The Fundamental Law of Memory Recall

Quantitative Life Science course, Final Presentation

Davide Maniscalco
July 16, 2021

Università degli Studi di Padova

Purposes

Università degli Studi di Padova

Presentation, analysis and results reproduction of the paper of Naim et.al. "Fundamental Law of Memory Recall"

- Theory
  - Theory mention from the paper
  - Study and exploration of the underlying theoretical framework, from the paper by Romani et.al. "Scaling Laws of Associative Memory Retrieval"
- Computational Results
  - Results from the paper
  - Reproduction of the results
  - Some additional analysis and ideas
- Results from experiments and conclusion

Results of Recall experiments: Study of Recall:
Experiments!

Typical experiments involve recalling randomly assembled lists of words in an arbitrary order after a brief exposure

- When the presented list becomes longer, the average number of recalled words grows but in a sublinear way
- In some studies, recall performance was found to exhibit a power-law relation to the number of presented words, but parameters of this relation were extremely variable across different experimental conditions.

? Impossible to describe memory recall with a universal mathematical law

NO!

In this paper it is shown that:

- Most of the **variability** in recall can be accounted for by measuring the **acquisition** and **maintenance** of information during the presentation phase of the experiment, and when that is controlled, recall itself is much more predictable
- Relation between the number of **items in memory** and the average fraction of it that can be successfully **recalled** is described by a **parameter-free analytical expression**

Items

1. Memory items are represented in the brain by overlapping random sparse neuronal ensembles in dedicated memory networks

$$\vec{\xi} = (0, 1, 1, \dots, 0, 1)$$

$$\xi_i = \begin{cases} 1 & \text{w.p. } f \\ 0 & \text{w.p. } 1 - f \end{cases}$$

Sparseness: $f \ll 1$

N: number of neurons of each item

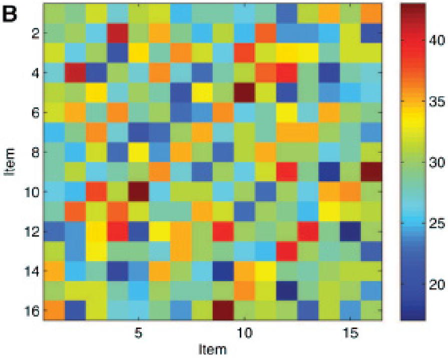
2. The next item to be recalled is the one with the largest **overlap** to the current one

Overlap (k,m) $= \sum_{i=1}^{N} \xi_i^k \xi_i^m = \vec{\xi^m} \cdot \vec{\xi^k}$

The Similarity Matrix

L: Number of items

$$S_{km} = \sum_{i=1}^{N} \xi_i^k \xi_i^m$$



Elements of the same row have positive correlation! After some algebra one finds

$$\langle S_{km} \rangle = N f^2$$
$$var(S_{km}) = N f^2 (1 - f^2)$$
$$cov(S_{km}, S_{kn}) = N f^3 (1 - f)$$
$$\varrho(S_{km}, S_{kn}) = \frac{f}{1+f} \tag{1}$$

In the limit of sparse encoding $(f \ll 1)$ we have that

$$\varrho(S_{km}, S_{kn}) \xrightarrow{f \ll 1} 0$$

That allows to replace the similarity matrix S_{km} with a **Random Symmetric Similarity Matrix**

The Model: The Recall Dynamic

Ingredients: The Similarity Matrix (SM) and the first recalled item

Dynamic: The next element being recalled is the one with the largest overlap with the current one

As soon as an element is recalled twice, we are in a loop, and no new elements can be recalled

Simulation: By repeating the process, one can calculate the average number of recalled items $\langle k \rangle$

Theory: As we are going to see now, a theoretical expression for $\langle k \rangle$ as a function of L can be found

