius})$,

fatter(Babe,Arnold),

fatter(Cornelius,Arnold).

In the company of these much fatter pigs, the farmer can no longer pick out Arnold by saying “the fat pig”. But, we think, he can pick him out by saying “the pig that isn’t fat”. When in this new context the farmer says “The pig that isn’t fat won a prize”, conveying *winner($\iota x (\text{pig}(x) \& \neg \text{fat-pig}(x))$)*, the accommodating hearer will choose a stricter sense of “fat pig”, so as to incorporate this conveyed meaning into the common ground. Understood in its new sense, the expression will not apply to Arnold, and so the prior presupposition *fat-pig(Arnold)* ought not to be carried over to the updated common ground. The common ground in this instance should not simply grow.

Sometimes, then, presuppositions must be excluded from the prior common ground when accommodating vague utterances. Even so, we seem

to retain as much common ground as we can. To elaborate our example somewhat, accommodating the speaker by denying Arnold the description ‘fat pig’ will not lead the neighbor to call into question unrelated matters, such as the curliness of Cornelius’ tail. The change involved seems to be *minimal* change. This consideration lead us to the following recipe for accommodation:

Take the conveyed meanings of an utterance. Then add as much of the prior common ground as you can without introducing inconsistency.

Notice that the accommodation which concerns us here does not require changes in the common ground because the world represented therein has changed. Changes are not required because a participant in the discourse was mistaken, either; rather, changes are required because vague language has come to be used differently. This suggests that accommodation ought not to lead us to reject any sentence that describes the world in purely non-vague terms. When we make minimal changes to regain consistency, only vague sentences should go. This suggests the following constraint on minimal change: a minimal change to a set of sentences will be one in which no vague sentences are unnecessarily deleted from it.¹² The recipe for accommodation then comes to this:

Take the conveyed meanings of an utterance. Add the non-vague sentences of the common ground. Finally, add as many of the remaining vague sentences as you can without introducing inconsistency.

To put this recipe precisely, the following definition is useful:

For any two sets A and S of sentences, $A \perp S$ is the set $\{T : S \subseteq T \subseteq A \cup S, T \not\models \perp, \text{ and for all } U, \text{ if } T \subseteq U \subseteq A \cup S \text{ and } U \not\models \perp, \text{ then } U = T\}$.¹³

¹² One might include in the common ground not only whatever the situation makes obvious and whatever has already been established or assumed in the conversation, but also things to which the participants are less firmly committed. These might come, say, from earlier conversations, or from more or less unreliable third parties. Accommodation should then sometimes allow such weakly entrenched non-vague sentences to be given up so as to make room for vague sentences. For now, though, we assume full commitment to all of the common ground.

¹³ This notion is familiar from work in belief revision. See (Alchourrón et al., 1985). Its application to the updating of common ground is natural given the similarity of the project: we describe how the addition of presuppositions to a set of previously accepted presuppositions leads to a modification of that set.

Let $g_{\text{vague}}(c)$ be the subset of $g(c)$ that includes just the vague sentences. Then $Cnv(s, c) \cup (g(c) \setminus g_{\text{vague}}(c))$ contains just the conveyed meanings of s and the non-vague part of the common ground. Consider any set g^* in $g_{\text{vague}}(c) \perp (Cnv(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)))$. On the current proposal, g^* has the following significance: it is one possible way of making minimal changes to $g(c)$ so as to accommodate the speaker's utterance of s . Since there can be more than one such g^* , we introduce a selection function Sel which, given as its argument a non-empty set of alternatives, returns a member of this set, and given no alternatives, returns the empty set. Intuitively speaking Sel has the task of choosing, from among a number of alternatives, the best way of updating the common ground. Finally, then, the proposal we shall consider for updating the common ground following an utterance is this:

$$g(c + s) = Sel(g_{\text{vague}}(c) \perp (Cnv(s, c) \cup (g(c) \setminus g_{\text{vague}}(c))))^{14}$$

What makes one way of accommodating better than another? What constraints can be placed on Sel ? We shall put off these questions until Section 6. Meanwhile, we shall illustrate the proposed model of accommodation by returning to the examples that motivated it. Then we shall consider a further example involving plural definite descriptions. In these examples the model will yield only one option for Sel . This leaves Sel with no choice, and us able to proceed without saying more about this function.

5. EXAMPLES

Consider the farmer's utterance, "The fat pig won a prize". Here, the farmer uses a vague definite description to refer to a pig, even though none satisfies the description. This anomalous situation brings about a change in common ground to one in which something does satisfy it.

EXAMPLE 1. There are just two pigs: a borderline fat pig and a runt. The common ground $g(c)$ contains the following sentences:

(1) $I(\text{fat-pig}(\text{Arnold}))$

(2) $D(\text{skinny-pig}(\text{Babe}))$

¹⁴ A reviewer suggests that this definition of revision is unintuitive, being sensitive to "the superficial way in which information is presented". As this reviewer points out, we could move to some form of "theory revision", if there is a problem here. See, for example, Carlos Alchourrón et al. (1985). But we do not think there is a problem. For a discussion of the advantages of a fine-grained notion of "belief bases", see, for example, Sven Ove Hansson (1992).

$$(3) \quad \forall x(\text{pig}(x) \rightarrow x = \text{Arnold} \vee x = \text{Babe})$$

$$(4) \quad \text{fatter}(\text{Arnold}, \text{Babe})$$

Now the farmer says s : ‘The fat pig won a prize’. Intuitively speaking, ‘the fat pig’ picks out Arnold, even though Arnold is a borderline fat pig. We show that the following sentences are entailed by $g(c+s)$: $\text{fat-pig}(\text{Arnold})$, $I(\text{fat-pig}(\text{Arnold}))$, $\neg\text{fat-pig}(\text{Babe})$.

DEMONSTRATION. $g(c) \setminus g_{\text{vague}}(c) = \{1, 2, 3, 4\}$. Since $\text{winner}(\iota x \text{ fat-pig}(x)) \in \text{Cnv}(s, c)$, $\text{Cnv}(s, c) \models \exists!x \text{ fat-pig}(x)$. Since $g_{\text{vague}}(c) = \{\}$, $g_{\text{vague}}(c) \perp (\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)))$ has just a single element: $\text{Cnv}(s, c) \cup \{1, 2, 3, 4\}$. This is then the set selected by Sel as the new common ground $g(c+s)$. In virtue of the constraint on models that the positive extensions of *skinny-pig* and *fat-pig* are disjoint, $\{\exists!x \text{ fat-pig}(x)\} \cup \{1, 2, 3, 4\} \models (\text{fat-pig}(\text{Arnold}) \ \& \ \neg \text{fat-pig}(\text{Babe}))$. Clearly, then, $g(c+s) \models \text{fat-pig}(\text{Arnold})$, $g(c+s) \models I(\text{fat-pig}(\text{Arnold}))$, and $g(c+s) \models \neg\text{fat-pig}(\text{Babe})$.

Ex. 1 illustrates a change in common ground that is simple growth: $g(c+s)$ is a superset of $g(c)$. The updated common ground reflects an adjustment of the positive extension of ‘*fat-pig*’ to include Arnold. Next, we consider a case where accommodation reflects an adjustment to both the positive and negative extensions of a predicate. There are two borderline fat pigs. The farmer uses the description ‘the fat pig’ to refer to the rounder one, but in doing so, brings about a common ground which contains the presupposition that the less round pig is *not* a fat pig.

EXAMPLE 2. The common ground $g(c)$ contains just the following sentences:

$$(1) \quad I(\text{fat-pig}(\text{Arnold}))$$

$$(2) \quad I(\text{fat-pig}(\text{Babe}))$$

$$(3) \quad \forall x(\text{pig}(x) \rightarrow x = \text{Arnold} \vee x = \text{Babe})$$

$$(4) \quad \text{fatter}(\text{Arnold}, \text{Babe})$$

Now the farmer says s : ‘The fat pig won a prize’. Here, too, ‘the fat pig’ ought to pick out Arnold. We show that the following sentences are entailed by $g(c+s)$: $\text{fat-pig}(\text{Arnold})$, $\neg\text{fat-pig}(\text{Babe})$.

DEMONSTRATION. It can be shown that $g(c + s) \models \exists!x \text{ fat-pig}(x) \cup \{1, 2, 3, 4\}$ and $g(c + s) \models \text{fat-pig}(\text{Arnold})$. (Analogous to above.) Since $\exists!x \text{ fat-pig}(x) \cup \{1, 2, 3, 4\} \models \neg \text{fat-pig}(\text{Babe})$, it also follows that $g(c + s) \models \neg \text{fat-pig}(\text{Babe})$.

So much for simple growth. We now return to the case that illustrates a shift in common ground, in Section 4. Arnold, a borderline fat pig, is in a pen with Babe and Cornelius, two very round pigs. The common ground includes the presupposition that all three pigs are fat pigs, but when the farmer uses the description ‘the pig that is not fat’ to pick out Arnold, it shifts.

EXAMPLE 3. The common ground $g(c)$ contains just the following sentences:

- (1) $I(\text{fat-pig}(\text{Arnold}))$
- (2) $\text{fat-pig}(\text{Arnold})$
- (3) $\text{fat-pig}(\text{Babe})$
- (4) $\text{fat-pig}(\text{Cornelius})$
- (5) $\exists x(\text{tail}(\text{Cornelius}, x) \& \text{curly-tail}(x) \& I(\text{curly-tail}(x)))$
- (6) $\forall x(\text{pig}(x) \rightarrow x = \text{Arnold} \vee x = \text{Babe} \vee x = \text{Cornelius})$
- (7) $\text{fatter}(\text{Babe}, \text{Arnold}) \& \text{fatter}(\text{Cornelius}, \text{Arnold})$

Now the farmer says s : ‘The pig that is not fat won a prize’. This time, the description that picks out Arnold is ‘the pig that is not fat’. We show that the following sentences are entailed by $g(c + s)$: $I(\text{fat-pig}(\text{Arnold}))$, $\neg \text{fat-pig}(\text{Arnold})$.

DEMONSTRATION. $g(c) \setminus g_{\text{vague}}(c) = \{1, 6, 7\}$. Since $\text{winner}(\iota x (\text{pig}(x) \& \neg \text{fat-pig}(x))) \in \text{Cnv}(s, c)$, $\text{Cnv}(s, c) \models \exists!x (\text{pig}(x) \& \neg \text{fat-pig}(x))$, and $\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)) \models \neg \text{fat-pig}(\text{Arnold})$. Therefore $\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)) \cup \{2\}$ is not satisfiable. Assuming, as is reasonable, that $\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)) \cup \{3, 4, 5\}$ is satisfiable, it follows that there is a single maximal subset of $g_{\text{vague}}(c)$ that is consistent with $\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c))$, namely $\{3, 4, 5\}$, and that $g_{\text{vague}}(c) \perp (\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)))$ has just a single element: $\text{Cnv}(s, c) \cup \{1, 3, 4, 5, 6, 7\}$. This is then the set selected by Sel

as the new common ground $g(c + s)$. Clearly, by the earlier observation, $g(c + s) \models I(\text{fat-pig}(\text{Arnold}))$ and $g(c + s) \models \neg \text{fat-pig}(\text{Arnold.})$

This example illustrates a shift in common ground. While the prior common ground includes the presupposition that Arnold is a fat pig, the updated common ground does not. It also illustrates the notion of minimal change. The updated common ground includes the non-vague sentences that are a part of the prior common ground, as well as vague sentences that are unaffected by the newly introduced presuppositions, such as 5.

In each of the examples considered above, the common ground is updated in such a way that only one thing comes to satisfy the description of the utterance. In this way, accommodation and the resolution of the referent of the description go hand in hand. Reference resolution need not accompany accommodation, however. In the next example, the farmer uses ‘the fat pigs’ to refer to a subset of the pigs that are present, but the neighbor cannot say which ones. Nonetheless, the prior common ground, which established that Arnold is a fat pig, is updated to include the presupposition that Babe is a fat pig too.

EXAMPLE 4. The common ground contains just the following sentences:

- (1) $\text{fat-pig}(\text{Arnold})$
- (2) $I(\text{fat-pig}(\text{Babe}))$
- (3) $I(\text{fat-pig}(\text{Cornelius}))$
- (4) $\text{skinny-pig}(\text{Doris})$
- (5) $\forall x(\text{pig}(x) \rightarrow x = \text{Arnold} \vee x = \text{Babe} \vee x = \text{Cornelius} \vee x = \text{Doris})$
- (6) $\text{fatter}(\text{Babe}, \text{Cornelius})$

Now the farmer says s : ‘The fat pigs won prizes’. We show that the following sentences are entailed by $g(c + s)$: $\text{fat-pig}(\text{Arnold})$, $\text{fat-pig}(\text{Babe})$, $\forall x(\text{fat-pig}(x) \rightarrow \text{winner}(x))$.

DEMONSTRATION. It can be seen that $g(c) \setminus g_{\text{vague}}(c) = \{2, 3, 5, 6\}$. $\text{Cnv}(s, c) \models \text{winners}(\iota_{\max} X \text{ fat-pigs}(X))$. Assuming, reasonably, that $\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)) \cup \{1, 4\}$ is satisfiable, it follows that $g_{\text{vague}}(c) \perp (\text{Cnv}(s, c) \cup (g(c) \setminus g_{\text{vague}}(c)))$ has just the single element: $\text{Cnv}(s, c) \cup \{1, 2, 3, 4, 5, 6\}$. This is then the set selected by Sel as the new common

ground $g(c + s)$. Since $\text{fat-pig}(\text{Arnold}) \in g(c + s)$, we have $g(c + s) \models \text{fat-pig}(\text{Arnold})$. In virtue of constraints on models, $\{\text{winners}(\iota_{\max} X \text{ fat-pigs}(X))\} \cup \{1, 2, 3, 4, 5, 6\} \models \text{fat-pig}(\text{Babe})$. Clearly, then, $g(c + s) \models \text{fat-pig}(\text{Babe})$. Finally, since $\text{winners}(\iota_{\max} X \text{ fat-pigs}(X)) \models \forall x (\text{fat-pig}(x) \rightarrow \text{winner}(x))$, it follows that $g(c + s) \models \forall x (\text{fat-pig}(x) \rightarrow \text{winner}(x))$.

It appears that the updated common ground is not committed either way on the question whether Cornelius is a fat pig. That is, $g(c + s) \not\models \text{fat-pig}(\text{Cornelius})$ and $g(c + s) \not\models \neg \text{fat-pig}(\text{Cornelius})$. Therefore the updated common ground does not settle the question of which pigs are picked out by the plural description ‘the fat pigs’.

6. ELABORATIONS

In many cases, one possible accommodation is better than the others. In Sect. 4, we introduced a selection function that tells us which one it is, but we have yet to say anything substantial about the choice. We have been able to get away with this because, in each of our examples so far, there was no real choice. There was just a single way to accommodate. In this section, we turn to cases in which there are several ways. The discussion brings out some factors in the choice between them.

EXAMPLE 5. The farmer and the neighbor are standing some distance from a feeder. Arnold is directly between them and the feeder. Babe is also between them and the feeder, but is much closer to it and a little off to one side. The farmer says, “The pig in front of the feeder won a prize”. We assume the arrangement of the pigs to be such that the neighbor understands that he means Babe.¹⁵

We assume that the positive extension of *in-front-of-the-feeder* includes everything within a cone extending from the feeder for some distance in the direction of the speakers, but that prior to the utterance it includes no pigs. The farmer’s utterance brings it about that the positive extension of *in-front-of-the-feeder* stretches to include a single pig. But it could stretch in two different ways. It could come to include Arnold, the pig at a greater distance from the feeder, or Babe, the one at a wider angle. Correspondingly, there will be two candidates for the selection function to choose between.

¹⁵ This seems reasonable if Arnold is sufficiently far away from the feeder, and Babe is sufficiently close.

We have assumed the position of the pigs to be such that, following the farmer's utterance, the common ground establishes that Babe is the unique pig in front of the feeder. The positive extension of *in-front-of-the-feeder* grows to include the pig at a wider angle rather than the one at a further distance. Somehow this change involves a lesser stretch. The selection function must be made sensitive to the fact that some "minimal" changes to the common ground reflect smaller changes in the usage of vague predicates than others. Here is another such case:

EXAMPLE 6. The farmer and his neighbor are standing some distance from a feeder. Arnold, a borderline fat pig, is directly in front of the feeder, and Babe, a definitely fat pig is slightly off to one side. The farmer says, "The fat pig in front of the feeder won a prize".

Here, again, there are two candidates to choose between, one reflecting an adjustment in the usage of 'fat pig', and the other an adjustment in the usage of 'in front of the feeder'. It seems that if Arnold is very close to being definitely fat, and if Babe is so much to the side that he is very near to being definitely *not* in front of the feeder, then the choice is clear – the updated common ground will establish that Arnold, not Babe, is the fat pig in front of the feeder. Again, it seems that among the different options for updating the common ground, the one that reflects the smallest degree of change in the usage of vague predicates is preferred.

It might happen that there are several possibilities for accommodation but no grounds for selecting among them. In the above case, suppose that Arnold were no closer to being definitely fat than Babe is to being definitely in front of the feeder. What accommodation takes place in these cases? Perhaps none at all. After all, the common ground includes only *shared* presuppositions – it does not change in ways that cannot be followed by both participants. Alternatively, a hearer might put off the choice between candidates for accommodation until more information becomes available.

We have described some of the possibilities for accommodating vague descriptions. We finish with a case in which we think accommodation is difficult or impossible, in that it stretches terms past their limits. The case we have in mind again involves three pigs, but these are no ordinary pigs. The first of them is immense, fatter than anything you have ever seen before. The second, though this seems scarcely possible, is fatter still. The third pig is truly, but truly obese – a real monster. And now the farmer says, "The fat pig is going to the fair".

One natural first response will be some sort of puzzlement. You cannot go calling just one pig fat, withholding this honorific from all except the one monster, since each pig is immense. Perhaps there has been some mistake, and the speaker does not mean for reference to be within this group of pigs after all? Perhaps something has gone wrong with the coordination of the common ground among the participants to the conversation? Perhaps one of the pigs is somehow more salient than the rest, and the hearer has not noticed? Perhaps the hearer is mistaken about which pigs are whose? Failing these and a few other possibilities, it seems to us, the farmer has just got to be joking. Whatever the case may be, the hearer ought not, without a blink, to accommodate the speaker by changing to a very exclusive sense of 'fat pig'. To do so would be to miss the difficulty with the coordination, or the joke, or whatever.

