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.

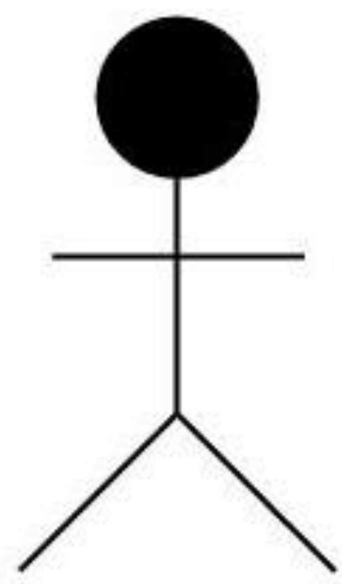
- Williamson: *…‘vague is vague’. Its everyday meaning is so diffuse that it can be the object of only the most desultory investigation*

- 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.

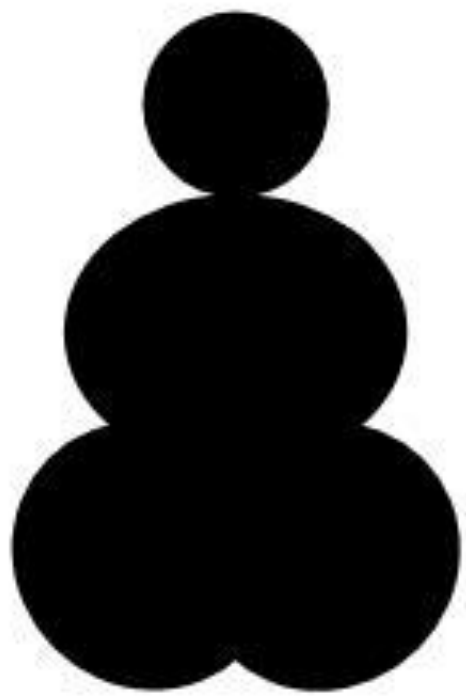
- Tolerance: There exists a degree of change on some relevant scale too small to make any seeming difference in the application of the term.

