Generating Vague Descriptions

Kees van Deemter
University of Aberdeen

- Vagueness: a challenge for NLP/NLG

- Simple things first: Vague Descriptions

- Problems

- Vagueness: still a challenge

Seminar, Ohio State University, 13 Oct. 2006

Generating Vague Descriptions

Kees van Deemter
University of Aberdeen

- Vagueness: a challenge for NLP/NLG
- Simple things first: Vague Descriptions
- Problems
- Vagueness: still a challenge

Acknowledgement: Richard Power

Seminar, Ohio State University, 13 Oct. 2006

Linguists/formal semanticists
usually think in **Boolean** terms: Y/N

But many NL statements admit boundary cases, e.g.

Many *statements are ...*
Linguists **usually** *think ...*
... *is* **near** *...*
... *is only a* **girl**

Understand meaning & use.

Linguists/formal semanticists
usually think in **Boolean** terms: Y/N

But many NL statements admit boundary cases, e.g.

Many *statements are ...*
Linguists **usually** *think ...*
... *is* **near** ...
... *is only a* **girl**

Understand meaning & use.

For now: focus on gradable/degree adjectives

Gradable adjectives (Quirk et al. 1972)

1. can be intensified by *very*

2. can take comparative and superlative form.

E.g., *small*, *interesting*

– *smaller*, *smallest*
– *more interesting*, *most interesting*

Degree adjectives are **acquired early**

Degree adjectives are among child's first dozens of words (Peccei 1994)

'Perceptual' context-dependence is typically understood at 2 years (Ebeling & Gelman 1994)

Degree adjectives are **highly frequent**

BNC's **10** most frequent adjectives:

[last], [other], new, good, old,
great, high, small, [different], large

2. Vagueness is a challenge (for NLG)

Problem 1.

Equivalence of Observationally Indifferent Entities (EOI, Kamp 1981):

If $\boldsymbol{x}$ is big and
$\boldsymbol{x} \sim \boldsymbol{y}$
then $\boldsymbol{y}$ is big

EOI leads to Sorites paradox.

(Cf. mathematical induction)

Problem 2. *Just Noticable Differences*

JNDs are widely used in psychophysics

But *Help Elements* imply unlimited discrimination (Dummett 1975):

$$\underbrace{\overbrace{\boldsymbol{h} \qquad\qquad\qquad \boldsymbol{x}}^{\sim}\overbrace{\qquad\qquad \boldsymbol{y}}^{\sim}}_{\prec}$$

Problem 3. *Why do speakers use vagueness?*

Natural Language Generation (NLG) offers
a useful perspective:

Suppose a generator has numerical input
(e.g., from measurements) then
how/when/why should it produce vague output?

NB: This is NLG from **raw data**
(not from complex logical forms)

Vague expressions in NLG

1. FOG system (Goldberg et al. 1994)

In: quantitative data (E.g., rainfall $= 45mm$)
Out: *'Heavy rain (fell on Tuesday)'*

– Context-sensitivity not modeled.
– EOI not modeled: *heavy rain* $= 30 - 50mm$
– No rules for deciding whether to use vagueness

2. SUMTIME system: similar, though user can define boundaries (Reiter and Sripada 2002)

3. DYD system: context-dependence is modelled.
'A famous sonata' $\Rightarrow$ more recordings
than average sonata (Van Deemter and Odijk 1997)

Simple things first!
– start with numerical KB
– generate vague descriptions (cf. Pinkal 1979)

TYPE(x)=TYPE(y)=dog
SIZE(x)=43cm
SIZE(y)=30cm

$x \rightsquigarrow$ *The large dog*

In the vet's surgery, with two dogs:
'The large dog has stomach problems'

'The large dog' is effectively a **crisp** description

'The 43 cm. dog' would be
– unnecessarily detailed (information **overload**), and
– hard to interpret (**measurement** vs. **comparison**!)

3. Generation of Referring Expressions

- Some referents don't have commonly-known names (trains, particles, trees, furniture, ...)

- Generation of Referring Expressions (GRE): computer "invents" a description

- Lots of simplifying assumptions, e.g.
  – description needs to be **'distinguishing'**

- Such assumptions have made GRE one of the most advanced areas of NLG.

This talk:

- **Any** GRE algorithm can be extended to *vague* descriptions (although one algorithm works most smoothly).

  van Deemter (2006). "Generating Referring Expressions that Involve Gradable Properties". *Computational Linguistics* 32 (2), 2006.