Our proposal does not allow accommodation to proceed normally in this example. To see what makes this example different from the previous ones, notice that the prior common ground will include sentences expressing that each of the pigs is *definitely* fat, which is to say, fat on all precisifications. Sentences within the scope of the operator '*D*', in keeping with the earlier exclusion of higher-order vagueness, are not vague. And since sentences expressing that all three of these pigs are definitely fat are jointly inconsistent with the presupposition that there is just one fat pig, it follows that accommodation is undefined in this case. There is no way at all to make minimal changes to the prior common ground so as to accommodate the speaker's utterance, and so our selection function finds itself with no alternatives to choose among. This is a formal reflection of the puzzlement that, as we see it, ought to prompt a hearer to wonder about coordination, jokes and the like.

Perhaps, after a few blinks, the hearer ought to understand that the farmer means the one monster pig anyway. If so, we think, the farmer's usage of 'the fat pig' might be understood in the way suggested by Donnellan's example of the usurper to the throne, discussed in the introduction. The usage is in this case not straightforwardly sincere. Alternatively, we might at this point bring in the comparison classes that we set aside earlier on, allowing accommodation to proceed by changing the comparison class that calibrates the adjective 'fat'. Whereas previously a pig had merely to be fat in relation to other pigs to qualify, after the recalibration it will have to be fat in relation to obese pigs. Similarly, the home handyman who usually gets by with his one adjustable wrench might sometimes need a larger one.

Assuming comparison classes can be brought in at this point, one might wonder why we do not recommend such a treatment for all of the ex-

amples. Since we have not seen a proposal along these lines developed in detail, a weighing of merits is likely to be unfair, both to our proposal and to its hypothetical competitor. But we can speculate. In the introduction, we gestured toward a theoretical reason to prefer a treatment such as ours: we expect that if we are to rely on comparison classes for everything, we will need to multiply them beyond the needs of the compositional semantics of noun phrases. Additionally, the case of the three definitely fat pigs may present an empirical difficulty for a treatment that involves comparison classes, without any alternative mechanism to balk at the farmer's talk about "the fat pig". Accommodation *ought* to be difficult in cases like this one. If changing comparison classes allows it to go without a hitch, then so much the worse for changing comparison classes.

REFERENCES

- Alchourrón, C., P. Gärdenfors, and P. Makinson: 1985, 'On the Theory of Logic Change: Partial Meet Functions for Contraction and Revision', *Journal of Symbolic Logic* **50**, 510–530.
- Beaver, D.: 1997, 'Presupposition', in J. Van Benthem and A. ter Meulen (eds.), *Handbook of Logic and Language*, Elsevier, Amsterdam.
- Donnellan, K.: 1966, Reference and Definite Descriptions', *Philosophical Review* **75**.
- Fine, K.: 1975, 'Vagueness, Truth, and Logic', *Synthese* **30**, 265–300.
- Gärdenfors, P.: 1988, *Knowledge in Flux*, MIT Press, Cambridge, Mass.
- Gazdar, G.: 1979a, *Pragmatics: Implicature, Presupposition and Logical Form*, Academic Press, New York.
- Gazdar, G.: 1979b, 'A Solution to the Projection Problem', in C.-K. Oh and D. A. Dinneen (eds.), *Syntax and Semantics II: Presupposition*, Academic Press, New York, 57–89.
- Grice, H. P.: 1967, 'Logic and Conversation', Unpublished MS of the William James Lectures, Harvard University, reprinted, in part, in P. Cole and J. L. Morgan (eds.) 1975, *Syntax and Semantics 3: Speech Acts*, Academic Press, New York.
- Hansson, S. O.: 1992, 'In Defense of Base Contraction', *Synthese* **91**, 239–45.
- Kamp, J. A. W.: 1975, 'Two Theories about Adjectives', in E. Keenan (ed.), *Formal Semantics of Natural Languages*, Cambridge University Press, Cambridge.
- Karttunen, L. and S. Peters: 1975, 'Conventional Implicature in Montague Grammar', *Syntax and Semantics* **11**, 1–56.
- Karttunen, L. and S. Peters: 1979, 'Conventional Implicature', *Syntax and Semantics* **11**, 1–56.
- Karttunen, L. and S. Peters: (19??) 'Requiem for Presupposition', in *Proceedings of the Third Annual Meeting of the Berkeley Linguistic Society*, 360–371.
- Levinson, S. C.: 1983, *Pragmatics*, Cambridge University Press, Cambridge.
- Lewis, D.: 1979, 'Scorekeeping in a Language Game', *The Journal of Philosophical Logic* **8**, 339–359. Reprinted in D. Lewis (1983) *Philosophical Papers*, Volume I. Oxford University Press, New York.
- Manor, R.: 1995, 'Pragmatic Considerations in Semantic Analyses', *Pragmatics and Cognition* **3**(2), 225–245.

- Marr, D.: 1982, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman, San Francisco.
- Morreau, M.: 1999, 'Supervaluation can Leave Truth-Value Gaps After All', *The Journal of Philosophy* **96**(3), 148–156.
- Pinkal, M.: 1983, 'On the Limits of Lexical Meaning', in R. Bauerle, C. Schwarze and A. von Stechow (eds.), *Meaning, Use and Interpretation* de Gruyter, Berlin.
- Pinkal, M.: 1995, *Logic and Lexicon*, Kluwer Academic Publishers, Dordrecht.
- Sorenson, R.: 1988, *Blindspots*, Clarendon Press, Oxford.
- Stalnaker, R.: 1974, 'Pragmatic Presuppositions', in M. K. Munitz and P. K. Unger (eds.), *Semantics and Philosophy*, NYU Press, New York, 197–213.
- Whitehead, A. N. and B. Russell: 1922, *Principia Mathematica*, 2nd edn., Cambridge University Press, Cambridge.
- Williamson, T.: 1994, *Vagueness*, Routledge, London.

Alice Kyburg
Department of Philosophy
University of Wisconsin Oshkosh
Oshkosh, WI 54901
kyburg@uwosh.edu

Michael Morreau
Department of Philosophy and UMIACS
University of Maryland College Park
College Park, MD 20742
mimo@umiacs.umd.edu

Interpreting Vague Utterances in Context

David DeVault and Matthew Stone

Department of Computer Science

Rutgers University

Piscataway NJ 08854-8019

David.DeVault@rutgers.edu, Matthew.Stone@rutgers.edu

Abstract

We use the interpretation of vague scalar predicates like *small* as an illustration of how systematic semantic models of dialogue context enable the derivation of useful, fine-grained utterance interpretations from radically underspecified semantic forms. Because dialogue context suffices to determine salient alternative scales and relevant distinctions along these scales, we can infer implicit standards of comparison for vague scalar predicates through completely general pragmatics, yet closely constrain the intended meaning to within a natural range.

1 Introduction

Modeling context and its effects on interpretation may once have seemed to call for an open-ended investigation of people's knowledge of the common-sense world (Hobbs et al., 1993). But research on the semantics of practical dialogue (Allen et al., 2001) now approaches dimensions of context systematically, through increasingly lightweight, factored models. The evolving state of real-world activity proceeds predictably according to background plans and principles of coordination (Rich et al., 2001). The status of the dialogue itself is defined by circumscribed obligations to ground prior utterances, follow up open issues, and advance real-world negotiation (Larsson and Traum, 2000). Finally, the evolving state of the linguistic context is a direct outgrowth of the linguistic forms interlocutors use and the linguistic relationships among successive utterances (Ginzburg and Cooper, 2001; Asher and Lascarides, 2003). These compatible models combine directly to characterize an aggregate information state that provides a general background for interpretation (Bunt, 2000).

We argue in this paper that such integrated models enable systems to calculate useful, fine-grained utterance interpretations from radically underspecified semantic forms. We focus in particular on vague scalar predicates like *small* or *long*. These predicates typify *qualitative* linguistic expression of *quantitative* information, and are thus both chal-

lenging and commonplace. Building on a multidimensional treatment of dialogue context, we develop and implement a theoretically-motivated model of vagueness which is unique in treating vague predicates as genuinely vague and genuinely context-sensitive, yet amenable to general processes of contextual and interpretive inference.

1.1 Semantic insights

We pursue our argument in the context of an implemented drawing application, FIGLET, which allows users to give English instructions to draw a caricature of an expressive face. Figure 1 shows a representative interaction with FIGLET; the user gives the successive instructions in (1):

- (1) a. Make two small circles.
- b. Draw a long line underneath.

Like Di Eugenio and Webber (1996), we emphasize that understanding such instructions requires contextual inference combining linguistic, task and domain knowledge. For example, consider the response to (1a) of placing circles so as to form the eyes of a new face. To recognize the possibility of drawing eyes exploits knowledge of the ongoing drawing task. To put the eyes where they belong in the upper part of the new face exploits domain knowledge. The response to (1b) adds the linguistic context as another ingredient. To identify where the line goes, the user uses the objects mentioned recently in the interaction as the understood spatial landmark for *underneath*. Figure 1 highlights the importance of using multidimensional representations of dialogue context in understanding instructions for quantitative domains.

We leverage this background context in our computational approach to vagueness. We model a vague utterance like *draw a long line* as though it meant *draw a line with, you know, length*. In this approach, vague predicates are completely underspecified; linguistic knowledge says nothing about how long something *long* is. Instead, vague language explicitly draws on the background knowl-

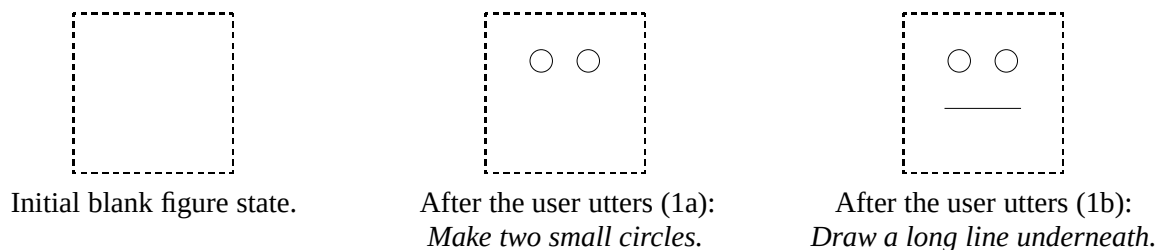


Figure 1: Motivating interaction: Vague instructions to draw a face.

edge already being applied in utterance interpretation. The user’s motivation in using *long* is to differentiate an intended interpretation, here an intended action, from alternative possibilities in context. Background knowledge already sets out the relevant ways to *draw a line*; *drawing a long line* means singling out some of them by the length of that new line. This model recalls dynamic theories of vague scalar predicates, such as the semantics of Kyburg and Morreau (2000), Barker (2002), or Kennedy (2003), but it is exactly implemented in FIGLET. The implementation capitalizes on the richness of current models of context to recover content for the *you know* of vagueness.

1.2 Overview

In Section 2, we motivate approaches to the semantics of vague scalar predicates that associate them with a presupposed standard of comparison. We illustrate how context can be understood to supply possible standards, and how pragmatic reasoning from utterances allows interlocutors to infer them. In Section 3, we establish a bridge to the general treatment of practical dialogue, by showing how multiple dimensions of context generally contribute to recognizing possible interpretations for underspecified utterances. Section 4 builds on Sections 2 and 3 to show how FIGLET exploits a rich model of utterance context to respond cooperatively to vague utterances like (1a) and (1b), while Section 5 details FIGLET’s actual implementation. We conclude in Section 6 by suggesting further challenges that vagueness still poses for computational semantics.

2 Vague standards in context

We adopt a view of vague predicates motivated by linguistic theory, particularly Kennedy’s approach (1999; 2003). We assume that gradable adjectives are associated with measurement functions mapping *individuals* to *degrees* on a *scale*. In FIGLET’s drawing domain, the relevant measurements pertain to spatial properties. For *long*, for example, the measurement maps individuals to their spatial lengths; for *small*, it maps individuals to degrees on an inverted scale of size.

Positive gradable adjectives compare measured degrees against a *standard* on the scale which is derived from context. For example, *long* says that an object’s length exceeds the threshold set by the current standard for length. Other forms, such as comparative adjectives or adjectives with explicit measure phrases, compare degrees differently.

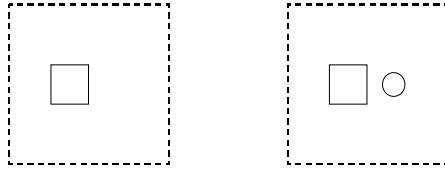
Importantly, grammar says *nothing* about how standards for positive gradable adjectives are derived. In other words, contra Cresswell (1977) and others, the interpretation of adjectives is not computed relative to a grammatically-specified “comparison class” of related objects. And, contra Oates et al. (2000) and Roy and Pentland (2002), the interpretation of adjectives need not require statistical knowledge about how objects with different measurements on a scale tend to be described. Instead, standards are derived directly from an evolving context by the general principles that govern pragmatic resolution of context dependence.

Kennedy synthesizes a range of evidence for this claim. Here we go further, and provide a formal, implemented model. We can sketch the evidence and our model by considering two key examples.

First, we illustrate that vagueness depends directly on specific contextually-relevant distinctions. Consider the session with FIGLET shown in Figure 2. The user has elected to draw two objects side-by-side. The initial context just contains a square. The user utters (2).

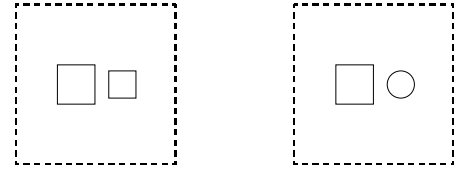
(2) Make a small circle.

To interpret (2) it doesn’t seem to help to appeal to general knowledge about how big circles tend to be. (It’s quite convoluted to even frame the idea in a sensible way.) Graff (2000) observes that standards often implicitly appeal to what we expect about particular individuals, not just what we know about similar individuals. In context, here, the user just seems to be asking for a circle vaguely smaller than the square. This is the interpretation FIGLET builds; to comply, FIGLET draws the circle an arbitrary but representative possible size. The point is that salient objects and actions inevitably set up meaningful dis-



Initial figure state. After the user utters (2).

Figure 2: Taking standards from context in (2): *Make a small circle.*



Initial figure state. After the user utters (3).

Figure 3: Disambiguating contextual standards in (3): *Make the small square a circle.*

tinctions in the context. Interlocutors exploit these distinctions in using vague adjectives.

Figure 3 illustrates that understanding vagueness is part of a general problem of understanding utterances. Figure 3 shows FIGLET's action in a more complex context, containing two squares of different sizes. We consider the user's instruction (3):

(3) Make the small square a circle.

FIGLET's action changes the smaller of the two squares. The standard behind this interpretation is implicitly set to differentiate the contextually-salient objects from one another; the natural resolution of (3) does not require that either square be definitely small (Kyburg and Morreau, 2000). In Figure 3, for example, there are different potential standards that would admit either both squares or neither square as small. However, we can rule out these candidate standards in interpreting (3). The user's communicative intention must explain how a unique square from the context can be identified from (3) using a presupposed *small* standard. If that standard is too big, both squares will fit. If that standard is too small, neither square will fit. Only when that standard falls between the sizes of the squares does (3) identify a unique square.

The examples in Figures 2 and 3 show two ways new standards can be established. Once established, however, standards become part of the evolving context (Barker, 2002). Old standards serve as defaults in interpreting subsequent utterances. Only if no better interpretation is found will FIGLET go back and reconsider its standard. This too is general pragmatic reasoning (Stone and Thomason, 2003).

3 Dimensions of context in interpretation

To cash out our account of contextual reasoning with vagueness, we need to characterize the context for practical dialogue. Our account presupposes a context comprising *domain and situation knowledge*, *task context* and *linguistic context*. In this section, we survey each of these dimensions of context, and show how they converge in the resolution of underspecification across a wide range utterances.

Domain and situation knowledge describes the commonsense structure of the real-world objects and actions under discussion. Practical dialogue restricts this otherwise open-ended specification to the circumscribed facts that are directly relevant to an ongoing collaboration. For example, in our drawing domain, individuals are categorized by a few types: types of shape such as *circles* and *squares*; and types of depiction such as *eyes* and *heads*. These types come with corresponding constraints on individuals. For example, the shape of a mouth may be a line, an ellipse, or a rectangle, while the shape of a head can only be an ellipse. These constraints contribute to interpretation. For instance, a head can never be described as *a line*, for example, since heads cannot have this shape.

Task context tracks collaborators' evolving commitment to shared goals and plans during joint activity. In FIGLET's drawing domain, available actions allow users to build figure parts by introducing shapes and revising them. Our experience is that users' domain plans organize these actions hierarchically into strategic patterns. For example, users tend to complete the structures they begin drawing before drawing elsewhere; and once they are satisfied with what they have, they proceed in natural sequence to a new part nearby. Task context plays a powerful role in finding natural utterance interpretations. By recording a plan representation and keeping track of progress in carrying it out, FIGLET has access to a set of candidate next actions at each point in an interaction. Matching the user's utterance against this candidate set restricts the interpretation of instructions based on the drawing already created and the user's focus of attention within it. For example, if the user has just drawn the right eye onto an empty face, they are likely to turn to the left eye next. This context suggests making a winking left eye in response to *draw a line*, an interpretation that might not otherwise be salient.

Linguistic context records the evolving status of pragmatic distinctions triggered by grammatical conventions. One role of the linguistic context is its contribution to distinguishing the prominent entities

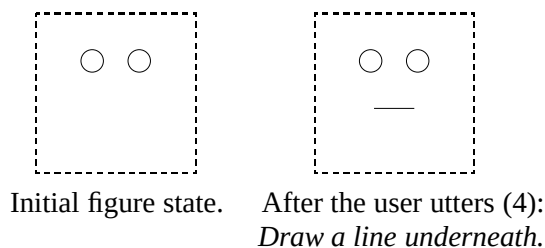


Figure 4: Context in instructions.

that can serve as the referents of pronouns and other reduced expressions. To see this, note that, as far as domain knowledge and task context go, the instruction *make it bigger* could apply to any object currently being created. If the figure is hierarchical, there will be many possibilities. Yet we typically understand *it* to refer specifically to an object mentioned saliently in the previous utterance. The linguistic context helps disambiguate *it*.

Figure 4 illustrates how the three different dimensions of context work together. It illustrates an interaction with FIGLET where the user has just issued an instruction to create two eyes, resulting in the figure state shown at the left in Figure 4. The user’s next instruction is (4):

(4) Draw a line underneath.

We focus on how the context constrains the position and orientation of the line.