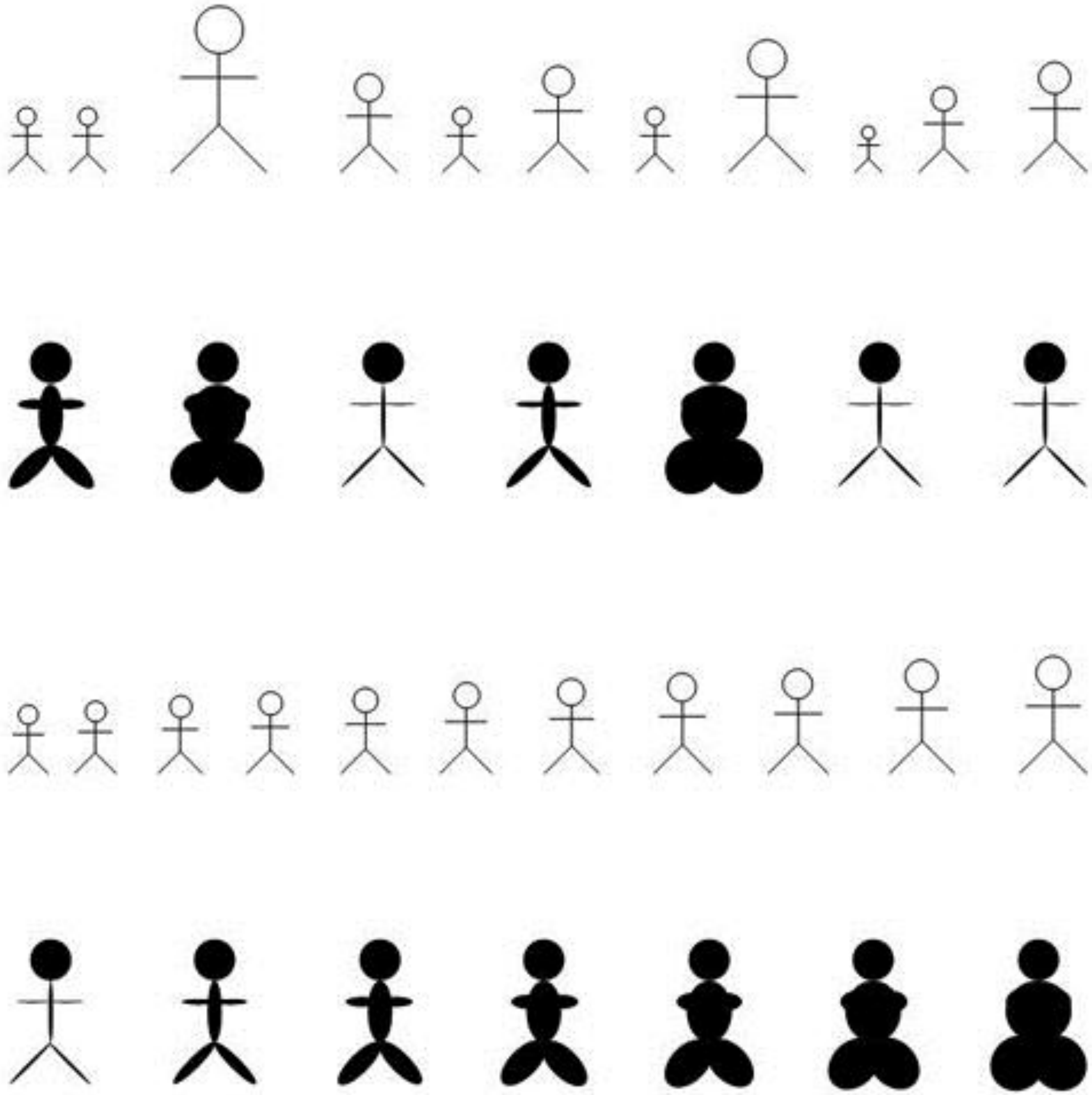
- A vague term is one whose application is subject to a principal of *Tolerance*
And for which there are *borderline cases,* borderline borderline cases, etc.

Is this stick figure tall?



Is this stick figure fat?

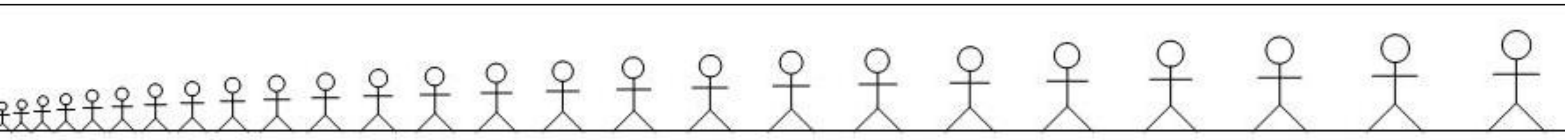




Issues With Vagueness

- ❖ 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.

