

NTS: $P(s_c | \bar{V}) > P(s_q | \bar{V})$
 [deaccented "has" is a cue for competitor]

$$\Leftrightarrow \tau_c q_{\bar{V}} > \tau_q p_{\bar{V}}$$

$$\Leftrightarrow \frac{q_{\bar{V}}}{p_{\bar{V}}} > \frac{\tau_q}{\tau_c}$$

any bias towards competitor must be counterbalanced by the evidence $q_{\bar{V}}/p_{\bar{V}} > 1$.

[we could naturally expect this to be the case] even if $\tau_q \neq \tau_c$ because of $[q_{\bar{V}} > p_{\bar{V}}]$

So assume:
 $\tau_q = \tau_c + \epsilon_r$
 $q_{\bar{V}} = p_{\bar{V}} + \epsilon_{\bar{V}}$ with $0 \leq \epsilon_r < \epsilon_{\bar{V}}$

Then:

$$\frac{p_{\bar{V}} + \epsilon_{\bar{V}}}{p_{\bar{V}}} > \frac{\tau_c + \epsilon_r}{\tau_c}$$

$$\Leftrightarrow (p_{\bar{V}} + \epsilon_{\bar{V}}) \cdot \tau_c > p_{\bar{V}} (\tau_c + \epsilon_r)$$

$$\Leftrightarrow p_{\bar{V}} \tau_c + \epsilon_{\bar{V}} \tau_c > p_{\bar{V}} \tau_c + p_{\bar{V}} \epsilon_r$$

$$\Leftrightarrow \epsilon_{\bar{V}} \tau_c > p_{\bar{V}} \epsilon_r$$

$$\Leftrightarrow \frac{\epsilon_{\bar{V}}}{\epsilon_r} > \frac{p_{\bar{V}}}{\tau_c}$$

true if $\epsilon_r < \epsilon_{\bar{V}}$ & $p_{\bar{V}} < \tau_c$
 both are natural assumptions
 also assumed here: $q_{\bar{V}} > p_{\bar{V}}$ & $\tau_q \geq \tau_c$ ✓

maybe assume that: $\tau_q = \tau_c + \epsilon_r$
 with $0 \leq \epsilon_r$ small & $p_{\bar{V}} + \epsilon_{\bar{V}} = q_{\bar{V}}$
 with $0 \leq \epsilon_r < \epsilon_{\bar{V}}$.

NTS: $P(s_q | V) > P(s_c | \bar{V})$
 [proof exists (?) for flat prior $\tau_q = \tau_c$ and identical string likelihood $p_{\bar{V}} = q_{\bar{V}}$]

$$\Leftrightarrow \frac{\tau_q p_V}{\tau_q p_V + \tau_c q_V} > \frac{\tau_c q_{\bar{V}}}{\tau_q p_{\bar{V}} + \tau_c q_{\bar{V}}}$$

$$\Leftrightarrow \frac{\tau_q p_V}{\tau_q p_V + \tau_c q_V} > \frac{\tau_c (1 - q_V - q_{\bar{V}})}{\tau_q (1 - p_V - p_{\bar{V}}) + \tau_c (1 - q_V - q_{\bar{V}})}$$

$$\Leftrightarrow \dots > \frac{\tau_c - \tau_c q_V - \tau_c q_{\bar{V}}}{\tau_q - \tau_q p_V - \tau_q p_{\bar{V}} + \tau_c - \tau_c q_V - \tau_c q_{\bar{V}}}$$

$$\Leftrightarrow \tau_q p_V (\tau_q - \tau_q p_V - \tau_q p_{\bar{V}} + \tau_c - \tau_c q_V - \tau_c q_{\bar{V}}) > (\tau_c - \tau_c q_V - \tau_c q_{\bar{V}}) \cdot (\tau_q p_V + \tau_c q_V)$$

$$\Leftrightarrow \tau_q^2 p_V - \tau_q^2 p_V^2 - \tau_q^2 p_V p_{\bar{V}} + \tau_q \tau_c p_V - \tau_q \tau_c p_V q_V - \tau_q \tau_c p_V q_{\bar{V}} > \tau_q \tau_c p_V + \tau_q \tau_c q_V - \tau_q \tau_c p_V q_V - \tau_c^2 q_V^2 - \tau_q \tau_c p_V q_{\bar{V}} - \tau_c^2 q_V q_{\bar{V}}$$

$$\Leftrightarrow \tau_q^2 p_V - \tau_q^2 p_V^2 - \tau_q^2 p_V p_{\bar{V}} > \tau_q \tau_c p_V - \tau_c^2 q_V^2 - \tau_c^2 q_V q_{\bar{V}}$$

$$\Leftrightarrow (\tau_q^2 - \tau_q \tau_c) p_V - \tau_q^2 p_V^2 - \tau_q^2 p_V p_{\bar{V}} > -\tau_c^2 q_V^2 - \tau_c^2 q_V q_{\bar{V}}$$

$$\Leftrightarrow (\tau_q^2 - \tau_q \tau_c + \tau_q \epsilon_r) p_V - \tau_q^2 p_V^2 - \tau_q^2 p_V p_{\bar{V}} > -(\tau_q^2 - 2\tau_q \epsilon_r + \epsilon_r^2) (q_V - q_V q_{\bar{V}})$$

$$\Leftrightarrow \tau_q \epsilon_r p_V - \tau_q^2 p_V^2 - \tau_q^2 p_V p_{\bar{V}} > (\dots)$$

okay, it is clear that any prior bias for s_q will pull closer towards s_q ; shows that, by mere likelihood, the same result is expected; so set: $\epsilon_r = 0$

