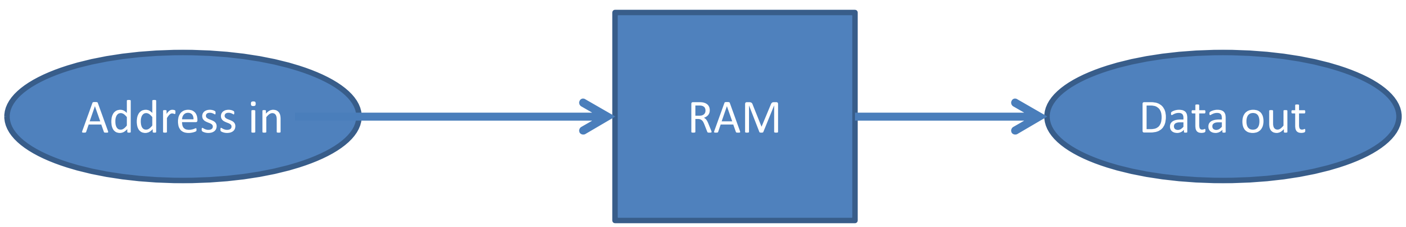
Models of memory

CS786

October 14th 2024

Random access memory

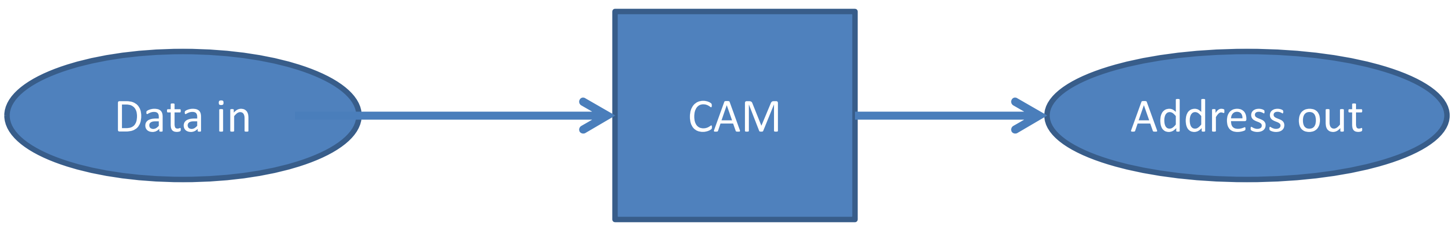
Computer memory can be randomly accessed given address locations



But human memory doesn't have consistent physical addressing. So how does it still work effectively?

Content addressable memory

Content addressable memories are used in some high-speed search operations



Human memory is closer to content addressable than random addressable

https://en.wikipedia.org/wiki/Content-addressable_memory

Retrieval depends on joint encoding

- “Neurons that fire together wire together” – Hebbian postulate
- Any memory encoding must store both the target to be stored and the cue that will trigger its retrieval
- Can’t store single items in memory
- But what are retrieval cues?
  - Hard to identify in natural settings
  - Hard to nail down even in experimental settings
  - In principle, could be any datum of experience

Encoding experiences in memory

- Not as straightforward as putting things in boxes and taking them out later
- How is the experience represented?
- How is it indexed?
- How is it retrieved?
- What factors affect encoding?

Environmental Effects on Encoding

- *Encoding specificity principle* proposes that the cues present during encoding serve as the best cues for retrieval
  - This is why elaborative rehearsal helps memory performance
  - Elaborative rehearsal → plant: tree :: sea: o____, generating targets helps in retention
- *State-dependent memory* is memory that depends upon the relationship of one’s physiological state at the time of encoding and at the time of retrieval
  - Relationship of smell to memory is a common literary trope, see c.f. Proust’s *Remembrance of things past*

Environmental Effects on Encoding

- *Mood-dependent memory* effects attest to the fact that memory is better when a person's mood is the same during encoding and retrieval
  - For example, if you are happy during encoding information, it is easier to retrieve that information if you are happy at the time of retrieval
- *Mood-congruence effect* is the fact that memory is better for experiences that are congruent with a person's current mood
  - For example, when we are sad it is easier to retrieve negative events in our lives

Encoding Specificity

- When conditions of retrieval are similar to conditions of encoding, retrieval is more likely to be successful
- You are more likely to remember things if the conditions under which you recall them are similar to the conditions under which you learned them

Encoding Specificity

- Context effects—environmental cues to recall
- State dependent retrieval—physical, internal factors
- Mood Congruence—factors related to mood or emotions

Encoding-dependent memory tricks

- The *spacing effect* (or distributed study effect) shows that your memory will improve if you study for an exam over an extended time interval rather than just a few days before the exam
  - Can be because studying in a diverse set of circumstances makes more retrieval cues accessible for encoding
- *Overlearning* is studying material past the point of initial learning, and has been demonstrated to aid in retrieval of that information
  - Can be because multiple retrieval cues are associated with the same target, aiding retrieval

Measuring memory performance

- *Recall* is a measure of retrieval that requires the reproduction of the information with essentially no retrieval cues
  - Typical measure = # of items successfully retrieved/# of items on list
- *Recognition* is a measure of retrieval that only requires the identification of the information in the presence of retrieval cues
  - Typical measure = d’ in n-AFC with n-1 non-targets
- *Relearning*, also called the savings method, is a measure of the amount of time saved when learning the information for a second time

Why do we forget?