Incremental Algorithm (Dale & Reiter 1995)

Target $= \boldsymbol{r}$

$\boldsymbol{\mathcal{P}}$ $=$ list of properties in order of preference

Algorithm delivers set $\boldsymbol{L} = \{\boldsymbol{P}_1, \ldots, \boldsymbol{P}_n\}$

such that $\boldsymbol{P}_1 \cap \ldots \cap \boldsymbol{P}_n \;=\; \{\boldsymbol{r}\}$.

- Start with first property in $\boldsymbol{\mathcal{P}}$
- If it's any good then include it
- Move on to next property in $\boldsymbol{\mathcal{P}}$
- And so on, until properties jointly characterise $\boldsymbol{r}$.

Incremental Algorithm: plural version (van Deemter 2000)

Replace $\{r\}$ by arbitrary set S:

$C :=$ Domain
$L := \emptyset$
For each $P_i \in P$ do
  If $S \subseteq [[P_i]]$ & $C \not\subseteq [[P_i]]$ then do
    $L := L \cup \{P_i\}$
    $C := C \cap [[P_i]]$
    If $C = S$ then Return L
Return *Failure*

Example of ‘crisp’ GRE, using IA:

Dog: $\{c_1, c_2, c_3, c_4, p_5\}$
White: $\{c_1, c_2, p_5\}$
Brown: $\{c_3, c_4\}$
British: $\{c_2, c_4, p_6\}$

To describe $\{r\} = \{c_4\}$,
$L = \{$Dog, Brown, British$\}$
‘The brown British dog’

To describe $S = \{c_2, c_4\}$,
$L = \{$Dog, British$\}$
‘The British dogs’

Limitation: Standard GRE algorithms do not treat context dependence

large chihuahua $\neq$ **large** dane

Can properties in KB be contextualised?

First idea: "when deciding whether large$(\boldsymbol{x})$, consider $\textsc{type}(\boldsymbol{x})$"

Obstacle 1: other properties in the description: Suppose $\textsc{type}(\boldsymbol{x}) =$ chihuahua. Then large chihuahua $\neq$ the large dog

Obstacle 2: numerals etc:
(See next slide)

Suppose Domain=$\{2cm, 5cm, 7cm, 9cm\}$

– *The large mouse* $= 9cm$
– *The two large mice* $= 7cm, 9cm$
– *The three large mice* $= 5cm, 7cm, 9cm$

This shows:

the (one) large $\neq$ the two large $\neq$ the three large

Root of the problem: KB cannot anticipate in what context a gradable concept will be used!

Context-dependent properties do not belong in a KB

Same observations affect formal theories

Theoretical accounts of vagueness focus on **comparison sets**:

“x is large with respect to A”: $\text{large}_A(x)$.

For example,

– larger than average over A
– larger than most in A, etc.

But standards are largely a matter of *fiat*

Kennedy 1999; van Deemter 2000; DeVault and Stone 2004

Experiments with human subjects

(very briefly!)

- By and large, readers understand
  'the ADJ' as 'the ADJ-est'
  E.g., 'the large mouse' = 'the 1 largest mouse'
- In fact, it's hard to find any differences in usage between

  'the ADJ' (base form)
  'the ADJ-er' (comparative form)
  'the ADJ-est' (superlative form)
- **Exceptions:** predicative and anaphoric uses

  'is the largest mouse in the house'

  'a large mouse' ... *'the large mouse'*

Eyetracking experiment (Sedivy et al. 1999)

- Materials: **1** target referent + **1** distractor

- E.g. Two cups of different sizes.
  Target = the tallest of the two,
  described as *'the tall cup'*

- Hearers spot target easily in all cases.
  (Low latency times.)

  Intrinsically small cup: first spotted after 554 ms.
  Intrinsically tall cup: first spotted after 538 ms.

- This suggests: **intrinsic size is hardly relevant**
  (It's all about comparison)

So: “Vague **descriptions**” are not so vague after all

Unclarity arises when numerals are omitted:

The large(st) mice: The largest $n =?$

Ambiguous between all values $n \geq 2$

sizes $3, 3, 5, 7, 8\ cm \;\Rightarrow$
The large(st) mice = $\{7, 8\}$ or $\{5, 7, 8\}$

sizes $3, 3, 3, 7, 8\ cm \;\Rightarrow$
The large(st) mice = $\{7, 8\}$

Caveat: All measurement is imprecise ...