$$\frac{p_V}{p_V + q_V} > \frac{q_{\bar{V}}}{p_{\bar{V}} + q_{\bar{V}}}$$

$$\Leftrightarrow p_V p_{\bar{V}} + p_V q_{\bar{V}} > p_V q_V + q_V q_{\bar{V}}$$

$$\Leftrightarrow p_V p_{\bar{V}} > q_V q_{\bar{V}}$$

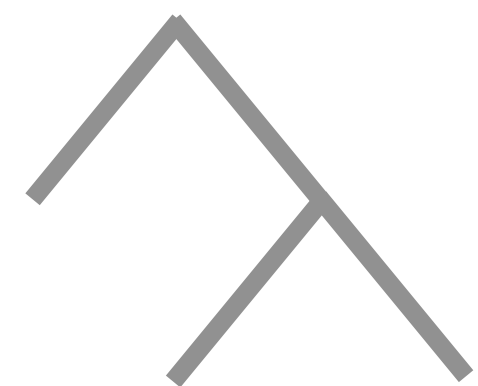
$$\Leftrightarrow \frac{p_{\bar{V}}}{q_V} > \frac{q_{\bar{V}}}{p_V}$$

assume:
 $q_{\bar{V}} = q_V + \epsilon_q$ $\epsilon_q > \epsilon_p$
 $p_V = p_{\bar{V}} + \epsilon_p$
 $(p_{\bar{V}} + \epsilon_p) p_{\bar{V}} > q_V (q_V + \epsilon_q)$
 $p_{\bar{V}}^2 + \epsilon_p p_{\bar{V}} > q_V^2 + \epsilon_q q_V$
 $p_{\bar{V}}^2 + \epsilon_p p_{\bar{V}} > q_V^2 + \epsilon_q q_V + \epsilon_p q_V$
 [producing V when adequate is less likely than producing \bar{V} when adequate] $p_V \leq q_{\bar{V}}$
 [producing V when inadequate is less likely than producing \bar{V} when inadequate] $p_{\bar{V}} > q_V$

Computational Pragmatics

Interpreting polite evaluations (chapter IX of problang.org)

Session 9



/lɪŋˈɡwɪstɪks/

$$[[\text{Joe}]] = \lambda e . \lambda w . \text{Joe}(e, w)$$

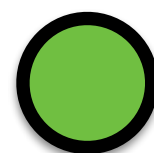
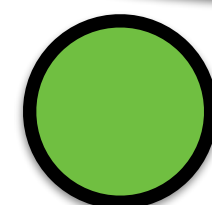
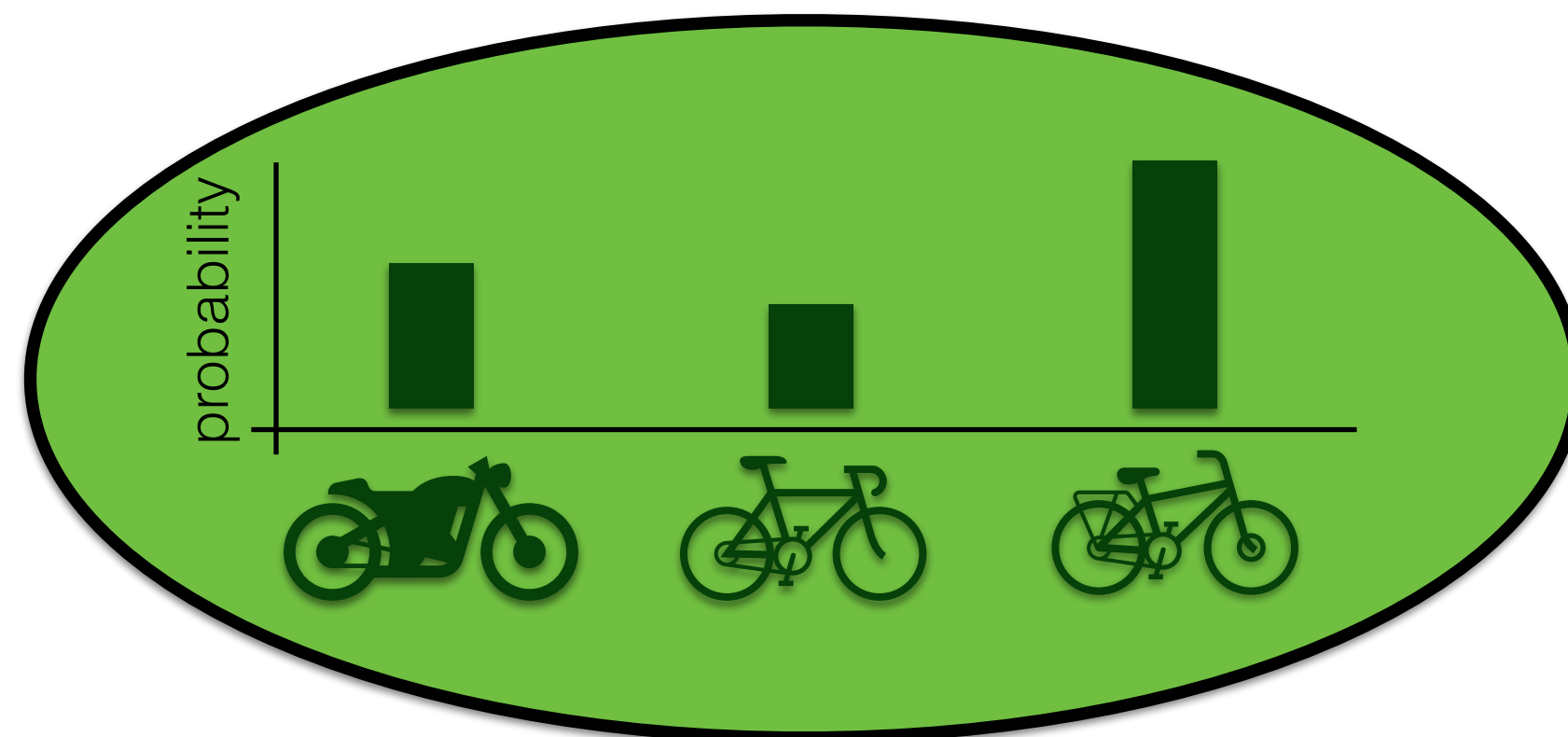
KNOWLEDGE OF LANGUAGE