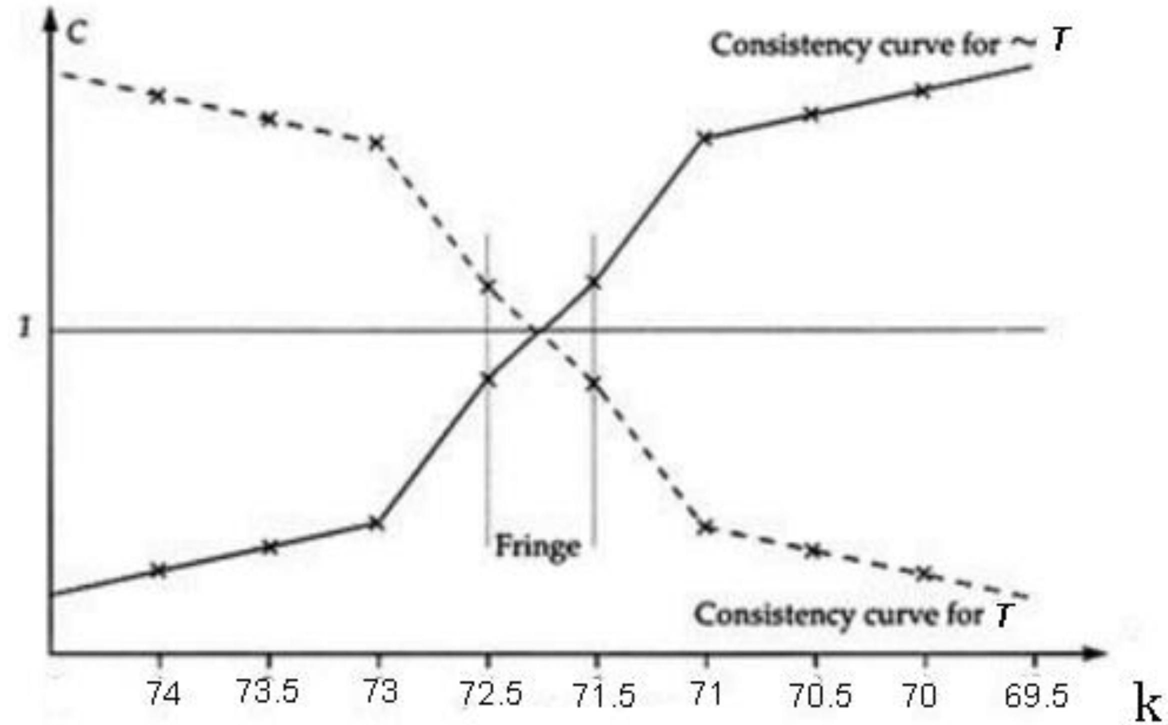


Sorites

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,

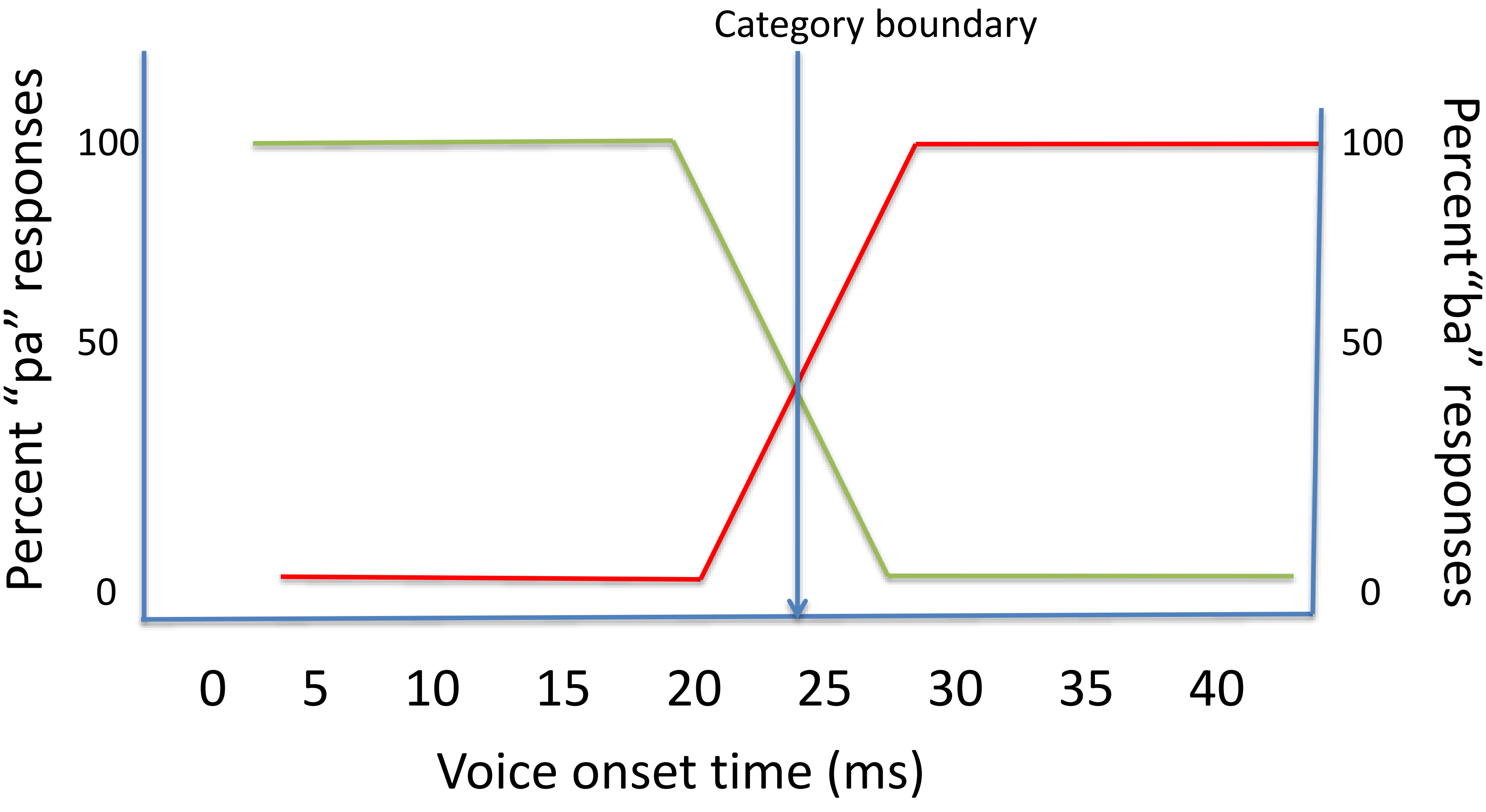
Dx_kT = A Person judged x_k tall.