Random Asymmetric SM: p_0

$\mathbf{p_0}$: Transition probability between any two items

$$p_0 = \frac{1}{L-1} \sim \frac{1}{L}$$

Prob. of coming back to any of the $m-1$ prev. recalled items:

$$(m-1)p_0$$

Prob. of recalling exactly k items

$$p(k;L) = (1-p_0)(1-2p_0)\dots(1-(k-2)p_0)\,(k-1)p_0$$

Let's go to the limit

$$1 \ll k \ll L$$

This allows us to write

$$1 - kp_0 \sim e^{-kp_0}$$

$$\begin{aligned} p(k;L) &= e^{-p_0} e^{-2p_0} \dots e^{-(k-2)p_0 (k-1)p_0} = \\ &= e^{-\frac{1}{2}(k-2)(k-1)p_0} (k-1)p_0 \sim \\ &\sim kp_0 e^{-\frac{k^2 p_0}{2}} = \\ &= \frac{k}{L} e^{-\frac{k^2}{2L}} \end{aligned}$$

And defining

$$x = k/\sqrt{L}$$

We can obtain the normalized PDF

$$p(x;L) = xe^{-\frac{x^2}{2}}$$

Random Asymmetric SM: $\langle k \rangle$

By definition

$$\langle x \rangle = \int_0^\infty x p(x; L) = \sqrt{\frac{\pi}{2}}$$

And therefore

$$\langle k \rangle = \sqrt{\frac{\pi L}{2}} \tag{2}$$

And by calculating

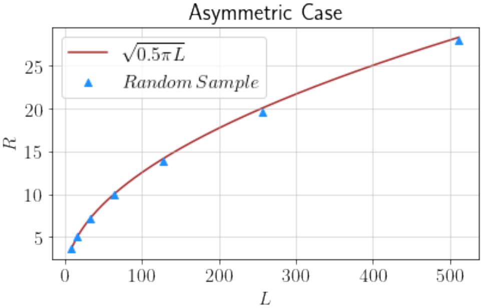
$$\langle k^2 \rangle = L \langle x^2 \rangle = 2L$$

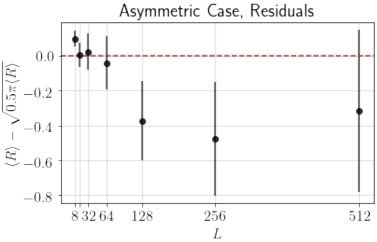
One can calculate the standard deviation also

$$\sigma(k) = \sqrt{\left(2 - \frac{\pi}{2}\right) L} \tag{3}$$

Random Asymmetric SM: Results

N = 10000





Problem: Due to the symmetry, the retrieval process quickly enters into a two-item loop

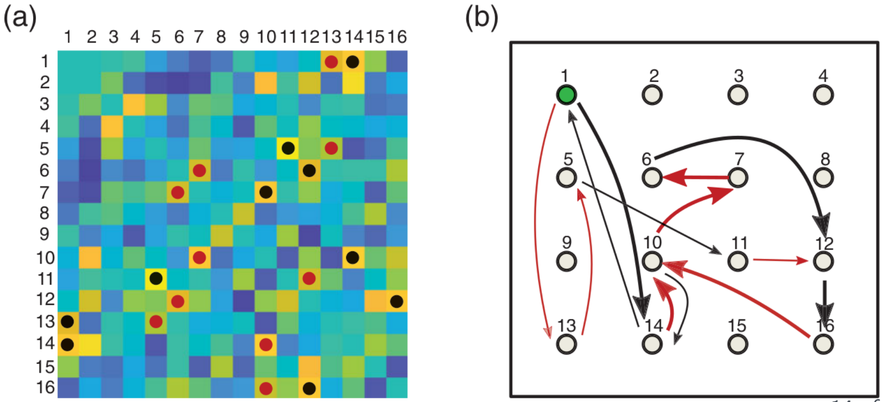
1. Let's consider the rows $\vec{S_n}$ and $\vec{S_k}$, and assume $S_{nk} = \max \vec{S_n}$
2. We have a two-item loop if $S_{nk} = S_{kn} = \max \vec{S_k}$
3. Let's define $\vec{S}_k^-$ as $\vec{S_k}$ without the S_{kn} element
4. We have the loop if and only if $S_{nk} > \max \vec{S}_k^-$ i.e. $\max \vec{S}_n > \max \vec{S}_k^-$
5. $\max \vec{S}_n$ random variable is the maximum of $L-1$ elements and $\max \vec{S}_k^-$ random variable is the maximum of $L-2$ elements

