Ling 8700: Lecture Notes 5
A Model of Ambiguity in Sentence Processing

We have seen how complex ideas can be encoded and decoded into sentences.

This lecture will describe how this process controls for ambiguity.

Contents

5.1 Ambiguity as superposition

If multiple targets are associated with the same cue vector:

target 1 (**N**) cue (e.g. the word ‘*that*’) target 2 (**C-bV**) cue (e.g. the word ‘*that*’)

$$\begin{bmatrix} .58 \\ .0 \\ .49 \\ .0 \\ .0 \\ .67 \\ .0 \end{bmatrix} \begin{bmatrix} .0 & .58 & .0 & .0 & .58 & .0 & .58 \end{bmatrix} + \begin{bmatrix} .0 \\ .50 \\ .50 \\ .0 \\ .50 \\ .0 \\ .50 \end{bmatrix} \begin{bmatrix} .0 & .58 & .0 & .0 & .58 & .0 & .58 \end{bmatrix}$$

$$= \begin{bmatrix} .0 & .33 & .0 & .0 & .33 & .0 & .33 \\ .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .28 & .0 & .0 & .28 & .0 & .28 \\ .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .38 & .0 & .0 & .38 & .0 & .38 \\ .0 & .0 & .0 & .0 & .0 & .0 & .0 \end{bmatrix} + \begin{bmatrix} .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .29 & .0 & .0 & .29 & .0 & .29 \\ .0 & .29 & .0 & .0 & .29 & .0 & .29 \\ .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .29 & .0 & .0 & .29 & .0 & .29 \\ .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .29 & .0 & .0 & .29 & .0 & .29 \end{bmatrix} = \begin{bmatrix} .0 & .33 & .0 & .0 & .33 & .0 & .33 \\ .0 & .29 & .0 & .0 & .29 & .0 & .29 \\ .0 & .57 & .0 & .0 & .57 & .0 & .57 \\ .0 & .0 & .0 & .0 & .0 & .0 & .0 \\ .0 & .29 & .0 & .0 & .29 & .0 & .29 \\ .0 & .38 & .0 & .0 & .38 & .0 & .38 \\ .0 & .29 & .0 & .0 & .29 & .0 & .29 \end{bmatrix}$$

when associative memory is cued using that vector, the result is both vectors, **superposed**:

$$\overbrace{\begin{bmatrix}.58\\.0\\.49\\.0\\.0\\.67\\.0\end{bmatrix}}^{\text{target 1 (N)}} + \overbrace{\begin{bmatrix}.0\\.50\\.50\\.0\\.50\\.0\\.50\end{bmatrix}}^{\text{target 2 (C-bV)}} = \overbrace{\begin{bmatrix}.58\\.50\\.99\\.0\\.50\\.67\\.50\end{bmatrix}}^{\text{combined target}} = \overbrace{\begin{bmatrix}.0&.33&.0&.0&.33&.0&.33\\.0&.29&.0&.0&.29&.0&.29\\.0&.57&.0&.0&.57&.0&.57\\.0&.0&.0&.0&.0&.0&.0\\.0&.29&.0&.0&.29&.0&.29\\.0&.38&.0&.0&.38&.0&.38\\.0&.29&.0&.0&.29&.0&.29\end{bmatrix}}^{\text{synaptic weights}} \overbrace{\begin{bmatrix}.0\\.58\\.0\\.0\\.58\\.0\\.58\end{bmatrix}}^{\text{cue (‘\textit{that}’)}}$$