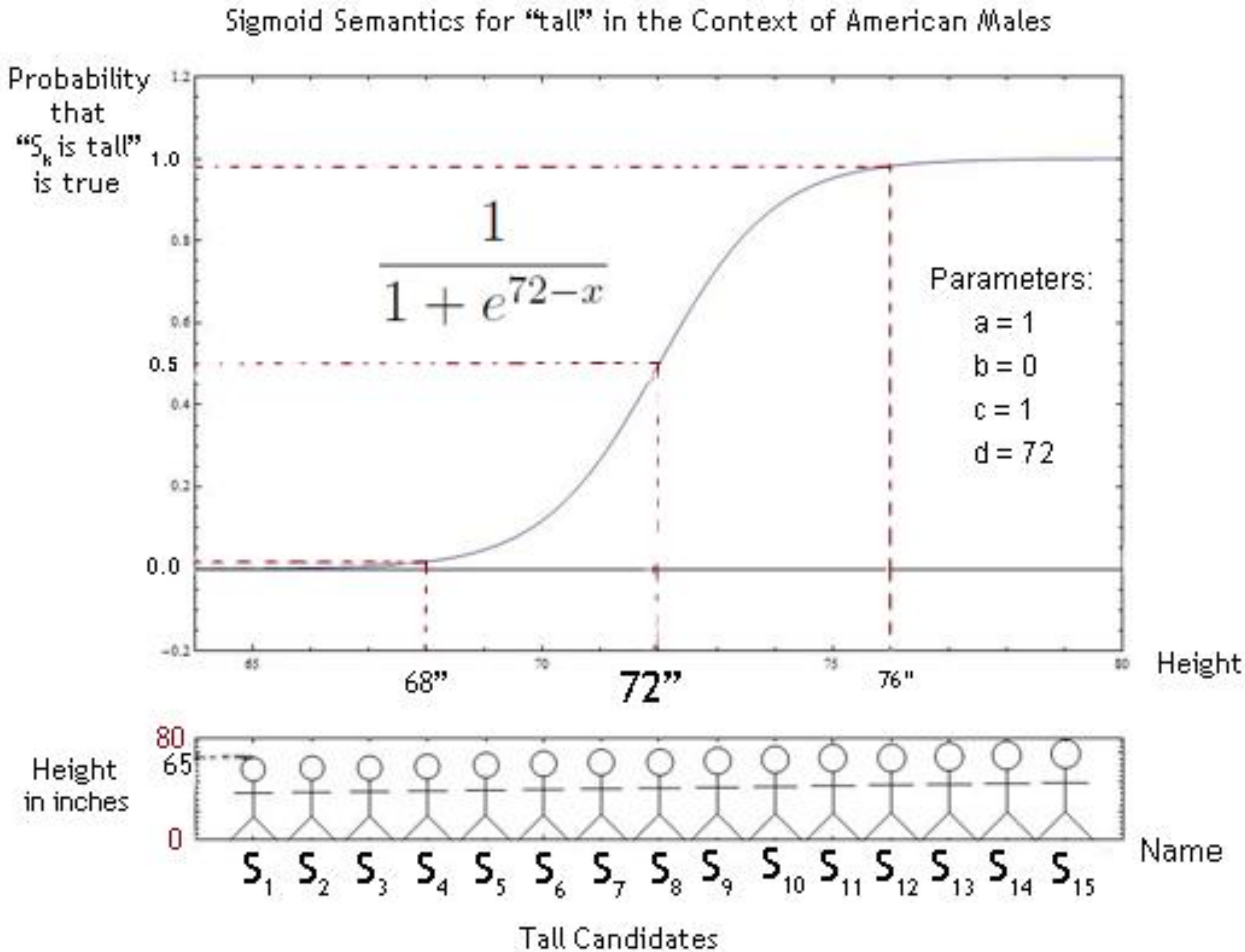
$Dx_k{\sim}T$ = A Person judged x_k not tall.