$$\max \vec{S}_n > \max \vec{S}_k^- \quad \text{w.p.} \sim 0.5$$

Random Symmetric SM

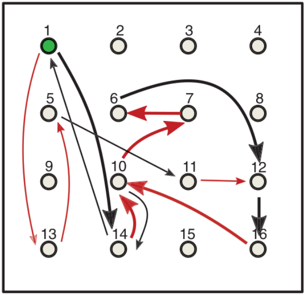
New dynamic!

We prevent the return to the just-retrieved item: when the most similar item is the one just retrieved, the second most similar item is chosen instead



We need to calculate three probabilities

- p_0 transition probability between two items
- p_1 the probability that retrieval proceeds along the previous trajectory in the opposite direction
- p_2 probability of retrieving a new item after an old one.



Calculation of p0, intuitive idea

In the Asymmetric case

$$p_0 \approx \frac{1}{L}$$

But here we prevent to go back to the just previously recalled item, that would happen with probability $1/2$

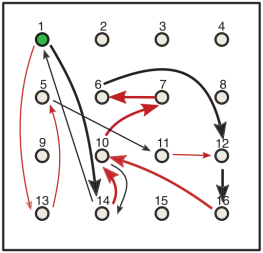
We therefore have

$$p_0 = \frac{1}{2L}$$

Calculation of p1, intuitive idea

We want $S_{nj} = S_{jn} = \max \vec{S}_j$ to be the maximum element of $\vec{S}_n$, i.e. that

$$\max \vec{S}_j > \max \vec{S}_n$$



$$p_1 = P(\max \vec{S}_j > \max \vec{S}_n \mid \max \vec{S}_n > \max \vec{S}_k)$$

$$p_1 = 1/3 \tag{4}$$

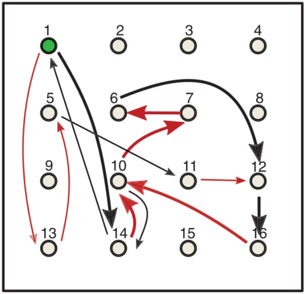
p_2 probability of retrieving a new item after an old one.

Every time we have two possibilities

1. Again go back to the previous item i (1)
2. Open a new path (with probability p_2)

After some algebra one finds

$$p_2 \approx 0.3$$



Random Symmetric SM: p_0, p_1, p_2

Summarizing:

$$p_0 = \frac{1}{2L} \quad p_1 = \frac{1}{3} \quad p_2 \approx 0.3$$

Average number of items recalled before the 1st repetition:

$$\langle \mathbf{k} \rangle_{\mathbf{1st}} \sim \sqrt{\frac{1}{p_0}} = \sqrt{\pi L}$$

Average number of opened cycles:

$$P(l) = \frac{2}{3}\left(\frac{1}{3}\right)^{l-1} \quad \longrightarrow \quad \langle \mathbf{l} \rangle = \frac{1}{1-p_1} = \frac{3}{2}$$

Average number of old items visited before a new cycle is opened

$$\frac{\mathbf{1}}{\mathbf{p_2}} \approx 3$$

In the Asymmetric case we found

$$\langle k \rangle = \sqrt{\frac{\pi}{2p_0}}$$

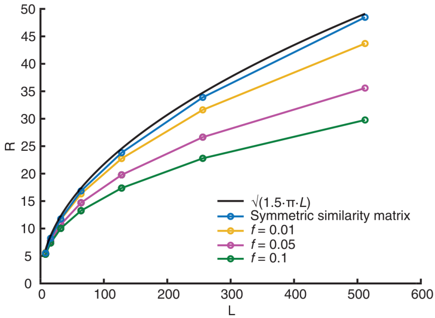
where $p_0 = 1/L$ was the probability to close the cycle (by going to a specific previous item). Here instead we have

