Modeling the Semantics of “Tall”

A probabilistic approach
“Tall” is a gradable adjective which is vague and context dependent.

Vagueness is a feature belonging to some linguistic items which is distinguished from, ambiguity, generality, and indefiniteness.

Williamson: ... ‘vague is vague’. Its everyday meaning is so diffuse that it can be the object of only the most desultory investigation.

A borderline case is a case where it is uncertain whether or not a word applies. A borderline borderline case is a case where it is uncertain whether or not it is uncertain whether a word applies.

Tolerance: There exists a degree of change on some relevant scale too small to make any seeming difference in the application of the term.

A vague term is one whose application is subject to a principal of Tolerance. And for which there are borderline cases, borderline borderline cases, etc.
Is this stick figure tall?
Is this stick figure fat?
The conditions under which an utterance is true are considered to be a fundamental component of the semantics of language.

Vagueness presents the issue of how to represent the truth conditions for an utterance which contains predicates of uncertain application.

A 7 ft. person is tall.
If a 7 ft. person is tall, a 6’11” person is tall.
If a 6’11” person is tall, a 6’10” person is tall.

A 4 ft. person is tall.
Max Black: Vagueness as Consistency of application

\[
\lim_{p_k \to \infty} \frac{m_k}{n_k}
\]

For a American male \( x \) of \( k'' \) height,

\[\text{D}_{x_k}T = \text{A Person judged } x_k \text{ tall.}\]

\[\text{D}_{x_k}\sim T = \text{A Person judged } x_k \text{ not tall.}\]

\[m_k = \text{Number of } \text{D}_{x_k}T, \quad n_k = \text{number of } \text{D}_{x_k}\sim T\]

\[p_k = m_k + n_k\]
Discrete-ish perception of events which vary along a continuum

![Graph showing the percentage of "pa" and "ba" responses varying with voice onset time (ms)].
Sigmoid Semantics for “tall” in the Context of American Males

Probability that "S_k is tall" is true

\[ \frac{1}{1 + e^{72-x}} \]

Parameters:
- a = 1
- b = 0
- c = 1
- d = 72

Height in inches

Tall Candidates
Vagueness as a Function of the Parameter $C$

- $c = e$
- $c = 1$
- $c = 0.25$
- $c = 0.000001$
General Sigmoid Function to Model Semantics of Vagueness and Context Dependency

\[
\frac{a - b}{1 + \left(\frac{e}{c}\right)d-x} + b
\]

Parameters:
- \(a\) = Upper asymptote
- \(b\) = Lower asymptote
- \(d\) = Symmetry point

Maximum Slope = \[
\frac{1}{4} (a - b) \left(1 + \log\left[\frac{1}{c}\right]\right)
\]

Set \(a = 1\), \(b = 0\), and \(d\) the Symmetry point in the middle of the range of indeterminate application.

**Hypothesis:**
Range of indeterminate application is related to the distribution of the universe of objects along a scale in the domain of application given the context.

"Tall" in the context of American males: Set \(d = 1\) standard deviation above mean.
“Tall” in the context of American Male

\[ P(T(S_k)) = \text{Probability that “} S_k \text{ is tall” is true.} \]