m_k = Number of Dx_kT , n_k = number of $Dx_k{\sim}T$

$p_k = m_k + n_k$

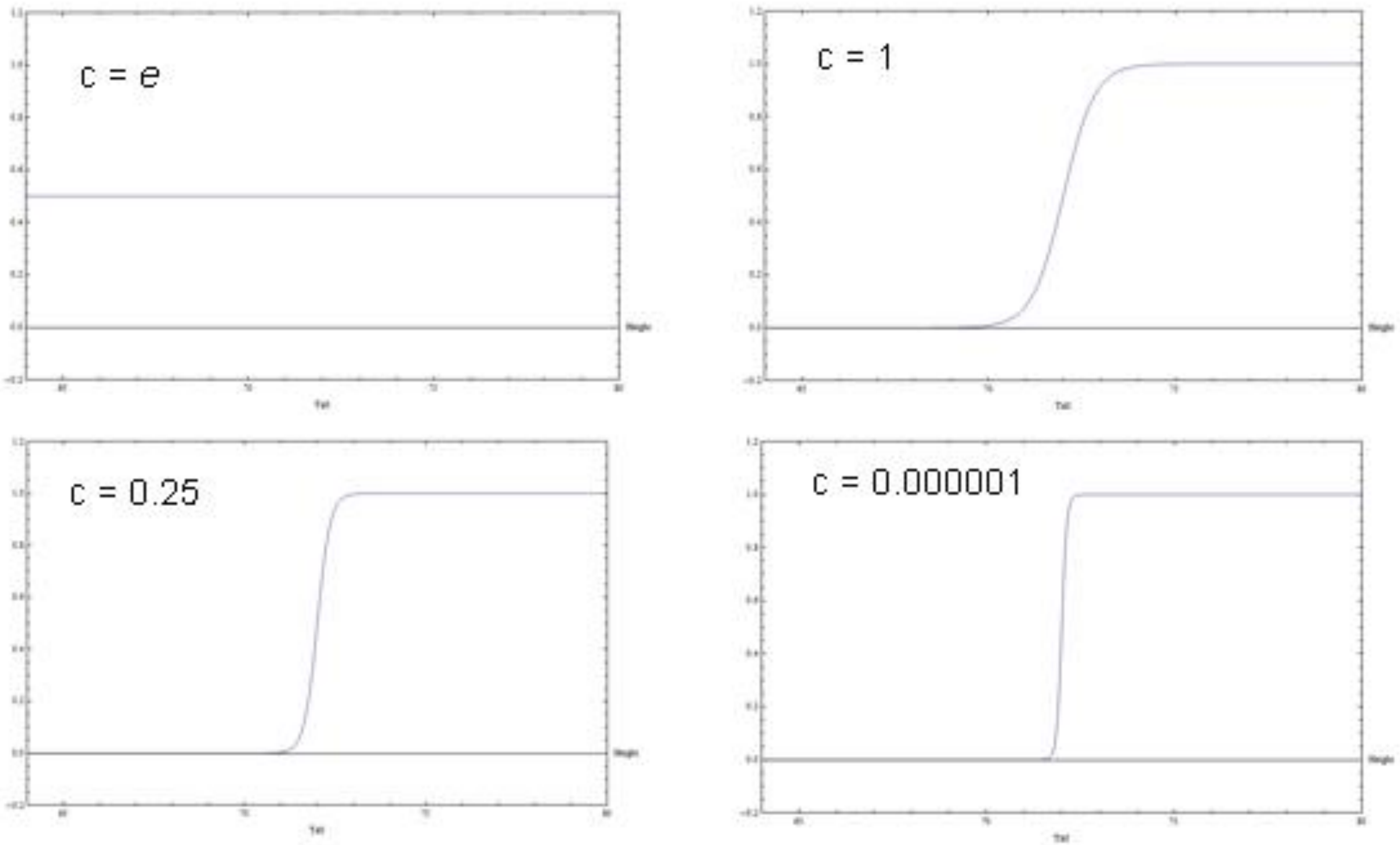
Discrete-ish perception of events which vary along a continuum