Linguistic context indicates that *underneath* should be understood as *underneath the eyes*. This provides one constraint on the placement of the line. Task context makes drawing the mouth a plausible candidate next action. Domain knowledge shows that the mouth can be a line, but only if further constraints on position, orientation and length are met. In understanding the instruction, FIGLET applies all these contextual constraints simultaneously. The set of consistent solutions—drawing a horizontal line at a range of plausible mouth positions below the eyes—constitutes the utterance interpretation. FIGLET acts to create the result in Figure 4 by choosing a representative action from this set.

4 Interpreting vague utterances in context

In our approach, the linguistic context stores agreed standards for vague predicates. Candidate standards are determined using information available from domain knowledge and the current task context. In FIGLET’s drawing domain, possibilities include the actual measurements of objects that have already been drawn. They also include the default domain measurements for new objects that task context says could be added. Setting standards by a measure-

ment is our shorthand for adopting an implicit range of compatible standards; these standards remain vague, especially since many options are normally available (Graff, 2000).

We treat the use of new candidate standards in interpretation as a case of presupposition accommodation (Bos, 2003). In presupposition accommodation, the interpretation of an utterance must be constructed using a context that differs from the actual context. When speakers use an utterance which requires accommodation, they typically expect that interlocutors will update the dialogue context to include the additional presumptions the utterance requires. We assume that all accommodation is subject to two Gricean constraints. First, we assume whenever possible that an utterance should have a uniquely identifiable intended interpretation in the context in which it is to be interpreted. Second, we assume that when interpretations in alternative contexts are available, the speaker is committed to the strongest one—compare Dalrymple et al. (1998). Inferring standards for vague predicates is a special case of this general Gricean reasoning.

The principles articulated thus far in Sections 2–4 allow us to offer a precise explanation of FIGLET’s behavior as depicted in Figure 1. The user starts drawing a face with an empty figure. In this domain and task context, *make two circles* fits a number of possible actions. For example, it fits the action of drawing a round head and its gaping mouth. However, in (1a), what the user actually says is *make two small circles*. The interpretation for (1a) must accommodate a standard for *small* and select from the continuum of size possibilities two new circles that meet this standard.

The standards in this context are associated with the size distinctions among potential new objects. The different qualitative behavior of these standards in interpretation can be illustrated by the standards set from possible new circular objects that are consistent with the face-drawing task. We can set the standard from the default size of an eye, from the default size of a mouth (larger), or from the default size of a head (larger still).¹ Because each standard allows all smaller objects to be created next, these standards lead to 1, 3, and 6 interpretations, respectively. So we recover the standard from the eye, which results in a unique interpretation.²

¹Since the default sizes of new objects reflect the relative dimensions of any other objects already in the figure, FIGLET’s default sizes are not generally equivalent to static comparison classes.

²Note that there are many potential sources of standards for *small* that FIGLET does not currently pursue. E.g. the average size of all objects already in the figure. We believe that general

In tandem with its response, FIGLET tracks the changes to the context. The task context is updated to note that the user has drawn the eyes and must continue with the process of creating and revising the features of the face. The linguistic context is updated to include the new *small* standard, and to place the eyes in focus.

This updated context provides the background for (1b), the user's next instruction *draw a long line underneath*. In this context, as we saw with Figure 4, context makes it clear that any response to *draw a line underneath* must draw the mouth. Thus, unlike in (1a), all the interpretations here have the same qualitative form. Nevertheless, FIGLET's Gricean reasoning can still adjust the standard for length to differentiate interpretations quantitatively, and thereby motivate the user's use of the word *long* in the instruction. FIGLET bases its possible standards for *length* on both actual and potential objects. It can set the standard from an actual eye or from the two eyes together; and it can set the standard from the default mouth or head. The mouth, of course, must fit inside the head; the largest standard is ruled out. All the other standards lead to unique interpretations. Since the length of the two eyes together is the strictest of the remaining standards, it is adopted. This interpretation leads FIGLET to the response illustrated at the right in Figure 1.

5 Implementation

We have implemented FIGLET in Prolog using CLP(R) real constraints (Jaffar and Lassez, 1987) for metric and spatial reasoning. This section presents a necessarily brief overview of this implementation; we highlight how FIGLET is able to exactly implement the semantic representations and pragmatic reasoning presented in Sections 2–4. We offer a detailed description of our system and discuss some of the challenges of building it in DeVault and Stone (2003).

5.1 Semantic representation

In FIGLET, we record the semantics of user instructions using *constraints*, or logical conjunctions of open atomic formulas, to represent the contextual requirements that utterances impose; we view these constraints as presuppositions that speakers make in using the utterance. We assume matches take the form of instances that supply particular domain representations as suitable values for variables. Stone (2003) motivates this framework in detail.

methods for specifying domain knowledge will help provide the meaningful task distinctions that serve as candidate standards for vague predicates on our approach, but pursuing this hypothesis is beyond the scope of this paper.

In (5a-d), we show the presuppositions FIGLET assigns to an utterance of *Make two small circles*, arranged to show the contributions of each individual word. In (5e), we show the contribution made by the utterance to an evolving dialogue; the effect is to propose that an action be carried out.

- (5) a. $simple(A) \wedge target(A, X) \wedge fits_plan(A) \wedge$
 $holds(result(A, now), visible(X)) \wedge$
 $holds(now, invisible(X)) \wedge$
- b. $number(X, 2) \wedge$
- c. $standard(small, S) \wedge$
 $holds(result(A, now), small(X, S)) \wedge$
- d. $number(X, multiple) \wedge$
 $holds(result(A, now), shape(X, circle))$
- e. $propose(A)$

We formulate these constraints in an expressive ontology. We have terms and variables for *actions*, such as A ; for *situations*, such as now and $result(A, now)$; for *objects*, such as X ; for standards for gradable vague predicates (scale-threshold pairs), such as S ; and for quantitative points and intervals of varying dimensionality, as necessary.

5.2 Pragmatic reasoning

Constraint networks such as (5a-e) provide a uniform venue for describing the various contextual dependencies required to arrive at natural utterance interpretations. Thus, the contextual representation and reasoning outlined in Sections 3 and 4 is realized by a uniform mechanism in FIGLET: specifications of how to reason from context to find solutions to these constraints.

For example, Section 3 described domain knowledge that links particular object types like *eyes* and *heads* with type-specific constraints. In our implementation, we specify real and finite constraints that individuals of each type must satisfy. In order for an individual e of type t to serve as part of a solution to a constraint network like (5a-e), e must additionally meet the constraints associated with type t . In this way, FIGLET requires utterance interpretations to respect domain knowledge.

Solving many of the constraints appearing in (5a-e) requires contextual reasoning about domain actions and their consequences. Some constraints characterize actions directly; thus $simple(A)$ means that A is a natural domain action rather than an abstruse one. Constraints can describe the effects of actions by reference to the state of the visual display in hypothetical situations; thus $holds(result(A, now), shape(X, circle))$ means that the individual X has a circular shape once action A is carried out. Constraints can additionally char-

acterize causal relationships in the domain; thus $target(A, X)$ means that action A directly affects X , and the constraints of (5a-d) together mean that carrying out action A in the current situation causes two small circles to become visible. These constraints are proved in FIGLET by what is in effect a planner that can find complex actions that achieve specified effects via a repertoire of basic domain actions.

Task context is brought to bear on interpretation through the $fits_plan(A)$ constraint of (5a). FIGLET uses a standard hierarchical, partially ordered plan representation to record the structure of a user's task. We specify the solutions to $fits_plan(A)$ to be just those actions A that are possible next steps given the user's current state in achieving the task. Since these task-appropriate actions can factor additional constraints into interpretation, enforcing the $fits_plan(A)$ constraint can help FIGLET identify a natural interpretation.

As discussed in Section 4, FIGLET records a list of current standards for vague scalar adjectives in the linguistic context. The constraint $standard(small, S)$ of (5c) connects the overall utterance interpretation to the available standards for *small* in the linguistic context. FIGLET interprets utterances carrying semantic constraints of the form $standard(vague_predicate, S)$ in one or two stages. In the first stage, the constraint is solved just in case S is the prevailing standard for *vague-predicate* in the linguistic context. If there is *no* prevailing standard for an evoked vague property, or if this stage does not yield a unique utterance interpretation, then FIGLET moves to a second stage in which the constraint is solved for any standard that captures a relevant distinction for *vague-predicate* in the context. If there is a strongest standard that results in a unique interpretation, it is adopted and integrated into the new linguistic context.

5.3 Parsing and Interpretation

Language understanding in FIGLET is mediated by a bottom-up chart parser written in Prolog. As usual, chart edges indicate the presence of recognized partial constituents within the input sequence. In addition, edges now carry constraint networks that specify the contextual reasoning required for understanding. In addition to finite instances (Schuler, 2001), these networks include real constraints that formalize metric and spatial relationships. Interpretation of these networks is carried out incrementally, during parsing; each edge thus records a set of associated candidate interpretations. Since domain reasoning can be somewhat time-intensive in our current implementation, we adopt a strategy

of delaying the solution of certain constraints until enough lexical material has accrued that the associated problem-solving is judged tractable (DeVault and Stone, 2003).

6 Assessment and Conclusion

In our approach, we specify a genuinely vague semantics: vague words evoke a domain-specific scale that can differentiate alternative domain individuals. To find a unique interpretation for a vague utterance, we leverage ordinary inference about the domain, task, and linguistic context to recover implicit thresholds on this scale.

We believe that further methodological advances will be required to evaluate treatments of vagueness in indefinite reference, such as that considered here. For example, obviously the very idea of a “gold standard” for resolution of vagueness is problematic. We believe that the best argument for a theory of vagueness in a language interface would show that naive users of the interface are, on the whole, likely to accept its vague interpretations and unlikely to renegotiate them through clarification. But the experiment would have to rule out confounding factors such as poorly-modeled lexical representation and context tracking as sources for system interpretations that users reject.

We intend to take up the methodological challenges necessary to construct such an argument in future work. In the meantime, while our current implementation of FIGLET exhibits the promising behavior discussed in this paper and illustrated in Figures 1–4, some minor engineering unrelated to language understanding remains before a fruitful evaluation can take place. As alluded to above, the tight integration of contextual reasoning and interpretation that FIGLET carries out can be expensive if not pursued efficiently. While our initial implementation achieves a level of performance that we accept as researchers (interpretation times of between one and a few tens of seconds), evaluation requires us to improve FIGLET's performance to levels that experimental participants will accept as volunteers. Our analysis of FIGLET indicates that this performance can in fact be achieved with better-regimented domain problem-solving.

Nevertheless, we emphasize the empirical and computational arguments we already have in support of our model. Our close links with the linguistic literature mean that major empirical errors would be surprising and important across the language sciences. Indeed, limited evaluations of treatments of vague definite reference using standards of differentiation or very similar ideas have been promising

(Gorniak and Roy, In Press). The computational appeal is that all the expensive infrastructure required to pursue the account is independently necessary. Once this infrastructure is in place the account is readily implemented with small penalty of performance and development time. It is particularly attractive that the approach requires minimal lexical knowledge and training data. This means adding new vague words to an interface is a snap.

Overall, our new model offers three contributions. Most importantly, of course, we have developed a computational model of vagueness in terms of underspecified quantitative constraints. But we have also presented a new demonstration of the importance and the feasibility of using multidimensional representations of dialogue context in understanding descriptions of quantitative domains. And we have introduced an architecture for resolving underspecification through uniform pragmatic mechanisms based on context-dependent collaboration. Together, these developments allow us to circumscribe possible resolutions for underspecified utterances, to zero in on those that the speaker and hearer could adopt consistently and collaboratively, and so to constrain the speaker's intended meaning to within a natural range.

Acknowledgments

We thank Kees van Deemter and our anonymous reviewers for valuable comments. This work was supported by NSF grant HLC 0308121.

References

- J. Allen, D. Byron, M. Dzikovska, G. Ferguson, L. Galescu, and A. Stent. 2001. Towards conversational human-computer interaction. *AI Magazine*, 22(4):27–37.
- N. Asher and A. Lascarides. 2003. *Logics of Conversation*. Cambridge.
- C. Barker. 2002. The dynamics of vagueness. *Linguistics and Philosophy*, 25(1):1–36.
- J. Bos. 2003. Implementing the binding and accommodation theory for anaphora resolution and presupposition. *Computational Linguistics*, 29(2):179–210.
- H. Bunt. 2000. Dialogue pragmatics and context specification. In H. Bunt and W. Black, editors, *Abduction, Belief and Context in Dialogue*, pages 81–150. Benjamin.
- M. Cresswell. 1977. The semantics of degree. In B. H. Partee, editor, *Montague Grammar*, pages 261–292. Academic.
- M. Dalrymple, M. Kanazawa, Y. Kim, S. Mchombo, and S. Peters. 1998. Reciprocal expressions and the concept of reciprocity. *Linguistics and Philosophy*, 21(2):159–210.
- D. DeVault and M. Stone. 2003. Domain inference in incremental interpretation. In *Proc. ICoS*.
- B. Di Eugenio and B. Webber. 1996. Pragmatic overloading in natural language instructions. *Int. Journal of Expert Systems*, 9(2):53–84.
- J. Ginzburg and R. Cooper. 2001. Resolving ellipsis in clarification. In *Proc. ACL*.
- P. Gorniak and D. Roy. In Press. Grounded semantic composition for visual scenes. *Journal of Artificial Intelligence Research*.
- D. Graff. 2000. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*, 28(1):45–81.
- J. Hobbs, M. Stickel, D. Appelt, and P. Martin. 1993. Interpretation as abduction. *Artificial Intelligence*, 63:69–142.
- J. Jaffar and J.-L. Lassez. 1987. Constraint logic programming. In *Proc. POPL*, pages 111–119.
- C. Kennedy. 1999. *Projecting the adjective: The syntax and semantics of gradability and comparison*. Garland.
- C. Kennedy. 2003. Towards a grammar of vagueness. Manuscript, Northwestern.
- A. Kyburg and M. Morreau. 2000. Fitting words: Vague words in context. *Linguistics and Philosophy*, 23(6):577–597.
- S. Larsson and D. Traum. 2000. Information state and dialogue management in the TRINDI dialogue move engine toolkit. *Natural Language Engineering*, 6:323–340.
- T. Oates, M. D. Schmill, and P. R. Cohen. 2000. Toward natural language interfaces for robotic agents. In *Proc. Agents*, pages 227–228.
- C. Rich, C. L. Sidner, and N. Lesh. 2001. COLLAGEN: applying collaborative discourse theory to human-computer interaction. *AI Magazine*, 22(4):15–26.
- D. Roy and A. Pentland. 2002. Learning words from sights and sounds: A computational model. *Cognitive Science*, 26(1):113–146.
- W. Schuler. 2001. Computational properties of environment-based disambiguation. In *Proc. ACL*, pages 466–473.
- M. Stone and R. H. Thomason. 2003. Coordinating understanding and generation in an abductive approach to interpretation. In *Proc. DiaBruck*, pages 131–138.
- M. Stone. 2003. Knowledge representation for language engineering. In A. Farghaly, editor, *A Handbook for Language Engineers*, pages 299–366. CSLI.

The Role of Ambiguity in the Sorites Fallacy

Kees van Deemter,
Institute for Perception Research (IPO)

1 Introduction

Semanticists¹ have often designed their theories as if all the predicates of natural languages make distinctions that are perfectly exact, in the sense that any individual to which the predicate is at all applicable must either belong to the denotation of the predicate or to its complement. Yet, it is well known that many words resist this simple picture. Predicates such as ‘long’ and ‘short’, for example, display a certain conceptual vagueness, in that the boundary between objects that belong to their denotation and objects that do not is often hard to determine. ‘Sorites’ or ‘chain argument’² is the ancient name for a family of paradoxes that exploit the vagueness of such expressions. Here is a representative example of a sorites argument:

Suppose you are to judge the heights of each of a long series of people, looking at them from a distance that makes a difference in height invisible as long as it amounts to less than, say, 1 cm. Of each of them, you are being asked whether they are short or not. The people are presented to you starting from the shortest, who is, by assumption, short, up to the tallest, who is very tall. The difference between subsequent persons is always less than 1cm, and therefore invisible to you.

Now if you decide that the first person (p_0) is short, you must also judge the next one (p_1) short, since you can perceive no difference between the two. But then, by the same token, the third person (p_2) must be short as well, and so on indefinitely. In particular, this makes also the last person (p_n) short. Yet, by assumption, p_n is not short. The argument can be summarized as follows:

1. p_0 is short, and p_n **is not short** (Assumptions).
2. If p_0 is short then p_1 is short.
3. Therefore, p_1 is short.

¹This paper is an outgrowth of work that was done during my 1993-1994 stay at CSLI, Stanford. A more elaborate description of this work is contained in [van Deemter 1994]. The current paper focuses on some issues that are not discussed in detail in that paper: (1) the role of ambiguity in my proposal, (2) the relation with an earlier proposal by Veltman and Muskens, and (3) the issue of vague predicates that are not context-dependent. Thanks are due to Johan van Benthem, Jan van Eijck, David Israel, Stanley Peters, Yoav Shoham, Martin Stokhof, Frank Veltman and Ed Zalta, all of which helped shape the ideas of this paper. I am grateful to the Netherlands Organization for Scientific Research (NWO) for funding, and to CSLI for being such a stimulating place.

²There exists some confusion about the word sorites. The philosopher Galenus (2nd century) used the Greek word *soreites* (cf. *soreuo* = *to amass*, or *to stack*) to designate the usually harmless situation in which a proof uses the conclusion of one syllogism as the premise of another. This is also called a polysyllogism. The sorites paradox might be described as a *vicious* polysyllogism. The name sor(e)ites suggests that the paradox arises through a number of inference steps, each of which is valid, but which are invalid in combination.

4. If p_1 is short then p_2 is short.
- (...)
- $2n+1$. Therefore, p_n **is short**. $\perp$

The reasoning is elementary, and consists of a long series of applications of Modus Ponens. The premises under (1) are unproblematic (for the right choice of individuals p_0 and p_n). The even-numbered premises are best seen as consequences of a very general ‘inductive’ premise saying

Inductive Premise: For all x and y , if x is short and y is indistinguishable³ from x , then y is short.

Many proposals to resolve this paradox have been put forward, most of them involving rather radical departures from classical logic.⁴ Before moving on to sketch my own proposed solution, which will be described in more detail in [van Deemter 1994], let me briefly point out what I take to be the requirements for a solution of the paradox. I believe a proper solution should

1. Explain the plausibility of the argument that leads to the paradox, while showing that the argument is nevertheless invalid.
2. Be consistent with what is known about the meaning of vague predicates. The solution should preferably be backed-up by empirical evidence.
3. Use as few *unmotivated* departures from classical logic as possible. A departure can only be motivated by an appeal to (2).

