

RSA math cheatsheet (material from Goodman and Frank 2012 – see paper for context)

Equation 1 (GF2012) – posterior probability

$$P(r_s|w, C) = \frac{P(w|r_s, C) \times P(r_s)}{\sum_{r' \in C} P(w|r', C) \times P(r')}$$

(i) r_s is an intended referent (element of C); (ii) C is a context (set of objects $r, r', \dots \in C$); (iii) w is the word uttered by the speaker; (iv) the **prior** is $P(r_s)$; (v) the **likelihood** is $P(w|r_s, C)$; and (vi) the **normalizing constant** is $\sum_{r' \in C} P(w|r', C) \times P(r')$

Equation 2 (GF2012) – likelihood for GF12 (derived from formulas below)

$$P(w|r_s, C) = \frac{|w|^{-1}}{\sum_{w' \in W} |w'|^{-1}}$$

- $|w|$ is the number of objects in C to which word w could(?) apply,
- and W is the set of words that apply to the intended referent r_s .

Equation S1 (GF12 supplement)

$$P(w|r_s, C) \propto e^{\alpha \times U(w; r_s, C)}$$

(i) the **context** $C = \{o_1, \dots, o_n\}$ is a set of objects; (ii) the **vocabulary** is a set of words $V = \{w_1, \dots, w_m\}$; (iii) each word $w \in V$ has a meaning $\llbracket w \rrbracket : C \rightarrow \{0, 1\}$; and (iv) words are chosen in proportion to their expected utility

Equation S2 – utility (GF12 supplement)

$$U(w; r_s, C) = I(w; r_s, C) - D(w)$$

where $I(w; r_s, C)$ is the informativeness of utterance w w.r.t. intended referent $r_s \in C$; and $D(w)$ is w 's cost (in C).

Surprisal (GF12 supplement), for a sample x from a known distribution $p(x)$

$$I_p(x) = -\log(p(x))$$

Utility and surprisal (GF12 supplement) – utility decreases with surprisal

$$I(w; r_s, C) = -I_{\tilde{w}_C}(r_s)$$

where \tilde{w}_C is the distribution over objects that would come from a literal interpretation of w in context C .

Equation S3 – literal listener model derivation (GF12, supplementary materials)

“if listeners interpret the utterance w literally, assigning zero probability to objects for which the word is false, they assign equal probability to each object consistent with w . This distribution over objects can be written:”

$$\tilde{w}_C(o) = \frac{1}{|w|} \text{ if } w(o) = \text{true, and otherwise } 0$$

Equation S4 – listener likelihood (GF12, supplement)

equation S4 follows from S1-S3, which is equivalent to equation 2 (aka “the size principle”)

$$P(w|r_s, C) = \frac{e^{-(-\log(|w|^{-1}))}}{\sum_{w' \in V \text{ s.t. } w'(r_s) = \text{true}} e^{-(-\log(|w'|^{-1}))}}$$