5.2 Propagation of ambiguity

If we take a (non-interfering) set of associations:

$$\overbrace{\begin{bmatrix}.58\\.0\\.58\\.0\\.0\\.58\\.0\end{bmatrix}}^{\text{target 1}} \overbrace{\begin{bmatrix}.0&.58&.0&.0&.58&.0&.58\end{bmatrix}}^{\text{cue 1}} + \overbrace{\begin{bmatrix}.0\\.50\\.50\\.0\\.50\\.0\\.50\end{bmatrix}}^{\text{target 2}} \overbrace{\begin{bmatrix}.0&.0&.58&.58&.0&.58&.0\end{bmatrix}}^{\text{cue 2}}$$

$$= \begin{bmatrix}.0&.33&.0&.0&.33&.0&.33\\.0&.0&.0&.0&.0&.0&.0\\.0&.33&.0&.0&.33&.0&.33\\.0&.0&.0&.0&.0&.0&.0\\.0&.0&.0&.0&.0&.0&.0\\.0&.33&.0&.0&.33&.0&.33\\.0&.0&.0&.0&.0&.0&.0\end{bmatrix} + \begin{bmatrix}.0&.0&.0&.0&.0&.0&.0\\.0&.0&.29&.29&.0&.29&.0\\.0&.0&.29&.29&.0&.29&.0\\.0&.0&.0&.0&.0&.0&.0\\.0&.0&.29&.29&.0&.29&.0\\.0&.0&.0&.0&.0&.0&.0\\.0&.0&.29&.29&.0&.29&.0\end{bmatrix} = \begin{bmatrix}.0&.33&.0&.0&.33&.0&.33\\.0&.0&.29&.29&.0&.29&.0\\.0&.33&.29&.29&.33&.29&.33\\.0&.0&.0&.0&.0&.0&.0\\.0&.0&.29&.29&.0&.29&.0\\.0&.33&.0&.0&.33&.0&.33\\.0&.0&.29&.29&.0&.29&.0\end{bmatrix}$$

and cue them with a combination of states, we get a proportional combination of targets:

$$\underbrace{\begin{bmatrix}.29\\.0\\.29\\.0\\.0\\.29\\.0\end{bmatrix}}_{\text{target 1}} + \underbrace{\begin{bmatrix}.0\\.50\\.50\\.0\\.50\\.0\\.50\end{bmatrix}}_{\text{target 2}} = \underbrace{\begin{bmatrix}.29\\.50\\.79\\.0\\.50\\.29\\.50\end{bmatrix}}_{\text{combined target}} = \underbrace{\begin{bmatrix}.0&.33&.0&.0&.33&.0&.33\\.0&.0&.29&.29&.0&.29&.0\\.0&.33&.29&.29&.33&.29&.33\\.0&.0&.0&.0&.0&.0&.0\\.0&.0&.29&.29&.0&.29&.0\\.0&.33&.0&.0&.33&.0&.33\\.0&.0&.29&.29&.0&.29&.0\end{bmatrix}}_{\text{synaptic weights}} \underbrace{\begin{bmatrix}.0\\.29\\.58\\.58\\.29\\.58\\.29\end{bmatrix}}_{\text{combined cue}}$$

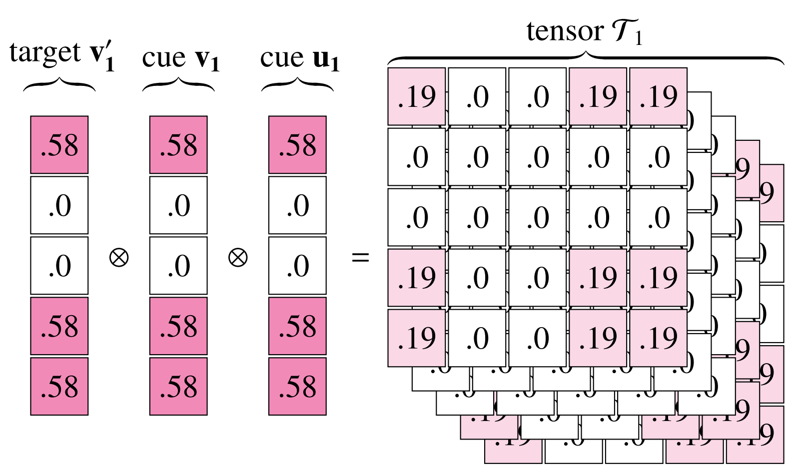
This is important because it allows ambiguity to propagate through a mental process.