My strategy in coping with these requirements will be to argue along the following lines: vague predication must always involve at least two arguments: an object judged *and* a set of elements with which this object is being compared. Consequently, vague predication is incomplete without a ‘comparison set’, in comparison with whose elements the predication must be understood. Nothing is small or large *per se*, but only in comparison with other things. The comparison sets are provided by the (linguistic and nonlinguistic) context of the utterance. Thus, the sorites argument is viewed as a piece of discourse during which a comparison set is built up. When the context-dependence of the vague expressions in the argument is taken into account, it will turn out that the Inductive Premise of the sorites argument becomes *ambiguous*. What makes sorites arguments plausible, in our view, is the following fact: In some of its interpretations, the inductive premise supports the argument, **but** then it fails to be true by virtue of meanings alone; in other versions, it is true by virtue of meanings alone, **but** then it does not support the argument.

³The term ‘indistinguishable’ is biased towards versions of the paradox in which the subsequent elements p_j , p_{j+1} are perceptibly indistinguishable in the relevant dimension. The paradox may – with slightly less cogency – also be formulated for cases in which p_j , p_{j+1} are distinguishable, but similar enough to make the inductive premise plausible. I can see no difference between genuine indistinguishability and mere similarity that would warrant different solutions of the associated versions of the paradox.

⁴Issue No.30 of the journal *Synthese* (1975) is still a classic collection of papers in the area.

The current approach to the paradox differs from most other accounts that I know of, in that it does not propose a change in the underlying logic. The logic is purely classical.⁵ On the other hand, the interpretation of non-logical constants changes, in that they acquire an additional argument place for a comparison set. In the sequel of this article, I will go on to talk about shortness or smallness as our paradigmatic case of a vague predicate, abbreviating ‘ x is short’ or ‘ x is small’ by the formula $S(x)$. With the comparison set A as an additional argument, I will write $S(x)^A$ to abbreviate ‘ x is small with respect to the comparison set A ’.

The structure of this paper is the following. In section 2, it is shown how context becomes a source of ambiguity. Section 3 applies the idea of contextualization to the crucial inductive premise of the sorites argument and shows how this premise becomes ambiguous in a very crucial way. Section 4 discusses ambiguity as a source of paradox, and section 5 summarizes, and reviews the relation between context-dependence and vagueness.

2 Context as a source of ambiguity

The idea that context is relevant to vagueness is hardly new. This holds especially for *local* context, such as the Noun in a ‘Pred + Noun’ construction. For example, in Ewan Klein’s analysis of ‘Some small elephant runs’, the set of elephants functions as a *comparison set* that somehow sets the standard for what counts as small [Klein 1980]. This may happen in one of various ways, for instance by calling all those elements small that are smaller than the average element of the comparison set. This contextual interpretation of *small* explains, for example, why a *small elephant* is not necessarily a *small animal*, since elephants are, on average, taller than the average animal.

Recent dynamic perspectives on discourse semantics can be taken to suggest the relevance of *global* context for the determination of denotations in general, whether they are vague or not [Kamp 1981], [Heim 1982]. For example, it has been argued that NPs of all major syntactic kinds allow ‘anaphoric’ interpretation, that is: interpretation with respect to a contextually determined subdomain of the universe of discourse [van Deemter 1992]. In the following piece of discourse, let D be the domain of discourse, and let $C \subseteq D$ be the set of people in the crowd, then, on one interpretation of the second sentence, the NP *only one man* is interpreted ‘with respect to’ C :

The crowd was approaching the castle; only one man carried a key.

In the (non-vague) case of a NP like *only one man*, interpretation ‘with respect to’ a subdomain comes down, roughly, to intersecting the Noun with the subdomain. Thus, only one element in the set $\| \text{Man} \| \cap C$ carried a

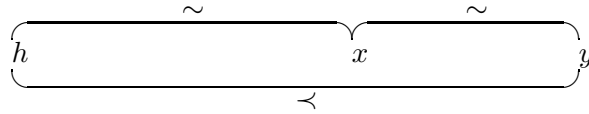
⁵In particular, the logic will be monotonic, since the addition of a new premise to an argument does not automatically change the comparison sets that are relevant for the sentences that occur in the argument.

key.⁶ Now suppose a context-dependent NP contains a vague Adjective, as does *one tall man* in

The crowd was approaching the castle; one tall man carried a key.

Then on at least one reading of the discourse, the man is tall ‘with respect to’ the comparison set C . This does not only imply that the man is a member of the crowd, but also that he was tall when compared to all the people in the crowd⁷. As in the case of local contextuality [Klein 1980], being tall ‘with respect to’ a comparison set may be defined in various nonequivalent ways, but what counts here is this: the standard of comparison that is relevant for a vague adjective can be dependent on the same sort of ‘global’ context as the domain of discourse that is relevant for the interpretation of a Noun Phrase in general.

Perhaps the most straightforward way to use the notion of global context in the treatment of the sorites paradox is the ‘dynamic’ account that was sketched by Veltman and Muskens [Veltman 1987]⁸. Veltman and Muskens make use of a construction that was first mentioned by Russell [Russell 1940], and elaborated by Goodman [Goodman 1966] and by Dummett [Dummett 1975]. The key idea is to view the elements that a context makes available as resources that an observer may use to distinguish between (for example) the heights of individuals. In the following example, the subject of the experiment cannot see a difference between the heights of x and y . On the other hand, he or she *can* see a difference in the ‘behavior’ of h towards the two. In particular, h is perceived as smaller than y (henceforth $h \prec y$) but h is not perceived as smaller than x , even though h actually is smaller than x (i.e. $\text{Size}(h) < \text{Size}(x)$):



The idea of Goodman and Dummett is that this difference in behavior between x and y allows the subject to tell x and y apart. One way to formalize this is to define

$$(x \prec y)^A \Leftrightarrow_{\text{Def}} x \prec y \vee \exists h \in A : (h \prec y \ \& \ \neg h \prec x) \vee (x \prec h \ \& \ \neg y \prec h).$$

⁶For a detailed analysis of how non-vague NPs are interpreted with respect to a contextually determined domain of discourse, see [Westerstahl 1985], [van Deemter 1992]. The latter also contains an analysis of how this approach can be extended to relational NPs such as the NP *a key* in the example discourse.

⁷We will disregard the question of how exactly local (*man*) and global (*crowd*) context interact, since dependence on local context is irrelevant for the sorites paradox, where vague adjectives appear without a Noun that they modify.

⁸The approach of Veltman and Muskens, described in [Veltman 1987], which is not widely available, was brought to my attention after finishing a first draft of [van Deemter 1994].

A simple motivation⁹ for this definition can be given as follows: Suppose $x \prec y$ is true if and only if $\text{Size}(y) - \text{Size}(x)$ exceeds some fixed amount. (It does not matter how Size is measured.) Then $(x \prec y)^A$ logically implies that $\text{Size}(x) < \text{Size}(y)$. Making use of this notion, one may define

$$(x \sim y)^A \Leftrightarrow_{\text{Def}} \neg((x \prec y)^A) \vee (y \prec x)^A$$

to express that A does not contain help-elements to distinguish between the heights of x and y . The relation $(x \sim y)^A$ will be called ‘Goodman/Dummett-style indistinguishability’¹⁰.

So, contextualizations exist not only for vague predicates, but also for the relation of indistinguishability. Veltman and Muskens put these contextualizations to use in the following way. First, they assume that, at each point in a discourse, a unique comparison set is relevant for predicates or relations occurring at that point. A comparison set is called *decisive* for a set of formulas if it contains all the individuals whose names occur in the formulas, and we will call the smallest comparison set that is decisive for a set of formulas Δ the comparison set that is *determined* by Δ . The notation A_Δ designates the union of A and the comparison set that is determined by Δ :

$$A_\Delta = A \cup \{p_i : p_i \text{ occurs somewhere in the sentences in } \Delta\}^{11}$$

Contextualization, for Veltman and Muskens, is automatic and need not appear explicitly in formulas. A comparison set is ‘read off’ a discourse, and this can be done in different ways. Veltman and Muskens assume that each object that occurs in a discourse enters the comparison set of any (occurrence of a) vague predication in the discourse. For this reason, truth is defined at the level of an entire discourse, rather than on the sentential level:

Truth of a discourse: A discourse Δ is true in a model M with respect to a comparison set $A \Leftrightarrow$ for every $\phi \in \Delta$, ϕ is true in M with respect to A_Δ .

Sorites’ inductive premise is not directly interpreted as a premise of the sorites argument, but rather as a constraint on models. For a model to be *coherent*, it should, among other things, satisfy this constraint:

Coherence constraint: If $(x \sim y)^A$ is true in M and M verifies $S(x)^A$, then M verifies $S(y)^A$ as well,

⁹See [Scott and Suppes 1958] and [Suppes and Zinnes 1963] for a precise treatment.

¹⁰Later on in this paper, I will also make use of the Goodman/Dummett construction. However, it should be stressed that Goodman/Dummett indistinguishability is by no means the only way to contextualize the notion of indistinguishability, nor is it the only one that has the formal properties that we will later need. See, for example, the notion of a *safety valve* that is formulated in section 5.

¹¹Note that this formula does not distinguish between an object and its name. This simplification will be used throughout this paper.

As a special case, consider the sorites argument presented in the Introduction, in which p_0 is small and p_n is not small. On any reasonable notion of logical consequence, the conclusion $S(p_n)$ will be a logical consequence of the following premises:

- 0. $S(p_0)$
- 1. $S(p_0) \rightarrow S(p_1)$,
- 2. $S(p_1) \rightarrow S(p_2), \dots$,
- n. $S(p_{n-1}) \rightarrow S(p_n)$.

Observe that, taken individually, each of the conditional premises is true in any coherent model. This is true because such a one-sentence discourse *determines* a comparison set that contains only the two elements occurring in it, and this comparison set does not contain enough help-elements to allow a distinction between the two.¹² However, when the premises co-occur in a logical argument, they should be viewed as parts of one and the same discourse. Together, they *determine* the comparison set that contains all of $p_0, \dots, p_n$. As a result, it is impossible that the set of premises is true, since there must be at least one pair p_j, p_{j+1} ($0 \leq j \leq n-1$) for which the comparison set $\{p_0, \dots, p_n\}$ does enforce a distinction, given that p_0 is small and p_n is not.

This solution, which explains the paradox as valid, but flawed by premises that may be true individually but not in combination, is appealing for various reasons. In particular, it is logically more conservative than most approaches, and it points to what I think is an important element of the solution of the paradox, namely contextuality. On the other hand, I believe it has certain drawbacks. First, consider a version of the paradox that makes use of one, universally quantified inductive premise, rather than of a long sequence of its instantiations. No treatment for quantified formulas is provided in [Veltman 1987], but it seems that the inductive premise would have to collect all its instantiations into the comparison set that is *determined* by it, and as a result, they would have to call this inductive premise *false*. So, one might ask, why does the paradox arise, if one of its premises is just false? Second, I believe that Veltman and Muskens do not quite take the similarities with anaphoric context-dependency serious enough. One thing that is crucial about anaphora is that it tends to create a resolution problem. This holds for pronominal anaphora, but to an even higher extent for the contextdependencies that are found in full NPs. In particular, a non-pronominal NP is systematically ambiguous between (1) various possible anaphoric readings, in which the NP appears restricted to an anaphoric antecedent, and (2) a non-anaphoric reading, in which the NP appears unrestricted. For example, in the above-mentioned example, the NP *only one man* can be restricted in its interpretation to the set of people in the crowd, but it can also be interpreted unrestricted, to say that only one man *in the world* carried

¹²More precisely, the premise i , saying that $S(p_{i-1}) \rightarrow S(p_i)$, determines the comparison set $A = \{p_{i-1}, p_i\}$, and $(p_{i-1} \sim p_i)^A$ is clearly true for this value of A . The coherence constraint will now imply the truth of the premise i .

a key. The anaphoric perspective predicts that the same ambiguity affects vague expressions and, more importantly, the relation of indistinguishability. According to this perspective, the context that is relevant at a certain point of the discourse cannot simply be ‘read off’ the clauses, as Veltman and Muskens have it. Comparison sets have to be added as explicit arguments before truth values can be assigned. Somewhat less centrally, an anaphoric perspective suggests that the comparison set that is determined by an occurrence of a formula is dependent on occurrences of formulas *earlier* in the discourse only, since objects that may be observed in the future are unlikely to affect present judgements.¹³

The most important consequence of the anaphoric perspective is that the inductive premise becomes ambiguous between readings in which indistinguishability is contextualized and readings in which it is not. This fact will lay at the core of our proposal for a solution of the sorites paradox.

3 The ambiguity of the inductive premise

The previous section has shown that when context is taken into account, the inductive premise of the sorites paradox becomes ambiguous between ‘contextual’ and ‘absolute’ readings. In the present paper, we will focus on one particular way to contextualize each of the two.¹⁴ Indistinguishability is contextualized as was proposed by Dummett and Goodman (cf. previous section). A vague predicate S is contextualized as in the following meaning postulate:

$$\textbf{Comparison postulate: } S(x)^A \Leftrightarrow |\{y \in A : y \succ x\}| \text{ is greater than } |\{y \in A : x \succ y\}|.$$

According to the Comparison Postulate, a predicate such as *small* involves a comparison between two sets: the set of elements in A that are visibly smaller than the element x that is judged (later abbreviated $K[x]^A$), and the set of elements in A that are visibly taller than it (abbreviated $G[x]^A$).

$$\underbrace{\dots\dots\dots}_{K[x]^A} \quad \dots\dots \quad x \quad \dots\dots \quad \underbrace{\dots\dots\dots}_{G[x]^A}$$

Whether or not x is small thus only depends on the cardinality of the two sets $K[x]^A$, $G[x]^A$. The elements in between $K[x]^A$ and x are not taken into account, and neither are those in between x and $G[x]^A$, to reflect the intuition that only those elements of A can matter for the truth of $S(x)^A$ that are visibly distinguishable from x .

Returning now to the inductive premise of the sorites argument, let A be

¹³As in pronominal anaphora, the possibility of very limited, forward-looking (i.e. kataphoric) reference cannot be excluded completely. But we will simplify and assume that the context-dependence of vague NPs is only backward-looking.

¹⁴For a systematic discussion of other ways to contextualize and their effect on the sorites paradox, see [van Deemter 1994].

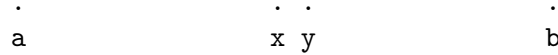
the comparison set that is ‘active’ before the inductive premise is processed. Further, we will simplify and assume that the comparison set for a clause that involves the predicate S consists of A plus all those elements that have appeared as the argument of S , earlier on in the discourse. The same assumption will be made with respect to the relation of indistinguishability, except that this relation can also be interpreted in an absolute, noncontextual way.¹⁵ As a result, the informal version of the inductive premise can be interpreted in either of the following ways:

Absolute version: $S(x)^A \ \& \ x \sim y \Rightarrow S(y)^{A \cup \{x\}}$.

Contextualized version: $S(x)^A \ \& \ (x \sim y)^{A \cup \{x\}} \Rightarrow S(y)^{A \cup \{x\}}$.

Actually, in both versions, the clause $S(y)^{A \cup \{x\}}$ can be simplified to $S(y)^A$, which is equivalent to the longer formula, since x can never show up in either $G^A[y]$ or $K^A[y]$. In the sequel of this section, I will use these contextualizations of Indistinguishability and Smallness to show that the inductive premise of the sorites argument is ambiguous between two classes of interpretations: plausible versions that are too weak to derive the paradox, and implausible ones that are strong enough to derive the paradox. As I will point out in section 4, this makes sorites an instance of a well-known type of fallacy, which rests on the ambiguity of a premise. A more formal account, in which this claim is generalized beyond the very specific definitions of ‘smallness with respect to a comparison set’ and ‘indistinguishability with respect to a comparison set’ of the present section is given in [van Deemter 1994].

First, the absolute version. It is easy to see that this version is strong enough to derive the paradoxical conclusion, since any element p_n can be linked to p_0 through a series $p_0 \sim p_1 \sim \dots \sim p_n$. On the other hand, the absolute version is not ‘analytical’: if a domain is large and dense enough, then it may be true that $S(x)^A$ and $x \sim y$, while $\neg S(y)^A$. Here is an example:



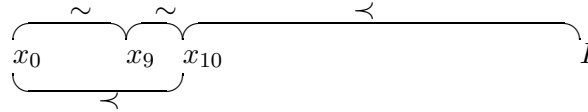
Now for the contextualized version. It may come as a surprise, but this version of the premise is *analytical*: it is true in all models that satisfy the definitions of the meanings of S and $\sim$ that were given so far. This can be seen as follows. Suppose $S(x)^A$ and $\neg S(y)^A$. Then the comparison principle (which amounts to counting elements, after all) implies that there has to exist at least one element $z \in A$ that behaves differently towards x than towards y , in that it is an element of $G^A[x]$ but not of $G^A[y]$, or in that it is an element of $K^A[y]$ but not of $K^A[x]$. But then z helps distinguish between x and y , contradicting the assumption that $(x \sim y)^A$.

¹⁵This asymmetry between S and $\sim$ is an immediate consequence of the earlier assumptions about the meanings of S and $\sim$.

So what does this mean? Is the sorites paradox a valid argument after all, provided a properly contextualized version of the inductive premise is chosen? No, of course not. We have only established that its crucial premise can be interpreted in such a way that it must be true in any domain. But, the premise loses so much of its power that no paradoxical conclusions can be derived anymore, as the following illustration clarifies.