Sensory memory
The senses momentarily register amazing detail

↓

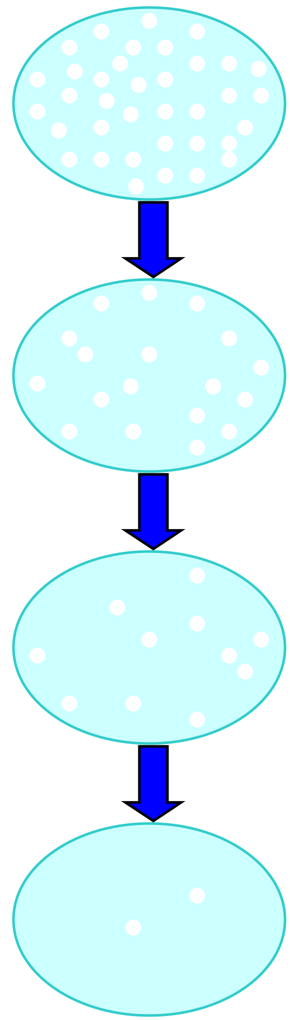
Short-term memory
A few items are both noticed and encoded

↓

Long-term storage
Some items are altered or lost

↓

Retrieval from long-term memory
Depending on interference, retrieval cues, moods, and motives, some things get retrieved, some don't



- Forgetting can occur at any memory stage

Why We Forget

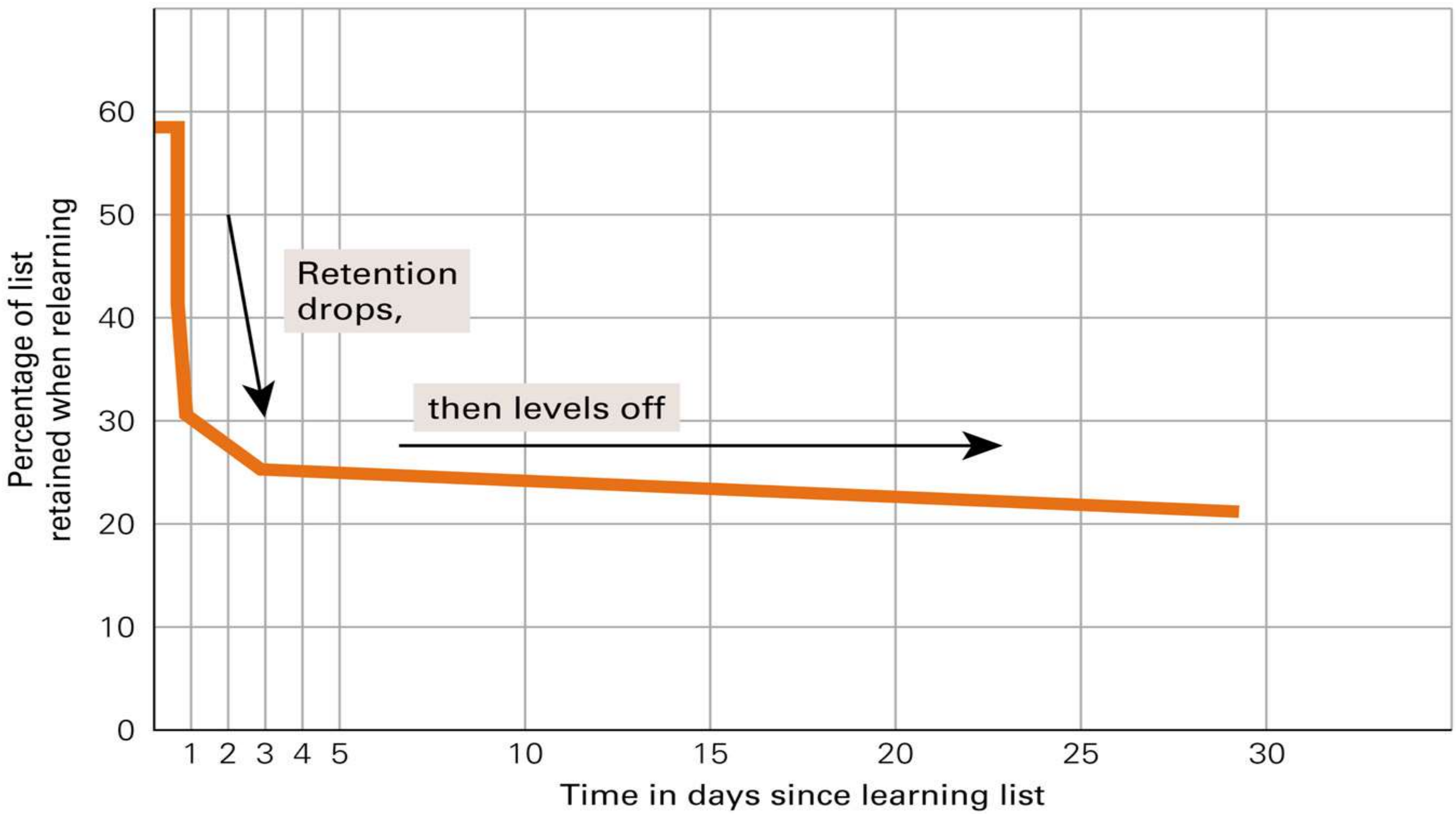
- *Encoding failure :* sometimes forgetting is not really forgetting, information never entered long-term memory in the first place
- *Storage decay theory* suggests that forgetting occurs because of a problem in the storage of the information
  - The biological trace of the memory gradually decays over time and the periodic usage of the information helps to maintain it in storage