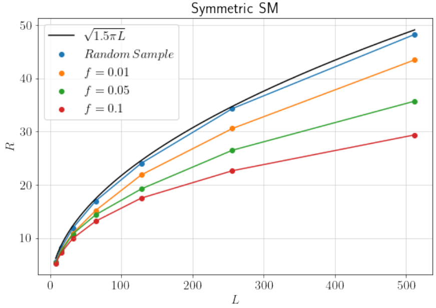
$$p_0(1 - p_1) = \frac{1}{2L}\frac{2}{3} = \frac{1}{3L}$$

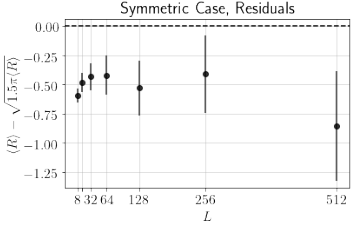
From which

$$R \equiv \langle k \rangle = \sqrt{\frac{3\pi L}{2}} \tag{5}$$

Random Symmetric SM: Results







Slow process. We want to build $\widetilde{S}_{km}$ with the same first and second order statistics of S_{km}. We define the normalized size of the k-th item as

$$M^k = \frac{1}{N} \sum_{i=1}^{N} \xi_i^k$$

that for large N can become

$$M^k \sim \mathcal{N}\left(f, \sqrt{\frac{f(1-f)}{N}}\right)$$

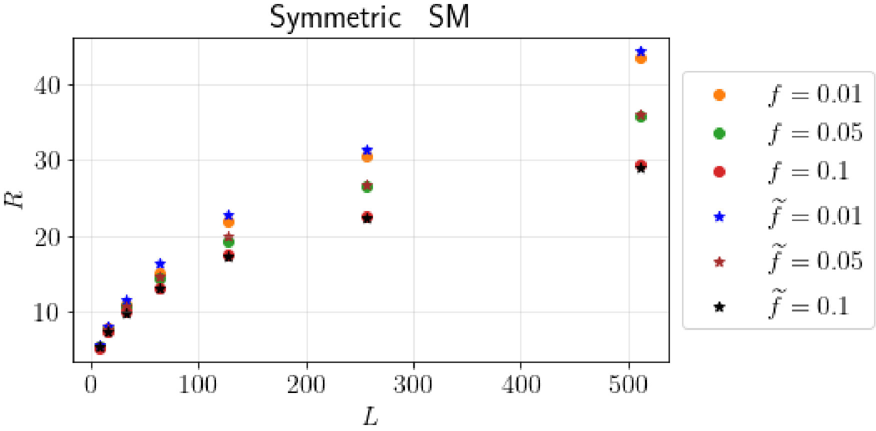
And in order to find the same 2nd order statistics one can define