For clarity, let us not deal with people, since we all have a complex history of acquaintances with actual people, while it is not clear how all of these affect our standards of comparison. (Memory issues become prevalent, etc.) Instead, we will talk about Martians, all of which have a height between 0.60 m and 1 m. Other dimensions than height will be disregarded. Assume a line-up consists of individuals that are just 1 mm apart, starting at 0.60 m. Let us number the Martians in the line-up with increasing height, so x_0 has height 0.60, and x_i has height $0.60 + i$ mm. A Just Notable Difference (JND) equals 10 mm. (So, Martians that differ 1 cm can just be told apart.) Now what is the standard of comparison for the first Martian that one encounters? Let us assume that the standard of comparison is something like the speaker's own height (i.e. $\text{Size}(I)$). In my case, $\text{Size}(I)$ is quite a bit greater than 1 m, so I would call the first Martian small:

0. $S(x_0)^A$, since $A = \{I\}$, while $G^A[x_0] = \{I\}$ and $K^A[x_0] = \phi$
 1. $S(x_1)^A$, since $A = \{I, x_0\}$, while $G^A[x_1] = \{I\}$ and $K^A[x_1] = \phi$
 -
 9. $S(x_9)^A$, since $A = \{I, x_0, \dots, x_8\}$, and $G^A[x_9] = \{I\}$ and $K^A[x_9] = \phi$
 10. $\neg S(x_{10})^A$, since $A = \{I, x_0, \dots, x_9\}$, and
- $G^A[x_{10}] = \{I\}$ and $K^A[x_{10}] = \{x_0\}$

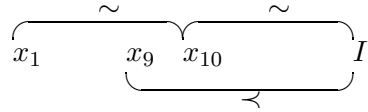


This shows that somehow, the sorites argument must break down at (or before) x_{10} . There are prospective applications of the premise from (0) to (1), from (1) to (2),..., and from (9) to (10), with different instantiations of x, y, A . For example, one gets from (0) to (1) by choosing $x := x_0, y := x_1$, and $A := \{I\}$, observing that $(x_0 \sim x_1)^{\{I\}}$, since both are distinguishably smaller than the speaker. This works until one arrives at step (10). By assumption, x_{10} is 1 mm taller than x_9 , causing it to be distinguishably taller than the initial element x_0 . As a result, $\neg((x_9 \sim x_{10})^{\{I, x_0, \dots, x_8\}})$, since $(x_0 \sim x_9)$, but $(x_0 < x_{10})$. In other words: after having inferred that x_9 must be small (relative to the context at hand), there happens to be no suitable object around to instantiate the antecedent of the inductive premise. Of course such an object might have been around; for example, an element $x_{10'}$ that is just 0.5 mm taller than x_9 , and then another after $x_{10'}$, and so on, approaching x_{10} as a limit. However, in that way one will always only infer

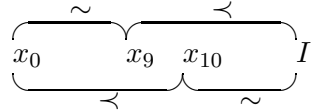
the smallness of elements that are smaller than x_{10} , and so no paradoxical conclusion is inferred. The value $\text{Size}(x_0) + \text{JND}$ functions as a limit: only x that are smaller than this value are inferentially indistinguishable from x_9 .

Observe that, in this example, elements stop being small at precisely the same time when they stop being inferentially indistinguishable from x_9 . This is a property of this particular example. Should our initial comparison set have consisted of two elements, I and an element I' that is just 1 mm taller than I , then x_{10} would also be inferentially indistinguishable from x_9 , but it would still be small, since $G^A[x_{10}] = \{I, I'\}$ and $K^A[x_{10}] = \{x_0\}$.

Starting from this example, it is easily seen that the paradox can break down in one of the following three ways: (i) the case illustrated above, in which the argument breaks down through an element x_i that has the same set $G^A[x_i]$ as its predecessor, but an enlarged $K^A[x_i]$; (ii) the same $K^A[x_i]$, but a diminished $G^A[x_i]$; or (iii) through an element that is located in such a way that $G^A[x_i]$ and $K^A[x_i]$ change at the same time. The second of the three situations is depicted in the following, where x_0 is absent, so I becomes the element that causes the breakdown:



The last of the three situations is depicted in the following:



So, why does the paradox break down when the contextualized version of the inductive premise is used? Well, firstly, we have assumed that a vague notion such as Small requires a comparison set. Now *whatever* one takes as the initial comparison set A_0 of finite size, some finite ‘norm’ will result, implying some upper limit on what is small with respect to that norm. As soon as an individual is judged that falls short of this norm (being too large in size), this individual fails to be small with respect to A_0 . Of course, the norm is not fixed forever, it will change when newly-judged individuals enter the comparison set. But, as long as the elements that are judged are *small* with respect to the current comparison set, they will only cause the norm to go *down*. In other words, adding a small element to a comparison set may sometimes cause a small element to become not-small, but never the other way round, as the reader may easily check:

$$\text{If } S(y)^A, \text{ then } \|S^{A \cup \{y\}}\| \subseteq \|S^A\|$$

(Here the extension of Small with respect to comparison set A is denoted $\|S^A\|$.) As a result, all the individuals $(x_0, x_1, \dots)$ that are added in the course

of a sorites argument first leave the norm of the initial comparison set intact (causing a break-down of the argument as soon as an individual turns up that is not small with respect to A_0 , as in (ii) and (iii)), and, in the somewhat longer run, they cause the norm to become even stricter (causing an even earlier break-down of the argument, as in (i)). As soon as comparison sets enter the picture, they are there to stay, causing vague predicates to become precise, and causing the sorites argument to break down.

We have seen that the inductive premise of the sorites argument is ambiguous between an absolute version, which is strong enough to support the paradox, but fails to be analytic, and a contextualized version, which is analytic, but fails to support the paradox. The following section will reflect on the implications of this situation for the evaluation of the paradox.

4 Ambiguity as a source of fallacy.

Which of the two are to blame for the paradox of sorites: the premises or the reasoning from the premises to the conclusion? Some approaches have analysed the paradox as a bit of logically sound reasoning from false premises, others have analysed it as a logically unsound reasoning from true premises. What is my own answer to this question?

According to the analysis that I have presented here, the form of the sorites argument is the following:

Premise p
 Premise q
 Conclusion r ,

where p equals the premise that p_0 is small, q is the inductive premise, and r equals the statement that p_n is small. Moreover, q is ambiguous between readings q_1 and q_2 , while the reasoning $p, q_1 \models r$ is classically valid but the reasoning $p, q_2 \models r$ is not. On the other hand, q_2 is true (in fact, it is analytic, but that does not make an essential difference now), and q_1 is not true. To illustrate, here is a very different instantiation of this general schema:

p = *This decision has been arrived at through a majority decision.*
 q = *Democratic decisions respect the Rule of Law.*
 r = *This decision respects the Rule of Law.*

Let us assume that the premise q is ambiguous between at least two readings, which make use of different conceptions of democracy. In one reading (q_1), *democratic* means ‘arrived at through a majority decision’ and in the other (q_2) ‘respecting the Rule of Law’. What makes this a fallacy in the narrow sense¹⁶ of a *plausible* piece of *invalid* reasoning, is the fact that one might get confused: q_2 is a logical truth, but fails to connect p and r ; on the

¹⁶See, for example, Mackie’s contribution on ‘Fallacies’ in the Encyclopedia of Philosophy [Mackie 1967].

other hand, q_1 does provide the required connection between p and r , but it is false. It is precisely this situation that makes the argument deceptive.

One is tempted to ask whether this fallacy has a name in traditional logic, which was always strongly preoccupied with fallacies. As it turns out, already Aristotle realized the role of ambiguity in fallacies, and he distinguished several kinds:¹⁷ *amphiboly* if the ambiguity involved grammatical constructions and *equivocation* or *ambiguity* if the ambiguity (modern usage) involved the meaning of a word. It is not obvious to me which of the two is more appropriate in the case of an anaphoric ambiguity – a type of ambiguity that seems to have been overlooked by traditional logicians – but the latter does not seem inappropriate given current usage of the word ‘ambiguity’.

Ultimately, of course, it is not terminology that matters. What does matter is the question of how *modern* logic looks at ambiguous premises. Until recently, modern logicians have hardly looked at ambiguity at all. This is excusable for logicians who work on foundations of mathematics or on the philosophy of science, where constructing unambiguous languages is of the essence. But surprisingly, it is also true for logicians who work on the semantics of natural language. An example is Thomason who wrote, explaining Montague’s notion of a ‘disambiguated language’:

“... there is no serious point to constructing an artificial language that is not disambiguated” [Thomason 1974].

As a result, there still are few formal tools to study logical consequence as a relation between sentences some of which may be ambiguous. But recently, some tentative efforts in this direction were made ([van Deemter 1991], [Poesio 1991]), [Reyle 1992]) and we will go on in the spirit of these. Assume, for simplicity, that every reading of every sentence of an ambiguous language L can be expressed unambiguously by some other sentence of L .¹⁸ Let a (syntactic) *disambiguation* d be a function that maps every unambiguous sentence of L to itself, and that maps every ambiguous sentence S of L to some sentence $d(S)$ that expresses one of the readings of S . In addition, let d also map *sets* of sentences of L to sets of unambiguous sentence, in the obvious way.

Given this setting, how can logical consequence be defined? As I have argued elsewhere¹⁹, several options may be justifiable. For example,

- 1a. $\Gamma \models_{\forall} \psi$ iff for all d, d' , $d(\Gamma) \models d'(\psi)$.
- 1b. $\Gamma \models_{\forall} \psi$ iff for all d there is a d' such that $d(\Gamma) \models d'(\psi)$.
- 2. $\Gamma \models_{\exists} \psi$ iff for some d, d' , $d(\Gamma) \models d'(\psi)$.

¹⁷The topic is discussed at length in Aristotle’s *Sophistici Enlenchi*, see [Ross 1923].

¹⁸If necessary, an extension L' of L may be constructed that is expressive enough to say everything in L unambiguously.

¹⁹See [van Deemter 1991]. The syntactic approach of the current paper (where d maps a sentence to its syntactic disambiguation) differs from the semantic account in that paper, but this difference is not essential for current purposes.

In addition to speaking of the relations $\models_{\exists}$ and $\models_{\forall}$ obtaining between premises and conclusion, we will also speak of an inference being $\text{valid}_{\exists}$ or $\text{valid}_{\forall}$. The first of the three definitions is the most direct implementation of the intuition that an inference is logically sound only if *nothing* can refute it. It says that ambiguous consequence requires logical consequence under all possible disambiguations. The second option, which has sometimes been advocated on the ground that it makes the relation of ambiguous consequence reflexive, transitive and monotonic [Reyle 1992], calls an argument valid iff for every disambiguation of the premises, a disambiguation of the conclusion can be found such that the combination makes the argument classically valid. Under the present circumstances, where the only relevant ambiguity resides in the *premises*, (1a) and (1b) collapse, of course. Option (2) is genuinely more liberal, calling an argument valid if it is valid under some disambiguation. Now as it turns out, these notions of logical consequence induce different answers to the question of whether the premises or the reasoning should be blamed for the sorites paradox. In particular, the first two ((1a) and (1b)) find fault with the reasoning, but the last one does not:

1. $p, q \not\models_{\forall} r$
2. $p, q \models_{\exists} r$.

According to the second notion of ambiguous consequence, the reasoning is valid.

So much for the soundness of the reasoning. How about the premises? Here a similar story can be told, since several notions of ‘ambiguous truth’ can be distinguished, including

1. ϕ is $\text{true}_{\forall}$ ($\text{false}_{\forall}$) in a model M iff for all d , $d(\phi)$ is true (false) in M .
2. ϕ is $\text{true}_{\exists}$ ($\text{false}_{\exists}$) in a model M iff for some d , $d(\phi)$ is true (false) in M .

These definitions can be extended to the case of *sets* of sentences in an obvious way. One has to conclude that the question of whether the sorites premises are true has no unique answer. Instead, they are (1) neither $\text{true}_{\forall}$ nor $\text{false}_{\forall}$, but (2) $\text{true}_{\exists}$. What does this mean? If (ambiguous) truth can be interpreted in such a way that the sorites premises are true, and if (ambiguous) logical consequence can be interpreted in such a way that the sorites reasoning comes out valid, then does that mean that there is nothing wrong with the sorites paradox? Of course not. There are disambiguations that make the reasoning valid and there are disambiguations that make the premises true, but there are no disambiguations that do both. The problem is that received terminology is not very good at making this point. Suppose one would dub an inference *good* if it is not only logically valid, but if its premises are true as well:

An inference ‘ p, q therefore r ’ is *good* iff it is logically valid and both p and q are true.

In the situation in which premises may be ambiguous, this definition can take at least 4 different forms, depending on which notion of ambiguous consequence and which notion of ambiguous truth are selected. Suppose, for instance, one selects $\models_{\exists}$ and $\text{true}_{\exists}$, then the definition makes our paradoxical argument *good*, which makes it hard to understand what is wrong with it; on the other hand, if one selects $\models_{\forall}$ and $\text{true}_{\forall}$, then the argument comes out *not good* for the double reason that it is not logically valid and its premises are not true, which makes it hard to see why the argument is at all plausible.

One way to look at this situation is the following: modern logic, in opposition to Aristotle, for example, has never made much fuss about the truth of premises. The notion of truth can simply be paired with that of logical consequence, in the manner of the definition that was just given. The two can be kept separate. On the other hand, as soon as one looks at ambiguous formulas, a simple pairing of truth and logical consequence tends to obfuscate some important issues: saying that the premises ‘can be true’ and that the argument ‘can be valid’ (i.e. on at least one interpretation) does not tell you whether the argument can be valid, having true premises *at the same time*. Here is a bit of simple terminology that might help to fill this gap, by introducing ‘disambiguation-dependent’ versions of truth and logical consequence:

Truth, relative to a disambiguation: ϕ is true_d (false_d) in a model M iff $d(\phi)$ is true (false) in M .

Logical consequence, relative to a disambiguation: $\Gamma \models_d \psi$ iff $d(\Gamma) \models d(\psi)$.

Goodness, relative to a disambiguation and a model: An argument ‘ $\Gamma \models \psi$ ’ is good_d in M iff $\Gamma \models_d \psi$, and all the elements of Γ are true_d in M .

Absolute *goodness* can then be defined as *goodness* with respect to at least one disambiguation:

An argument $\Gamma \models \psi$ is *good* in M iff $\exists d$ such that $\Gamma \models \psi$ is good_d in M .

If the relation $\Gamma \models_d \psi$ holds between premises and conclusion, we will also say that the argument is valid_d . This terminology allows the following analysis of the paradox: it is valid_d for some d and its premises are true_d for some d ; however, there is no d such that both conditions are met. Consequently, the argument is not *good*. The meta-level claim that I want to make can now be expressed as the claim that any argument that is $\text{valid}_{\exists}$, whose premises are $\text{true}_{\exists}$, but that is not *good* is a Fallacy of Ambiguity. The more specific claim is that the sorites argument instantiates this schema.

5 Conclusion: Vagueness and Context-dependence

In the above I have sketched an approach to the sorites paradox of vagueness in which the notion of context-dependence plays a central role. This raises

the question²⁰ of whether vagueness must always involve context-dependence and how possible counterexamples affect my proposal. Does my solution only work for *some* vague predicates, leaving versions of the paradox that use others unexplained?

First, it is at least conceivable that there exist vague predicates that are not context-dependent. For example, their vagueness may result from perceptual or judgemental limitations, while the standards for comparison are invariable. A tentative example would be the vague predicate *healthy*. Arguably, whether a person is healthy is not dependent on the context in which the person's health is considered, but only on considerations related to the requirements of survival and procreation. But let us stay away from the complexities of real life and define an artificial predicate *smoll* with a completely absolute meaning:

$$Smoll(x) \Leftrightarrow_{Def} Size(x) < Size(I).$$

This may not make *smoll* completely independent of context, but it does make it insensitive to the 'discourse related' growth of context that we have used in much of this article. Yet, *smoll* is vague, since perceptual limitations make it impossible, for any perceiver, to determine for each possible x whether $Smoll(x)$. So how about this predicate *smoll*? Does it give rise to a sorites paradox, and if so, in how far can this version of sorites be explained using the strategy of the present paper?

It may well be that a believable sorites argument can be set up using the predicate *smoll*. Using the definiens of *smoll*, the inductive premise says

Inductive Premise: For all x and y , if $Size(x) < Size(I)$ and y is indistinguishable from x , then $Size(y) < Size(I)$.

It might be that this version of the premise lacks plausibility, and in that case there is nothing to worry about. But *if* the premise is accepted, then the paradox does come up and the question is: can we solve it? The answer is yes, I think.

Recall our strategy: find (1) interpretations of the inductive premise that are analytic but too weak to support sorites, and (2) interpretations that do support the paradox but are not analytic. Let us try to do this and define, as before, the contextualized version $(x \sim y)^A$ in Dummett/Goodman style, while uncontextualized $x \sim y$ is equivalent to something of the form

$$| Size(y) - Size(x) | < n.$$

Then what we see is the usual pattern:

1. $Smoll(x) \ \& \ x \sim y \Rightarrow Smoll(y)$.
2. $Smoll(x) \ \& \ (x \sim y)^A \Rightarrow Smoll(y)$.

²⁰Both Martin Stokhof and Yoav Shoham have insisted that I try to answer this question.

That (1) is non-analytic and supportive is obvious. To determine the status of (2), one has to assume that $I \in A$, which seems natural enough.²¹ Consequently, (2) becomes analytic for perfectly familiar reasons: suppose $S(x)$ and $\neg S(y)$, then $x < \text{Size}(I)$, but not $y < \text{Size}(I)$, which implies $\neg(x \sim y)^A$.²² This shows that a vague predicate need not be dependent on discourse context to give rise to the kinds of ambiguities that we have been concerned with in this paper, and which help us to understand the sorites paradox.

This example may be put aside on the grounds that *smoll* is context-dependent, in a way, since a person's smollness depends on the size of the speaker. I can see two counterarguments to this objection. Firstly, it can be argued that absolute judgements involving a vague predicate can always be re-analysed as involving an implicit comparison with other objects. (For example, being unhealthy may be analysed as being less healthy than the unhealthiest person who is still healthy.) Secondly, one can move in the opposite direction and come up with a completely non-contextdependent version of the inductive premise, which is still both analytical and non-supportive. Suppose, for example, $\text{Size}(I) = 1.70m$, then the following definition makes the artificial predicate *smull* coextensive with *smoll*:

$$\text{Smull}(x) \Leftrightarrow_{Def} \text{Size}(x) < 1.70m.$$

A version of the inductive premise that makes use of the predicate *smull* and that is supportive/non-analytic is easy to find. For example, a variant of (1) (above) does the job. But this time around, an analytic version of the inductive premise requires a different notion of indistinguishability than the one proposed by Goodman/Dummett, since $\text{Smull}(x) \& \neg \text{Smull}(y)$ does not imply the existence of z such that $x \prec z$ but $y \not\prec z$. However, contextualizing indistinguishability in the Goodman/Dummett style is by no means the only way to obtain analytic versions of the inductive premise²³.

Here comes a somewhat trivial, but instructive example. Let $S(x)^A$ be any reasonable notion of smallness (predicated of a suitable object x), in which the comparison set A may play a real role or not.²⁴ Let $(x \sim y)^A$ be any reasonable relation of indistinguishability between suitable relata x and y . Then one might modify $(x \sim y)^A$ by applying a kind of *safety valve* to the relation, naming the result $(x \sim y)_S^A$. This novel version of the indistinguishability relation, which might be called '*S*-safe indistinguishability', is defined as follows:

²¹For example, it is reasonable to assume that *if* the speaker is not an element of the comparison set that is active when processing of the premise starts, *then* the clause $S(x)$ (which involves, under the present definition of S , a comparison between x and the speaker) adds the speaker to the comparison set.

