REM 1997 Derivations

September 15, 2025

This document derives REM in details for future reference and to make it easier to understand and extend the model if needed. It is based on Greg Cox 2024 Dream paper.

The goal of the model is to compute the posterior probability that to a given probe, there is a single matching trace in memory versus no matching traces.

Definitions

For a given test probe, let:

- n: number of memory traces that the probe is compared to
- S_j: the event that trace j is an image stored for the test probe (S-image)- trace and probe match
- N_j: the event that trace j is an image stored for some word other than the test probe (D-image)- trace and probe mismatch
- K_i^M: the number of matching features between the test probe and memory trace i
- K_i^N: the number of mismatching features between the test probe and memory trace i

K_i^M and K_i^N are basically the date (D) in the model, which is the result of comparing the test probe to memory traces

The model assumes that (1) there can only be a maximum of one trace matching the probe and (2) traces and features are independent

Derivations

$$\frac{P(O|D)}{P(N|D)} = \frac{\sum_{i=1}^{n} P(S_i|K_i^M, K_i^N) \prod_{j \neq i} P(N_j|K_j^M, K_j^N)}{\prod_i P(N_i|K_i^M, K_i^N)}$$

$$= \sum_{i=1}^{n} \frac{P(S_i|K_i^M, K_i^N)}{P(N_i|K_i^M, K_i^N)} \cancel{\frac{\prod_{j \neq i} P(N_j|K_j^M, K_j^N)}{\prod_{j \neq i} P(N_j|K_j^M, K_j^N)}}$$

Using Bayes theorem:

$$= \sum_{i=1}^{n} \frac{P(K_i^M, K_i^N|S_i)P(S_i)}{P(K_i^M, K_i^N|N_i)P(N_i)}$$

Since all traces are equally likely to match the probe, $S_j = \frac{1}{n}$ and $N_j = 1 - S_j$

They approximately go to (not exactly):

$$= \frac{1}{n} \sum_{i=1}^{n} \frac{P(K_i^M, K_i^N|S_i)}{P(K_i^M, K_i^N|N_i)}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \lambda_i$$

Now Let's define each of the four probabilities in the λ equation:

$P(M|S) = c + (1-c)g$, a feature match if the whole trace matches
$P(N|S) = (1-c)(1-g)$, a feature mismatch if the whole trace matches
$P(M|D) = g$, a feature match if the whole trace mismatches
$P(N|D) = (1-g)$, a feature mismatch if the whole trace mismatches

Substitute in the previous equation:

$$\frac{P(O|D)}{P(N|D)} = \frac{1}{n} \sum_{i=1}^{n} \frac{[c + (1-c)g]^{K_i^M} \; [(1-c)(1-g)]^{K_i^N}}{g^{K_i^M} (1-g)^{K_i^N}}$$

$$= \frac{1}{n} \sum_{i=1}^{n} (1-c)^{K_i^N} \left[\frac{c + (1-c)g}{g} \right]^{K_i^M}$$