Sigmoid Semantics for "tall" in the Context of American Males
Probability that "S_k is tall" is true
1.0
0.5
0.0
$$\frac{1}{1+e^{72-x}}$$
Parameters:
a = 1
b = 0
c = 1
d = 72
68"
72"
76"
Height
80
65
Height in inches
0
Name
S1 S2 S3 S4 S5 S6 S7 S8 S9 S10 S11 S12 S13 S14 S15
Tall Candidates

Vagueness as a Function of the Parameter C



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 =

$$\frac{1}{4}\,(a-b)\,\left(1+\mathrm{Log}\left[\frac{1}{c}\right]\right)$$

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

$P(T(S_k))$ = Probability that “S_k is tall” is true.

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

What information do we get when someone says “John is tall.” ? We can interpret the sigmoid function as a cumulative distribution function.

P(John’s height ≤ x)

$$\frac{d}{dx}\left(\frac{1}{1+e^{72-x}}\right) = \frac{e^{72+x}}{(e^{72}+e^{x})^2}$$

$$\int_{-\infty}^{\infty} \frac{e^{72+x}}{(e^{72}+e^{x})^2}\,\mathrm{d}x = 1$$

Ernest Adam’s Probability Logic (1998)

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 ϕ, and ψ, $P(\phi \rightarrow \psi) = P(\psi \mid \phi)$. Further it is fair to assume that $P(T(S_k)) \leq P(T(S_{k+1}))$. Also,

(i) $P(T(S_{k+1}) \mid T(S_k)) = 1.$

By the definition of conditional probability,

(ii) $P(T(S_{k+1}) \;\&\; T(S_k)) = P(T(S_k)) \; P(T(S_{k+1}) \mid T(S_k))$.

(iii) $P(T(S_k) \;\&\; T(S_{k+1})) = P(T(S_{k+1})) \; P(T(S_k) \mid T(S_{k+1}))$.

By (i) and (ii),

(iv) $P(T(S_{k+1}) \;\&\; T(S_k)) = P(T(S_k)) * 1 = P(T(S_k))$

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

(v) $P(T(S_k)) = P(T(S_{k+1})) \; P(T(S_k) \mid T(S_{k+1}))$.

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

(vi) $P(T(S_k) \mid T(S_{k+1})) = \dfrac{P(T(Sk))}{P(T(Sk_{+1}))}$.

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.