Highly recommended extra work: watch Inside Out

A classic relearning experiment

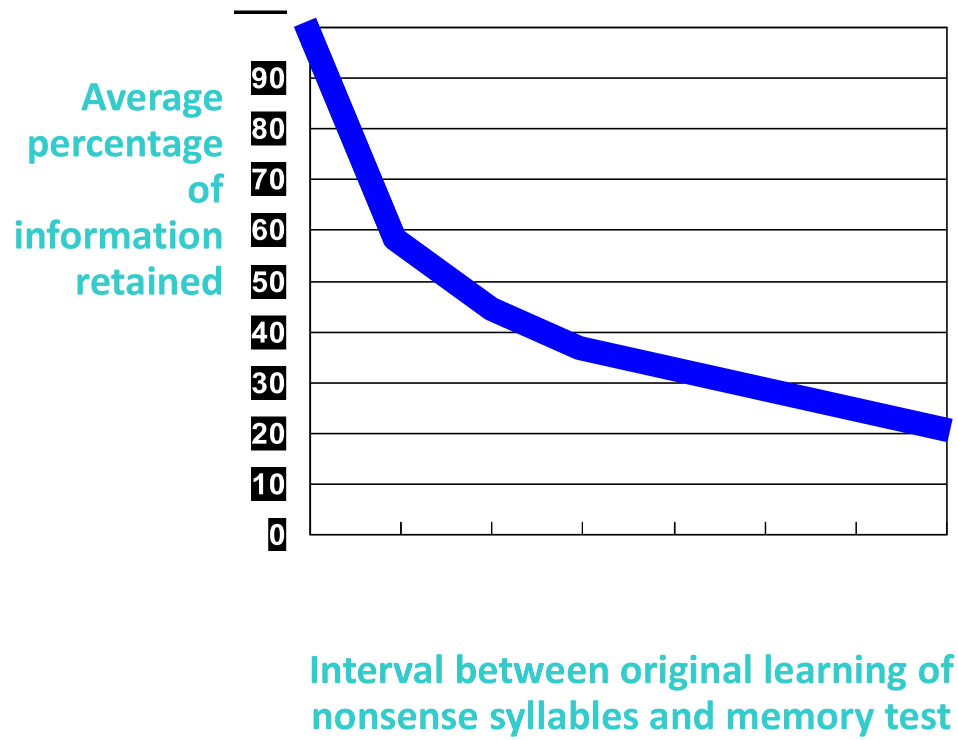
- Ebbinghaus conducted the first experimental studies on human memory more than 100 years ago using the relearning method.
- He would study a list of nonsense syllables until he could correctly recite the complete list without any hesitations. He then put the list aside and waited some period of time and then relearned the list to the same criterion.
- To get a measure of learning, he computed a savings score – the reduction in the number of trials it took him to reach criterion.
- Result? The “forgetting curve” reveals that most forgetting occurs in the first two days after learning material.