²²Note that this situation is exactly analogous to the one in the illustration of section 2 (part (ii)), where the paradox breaks down because $x_{10} \sim I$.

²³For a general discussion of how analytic/nonsupportive, as well as non-analytic/supportive versions of the inductive premise can be obtained, regardless of overly specific assumptions about the meanings of S and $\sim$, see [van Deemter 1994].

²⁴The set A does not 'play a real role' in the definition if, for all A, A' , it holds that $\forall x \forall y ((x \sim y)^A \Leftrightarrow (x \sim y)^{A'})$.

$$(x \sim y)_S^A \Leftrightarrow_{Def} (x \sim y)^A \ \& \ (S(x)^A \leftrightarrow S(y)^A).$$

In other words, x and y are ‘ S -safe indistinguishable’ (under given definitions of S and indistinguishability) iff they are not only indistinguishable from each other, but either both have the property S or neither of them has the property S . It is easy to see that this notion of indistinguishability can never bring one from an object that the property S to one that does not have the property S .²⁵ Consequently, the resulting version of the inductive premise becomes analytic:

Inductive premise: For all x and y , if $S(x)^A$ and $(x \sim y)_S^A$ then $S(y)^A$.

Also, it should be obvious that this version of the premise does not support the sorites paradox. One thing that is interesting about this construction is that it has the kind of effect that the Goodman/Dummett construction also has, but independently of the definition of S .²⁶ In particular, it does not presuppose that S is context-dependent in any way. For example, the reader may easily check that it leads to analytical/nonsupportive S if combined with the predicate *Smull*.

To conclude, does all this amount to a solution of the sorites paradox? I believe that it does succeed in answering the *first* requirement that was formulated in the Introduction, namely of explaining the invalidity of the paradox while at the same time accounting for its plausibility. Also, the approach is very much in the spirit of the *third* requirement, since its departures from classical logic are extremely modest. The only potential problem that I can see has to do with the *second* requirement, of empirical adequacy. It might be argued that my proposal denies the vagueness of the predicates involved, replacing vagueness by context-dependence and – in its wake – ambiguity. Consequently, it does not reflect an intuition that has sometimes been put forward: that an object a that is smaller than a certain small object b is small *to a higher degree* than b . It seems that doing full justice to this purported intuition of graduality must necessarily lead to a highly nonclassical logic²⁷. I see no technical obstacle for superimposing graduality – in the form of multiple truth values – on my own proposal. What I have tried to show is that such a drastic move is not required for the solution of the sorites paradox.

REFERENCES

[Barwise 1985] Noun Phrases, Generalized Quantifiers and Anaphora. CSLI Stanford, Report No. CSLI-86-52, CSLI, Stanford University.

²⁵Note that if $S(x)^A$ is defined as in the Comprison Principle, then the safety valve leads to a less drastic sharpening of indistinguishability than the Goodman/Dummett construction: $(x \sim y)^A$ implies, but is not implied by $(x \sim y)_S^A$.

²⁶Recall that the Goodman/Dummett construction will only lead to an analytical version of the inductive premise under some highly specific assumptions about the meaning of S .

²⁷A well-known example of a ‘gradual’ logic is fuzzy logic [Zadeh 1975], which is notorious for its drastic deviations from classical logic.

- [**Carter 1987**] Interpreting anaphors in natural language textst, Ellis Horwood, Wiley & Sons, New York.
- [**van Deemter 1991**] The ambiguous logic of ambiguity, In: Proc. First CLIN-meeting, University of Utrecht, Oct. 1990. OTS-WP-CL-91-001
- [**van Deemter 1992**] Towards a Generalization of Anaphora. In Journal of Semantics 9: pp. 27-51.
- [**van Deemter 1994**] The sorites fallacy and the context-dependence of vague predicates. Ms., Institute for Perception Research. Submitted for publication in Kanazawa, Pinon and de Swart (Eds.), 'Quantifiers, deduction, and context' (tentative title).
- [**Dummett 1975**] Wang's Paradox, In: Synthese 30, pp.301-324.
- [**Goodman 1966**] The structure of appearance. Bobbs-Merrill Publ.
- [**Groenendijk and Stokhof 1991**] Dynamic Predicate Logic, In: Linguistics and Philosophy 14 (1991), pp. 39-100.
- [**Heim 1982**] The semantics of definite and indefinite Noun Phrases. Ph.D.Thesis at the Univ. of Mass., Amherst, Mass.
- [**Kamp 1981**] A theory of truth and semantic interpretation. In "Formal Methods in the Study of Language", J.Groenendijk, T.Janssen, M.Stokhof (Eds.) Mathematical Centre Tracts No.136.
- [**Klein 1980**] A semantics for positive and comparative adjectives. In Linguistics and Philosophy No.4, 1980.
- [**Mackie 1967**] "Fallacies", Entry in P.Edwards (Ed.) The Encyclopedia of Philosophy. Macmillan, New York.
- [**Poesio 1991**] Scope ambiguity and Inference, Technical Report 389, University of Rochester, New York, July 1991.
- [**Reyle 1992**] Dealing with scope ambiguities by underspecification: A first order calculus for unscoped representations. In: Proc. Amsterdam Colloquium 1991.
- [**Ross 1923**] Sir David Ross, Aristotle. Methuen & Co. LTD, London.
- [**Russell 1940**] B.Russell, An Inquiry into Meaning and Truth, George Allen and Unwin, London.
- [**Scott and Suppes 1958**] P.C Suppes and D.Scott, Foundational aspects of theories of measurement, In: Journal of Symbolic Logic 23, pp.113-128.
- [**Suppes and Zinnes 1963**] P.C Suppes and J.L Zinnes, Basic Measurement Theory, In: Handbook of Mathematical Psychology, Vol. 1. Wiley, New York, pp.1-76.

- [**Thomason 1974**] R.H.Thomason, Formal Philosophy; Selected Papers of Richard Montague. Yale University Press, New Haven and London.
- [**Veltman 1987**] Frank Veltman, Logische Analyse, Course material for the University of Amsterdam, Dept. of Philosophy.
- [**Westerstahl 1985**] Determiners and context sets, in: J.van Benthem and A.ter Meulen (Eds.)Generalized Quantifiers in Natural Language, Foris, GRASS-4.
- [**Zadeh 1975**] Fuzzy Logic and Approximate Reasoning. In: Synthese 30, pp.407-428.



University of Brighton

ITRI-00-31 **Generating Vague Descriptions**

Kees van Deemter

December, 2000

Also published in Proceedings of First Conference on Natural Language Generation (INLG2000), pp. 179-185.

Information Technology Research Institute Technical Report Series

ITRI, Univ. of Brighton, Lewes Road, Brighton BN2 4GJ, UK
TEL: +44 1273 642900 EMAIL: firstname.lastname@itri.brighton.ac.uk
FAX: +44 1273 642908 NET: <http://www.itri.brighton.ac.uk>

Generating Vague Descriptions

Kees van Deemter

ITRI, University of Brighton

Lewes Road, Watts Building

Brighton BN2 4GJ, United Kingdom

Kees.van.Deemter@itri.brighton.ac.uk

Abstract

This paper deals with the generation of definite (i.e., uniquely referring) descriptions containing semantically vague expressions ('large', 'small', etc.). Firstly, the paper proposes a semantic analysis of vague descriptions that does justice to the context-dependent meaning of the vague expressions in them. Secondly, the paper shows how this semantic analysis can be implemented using a modification of the Dale and Reiter (1995) algorithm for the generation of referring expressions. A notable feature of the new algorithm is that, unlike Dale and Reiter (1995), it covers plural as well as singular NPs. This algorithm has been implemented in an experimental NLG program using PROFIT. The paper concludes by formulating some pragmatic constraints that could allow a generator to choose between different semantically correct descriptions.

1 Introduction: Vague properties and Gradable Adjectives

Some properties can apply to an object to a greater or lesser degree. Such continuous, or *vague* properties, which can be expressed by, among other possibilities, gradable adjectives (e.g., 'small', 'large', e.g. Quirk et al. 1972 sections 5.5 and 5.39), pose a difficult challenge to existing semantic theories, theoretical as well as computational. The problems are caused partly by the extreme *context-dependence* of the expressions involved, and partly by the resistance of vague properties to discrete mathematical modeling (e.g., Synthese 1975, Pinkal 1995). The weight of these problems is increased by fact that vague expressions are ubiquitous in many domains.

The present paper demonstrates how a Natural Language Generation (NLG) program can be enabled to generate uniquely referring descriptions containing one gradable adjective, *despite* the vagueness of the adjective. Having presented a semantic analysis for such vague descriptions, we describe the semantic core of an NLG algorithm that has numerical data as input and vague (uniquely referring) descriptions as output.

One property setting our treatment of vagueness apart from that in other NLG programs (e.g. Goldberg 1994) is that it uses vague properties for an exact task, namely the ruling out of distractors in referring expressions (Dale and Reiter 1995). Another distinctive property is that our account allows the 'meaning' of vague expressions to be determined by a combination of linguistic context (i.e., the Common Noun following the adjective) and nonlinguistic context (i.e., the properties of the elements in the domain).

2 The Meaning of Vague Descriptions

Several different analyses are possible of what it means to be, for example, 'large': larger than average, larger than most, etc. But there is not necessarily just one correct analysis. Consider a domain of four mice, sized 2,5,7, and 10cm.¹ In this case, for example, one can speak of

1. *The large mouse*
(= *the one whose size is 10cm*), and of
2. *The two large mice*
(= *the two whose sizes are 7 and 10cm*).

Clearly, what it takes to be large has not been written in stone: the speaker may decide that 7cm is enough (as in (2)), or she may set the standards higher (as in (1)). A numeral (explicit, or implicit as in (1)), allows the reader to make inferences about the standards employed by the speaker.² More precisely, it appears that in a definite description, the absolute form of the adjective is semantically equivalent with the superlative form:

The n large mice – *The largest n mice*
The large mice – *The largest mice*
The large mouse – *The largest mouse*.

¹For simplicity, the adjectives involved will be assumed to be one-dimensional. Note that the degree of precision reflected by the units of measurement affects the descriptions generated, and even the objects (or sets) that *can* be described, since it determines which objects count as having the same size.

²Thanks are due to Matthew Stone for this observation.

This claim, which has been underpinned by a small experiment with human subjects (see Appendix), means that if a sentence containing one element of a pair is true then so is the corresponding sentence containing the other. There are bound to be differences between the two forms, but these will be taken to be of a *pragmatic* nature, having to do with felicity rather than truth (see section 5.2).

An important qualification must be made with respect to the analysis that we propose: to simplify matters, we assume that the *entire* domain of relevant individuals is available and that it is this domain alone which is taken into account when the adjective is applied. In the case of the example above, this means that all mice are irrelevant except the four that are mentioned: no other knowledge about the size of mice is assumed to be available.³

2.1 A Formal Semantics for Vague Descriptions

Let us be more precise. In our presentation, we will focus on the adjective ‘large’, without intended loss of generality. For simplicity, ‘large’ will be treated as semantically one-dimensional.

i. ‘**The largest n mouse/mice**’. Imagine a set C of contextually relevant animals. Then the NP ‘The largest n mouse/mice’ ($n > 0$) presupposes that there is an $S \subseteq C$ that contains n elements, all of which are mice, and such that (1) $C - S \neq \emptyset$ and (2) every mouse in $C - S$ is smaller than every mouse in S . If such a set S exists then the NP denotes S . The case where $n = 1$, realized as ‘The [Adj]-est [CN_{sg}]’ (*sg* = singular), falls out automatically.

ii. ‘**The largest mice**’. This account can be extended to cover cases of the form ‘The [Adj]-est [CN_{pl}]’ (*pl* = plural), where the numeral n is suppressed: these will be taken to be ambiguous between all expressions of the form ‘The [Adj]-est n [CN]’ where $n > 1$. Thus, in a domain where there are five mice, of sizes 4,4,4,5,6 cm, the only possible value of n is 2, causing the NP to denote the two mice of 5 and 6 cm size.

iii. ‘**The n large mouse/mice**’. We analyse ‘The n [Adj] [CN]’ ($n > 0$) as semantically equivalent with the corresponding NP of the form ‘The [Adj]-est n [CN]’. ‘The two large mice’, for example, denotes a set of two mice, each of which is bigger than all other contextually relevant mice.

iv. ‘**The large mice**’. Expressions of this form can be analysed as being of the form ‘The n [Adj] [CN]’ for some value of n . In other words, we will take

them to be ambiguous or unspecific – the difference will not matter for present purposes – between ‘The 2 large mice’, ‘The 3 large mice’, etc.

3 Generation of Crisp Descriptions

Generation of descriptions covers a number of tasks, one of which consists of finding a set L of properties which allows a reader to pick out a given unique individual or set of individuals. The state of the art is discussed in Dale and Reiter (1995), who present a computationally tractable algorithm for characterizing *individuals*. This algorithm (henceforth D&R), deals with vague properties, such as *size*, to some extent, but these are treated as if they were context-independent: always applying to the same sets of objects.

In many cases, generating vague descriptions involves generating a plural and no generally accepted account of the generation of *plural* descriptions has been advanced so far. In the following section, therefore, a generalization or D&R will be offered, called D&R_{Plur}, which focuses on *sets* of individuals. Characterization of an *individual* will fall out as a special case of the algorithm.

3.1 Plural Descriptions: Dale and Reiter generalized

The *properties* which form the basis of D&R_{Plur} are modeled as pairs of the form $\langle \text{Attribute}, \text{Value} \rangle$. In our presentation of the algorithm, we will focus on complete properties (i.e., $\langle \text{Attribute}, \text{Value} \rangle$ pairs) rather than attributes, as in Dale and Reiter (1995), since this facilitates the use of set-theoretic terminology. Suppose S is the ‘target’ set of individuals (i.e., the set of individuals to be characterized) and C (where $S \subseteq C$) is the set of individuals *from which* S is to be selected.⁴ Informally – and forgetting about the special treatment of head nouns – what happens is the following: The algorithm iterates through a list P in which the properties appear in order of ‘preference’; for each attribute, it checks whether specifying a value for that attribute would rule out at least one additional member of C ; if so, the attribute is added to L , with a suitable value. (The value can be optimized using some further constraints but these will be disregarded here.) Individuals that are ruled out by a property are removed from C . The process of expanding L and contracting C continues until $C = S$. The properties in L can be used by a linguistic realization module to produce NPs such as ‘The white mice’, ‘The white mice that are pregnant’, etc. Schematically, the algorithm goes as follows: (*Notation*: Given a property Q , the set of objects that have the property Q is denoted $[[Q]]$.)

³In other words, only *perceptual* context-dependence is taken into account, as opposed to *normative* or *functional* context-dependence Ebeling and Gelman (1994).

⁴Note that C *contains* r , unlike Dale and Reiter’s ‘contrast set’ C , which consists of those elements of the domain from which r is set apart.

```

 $L := \emptyset$  {#  $L$  is initialized to the empty set #}
For each  $P_i \in P$  do
  If  $S \subseteq [[P_i]]$  &  $C \not\subseteq [[P_i]]$  {# Adding  $P_i$ 
    would remove distractors from  $C$  #}
  then do
     $L := L \cup \{P_i\}$  {# Property  $P_i$  is added
      to  $L$  #}
     $C := C \cap [[P_i]]$  {# All elements outside
       $[[P_i]]$  are removed from  $C$  #}
    If  $C = S$  then Return  $L$  {# Success #}
  Return Failure {# All properties in  $P$  have been
    tested, yet  $C \neq S$  #}

```

‘Success’ means that the properties in L are sufficient to characterize S . Thus, $\bigcap \{[[P_i]] : P_i \in L\} = S$. The case in which S is a singleton set amounts to the generation of a singular description: $D\&R_{Plur}$ becomes equivalent to $D\&R$ (describing the individual r) when S in $D\&R_{Plur}$ is replaced by $\{r\}$.

$D\&R_{Plur}$ uses hill climbing: an increasingly good approximation of S is achieved with every contraction of C . Provided the initial C is finite, $D\&R_{Plur}$ finds a suitable L if there exists one. Each property is considered at most once, in order of ‘preference’. As a consequence, L can contain semantically redundant properties – causing the descriptions to become more natural, cf. Dale and Reiter 1995 – and the algorithm is polynomial in the cardinality of P .

Caveats. $D\&R_{Plur}$ does not allow a generator to include collective properties in a description, as in ‘the two *neighbouring* houses’, for example. Furthermore, $D\&R_{Plur}$ cannot be employed to generate conjoined NPs: It generates NPs like ‘the large white mouse’ but not ‘the black cat *and* the large white mouse’. From a general viewpoint of generating descriptions, this is an important limitation which is, moreover, difficult to overcome in a computationally tractable account. In the present context, however, the limitation is inessential, since what is crucial here is the interaction between an Adjective and a (possibly complex) Common Noun following it: in more complex constructs of the form ‘NP and the Adj CN’, only CN affects the meaning of Adj.⁵ There is no need for us to solve the harder problem of finding an efficient algorithm for generating NPs uniquely describing arbitrary sets of objects, but only the easier problem of doing this whenever a (nonconjunctive) NP of the form ‘the Adj CN’ is possible.

4 Generation of Vague Descriptions

We now turn our attention to extensions of $D\&R_{Plur}$ that generate descriptions containing the expression

⁵In ‘The elephant and the big mouse’, for example, the mouse does not have to be bigger than *any* elephant.

of one vague property. Case i of section 2.1, ‘The largest n chihuahuas’ will be discussed in some detail. All the others are minor variations.

Superlative adjectives. First, ‘The largest chihuahua’. We will assume that **size** is stored (in the KB that forms the input to the generator) as an attribute with exact numerical values. We will take them to be of the form n *cm*, where n is a positive natural number. For example,

```

type = dog, chihuahua
colour = black, blue, yellow
size = 1cm, 2cm, ..., 10cm.

```

With this KB as input, $D\&R$ allows us to generate NPs based on $L = \{\text{yellow, chihuahua, 9cm}\}$, for example, exploiting the number-valued attribute **size**. The result could be the NP ‘The 9cm yellow chihuahua’, for example. The challenge, however, is to generate superlatives like ‘The *largest* yellow chihuahua’ instead.