GENERAL WORLD KNOWLEDGE



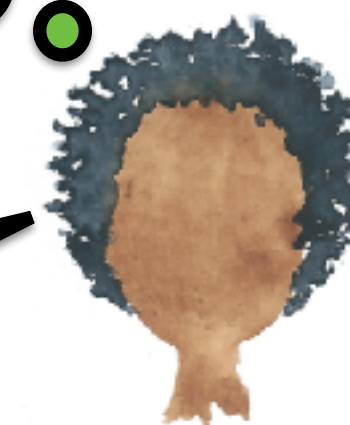
ACTUAL CONTEXT OF CONVERSATION



PRAGMATIC INTERPRETER



BIKE!

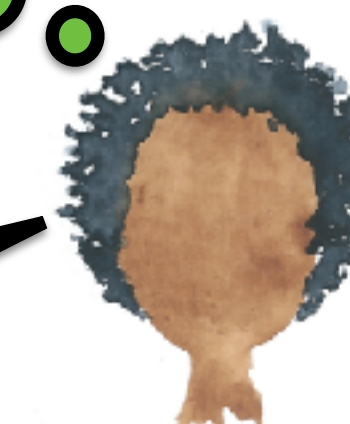


ACTUAL SPEAKER

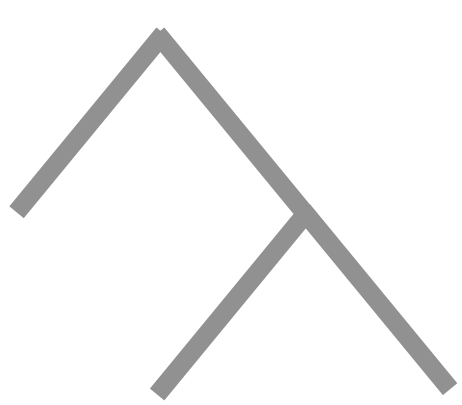
MODEL OF SPEAKER BEHAVIOR



MOTOR!