‘Proof of concept’ system

VAGUE was implemented in SICSTUS PROLOG by Richard Power

Load KB, then ask VAGUE to describe a referent

Sample outputs: (including some stilted ones)

‘The largest one among the white mice’

‘The white mice whose size is 3cm’

‘The fast ones among the four largest ones among the white animals’

A *sketch* of how VAGUE works:

Incremental Algorithm modified: vague descriptions

Let *size* have numerical values, e.g,

$$\mathit{size} = 10cm, 20cm, \ldots$$

This allows us to generate a description
based on $L = \{$ yellow, chihuahua, $30cm$ $\}$
iff r is the only yellow chihuahua with size $30cm$

Inference step:
If $30cm$ is maximal among yellow chihuahuas
then replace $30cm$ by $\mathit{largest}_1$:

$L = \{$ yellow, chihuahua, largest_1 $\}$

Plural descriptions

The largest n chihuahuas

What if the n chihuahuas have different sizes?

Compile a new KB, with information of the form SIZE$(x) > a$ (for a in old KB). First generate

$$L = \{P_1, .., P_m, Q\}, \text{ where}$$
$$Q = \lambda x : \text{SIZE}(x) > a.$$

Inference step:

Replace Q by *largest*$_n$, where $n = \|S\|$

(the largest n objects in $P_1 \cap ... \cap P_m$)

Output: *'the largest n $P_1 \cap ... \cap P_m$'*

Example

'The largest $\mathbf{2}$ *chihuahuas'*

TYPE = chihuahua: $\boldsymbol{c1, c2, c3, c4}$

TYPE = poodle: $\boldsymbol{p5}$

SIZE = 30cm: $\boldsymbol{c1}$

SIZE = 50cm: $\boldsymbol{c2}$

SIZE = 80cm: $\boldsymbol{c3}$

SIZE = 90cm: $\boldsymbol{c4, p5}$

New KB

TYPE = chihuahua: $\boldsymbol{c1, c2, c3, c4}$
TYPE = poodle: $\boldsymbol{p5}$

SIZE $>$80cm: $\boldsymbol{c4, p5}$
SIZE $>$50cm: $\boldsymbol{c4, c3, p5}$
SIZE $>$30cm: $\boldsymbol{c4, c3, c2, p5}$

Suppose $\boldsymbol{S = \{c3, c4\}}$

Then $\boldsymbol{L = \{chihuahua, > 50cm\}}$

After revision: $\boldsymbol{L' = \{chihuahua, Largest_2\}}$

Combinations of vague adjectives:

TYPE= chihuahua: $\boldsymbol{c1, c2, c3, c4}$

TYPE= poodle: $\boldsymbol{p5}$

SIZE $>$80cm: $\boldsymbol{c4, p5}$

SIZE $>$50cm: $\boldsymbol{c4, c3, p5}$

SIZE $>$30cm: $\boldsymbol{c4, c3, c2, p5}$

WEIGHT$<$50kg: $\boldsymbol{c1, p5}$

WEIGHT$<$80kg: $\boldsymbol{c2, c1, p5}$

WEIGHT$<$90kg: $\boldsymbol{c3, c2, c1, p5}$

$$\boldsymbol{\{c3\} \;=\; \{Chihuahua, Largest_2, Lightest_1\}}$$

$$L = \{Chihuahua, Largest_2, Lightest_1\}$$

Possible wordings include

a. *'The chihuahua that's larger than 50cm but lighter than 90kg'*

b. *'The lightest one of the largest two chihuahuas'*

VAGUE chooses (**b.**) (never using inequalities)

Lots of other choices ...

and linguistics is of remarkably little help

1. When to use ‘exact’ measures, when adjectives?
– *‘The 3cm mouse’* vs.
– *‘The* small *mouse’*

2. When to omit the numeral?
– *‘The small(est) mice’* vs.
– *‘The* three *small(est) mice’*

3. Use superlatives, comparatives, or base forms?
– *‘The smallest mouse’* vs.
– *‘The smaller mouse’* vs.
– *‘The small mouse’*

Answers implemented in VAGUE:

1. ‘Exact’ measures when one measure suffices, while one *adjective* is not enough:

– *‘The large mouse’*
– *‘The 3cm mouse’* (neither largest nor smallest)

