Computational notes on the Model of Agrammatic Aphasia by Stocco & Crescentini (2005)

Andrea Stocco

25th July 2005

1 Introduction

1.1 The model

This document contains some computational notes on the model of agrammatic aphasia developed by Andrea Stocco & Cristiano Crescentini.

1.2 Relevant features of the ACT-R architecture

In ACT, activation is spread over the N elements of the goal chunk. Each fraction W/N of the total attention amount W flows to the associated chunks through interassociative links S. The value of S is also a measure of the association between two chunks.

Each chunk i has a base-level activation B_i which only depends on its past history (i.e., latency and frequency of its past encodings and retrievals). Ther total amount of activation is given by $B_i + \sum_{j=1}^{N} W_j S_{i,j}$, where j are individual slots in the goal chunk. Of course, the latter term is null whenever the goal is unrelated to the retrievable chunk.

Now, let us consider the activation of the default thematic grid (D), B_D, and that of the non-default, reversed grid (R), B_R. We pose that $B_D \gg B_R$.

In our simplified domain, there is only one slot of the goal chunk that is associated to the grid, and is the one holding the verb. Therefore, contextual activation only amounts to WS/N.

Given the values of S, N, B_D and B_R, agrammatic parsing occurs whener a pathological value of W does not allow B_R to overcome B_D. We will refer to this pathological value as ${}^{\dagger}W$. It must be such that:

$$B_D \approx B_R + \frac{{}^{\dagger}WS}{N}$$

and, therefore:

$${}^{\dagger}W \approx \frac{N(B_D - B_R)}{S}$$

In normal parsing, attentional resources are greater that in agrammatic patients. We will refer to this non-pathological value simply as W, and pose that $W \gg {}^{\dagger}W$. This implies that:

$$W \gg \frac{N(B_D - B_R)}{S}$$

and that, therefore

$$B_D \ll B_R + \frac{WS}{N}$$

so that now the non-default reversed grid will overcome the default one.

1.3 More precise estimates

How much should B_D be greater than B_R, and W of ${}^{\dagger}W$? Values should be choosen according to language-specific features, such the relative frequencies of the two grids in a corpus of sentences. However, certain boundaries may be easily calculated.

First, if B_D has to be effectively stronger that B_R, it must overcome normally-distributed noise in base level activation. Let us choose a fixed noise component, ε, according to our desired probability of having B_D retrieved over B_R in random contrasts. For instance, if we need B_D to be retrieved 95% of the times, we will choose a value of ε that is at least equal to twice the standard deviation of noise[1].

Once this cutoff has been fixed, you can impose $B_D = B_R + \varepsilon$. Note that this implies that $\varepsilon = B_D - B_R$. It follows immediately that:

$${}^{\dagger}W = \frac{N\varepsilon}{S}.$$

Now, the problem is to estimate W so that, in the opportune cases, B_R would be retrieved over B_D. A different cutoff probability could be chosen but, for simplicity's sake, let us take the usual ε. This means that, when contextually activated, the activation of R should be equal to $B_D + \varepsilon$:

$$B_R + \frac{WS}{N} = B_D + \varepsilon$$

But $B_D = B_R + \varepsilon$:

$$B_R + \frac{WS}{N} = B_R + 2\varepsilon$$

$$\frac{WS}{N} = 2\varepsilon$$

$$W = \frac{2N\varepsilon}{S}$$

[1] To be exact, a value of ε that is equal to 1.96 times the standard deviation, plus the mean of the distribution—which is usually zero

1.4 An alternative view

In designing the model, we choose an alternative view. We simply set an arbitrary difference between the values of B_D and B_R, and indicate this as δ. It is obvious that, in this case, $^{\dagger}W = N\delta/S$. In this case, to obtain the desired performance, it was sufficient to set $W =^{\dagger} W + N\varepsilon/S$.