**ALTERNATIVE UTTERANCES
& INTERPRETATIONS**


 /lɪŋ'gwɪstɪks/

$[[\text{Joe}]] = \lambda e . \lambda w . \text{Joe}(e, w)$

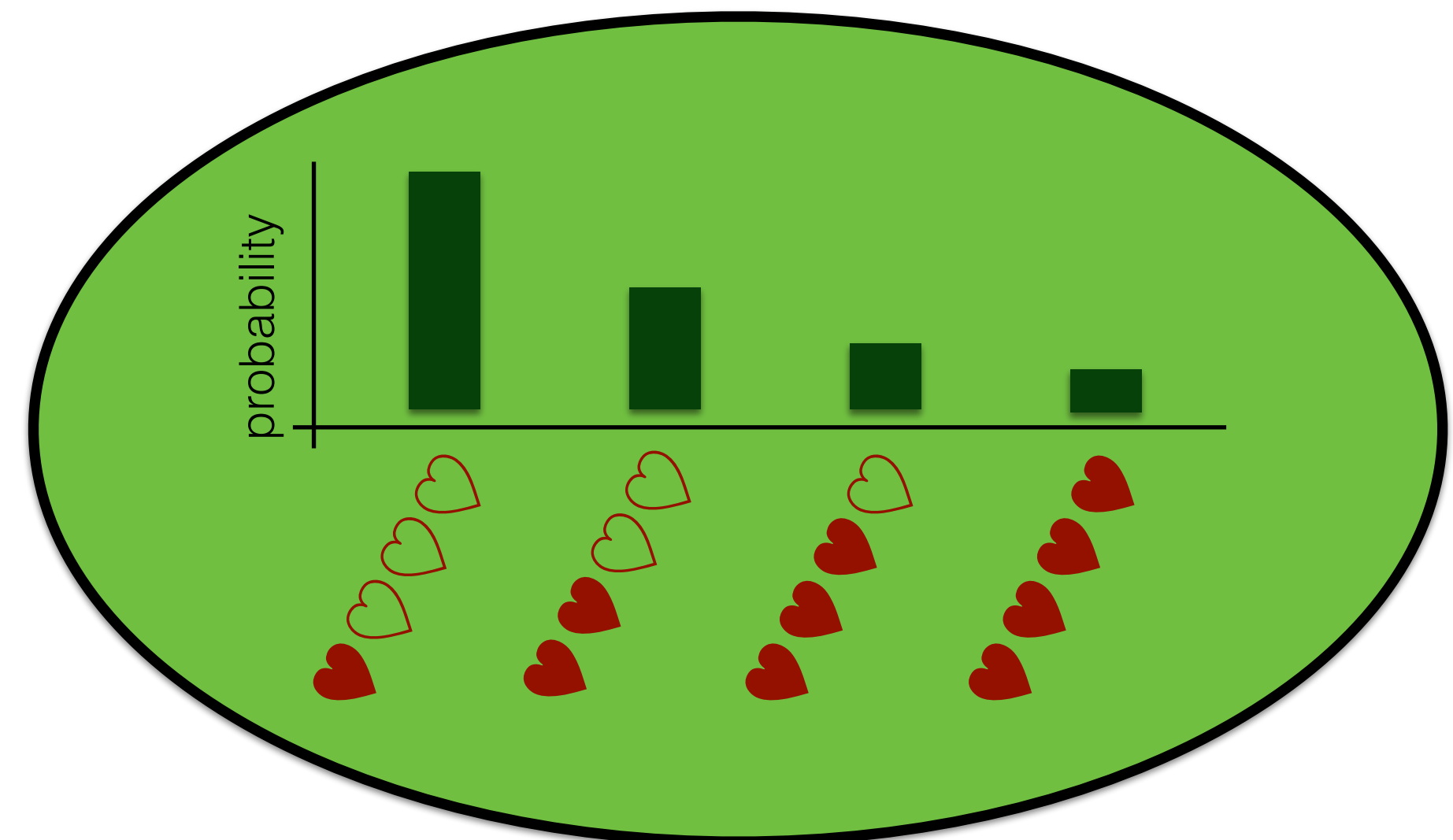