2. Only the last-added adjective goes without numeral

– *‘The tall ones among the 5 large mice’*

3. Base form iff the size of the gap is ‘sufficient’ (to be determined interactively)

– This is where Kamp’s EOI returns
– Many small experiments (Paper at INLG-2004)

Algorithm (using IA)

1. Replace equalities by inequalities in KB

2. Determine preference order
(Default: gradables come last)

3. Run IA_{Plur}

4. Apply **Inferences**

5. Perform Linguistic Realisation

Inferences:

– Replace combinations of inequalities by one exact Value

– Replace inequalities by properties that involve cardinality

– *(etc.)*

Algorithm (using any GRE algorithm G)

1. Replace equalities by inequalities in KB

2. Apply G

3a. Impose linear ordering on properties generated by G. *(Default: gradables come last)*

3b. Delete superfluous inequalities

4. Apply **Inferences**

5. Perform Linguistic Realisation

4. Possible extensions

1. Beyond degree adjectives

2. References with pointing

3. Degrees of salience

1. Beyond degree adjectives

Maybe *‘British’* and *‘brown’* are vague after all ..

1. Suppose x is closer to prototypical brown than y.
Then *‘the brown dog’* $= x$

2. x is *‘the British dog’*

OWNER(x)= John; OWNER(y)= Sarah
NATIONALITY(John)= NATIONALITY(Sarah)=UK
TYPE(x)= German shepherd
TYPE(y)= Border collie

3. Nouns like *‘girl’*, *‘academic’*, ...

Maybe vagueness is the rule not the exception
(e.g. Prototype Theory, Rosch et. al. 1976)

‘the academic’ $=$ element most typical of academics

2. Beyond text. Suppose `W` denotes women:

```
[W]                 [W]  [W]
 a    b    c    d    e    f    g



                 ^^^^^^^^^^
                 ||||||||||
```

'The woman' (+ pointing) $=$ `e`

Treatment: 'vaguify' Van der Sluis & Krahmer
(measure proximity to center of pointing)

3. Salience.

Some distractors are more salient than others

the red mouse = the most salient red mouse in D (Krahmer & Theune 2002).

Observe: Salience itself is gradable: an implicit gradable adjective. **Treatment:**

- Model degrees of salience numerically: $\textsc{salience}(x) \in \mathcal{N}$
- Test whether new property removes any distractors *that are at least as salient as* r.
- Stop when no distractors are left *that are at least as salient as* r.
- Do not *realise* salience in words.

5. A problem: Multidimensionality

1. Iterations of degree adjectives:

'The large hairy dog'

- The dog that is largest & hairiest?
- The largest dog that is hairier than average?
- The dog who scores highest on 'large & hairy'?

2. Recursive use of degree adjectives:
'The large dog in the small barn'

When vague properties are combined,
expect trouble!

3. We saw that **salience** acts
like an implicit gradable adjective

So: Problems in combination with other gradables:
‘the large mouse’ =?=

– the largest mouse that’s salient enough?
– the most salient mouse that’s large enough?

Even without overt vagueness:
‘the railway station’ (size or importance?)

Overall conclusion regarding GRE:

Referential **success** becomes hard to predict!

Conclusions

Some successes but problems remain:

1. Multidimensionality (Cf. Masthoff 2004)

2. Is *‘the large mice’* understood correctly?

(*Possible answer*: perceptual grouping (Thorisson 1994; Gatt 2006a, 2006b.)

3. What about other sources of vagueness? E.g. vague quantifiers: *many*, *few*.

(*Possible answer*: Moxey & Sanford’s work on speaker’s expectations and goals)

Relevant for adjectives and nouns too:
‘The large dog was barking’
‘This is no civil war’

4. Vague **descriptions** are exceptional. Elsewhere, **vagueness implies loss of information**

   Experiments: Vague expressions are understood differently by different people (e.g. Toogood 1984)

   Informativity and digestibility need to be balanced

   Compare: the term **'ease of use'** in HCI
   masks what an interface can achieve

4. Vague **descriptions** are exceptional. Elsewhere, **vagueness implies loss of information**

   Experiments: Vague expressions are understood differently by different people (e.g. Toogood 1984)

   Informativity and digestibility need to be balanced

   Compare: the term **‘ease of use’** in HCI masks what an interface can achieve

   More empirical work is needed!