$$\widetilde{S}_{km} = M^k M^m Z^{km}$$

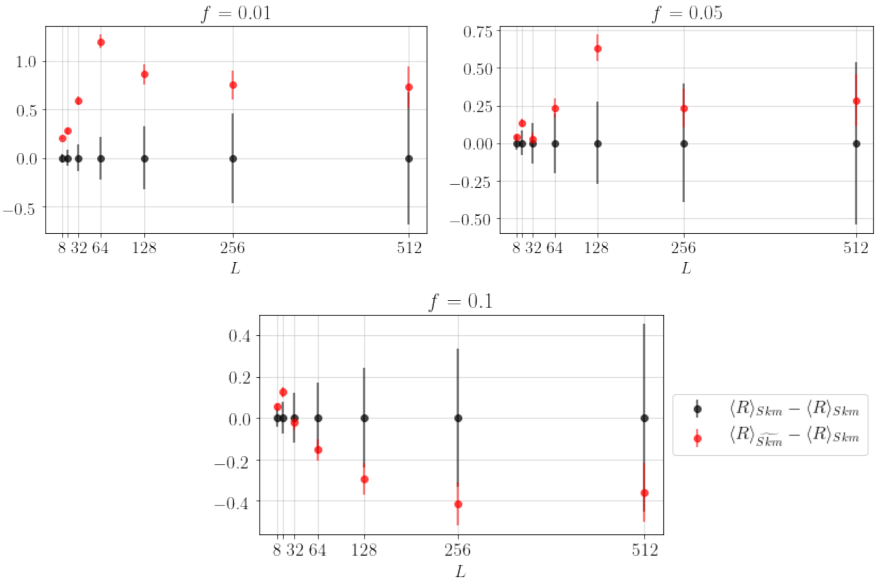
with the symmetric matrix

$$Z^{km} \sim \mathcal{N}\left(1, \frac{1-f}{f\sqrt{N}}\right)$$

One can directly sample $\widetilde{S}_{km}$ in place of building the items to build S_{km}



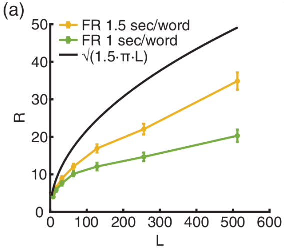
Random Symmetric SM: Equivalent Stat



Model Correction

$$R \approx \sqrt{\frac{3\pi}{2} L}$$

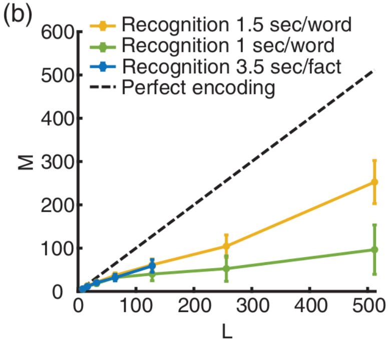
- **No free parameters** $\longrightarrow$ Result of the assumed recall mechanism
- **Universality** $\longrightarrow$ But people differ in terms of their memory effectiveness depending, e.g., on their age and experience.
- At odds with previous experimental studies showing that performance in a free recall task strongly depends on the **experimental protocol** (e.g. presentation rate during the acquisition stage)



Model Correction

It seems reasonable that *acquisition* depends on various factors, such as attention, age of participants, acquisition speed.
So, one should correct the Eq. replacing L with $\mathbf{M}$, **the number of words in memory after the whole list is presented**

$$R \approx \sqrt{\frac{3\pi}{2} M} \tag{6}$$



Model Correction: Recognition Experiment

Recognition experiment:

- Presentation of either words or short sentences expressing common knowledge facts
- At the end of presentation to each participant were shown a pair of words, one from the list just presented and one randomly chosen lure

c fraction of correctly recognized words
M average number of words remaining in memory

Number of errors:

$$L(1-c)$$

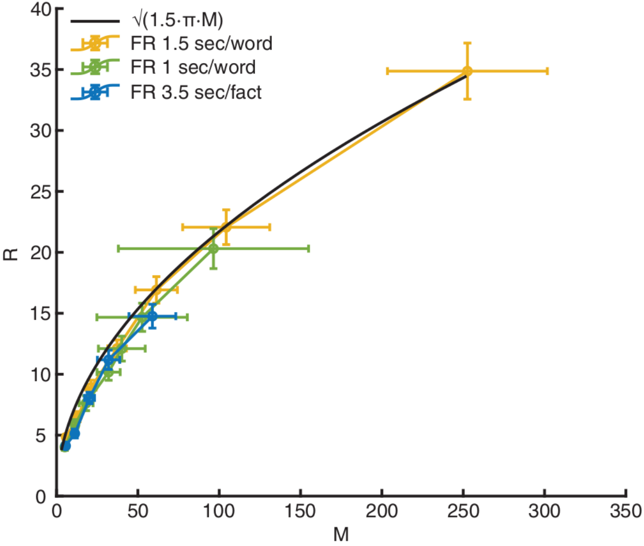
Words really not present in memory

$$2L(1-c)$$

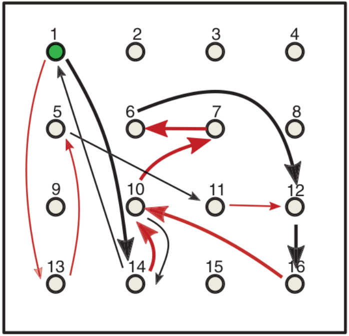
Words effectively remained in memory

$$M = L(2c-1) \tag{7}$$

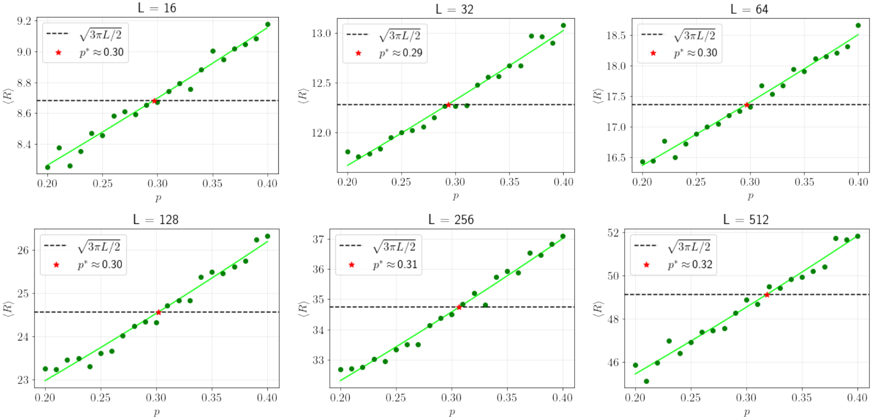
Model Correction: Final Result



1. Same dynamic as before until an element is reached for the second time
2. Jump to a never visited element with probability p^*
3. The process stops when either:
   - A jump fails
   - All the elements are retrieved



New Framework: Results



$$p* \lesssim p_1 = \frac{1}{3}$$

Conclusions

- The designed recall dynamic for a random symmetric SM of size x exhibits the relation
$$R \approx \sqrt{\frac{3\pi}{2}x}$$
- For large N, large L, sparse encoding $f \ll 1$, the SM son of our neuronal representation approximates previous case
- Data confirm with remarkable precision the analytical, parameter-free expression
$$R \approx \sqrt{\frac{3\pi}{2}M}$$
- Idea about the equivalence between paper's model and the jump model

APPENDIX

Random Symmetric SM: p_0

Calculation of p_0

Notation: now with $\vec{S}_n$ we indicate the *relevant elements* of the n-th row vector of the similarity matrix

Let's consider k current element and n an already visited element

$$
\begin{aligned}
p_0 &= P(S_{kn} = \max \vec{S}_k \mid S_{nk} < \max \vec{S}_n) = \\
&= P(S_{kn} = \max \vec{S}_k \mid S_{kn} < \max \vec{S}_n) = \\
&= \frac{P(S_{kn} = \max \vec{S}_k, S_{kn} < \max \vec{S}_n)}{P(S_{kn} < \max \vec{S}_n)} = \\
&= \frac{P(S_{kn} = \max \vec{S}_k) P(\max \vec{S}_k < \max \vec{S}_n)}{P(S_{kn} < \max \vec{S}_n)}
\end{aligned}
$$