5.3 Resolution of ambiguity

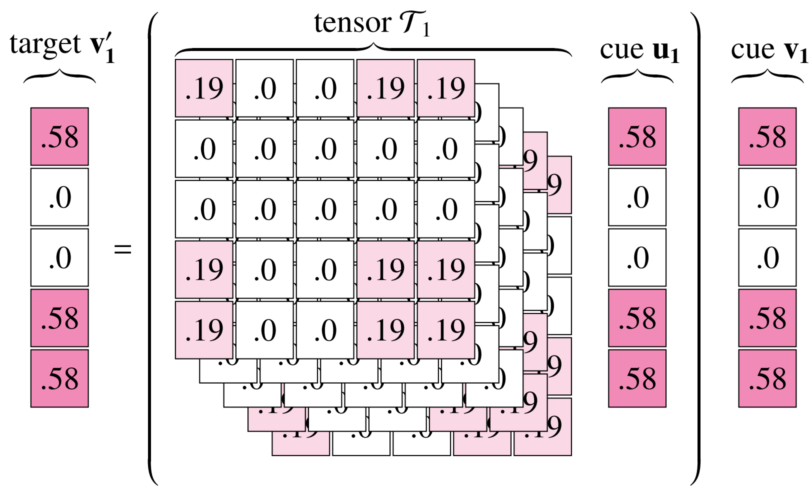
Recall the outer product of two vectors produces a matrix with pointwise products:

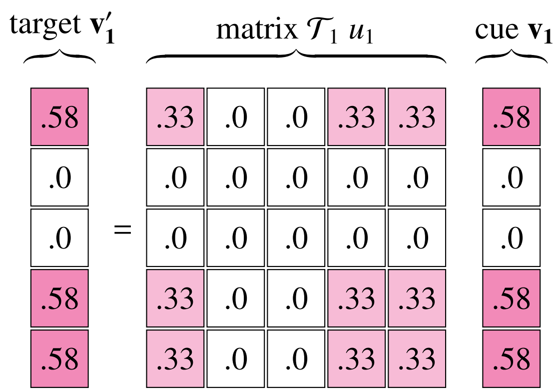
$$\underbrace{\begin{bmatrix}.0\\.71\\.0\\.0\\.71\end{bmatrix}}_{\text{target } \mathbf{v_1}} \underbrace{\begin{bmatrix}.0&.71&.0&.0&.71\end{bmatrix}}_{\text{cue } \mathbf{u_1}} = \underbrace{\begin{bmatrix}.0&.0&.0&.0&.0\\.0&.50&.0&.0&.50\\.0&.0&.0&.0&.0\\.0&.0&.0&.0&.0\\.0&.50&.0&.0&.50\end{bmatrix}}_{\text{associations } \mathbf{M_1}} \tag{1}$$

This generalizes to triples of vectors as a **tensor product**:



Targets are then cued by multiplication (left-associative):