<table>
<thead>
<tr>
<th>Name</th>
<th>Height</th>
<th>( P(T(S_k)) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_0 )</td>
<td>5 ft. 4 in.</td>
<td>0.00033535</td>
</tr>
<tr>
<td>( S_1 )</td>
<td>5 ft. 5 in.</td>
<td>0.000911051</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>5 ft. 6 in.</td>
<td>0.00247262</td>
</tr>
<tr>
<td>( S_3 )</td>
<td>5 ft. 7 in.</td>
<td>0.00669285</td>
</tr>
<tr>
<td>( S_4 )</td>
<td>5 ft. 8 in.</td>
<td>0.0179862</td>
</tr>
<tr>
<td>( S_5 )</td>
<td>5 ft. 9 in.</td>
<td>0.0474259</td>
</tr>
<tr>
<td>( S_6 )</td>
<td>5 ft. 10 in.</td>
<td>0.119203</td>
</tr>
<tr>
<td>( S_7 )</td>
<td>5 ft. 11 in.</td>
<td>0.268941</td>
</tr>
<tr>
<td>( S_8 )</td>
<td>6 ft.</td>
<td>0.5</td>
</tr>
<tr>
<td>( S_9 )</td>
<td>6 ft. 1 in.</td>
<td>0.731059</td>
</tr>
<tr>
<td>( S_{10} )</td>
<td>6 ft. 2 in.</td>
<td>0.880797</td>
</tr>
<tr>
<td>( S_{11} )</td>
<td>6 ft. 3 in.</td>
<td>0.952574</td>
</tr>
<tr>
<td>( S_{12} )</td>
<td>6 ft. 4 in.</td>
<td>0.982014</td>
</tr>
<tr>
<td>( S_{13} )</td>
<td>6 ft. 5 in.</td>
<td>0.993307</td>
</tr>
<tr>
<td>( S_{14} )</td>
<td>6 ft. 6 in.</td>
<td>0.997527</td>
</tr>
<tr>
<td>( S_{15} )</td>
<td>6 ft. 7 in.</td>
<td>0.999089</td>
</tr>
<tr>
<td>( S_{16} )</td>
<td>6 ft. 8 in.</td>
<td>0.999665</td>
</tr>
</tbody>
</table>

What information do we get when someone says “John is tall.”? We can interpret the sigmoid function as a cumulative distribution function. \( P(\text{John’s height} \leq x) \)

\[
\frac{d}{dx} \left( \frac{1}{1 + e^{72-x}} \right) = \frac{e^{72+x}}{(e^{72} + e^x)^2} \quad \int_{-\infty}^{\infty} \frac{e^{72+x}}{(e^{72} + e^x)^2} \, dx = 1
\]
Adam’s phrases Kolgomorov’s axioms in terms of propositional logic:
Let φ and ψ be propositions. Then:

(1) $0 \leq P(\phi) \leq 1$.
(2) If φ is logically true then $P(\phi) = 1$.
1) If φ logically implies ψ then $P(\phi) \leq P(\psi)$.
2) If φ and ψ are logically inconsistent then $P(\phi \& \psi) = P(\phi) + P(\psi)$

The uncertainty of a proposition φ, is written $U(\phi)$.
(5) $U(\phi) = 1 - P(\phi) = P(\sim\phi)$.

From these axioms and classical first order logic follows a generalization of classical validity:

(6) The uncertainty of the conclusion of a valid inference cannot exceed the sum of the uncertainties of the premises.
A Short Proof

We will assume that for any propositions \( \Phi \), and \( \Psi \),
\[
P(\Phi \rightarrow \Psi) = P(\Psi | \Phi).
\]
Further it is fair to assume that
\[
P(T(S_k)) \leq P(T(S_{k+1})).
\]
Also,

(i) \[
P(T(S_{k+1}) | T(S_k)) = 1.
\]

By the definition of conditional probability,

(ii) \[
P(T(S_{k+1}) \& T(S_k)) = P(T(S_k)) P(T(S_{k+1}) | T(S_k)).
\]

(iii) \[
P(T(S_k) \& T(S_{k+1})) = P(T(S_{k+1})) P(T(S_k) | T(S_{k+1})).
\]

By (i) and (ii),

(iv) \[
P(T(S_{k+1}) \& T(S_k)) = P(T(S_k)) \times 1 = P(T(S_k))
\]

Substituting the right side of (iv) into the left side of (iii),

(v) \[
P(T(S_k)) = P(T(S_{k+1})) P(T(S_k) | T(S_{k+1})).
\]

Dividing both sides of (v) by \( P(T(S_{k+1})) \),

(vi) \[
P(T(S_k) | T(S_{k+1})) = \frac{P(T(S_k))}{P(T(S_{k+1}))}.
\]
Two Challenges: Composition and Generalization

- What about predicates like “fat” or “red” which don’t vary along a single immediately apparent continuum?
- How would this sort of lexical semantics work with composition?
- We can imagine how sets interact but how are the sigmoid functions of two vague predicates going to interact?
- Idea: “Very tall” shifts the function one standard deviation forward. “Sort of tall” shifts the function one standard deviation backward.