There are several ways in which this challenge may be answered. One possibility is to replace an exact value like *9cm*, in L , by a superlative value whenever all distractors happen to have a smaller size. The result would be a new list $L = \{\text{yellow, chihuahua, largest}_1\}$, where ‘largest₁’ is the property ‘being the unique largest element of C ’. This list can then be realized as a superlative NP. We will present a different approach that is more easily extended to plurals, given that a plural description like ‘the 2 large mice’ does not require the two mice to have the same size.

Suppose **size** is the only vague property in the KB. Vague properties are less ‘preferred’ (in the sense of section 3.1) than others (Krahmer and Theune 1999).⁶ As a result, when they are taken into consideration, all the other relevant properties are already in L . For instance, assume that this is the KB, and that the object to be described is c_4 :

```

type( $c_1, c_2, c_3, c_4$ ) = chihuahua
type( $p_5$ ) = poodle
size( $c_1$ ) = 3cm
size( $c_2$ ) = 5cm
size( $c_3$ ) = 8cm
size( $c_4$ ) = size( $p_5$ ) = 9cm

```

At this point, *inequalities* of the form $\text{size}(x) > n$ *cm* are added to the KB. For every value of the form n *cm* occurring in the *old* KB, all inequalities of the form $\text{size}(x) > n$ *cm* are added whose truth *follows* from the old KB. Inequalities are more

⁶Note, by contrast, that vague properties tend to be *realized first* (Greenbaum et al. 1985, Shaw and Hatzivassiloglou 1999). Surface realization, however, is not the topic of this paper.

preferred than equalities, while logically stronger inequalities are more preferred than logically weaker ones.⁷ Thus, in order of preference,

$$\begin{aligned} \text{size}(c_4), \text{size}(p_5) &> 8cm \\ \text{size}(c_3), \text{size}(c_4), \text{size}(p_5) &> 5cm \\ \text{size}(c_2), \text{size}(c_3), \text{size}(c_4), \text{size}(p_5) &> 3cm. \end{aligned}$$

The first property that makes it into L is ‘chihuahua’, which removes p_5 but not c_4 from the context set. (Result: $C = \{c_1, \dots, c_4\}$.) Now size is taken into account, and the property $\text{size}(x) > 8cm$ singles out c_4 . The resulting list is $L = \{\text{chihuahua}, > 8cm\}$. This implies that c_4 is *the only* chihuahua in the KB that is greater than $8cm$ and consequently, the property $\text{size}(x) > 8cm$ can be replaced, in L , by the property of ‘being larger than all other elements of C ’. The result is a list that may be written as $L = \{\text{chihuahua}, \text{largest}_1\}$, which can be employed to generate the description ‘the largest chihuahua’.

Plurals can be treated along analogous lines. Suppose, for example, the facts in the KB are the same as above and the target set S is $\{c_3, c_4\}$. Its two elements share the property $\text{size}(x) > 5cm$. This property is exploited by $D\&R_{Plur}$ to construct the list $L = \{\text{chihuahua}, > 5cm\}$. Analogous to the singular case, the inequality can be replaced by the property ‘being a set all of whose elements are larger than all other elements of C ’ (largest_n , for short), leading to NPs such as ‘the largest chihuahuas’. Optionally, the numeral may be included in the NP (‘the *two* largest chihuahuas’).

– ‘**Absolute**’ adjectives. The step from the superlative descriptions of case **i** to the analogous ‘absolute’ descriptions is a small one. Let us first turn to case **iii**, ‘The n large mouse/mice’. Assuming the correctness of the semantic analysis in section 2, the NP ‘The n large mouse/mice’ is semantically equivalent to the one discussed under **i**. Consequently, an obvious variant of the algorithm that was just described can be used for generating it. (For pragmatic issues, see section 5.2)

Finally, case **iv**, ‘The large mice’. Semantically, this does not introduce any new problems, since it is to case **iii** what case **ii** is to case **i**. According to the semantic analysis of section 2.1, ‘The large mice’ should be analysed just like ‘The n large mouse/mice’, except that the numeral n is suppressed. This means that a simplified version (i.e., without a cardinality check) of the algorithm that takes care of case **iii** will be sufficient to generate descriptions of this kind.

⁷E.g., $\text{size}(x) > m$ is preferred over $\text{size}(x) > n$ iff $m > n$. The preference for *inequalities* causes the generator to avoid the mentioning of measurements unless they are needed for the identification of the target object.

5 Conclusions and loose ends

We have shown how vague descriptions can be generated that make use of one vague property. We believe our account to be an instructive model of how the ‘raw data’ in a standard knowledge base can be presented in English expressions that have a very different structure. The numerical data that are the input to our algorithm, for example, take a very different form in the descriptions generated, and yet there is, in an interesting sense, no loss of information: a description has the same reference, whether it uses ‘exact’ information (‘The $3cm$ mouse’) or ‘vague’ information (‘The large mouse’).⁸

5.1 Limitations of the semantic analysis

Our proposal covers the generation of vague descriptions ‘*from absolute values*’, which is argued in Dale and Reiter (1995, section 5.1.2) to be most practically useful. When vague input is available (e.g., in the generation component of a Machine Translation system, or in WYSIWYM-style generation (Power and Scott 1998)), simpler methods can be used. Our own account is limited to the generation of definite descriptions and no obvious generalization to indefinite or quantified NPs exists. Other limitations include

- Descriptions that contain properties for other than individuating reasons (as when someone asks you to clean ‘the dirty table cloth’ when only one table cloth is in sight). This limitation is inherited directly from the $D\&R$ algorithm that our own algorithm extends.
- Descriptions containing more than one vague property, such as ‘The fat tall bookcase’, whose meaning is more radically unclear than that of definite descriptions containing only one vague term. (The bookcase may be neither the fattest nor the tallest, and it is not clear how the two dimensions are weighed.)
- Descriptions that rely on the salience of contextually available objects. Krahmer and Theune (1998) have shown that a contextually more adequate version of $D\&R$ can be obtained when degrees of salience are taken into account. Their account can be summarized as analysing ‘the black dog’ as denoting the unique most salient object in the domain that is both black and a dog. (Generalizations of this idea to $D\&R_{Plur}$ are conceivable but nontrivial since not all elements of the set S have to be equally salient.) Our own extensions of $D\&R$ (and perhaps $D\&R_{Plur}$) could be ‘contextualized’ if the

⁸This may be contrasted with the vague expressions generated in (Goldberg et al. 1994), where there is a real – and intended – loss of information. (E.g., ‘Heavy rain fell on Tuesday’, based on the information that the rainfall on Tuesday equalled $45mm$.)

role of salience is changed slightly: focusing on the singular case, the algorithm can, for example, be adapted to legislate that ‘the large(est) mouse’ denotes the largest of all those mice that are salient (according to some standard of salience). Note that this analysis predicts *ambiguity* when the largest mouse that is salient according to one standard is smaller than the largest mouse that is salient according to a more relaxed standard. Suppose, for example,

Salient (strict):

m1 (2cm), *m2* (5cm)

Salient (relaxed):

m1 (2cm), *m2* (5cm), *m3* (7cm);

then ‘the large(est) mouse’ may designate either *m2* or *m3* depending on the standards of salience used. What this illustrates is that salience and size are both vague properties, and that – as we have seen under point b – combining vague properties is a tricky business.

5.2 Pragmatics

An experimental PROFIT (Erbach 1995) program has implemented the algorithms described so far, generating different descriptions, each of which would allow a reader/hearer to identify an object or a set of objects. But of course, an NLG program has to do more than determine under what circumstances the use of a description leads to a true statement: an additional problem is to choose the most appropriate description from those that are semantically correct. This makes NLG an ideal setting for exploring issues that have plagued semanticists and philosophers when they studied the meaning of vague expressions, such as whether it can be *true* for two objects *x* and *y* which are indistinguishable in size that *x* is large and *y* is not (e.g. Synthese 1975). The present setting allows us to say that a statement of this kind may be true yet infelicitous (because they conflict with certain pragmatic constraints), and consequently to be avoided by a generator.

As for the choice between the ‘absolute’/superlative forms of the gradable adjective, we conjecture that the following constraints apply:

- C1. **Distinguishability.** Expressions of the form ‘The (*n*) large [CN]’ are infelicitous when the *smallest* element of the designated set *S* (named *x*) and the *largest* CN smaller than all elements of *S* (named *y*) are perceptually indistinguishable.
- C2. **Natural Grouping.** Expressions of the form ‘The (*n*) large [CN]’ are better avoided when the difference in size between *x* and *y* is ‘comparatively’ small. One way of making this precise is by requiring that the difference between *x* and

y cannot be smaller than that between either *x* or *y* and one of their neighbouring elements. Consider, for example, a domain consisting of mice that are 1cm, 1cm, 2cm, 7cm, 9cm and 9cm large; then C2 predicts that the only felicitous use of ‘the large mice’ refers to the largest three of the group.

- C3. **Minimality.** Otherwise, preference is given to the absolute form. This implies that when objects of only two sizes are present, and the difference is perceptually distinguishable, the absolute form is preferred over the superlative form. (For example, in a domain where there are two sizes of pills, we are much more likely to speak of ‘the large pills’ than of ‘the largest pills’.) In languages in which the superlative form is morphologically more complex than the absolute form, constraint C3 can be argued to follow from general Gricean principles (Grice 1975)).

As for the presence/absence of the *numeral*, we conjecture that the disambiguating numeral (as in ‘the *n* large mice’ or ‘the *n* largest mice’) can be omitted under two types of circumstances: (1) when any ambiguity resulting from different values of *n* is likely to be inconsequential (see Van Deemter and Peters (1996) for various perspectives); (2) when the domain allows only one ‘natural grouping’ (in the sense of C2). Before and until a more accurate version of the notion of a natural grouping is available (perhaps using fuzzy logic as in Zimmermann 1985), generators could be forbidden to omit the numeral, except in the case of a definite description in the singular.

Appendix: A Supporting Experiment

Human subjects were asked to judge the correctness of an utterance in a variety of situations. The experiment was set up to make plausible that, in a situation in which only perceptual context-dependence (see section 1) is relevant, expressions of the form ‘the *n* large CN’ can be used whenever certain simple conditions are fulfilled. Note that this ($\Rightarrow$) direction of the hypothesis is most directly relevant to the design of a generator, since we expect a generator to avoid mistakes rather than always use an expression whenever it is legitimate.

Hypothesis ($\Rightarrow$): In a situation in which the domain *D* represents the set of perceptually relevant objects, an expression of the form ‘the *n* large CN’ (where $n \geq 1$), can be used to refer to a set *S* of cardinality *n* if all objects in *D* – *S* are smaller than any of the *n*.

The experiment explores whether ‘the n large CN’ can refer to the n largest objects in the domain, whether or not this set of objects is held together by spatial position or other factors. Subjects were presented with 26 different situations, in each of which they had to say whether the sentence

The two high numbers appear in brackets

would constitute a *correct* utterance. The literal text of our question was:

Suppose you want to inform a hearer *which numbers in a given list appear in brackets*, where the hearer knows what the numbers are, but not which of them appear in brackets. For example, the hearer knows that the list is 1 2 1 7 7 1 1 3 1. You, as a speaker, know that only the two occurrences of the number 7 appear in brackets: 1 2 1 (7) (7) 1 1 3 1. Our question to you is: Would it be *correct* to convey this information by saying “The two high numbers appear in brackets”? (...).

All subjects were shown the 26 situations in the same, arbitrary, order. Each situation presented to the subjects contained a list of nine numbers. In 24 cases, the lists had the following form:

1 1 1 x y z 1 1 1,

where each of x, y, z equalled either 6 or 9, and where there were always two numbers among x, y, z that appear in brackets. In 16 out of 24 cases, the two bracketed positions are right next to each other, allowing us to test whether spatial contiguity plays any role. Subjects were presented with two additional situations, namely 1 1 1 (6) 1 (7) 1 1 1 and 1 1 1 (7) 1 (6) 1 1 1 in which, unlike the other 24 situations, the two largest numbers are not equally large, to make sure that the descriptions do not require the elements in their denotation to be similar in that respect. Our questions were presented via email to 30 third-year psychology/cognitive science students at the University of Durham, UK, all of whom were native speakers of English and ten of which responded.

Results: Eight subjects responded in exact conformance with the analysis of section 2.1, marking all and only those five sequences in which *the highest 2* numbers appeared in brackets. Only two subjects deviated slightly from this analysis: one of the two (subject 9) described all the expected situations as ‘correct’ plus the two cases in which two contiguous 6-es appeared in brackets; the other subject (subject 10) appears to have made a typing error, confusing

two subsequent situations in the experiment.⁹ All other responses of subjects 9 and 10 were as predicted. This means that all subjects except subject 10 were consistent with our ‘ $\Rightarrow$ ’ hypothesis. The experiment suggests that the converse of the hypothesis might also be true, in which it is claimed that expressions of the form ‘the n large CN’ *cannot* be employed to refer to the set S *unless* S consists of the n largest objects in D :

Hypothesis ($\Leftarrow$): In a situation in which the domain D represents the set of perceptually relevant objects, an expression of the form ‘the n large CN’ (where $n \geq 1$), can *only* be used to refer to a set S of cardinality n if all objects in $D - S$ are smaller than any of the n .

Again disregarding subject 10, eight out of nine subjects act in accordance with Hypothesis $\Leftarrow$, while only one appears to follow a somewhat more liberal rule. Given these findings, it appears to be safe to build a generator that implements *both* hypotheses, since none of our subjects would be likely to disagree with any of the descriptions generated by it.

This experiment has evident limitations. In particular, it has no bearing on the pragmatic constraints suggested in section 5.2, which might be tested in a follow-up experiment.

Acknowledgements

Thanks are due to: Richard Power for discussions and implementation; Emiel Krahmer, Ehud Reiter and Matthew Stone for comments on an earlier draft; Hua Cheng for observations on linguistic realization; Rosemary Stevenson and Paul Piwek for their help with the experiment described in the Appendix.

6 References

- Dale and Reiter 1995. R. Dale and E. Reiter. Computational Interpretations of the Gricean Maxims in the Generation of Referring Expressions. *Cognitive Science* 18: 233-263.
- Ebeling and Gelman 1994. Ebeling, K.S., Gelman S.A. 1994. Children’s use of context in interpreting “big” and “little”. *Child Development* 65(4): 1178-1192.
- Erbach 1995. G. Erbach. Web page on the PROFIT

⁹The situations that we suspect to have been confused are 1 1 1 (9) (9) 9 1 1 1, which was marked as correct (although, remarkably, none of the other ‘three nines’ situations were marked as correct) and 1 1 1 (9) (9) 6 1 1 1.

programming language, <http://coli.uni-sb.de/erbach/formal/profit/profit.html>.

- Goldberg et al. 1994. E. Goldberg, N. Driedger, and R. Kitteridge. Using Natural-Language Processing to Produce Weather Forecasts. *IEEE Expert* **9** no.2: 45-53.

- Greenbaum et al. 1985. "A Comprehensive Grammar of the English Language". Longman, Harlow, Essex.

- Grice 1975. P. Grice. Logic and Conversation. In P. Cole and J. Morgan (Eds.), "Syntax and Semantics: Vol 3, Speech Acts": 43-58. New York, Academic Press.

- Krahmer and Theune 1999. E. Krahmer and M. Theune. Generating Descriptions in Context. In R. Kibble and K. van Deemter (Eds.), Procs. of workshop The Generation of Nominal Expressions, associated with the 11th European Summer School in Logic, Language, and Information (ESSLLI'99).

- Pinkal 1995. M. Pinkal. "Logic and Lexicon". Oxford University Press.

- Power and Scott 1998. R. Power and D. Scott. Multilingual Authoring using Feedback Texts. In Proc. COLING/ACL, Montreal.

- Quirk et al. 1972. R. Quirk, S. Greenbaum, and G. Leech. "A Grammar of Contemporary English". Longman, Harlow, Essex.

- Shaw and Hatzivassiloglou 1999. Ordering Among Premodifiers. In Procs. of ACL99, Univ. Maryland.

- Synthese 1975. Special issue of the journal *Synthese* on semantic vagueness. *Synthese* **30**.

- Van Deemter and Peters 1996. K. van Deemter and S. Peters (Eds.) "Semantic Ambiguity and Underspecification". CSLI Publications, Stanford.

- Zimmermann 1985. H. J. Zimmermann. "Fuzzy Set Theory - and its Applications". Kluwer Academic Publishers, Boston/Dordrecht/Lancaster.

Finetuning NLG through experiments with human subjects: the case of vague descriptions

Kees van Deemter

ITRI, University of Brighton, Lewes Road, Brighton BN2 4GJ,
Kees.van.Deemter@itri.bton.ac.uk,
WWW home page: <http://www.itri.brighton.ac.uk/~Kees.van.Deemter>

Abstract. This discussion paper describes a sequence of experiments with human subjects aimed at finding out how an NLG system should choose between the different forms of a gradable adjective. This case study highlights some general questions that one faces when trying to base NLG systems on empirical evidence: one question is what task to set a subject so as to obtain the most useful information about that subject, another question has to do with differences between subjects.

1 Introduction

NLG systems express information in linguistic form. As these systems become more and more sophisticated, a key problem is choosing between the different ways in which the same information can be expressed. Although intuition and general linguistic principles can go a long way, the problem is essentially an empirical one that can only be resolved by empirical study.

How can we *find out* whether one way of expressing a given input is better than another? One option is to look at corpora. But large corpora, such as the BNC, tend not to come with extensive semantic annotation, and even if they do, the domains with which the texts in the corpus are concerned tend not to have been formalised. (In NLG *jargon*: We don't know the input from which the corpus was generated.) In many cases, this is a crucial limitation, necessitating an alternative method. An alternative is to 'generate' corpora by logging human users' verbal behaviour during a controlled experiment (e.g., [2], [7]), resulting in corpora whose domains can be explicitly modelled, and where the truth value of all the relevant domain facts is known. The present paper focusses mostly on a slightly simpler method¹, by letting human subjects choose between the different ways in which the input information may be expressed. Rather than producing a corpus of extended discourse/dialogue, such experiments can tell us, at least in principle, which expression is favoured by subjects: the fact that explicit judgements are

¹ An partial exception is our third experiment, described in section 5, which could be seen as generating a corpus.

obtained from subjects might be even thought an advantage. We will focus on issues raised by previous work on referring expressions, focussing on the question how one might inform the generator’s choice between the different forms of a gradable adjective in descriptions such as

- a. – ‘*the fat chihuahua*’ [base form]
- b. – ‘*the fatter chihuahua*’ [comparative form]
- c. – ‘*the fattest chihuahua*’ [superlative form]

After discussion of the experiments that we have done to inform this choice, we will discuss two problems: one problem has to do with experimental methodology (e.g., how does one find out what constitutes the optimal use of language), another problem has to do with differences between human subjects’ preferences.