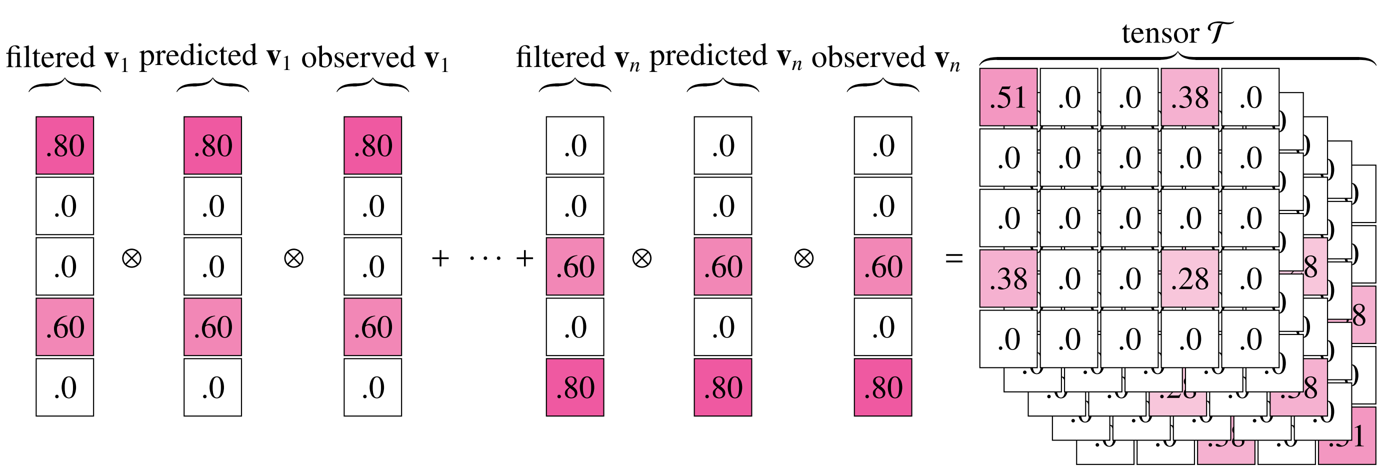




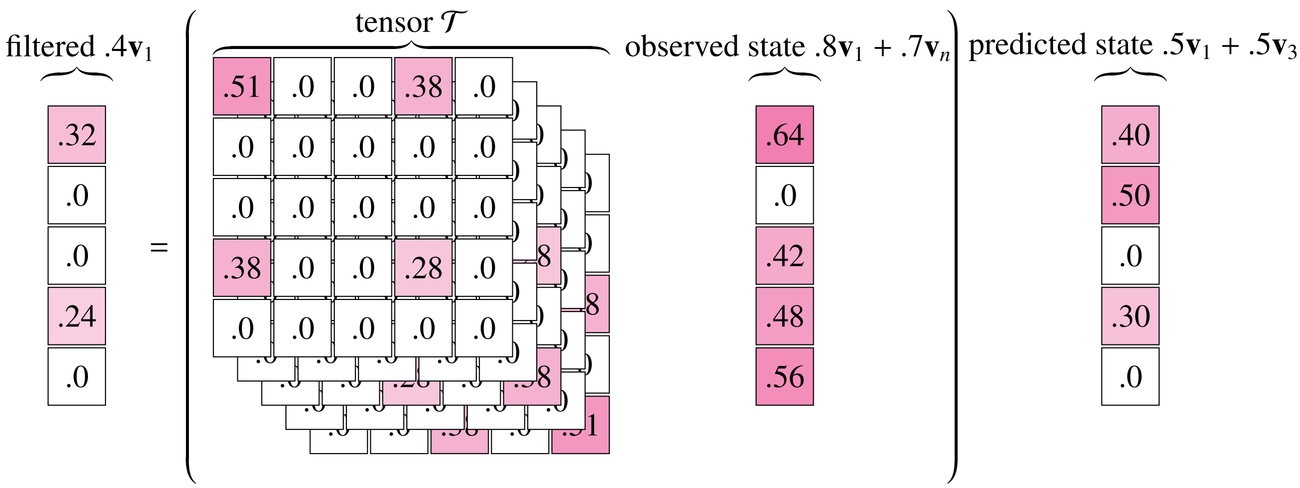
This can be implemented with 'joint features'.

This gives us a means to combine 'top-down' predictions with 'bottom-up' observations…

Build auto-associations of all states:



Then cue on observed state to pick out compatible component of mixed source state:



Note the magnitude of the target is reduced compared to the source.

This reduction correlates with reading time delays (‘surprisal’) on encountering unpredicted words.

(It may take time proportional to the reduction to ‘amp up’ this state to unit magnitude.)