KNOWLEDGE OF LANGUAGE



GENERAL WORLD KNOWLEDGE



ACTUAL CONTEXT OF CONVERSATION

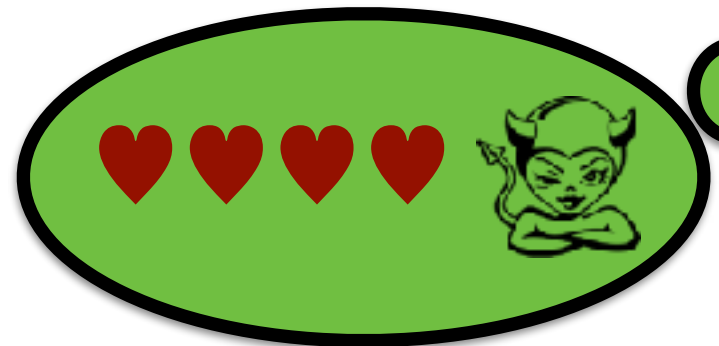







ACTUAL SPEAKER

MODEL OF SPEAKER BEHAVIOR



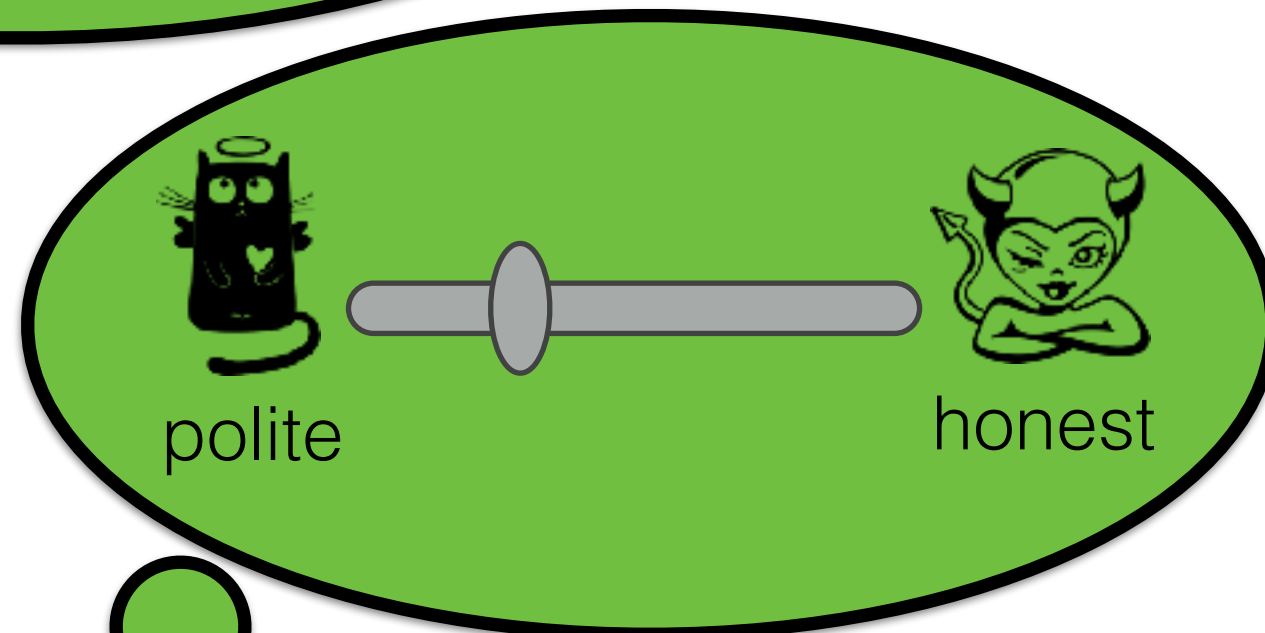
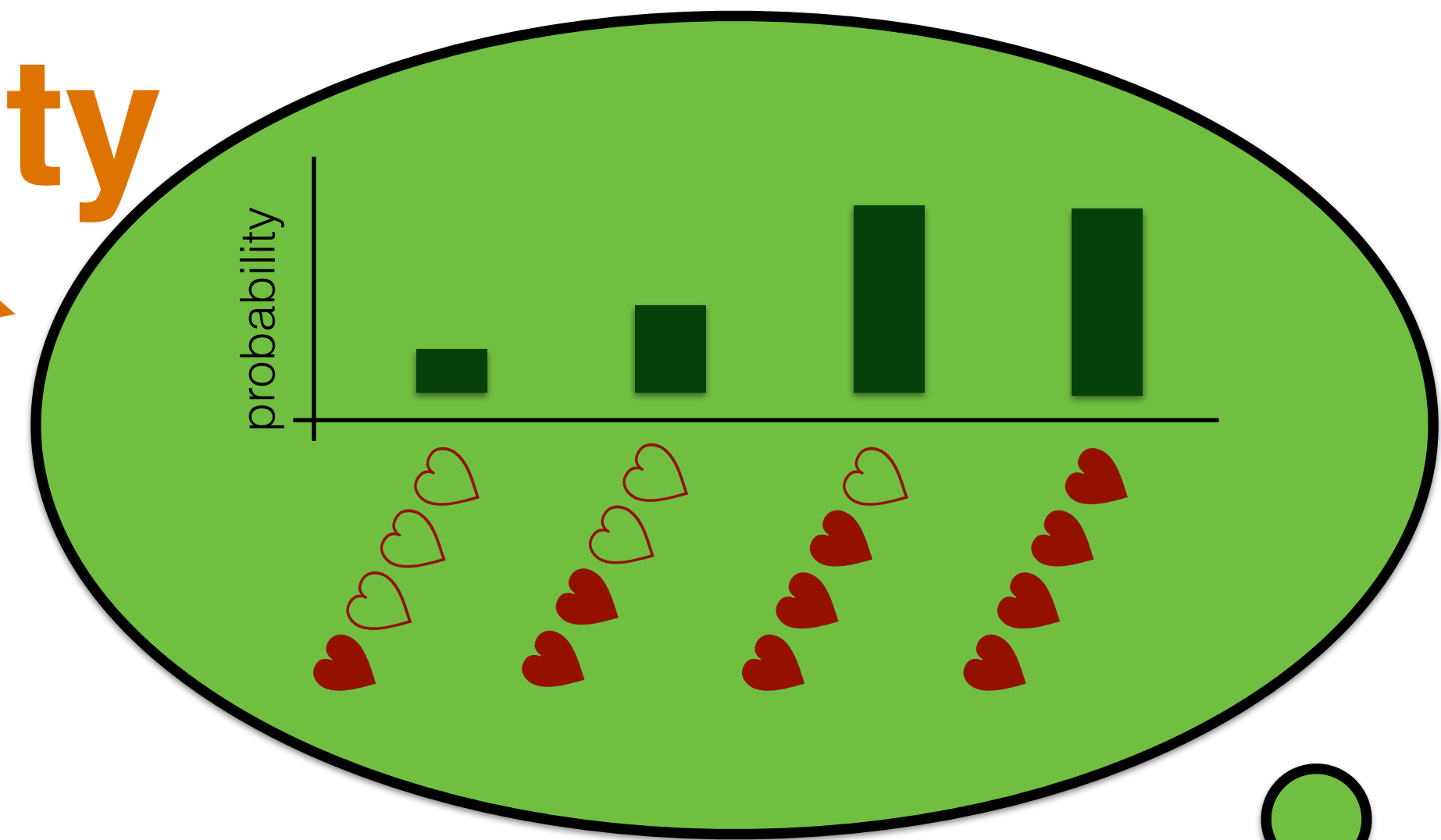
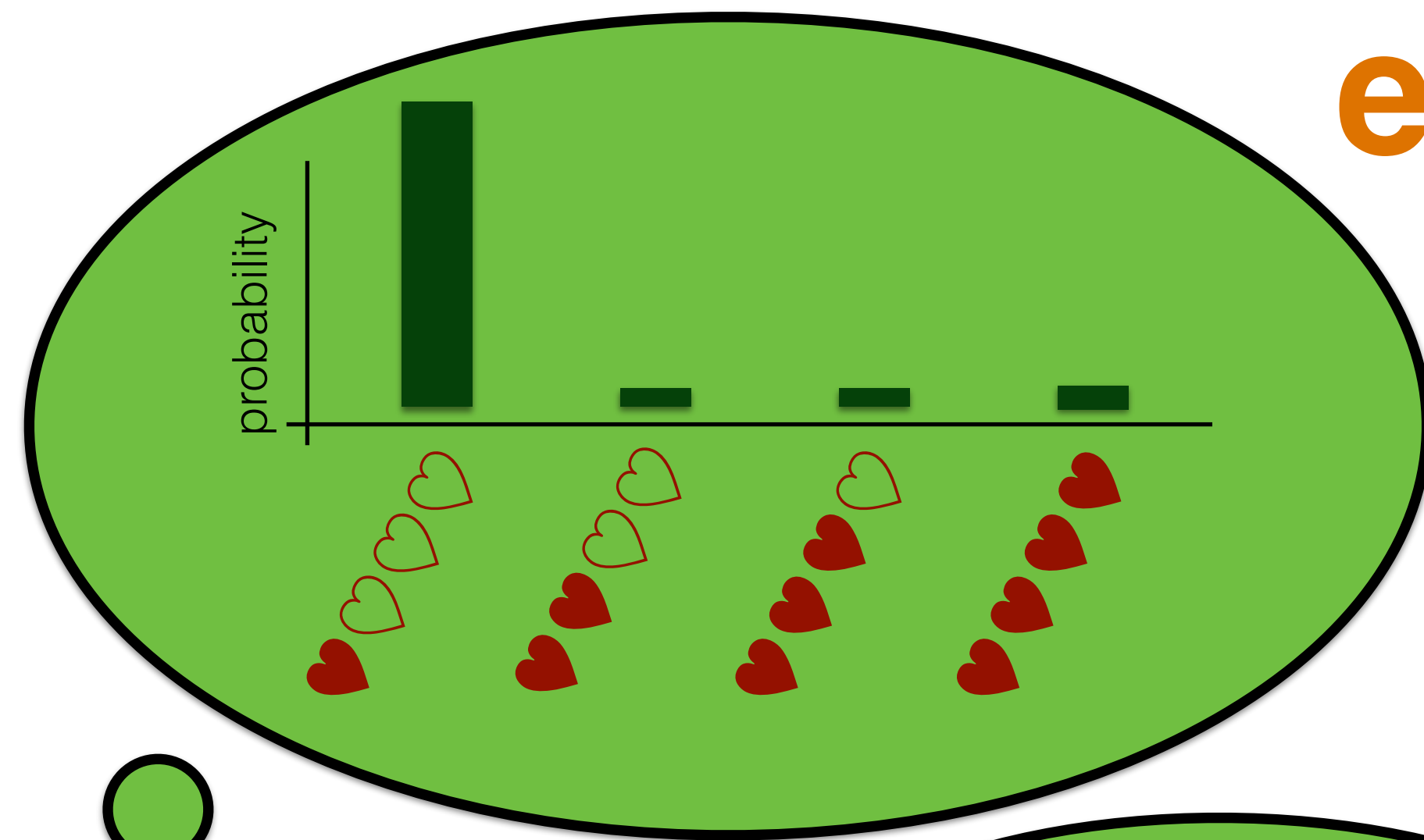