2 Generating vague descriptions; research hypotheses

It was reported in [12] how the module of an NLG system that generates referring expressions can use numerical information in the input to the generator to produce referring expressions like a-c above. Generating such *vague descriptions* is nontrivial partly because properties like ‘fat’, ‘fatter’, ‘fattest’ are not simply true or false of an animal: the fattest chihuahua is not necessarily the fattest dog, for example, or the fattest animal. Unlike the numerical properties that measure the animals’ waistlines, properties like these are context dependent, and this has to be taken into account by the generation algorithm. If, for example, the knowledge base were to simply list *all* hippos and *no* chihuahuas as fat, then this would make it impossible to ever refer to a dog as ‘the fat chihuahua’, and this could even mean that the referent in question cannot be identified uniquely at all (i.e., if there happens to be no combination of other properties that distinguishes it from everything else in the domain). The problems surrounding gradable adjectives have been studied widely (e.g., [3], [10], [4]), but little of this is directly applicable to NLG, let alone to the choice between the different forms of the gradable adjective in descriptions such as a-c above.

In addition to presenting an algorithm for generating vague descriptions, [12] also described an informal experiment that supported the idea that for large classes of adjectives², the three forms a,b,c (above) of the vague description are *semantically equivalent* in the sense that, in those cases where they are used referentially, they are often interpreted as having the same reference.³ The experiment involved numbers; the outcome suggested strongly that, normally, ‘the

² *Evaluative* adjectives, such as ‘brilliant’ or ‘stupid’ are excluded [1].

³ Notable exceptions arise when the description refers anaphorically or in combination with other gradable adjectives. For example, if an entity has been introduced as ‘a large dog’ then this legitimises later references to it as ‘the large dog’ even where other, larger dogs are in evidence. The second type of exception arises when, for example, ‘the large dog’ is juxtaposed with another vague expression such as ‘an even larger dog’. For background and elaboration, see [13].

large number’ is the same as ‘the larger number’, which is the same as ‘the largest number’. What this experiment did *not* tell us is how to choose between a-c: for all we know, there might be situations in which the use of the superlative, for example, would be highly unnatural (even though it would tend to be interpreted correctly, as the experiment suggested). It is questions of the latter kind that the present paper addresses.

Intuition, supported by pilot experiments and informal study of corpora, suggested to us that there *are* situations where the base form is preferred over the superlative. We will henceforth refer to the difference in the relevant dimension between the intended referent and the nearest relevant object (the nearest ‘distractor’ removed by the adjective) as the *gap*. This gap itself can have different sizes. Roughly in line with Gricean principles, we hypothesised that when the gap is large, the base form is preferred; in these cases it would be unnecessarily elaborate to say ‘the largest number’. Conversely, when the gap is small, we expected the superlative to be preferred. (In such cases, to use the base form sounds like an exaggeration.)

Hypotheses:

1. Small gap $\Rightarrow$ Superlative > Base
2. Large gap $\Rightarrow$ Base > Superlative

(*How* large a gap has to be to switch from superlative to base form would have to be assessed separately.) Our first experiment did not take the comparative form into account, but this was added in later experiments. All experiments asked whether subjects considered themselves native speakers of English, fluent, or none of the two; no major differences between the three groups were found.

3 First experiment: picking a numeral

25 academics at the University of Brighton were shown pairs of numerals, one of which appeared in brackets. We chose numerals because numbers come in all possible ‘sizes’, so when assessing the size or height of a numeral, the only sensible comparison appears to be with other numerals that the experiment has presented. (Previously seen numerals can be ignored.) The gap between the two numbers was either large (in this case: gap size 7) or small (gap size 1). Eight different patterns involving the numbers 1, 2, 8, 9 were offered, depending on whether the gap was large or small, in both possible orders, and involving numbers that are somewhat higher (2 and 9 when the gap is large; 8 and 9 when the gap is small) or lower (1 and 8 when the gap is large; 1 and 2 when the gap is small). They were asked

Which of the following statements offers the most natural description of this sequence?

- ‘the large number appears in brackets’, or
- ‘the largest number appears in brackets’?

Thus, when confronted with the patterns [2 (9)], [(9) 2], [1 (8)], or [(8) 1], we expected to find that more subjects prefer the base form (because the gap is large); when confronted with the patterns [8 (9)], [(9) 8], [2 (1)], or [(1) 2], we expected to find that more subjects prefer the superlative.

Outcomes: The number of choices for each of the different types of adjective can be summarised in a contingency table:

	Base	Superlative
Large Gap	25	75
Small Gap	1	99

The table shows the gap size and adjective type to be non-independent of each other (Chi-square = 25.464 with $df = 1$ at $p = 0.001$.) More specifically, the experiment overwhelmingly supported Hypothesis 1. (Only one of the 25 subjects preferred the base form, and only on one occasion.) Equally clearly, however, the experiment failed to confirm Hypothesis 2: in fact, the superlative was chosen in the majority of cases even where the gap was large. *Post hoc*, the experiment suggests that base-form adjectives occur more often when the gap is large than when it is small. (Using the t-test, for example, this is highly significant, with $p = .008361$ and $df = 24$.) This pattern was confirmed by later experiments.

Note that there was a striking difference between two groups of subjects: all the people in the IT faculty, except one, consistently chose the superlative, while people at the School of Languages showed more variety, including some who chose the base form *if and only if* the gap was large. We will return to the issue of differences between groups of subjects in section 7.2.

4 Second experiment: picking a number from a sequence

Confronted with this failure to confirm Hypothesis 2, it is not difficult to find excuses: maybe the gap between the numbers in the experiment had not been large enough, or maybe the question that was asked of the subjects (containing a possibly theory-laden expression like ‘natural’) was unclear; also, subjects were not offered the possibility of choosing the comparative form (‘the larger number’), which is sometimes thought to be preferable when there are only two things to compare. To remove these obstacles, we did the following experiment, which varies on the same numerical theme.

Fourteen subjects (participants at the 2001 EACL conference in Toulouse) were shown sequences of two, three or four numbers, one or two of which appeared in brackets. There were 18 patterns, including, for example, [1 (59)] (large gap), [1 (59) (58)] (large gap), [55 (59) 54] (small gap), [(59) 55 (59) 55] (small gap), and [(59) 1 (59) 0] (large gap). Subjects were asked, in each of these cases, to say *Which description do you consider most likely to have been produced by a native speaker?*, where they could choose between

- ‘the large numbers appear in brackets’ or
- ‘the larger numbers appear in brackets’ or
- ‘the largest numbers appear in brackets’?

Based on the previous experiment, we refined our hypotheses. Given this subjects were offered a choice between three options, it seemed reasonable to compare the numbers of base forms with each of the two other forms separately:

Hypotheses:

- 1a. Small gap $\Rightarrow$ Comparative $>$ Base
- 1b. Small gap $\Rightarrow$ Superlative $>$ Base
- 2a. Large gap $\Rightarrow$ Base $>$ Comparative
- 2b. Large gap $\Rightarrow$ Base $>$ Superlative

Outcomes: The data can be summarised as follows:

	Base	Comparative	Superlative
Large Gap	43	37	42
Small Gap	2	65	59

Chi-square suggests a dependency between gap size and adjective type (at $p = 0.001$, $df = 2$, Chi-square = 47.851). Using a t-test, we see that, in line with the first experiment, Hypotheses 1a and 1b were confirmed but, crucially, Hypotheses 2a and 2b were not. The following table lists the number of choices for each of the three types of adjectives, for each of those 9 patterns of numbers where the gap was large. (The gap was small in patterns **b,c,d,e,h,i,k,m,q**.)⁴

[Large Gap]	Base	Comparative	Superlative
Pattern a	3	8	3
Pattern f	4	1	9
Pattern g	3	3	8
Pattern j	6	5	3
Pattern l	4	6	4
Pattern n	5	5	3
Pattern o	4	2	7
Pattern p	7	2	5
Pattern r	7	5	2

Focussing on Hypothesis 2a, we find that the two-tailed P value is not significant at 0.4774 ($t = 0.7276$, $df = 16$). Hypothesis 2b is even further off the mark with $P = 0.9120$ ($t = 0.1123$). As before, there were substantial differences between subjects, including some who used the base form *if and only if* the gap was small. Also as before, the base form was used much more often when the gap was large than when the gap was small.

⁴ The numbers for patterns **n** and **o** do not add up to 14 because one subject had failed to enter a clearly legible choice.

5 Third experiment: describing triangles

Going by the two experiments reported so far, it appears that superlatives are preferred, even in situations where the gap is very large. Again, one can think of excuses. *Asking* subjects about their preferences may not be the best way of assessing them: perhaps speakers cannot always tell how they use language (just like most of us cannot say how they walk). Also, numbers have their peculiarities: contrary to what was assumed in section 2, one might argue that some numbers (the number 0, for example, given that we only used non-negative numbers) *are* intrinsically small. For these reasons, we shifted the subject matter from numbers to geometrical shapes, and we designed the experiment in such a way that subjects could produce their own referring expressions (instead of being offered a forced choice). 34 Subjects (students of an HCI module at the University of Brighton) were shown a piece of paper showing two isosceles, and approximately equilateral triangles,⁵ one of which was circled. To encourage the use of size-related descriptions (like ‘the large triangle’, ‘the biggest of the two figures’, but unlike ‘the figure on the top left’), the instructions asked subjects to imagine themselves on the phone to someone holding a copy of the same sheet of paper, but not necessarily with the same orientation (e.g., possibly upside down). To discourage lengthy descriptions, the space for text was limited:

Q: Which triangle on this sheet
was circled?
A: The triangle

Each of the 34 subjects saw six pairs of triangles, one of which was smaller than the other. Triangles came in three sizes, so the ‘gap’ between the two triangles could either be large (when the smallest of the three was juxtaposed to the largest of the three) or small (when the mid-size triangle was juxtaposed to either the smallest one or the largest one). The same hypotheses were used as in the previous experiment (see section 4).

Results: Rather pleasingly, all except one subject used brief, size-related descriptions throughout. There were a few occurrences of descriptions like ‘the top triangle’, ‘the bottom triangle’, ‘the tiny triangle’, yet the data allow a remarkably simple summary:

- 12 subjects chose **superlatives** throughout (i.e., 6 items)
- 6 subjects chose **comparatives** throughout
- 3 subjects chose **Base form** throughout
- 6 subjects oscillated between Superlative and Comparative
- 6 subjects oscillated between Base form and some other form

⁵ Only two triangles were shown at a time, since the use of larger sequences in the second experiment had paid no obvious benefit. The lengths of the bases of the triangles were *5mm*, *8mm*, and *16mm* respectively.

As before, Hypotheses 2a and 2b were not confirmed (while Hypotheses 1a and 1b were confirmed). Once again, these findings appear to suggest that superlatives are preferred regardless of circumstance.

6 Fourth experiment: recognising human behaviour

It seemed plausible that the fact of being under scrutiny (being tested!) might have encouraged ‘pedantic’ behaviour in our subjects, causing this unexpected run of results. We therefore made one more attempt at getting at subjects’ normal behaviour, using a procedure reminiscent of the Turing test, which involves letting subjects judge whether a given pattern of behaviour is likely to have been produced by a real person as opposed to, for example, a computer or a linguist [11]. We showed results (or, in one case, fake results) from the previous experiment to 14 students of the School of Languages, at the University of Brighton. (Each subject in the third experiment gave answers to six questions of the form *Which triangle (...)?*, so each subject produced a sequence of six answers, e.g., **smallest large smaller smaller larger larger**.) We told our new subjects that each output was produced by one of our subjects (**a**, **b**, **c**, and **d**) in an earlier experiment:

a: smallest largest smallest smallest largest largest
b: smaller larger smaller smaller larger larger
c: small large smaller smaller larger larger
d: smaller larger small small large large

(Subjects were also shown the geometrical patterns on which the four sets of judgments were based.) Subjects **a** and **b** show completely uniform output; **c** mixes base forms and comparatives in a manner consistent with our Hypotheses; **d** (constructed by hand, because no such output had been produced by a real subject) reverses the patterns produced by **c**, using a base form where **c** had produced a comparative and conversely. We asked:

*Which of the four sequences of responses is most likely to have been produced by a real person, in your opinion? Circle one of the options: Subject **a**, **b**, **c**, **d**.*

In line with our earlier expectations, we hypothesised that **c** would be favoured over all others: that (i) **c** > **a**, that (ii) **c** > **b**, and that (iii) **c** > **d**

Results. The data are as follows:

- 9 subjects thought **b** was most likely
- 4 subjects thought **c** was most likely
- 1 subject thought **a** was most likely
- 0 subjects thought **d** was most likely

On the positive side, this suggests that a generator programmed to behave in accordance with our Hypotheses (leading to output as in **c**) would be judged

to be quite natural. But the more striking conclusion is surely that subjects rated completely uniform behaviour even more highly. The number of subjects partaking in this fourth experiment may be too small to warrant firm conclusions but the experiment clearly does little to confirm Hypothesis (ii).

7 Discussion

Two themes emerge: the question how to set up experiments in such a way that their outcomes reflect the linguistic preferences of the subjects; and the question how to cope with differences between subjects.

7.1 Eliciting ‘natural’ use of language

Our experiments suggest that vague descriptions in the base form are almost always dispreferred in comparison with comparatives and/or superlatives. To us at least, this seems counterintuitive.⁶ Otherwise, the experiments diverge. For example, superlatives came out on top of the third experiment, but were trumped by comparatives in the fourth. Note that such surprises are not easily explained along the lines of [5]. There, Oberlander invokes an asymmetry between speakers’ inclinations and hearer’s expectations, because speakers may be subject to factors that hearers cannot be aware of before the utterance is made (e.g., its propositional content). Our experiments, however, appear to make subjects aware of all the factors that writers/speakers would be aware of. (More specifically, the third and fourth experiments gave their subjects exactly the same information about the triangles and their properties.)

It is, of course, well known that questionnaire-type experiments are sensitive to subtle differences in formulations, and it is possible that our experiments have failed to capture spontaneous language use. One might therefore insist on naturalistic settings where subjects are given immersive real-world tasks to make them less self-conscious, and experiments of this kind (involving the narration of bedtime stories for children) are currently being prepared. The analysis of such experiments (involving unconstrained language use) tends to be difficult and time consuming, however. More importantly, one can always doubt the validity of the experiment by asking whether the task was immersive *enough*.

Perhaps most worryingly, our experiences cast doubt on previous findings: if questions of the form *Which description do you consider most likely to have been produced by a native speaker?* lead to outcomes that misrepresent writers/speakers’ inclinations, then how do we know that the answers to other,

⁶ As was explained in section 1, it would be no simple matter to test our hypotheses using a corpus like the BNC: even though base-form descriptions *are* much more frequent there than the two other types of descriptions, it would be no trivial matter to filter out attributive, anaphoric, and other uses which, collectively, are much more frequent than the referring expressions on which this paper has focussed.

similar questions that psycholinguists ask (e.g., ‘How would you continue this sentence?’ [9]; similarly for other types of questions relating to generation of referring expressions, e.g., [6]) are safe?

7.2 Honouring style differences

In recent work, Reiter and Sripada have argued that corpora can be a slippery basis on which to base the decisions of a generator [8]. Our own findings suggest a similar conclusion in connection with controlled experiments: one ignores style differences at one’s peril. The underlying point is a simple mathematical one: Suppose a given input can be expressed in three ways: A , B , and C , where there are two different sub-styles, say **1** and **2**. Suppose expression A is most frequent in none of the two styles, while occupying a very solid second place in both:

Substyle **1**: $B > A >> C$

Substyle **2**: $C > A >> B$

Overall: $A > B > C$

If the construction of a generator was based on overall frequencies, it would express the input as A , the most frequent expression overall. But this would cause the generator to be a poor imitator of anyone’s language use, since no existing substyle would be modelled. It would be better to choose either B or C since, at least, this would capture what is most frequent in one substyle.

8 Conclusion

When gathering data to inform natural language generation, controlled experiments are a natural complement to the study of corpora. This discussion paper has used a sequence of experiments to highlight two types of problems with the experimental approach, one of which (section 7.1) is specific to the experimental approach, while the other echoes similar problems in the study of language corpora (section 7.2).

9 Acknowledgments

Support for the TUNA project from the EPSRC under grant number GR/S13330 is gratefully acknowledged. Further thanks are due to Paul Piwek and to my TUNA colleagues: Albert Gatt, Emiel Krahmer, Richard Power, and Sebastian Varges, all of whom have contributed to this paper in some way or other.

10 References

1. Bierwisch, M.: The semantics of gradation. In M.Bierwisch and E.Lang (Eds.) *Dimensional Adjectives*. Berlin, Springer Verlag, (1989) pp.71-261.

2. Jordan, P.W.: Can Nominal Expressions Achieve Multiple Goals?: An Empirical Study. Procs. of ACL-2000, Hong Kong (2000).
3. Kamp, H.: Two theories about adjectives. In “Semantics for natural language”, ed. E. Keenan. Cambridge University Press (1975).
4. Klein E.: A semantics for positive and comparative adjectives. *Linguistics and Philosophy* **4** (1980).
5. Oberlander, J.: Do the right thing ... but expect the unexpected. Computational Linguistics, **24** (3), (1998) 501-507.
6. Pechmann, Th.: Incremental Speech Production and Referential Overspecification. *Linguistics* **27** (1989) 98-110.
7. Piwek, P. Cremers, A.: Dutch and English Demonstratives: A Comparison. Language Sciences **18** (3-4), (1996) pp. 835-851.
8. Reiter, E. Sripada, S.: Should Corpora Texts Be Gold Standards for NLG? In Proceedings of INLG-02 (2002) pages 97-104.
9. Stevenson, R.J., Crawley, R.A., and Kleinman, D.: Thematic roles, focus and the representation of events. *Language and Cognitive Processes* **9** (1994) 519-548.
10. Synthese. Special issue of the journal *Synthese* on semantic vagueness. *Synthese* **30** (1975).
11. Turing, A.: Computing Machinery and Intelligence. *Mind* **59**, No. 236 (1950) pp. 433-460.
12. van Deemter, K.: Generating Vague Descriptions. In Procs. of Int. Conf. on Natural Language Generation (INLG-2000) (2000) Mitzpe Ramon.
13. van Deemter, K.: Computational Generation of Vague Descriptions. Manuscript, ITRI, University of Brighton.