Random Symmetric SM: p_0

$$p_0 = \frac{P(S_{kn} = \max \vec{S}_k) P(\max \vec{S}_k < \max \vec{S}_n)}{P(S_{kn} < \max \vec{S}_n)}$$

$$P(S_{kn} = \max \vec{S}_k) = \frac{1}{L-2} \approx \frac{1}{L}$$

$$P(\max \vec{S}_k < \max \vec{S}_n) = \frac{1}{2}$$

$$P(S_{kn} < \max \vec{S}_n) = 1 - P(S_{kn} = \max \vec{S}_n) \approx 1 - \frac{1}{L} \approx 1$$

And by putting together:

$$p_0 \approx \frac{1}{2L} \tag{8}$$

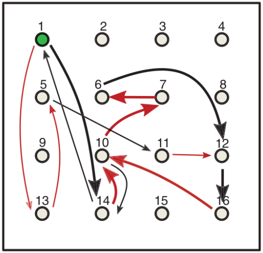
Random Symmetric SM: p_1

Calculation of p_1

$\mathbf{k}$ current element (16)
$\mathbf{n}$ old element, recalled after k (10) $\longrightarrow$ $S_{kn} = \max \vec{S}_k$
$\mathbf{j}$ element preceding n in the first path (14) $\longrightarrow$ $S_{jn} = \max \vec{S}_j$

We have two possibilities

1. Previous $n \to l$ was strong — Loop!
2. Previous $n \to l$ was weak — The strong is the one coming back!



Random Symmetric SM: p_1

We want $S_{nj} = S_{jn} = \max \vec{S}_j$ to be the maximum element of $\vec{S}_n$, i.e. that

$$\max \vec{S}_j > \max \vec{S}_n$$

(remember that S_{nj} is not a member of $\vec{S}_n$)!

We are moving from k to n, and therefore we must impose the condition that before we didn't move from n to k:

$$\max \vec{S}_n > \max \vec{S}_k$$

Therefore

$$p_1 = P(\max \vec{S}_j > \max \vec{S}_n \mid \max \vec{S}_n > \max \vec{S}_k)$$

$$p_1 = 1/3 \tag{9}$$

Random Symmetric SM: p_2

We have come back to n (10) and now we start retrieving old items: now we are at j (14). Every time we have two possibilities

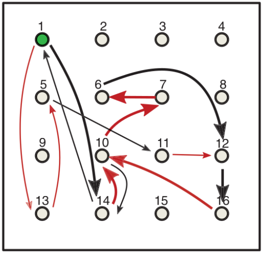
1. Again go back to the previous item i (1)
2. Open a new path (with probability p_2)

We have that

$$S_{jn} = \max \vec{S}_j \quad S_{ij} = \max \vec{S}_i$$

We can't go back to n; therefore our next element is either given by

$$\max{}_2 \vec{S}_j \quad \text{or} \quad S_{ji} = \max \vec{S}_i$$



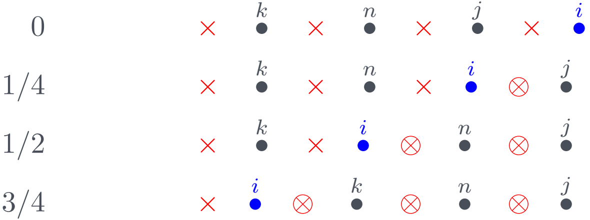
Random Symmetric SM: p_2

Since we want to open a new path, we *don't want* to go back to i, i.e. we want

$$\max_2 \vec{S}_j > \max \vec{S}_i = S_{ji}$$

With the conditions:

$$p_2 = P(\max_2 \vec{S}_j > \max \vec{S}_i \mid \max \vec{S}_j > \max \vec{S}_n, \max \vec{S}_n > \max \vec{S}_k)$$



$$p_2 = 1/4(0 + 1/4 + 1/2 + 3/4) = 3/8$$

$$p_2 \approx 0.3 \tag{10}$$

Notes about the Experiments

- To avoid practice effects, each participant performed a single free recall trial with a randomly assembled list of words of a given length
- Each participant performed one **recognition** and one **recall** trial with lists of the same number of words (but different words between recognition and recall)