ALTERNATIVE UTTERANCES & INTERPRETATIONS

PRAGMATIC INTERPRETER

epistemic utility

informativity
KL-divergence



linear
combination

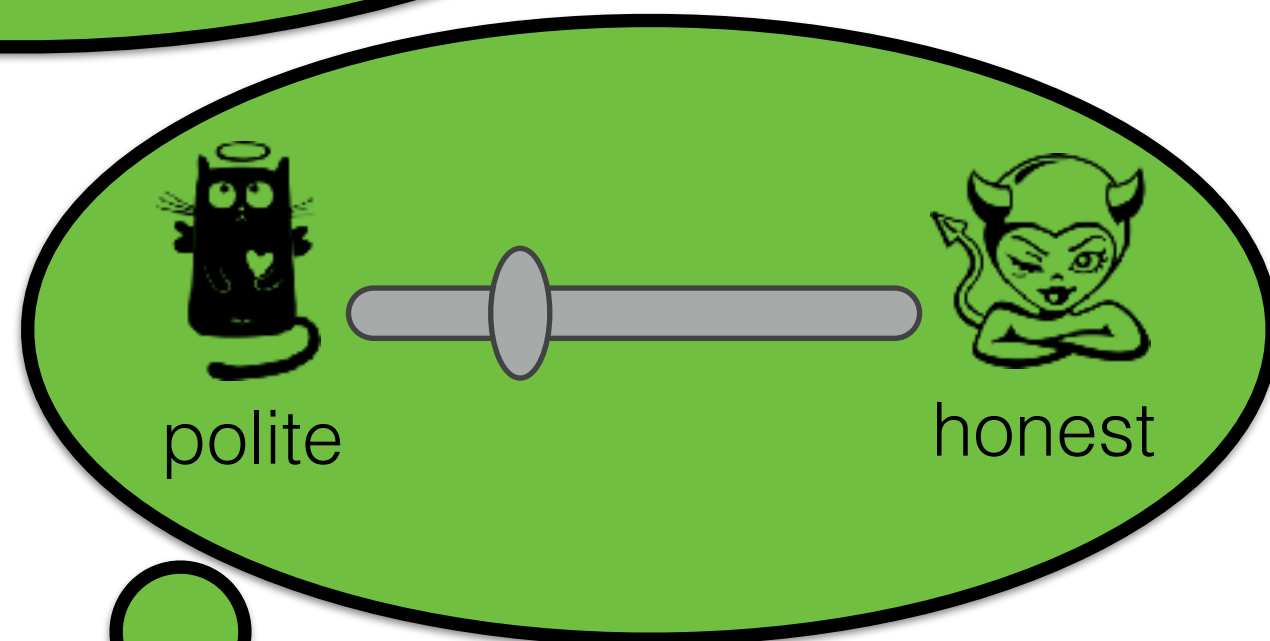
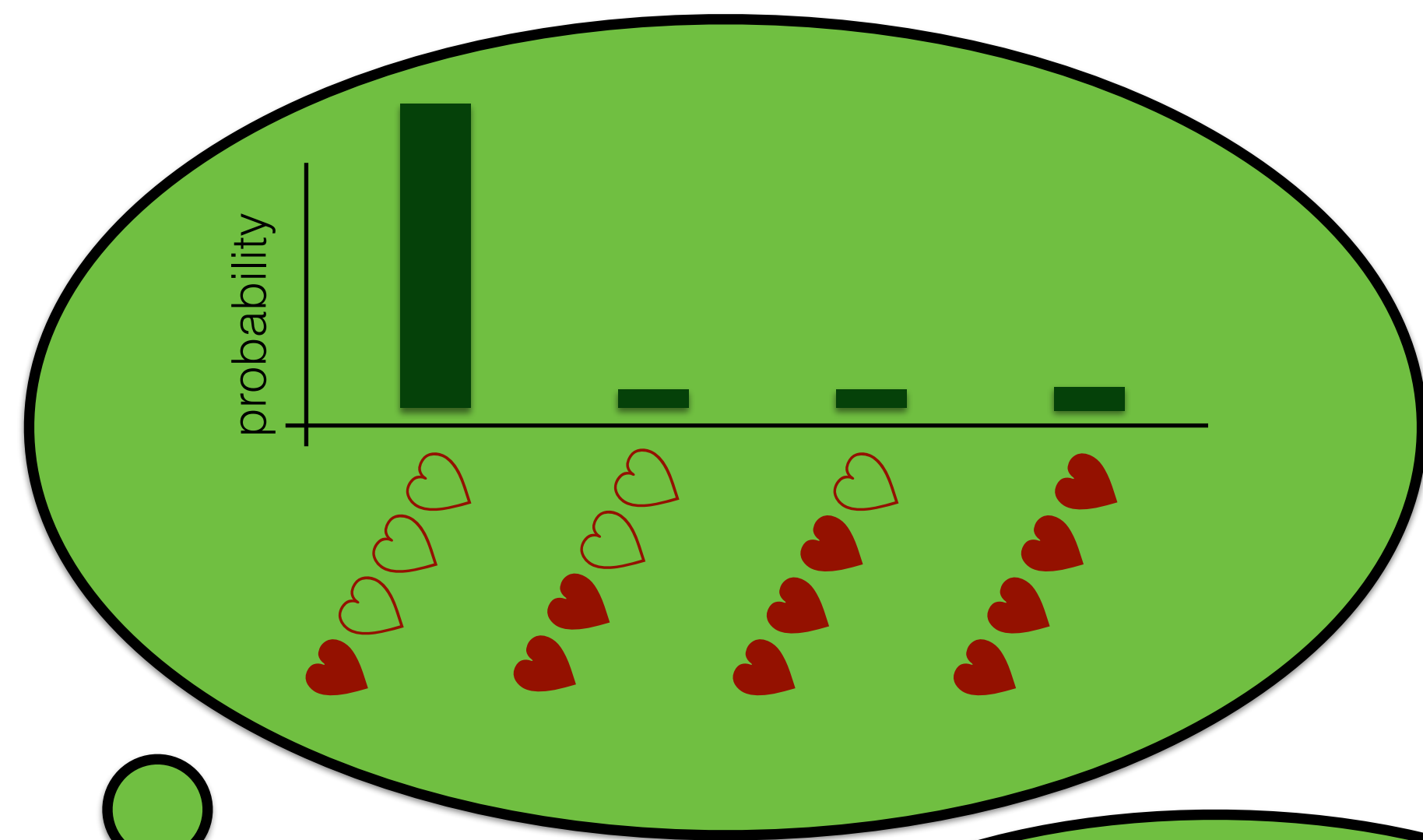
social utility



PRAGMATIC SPEAKER

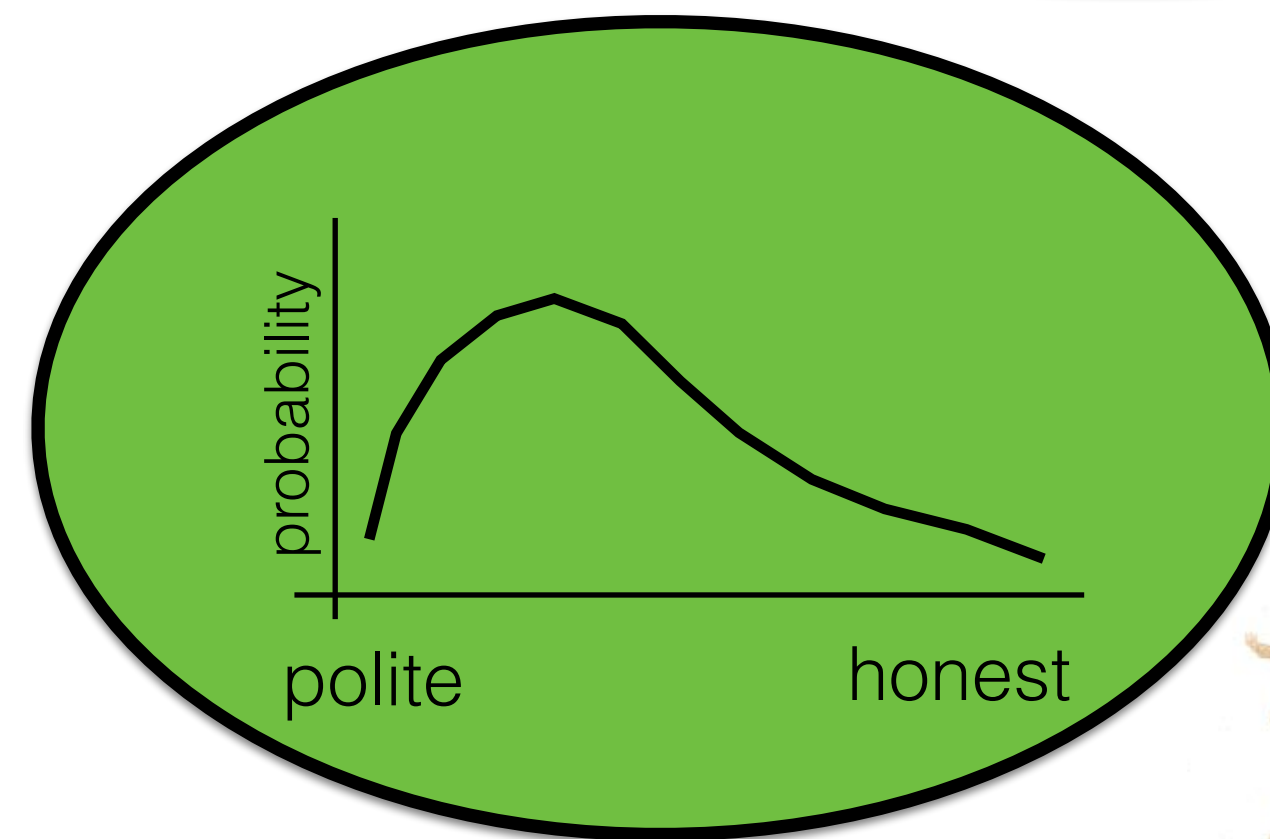
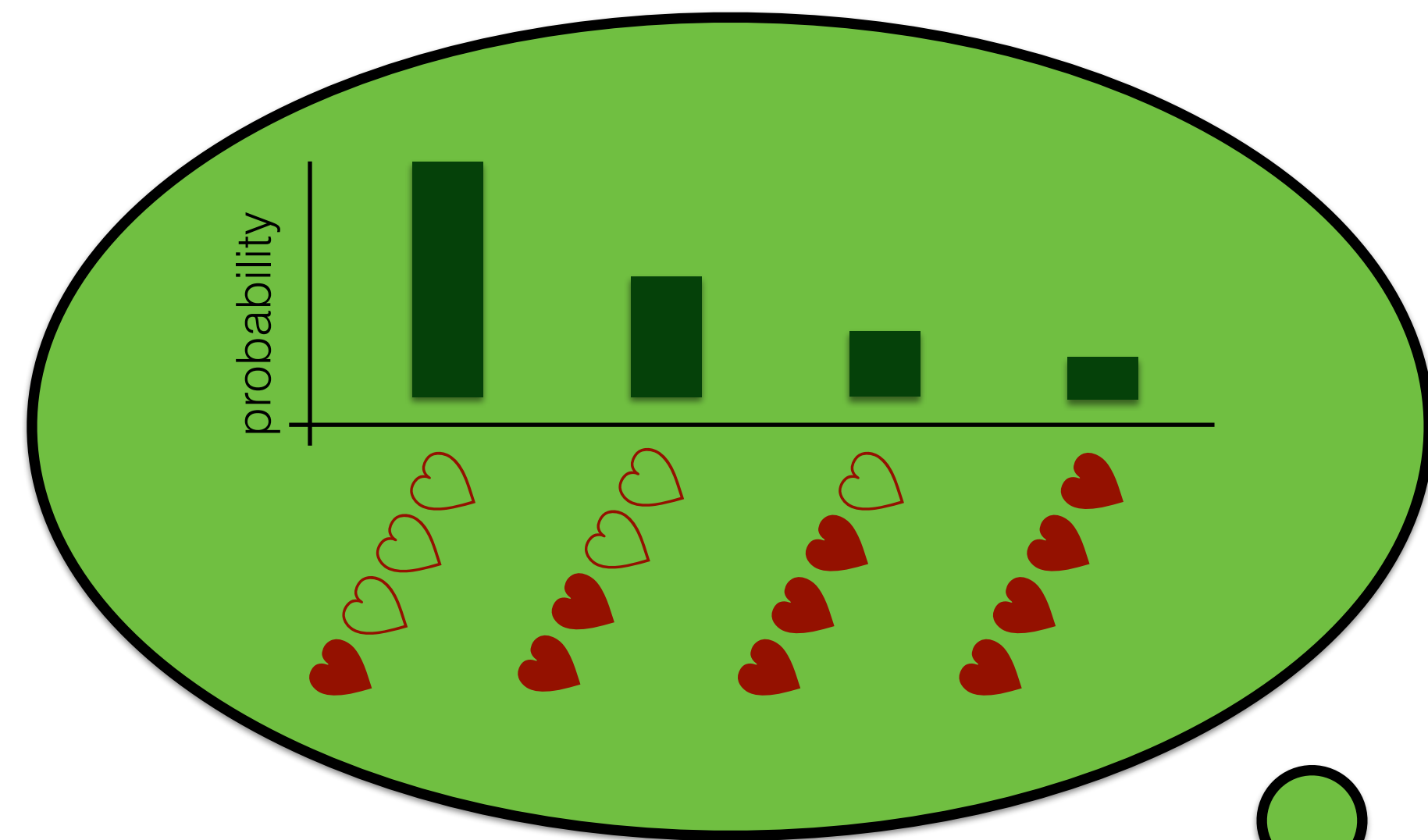


LITERAL INTERPRETER



PRAGMATIC SPEAKER

“good”



PRAGMATIC INTERPRETER



**modeling polite
language use**

Vanilla RSA

$$P_{L_0}(s \mid u) = P(s \mid \llbracket u \rrbracket)$$

$$P_{S_1}(u \mid s) \propto \exp \left(\alpha \left(\log P_{L_0}(s \mid u) - \mathbf{C}(u) \right) \right)$$

$$P_{L_1}(s \mid u) \propto P(s) P_{S_1}(u \mid s)$$

polite language model

$$P_{L_0}(s \mid u) = P(s \mid \llbracket u \rrbracket)$$

$$P_{S_1}(u \mid s, \varphi) \propto \exp \left(\alpha U(s, u, \varphi) \right)$$

$$U(s, u, \varphi) = \underbrace{\varphi \log P_{L_0}(s \mid u)}_{\text{epistemic util}} + (1 - \varphi) \underbrace{\sum_{s'} P_{L_0}(s' \mid u) V(s')}_{\text{social util}}$$

$$P_{L_1}(s, \varphi \mid u) \propto P(s) P(\varphi) P_{S_1}(u \mid s, \varphi)$$

$$P_{S_2}(u \mid s, \varphi) \propto \exp \left(\alpha' \left(\log P_{L_1}(s, \varphi \mid u) \right) \right) \text{ “self-representational speaker”}$$