Forgetting Curve for Long-Term Memory



Decay Theories

- Memories fade away or decay gradually if unused
- Time plays critical role
- Ability to retrieve info declines with time after original encoding
- Lately subsumed quite well by cue-dependent forgetting theories



Why We Forget

- *Cue-dependent theory* says we forget because the cues necessary for retrieval are not available
  - The information is in memory, but we cannot access it
  - This theory is analogous to knowing a book is in the library but you cannot access it because the library lacks call numbers
- *Interference theory* proposes that other similar information interferes and makes the forgotten information inaccessible

Types of Interference

- *Proactive interference* occurs when information you already know makes it hard to retrieve newly learned information
- *Retroactive interference* occurs when information you just learned makes it hard to retrieve old information

Types of Interference

- Think about changing phone numbers after having a certain number for many years. When asked for your new phone number, remembering the old one interferes with retrieving the new one.
  - This is *proactive interference*
- Now think about being at a party with many people you don't know. You meet someone whom you want to talk to later, but after meeting her, you are introduced to many more people. Now, you cannot remember her name.
  - This is *retroactive interference*

Modeling these facts together

SEARCH OF ASSOCIATIVE MEMORY

Elements of memory models

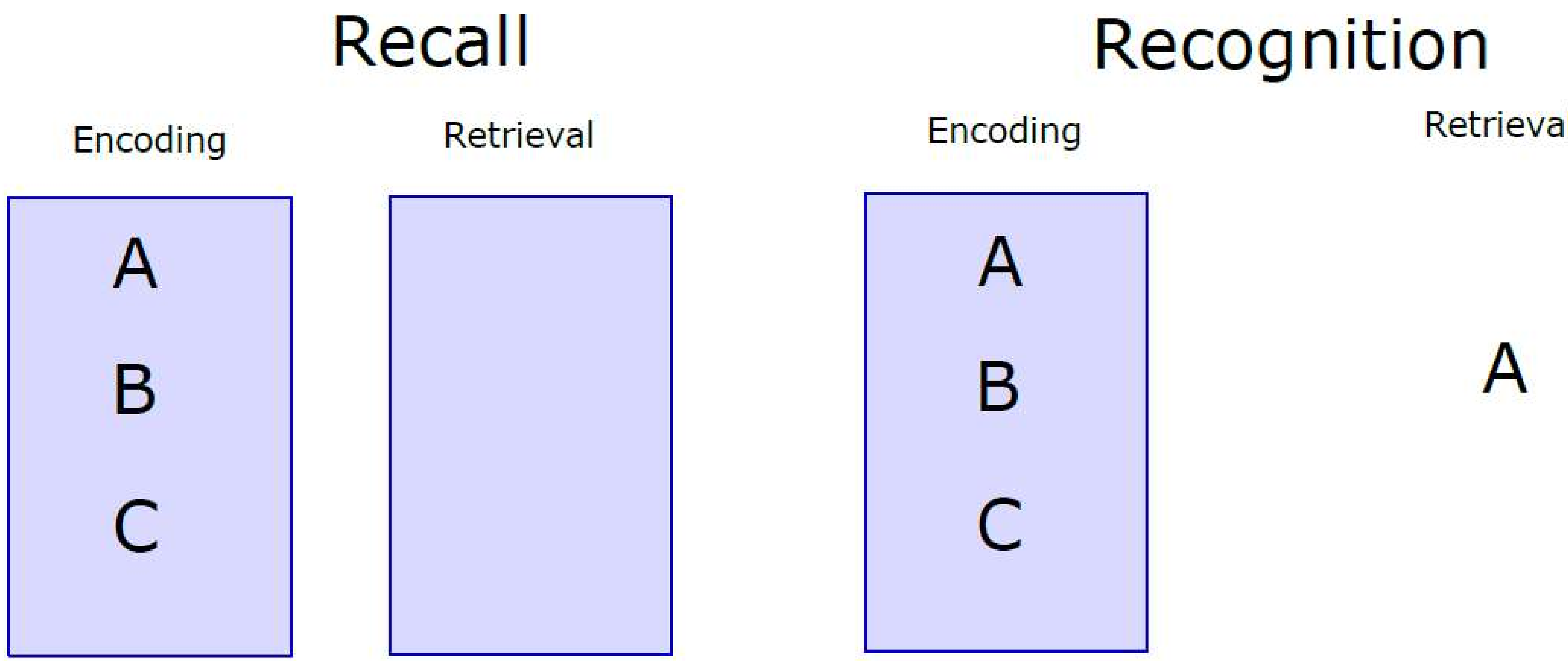
- Search
  - To explain list-length and fan effects
- Direct access
  - To explain rapid true negatives in recognition
- Implicit recognition
  - To explain the mind’s solution to the correspondence problem

https://pdfs.semanticscholar.org/7ebc/6fb776eb2250da7c925946424b9fee37097c.pdf

Modeling retrieval

- Search of associative memory (SAM) was the first modern computational model of memory retrieval (Gillund & Shiffrin, 1984)
  - Assumes that information is stored in a distributed manner as memory ‘images’
    - Images contain item, associative and contextual information
  - Retrieval is modeled as a cue-dependent process
    - Whether an image is retrieved or not is a function of the associative strengths of the retrieval cues to that image
  - Single process model of recall and recognition

Recognition and recall



SAM basic idea

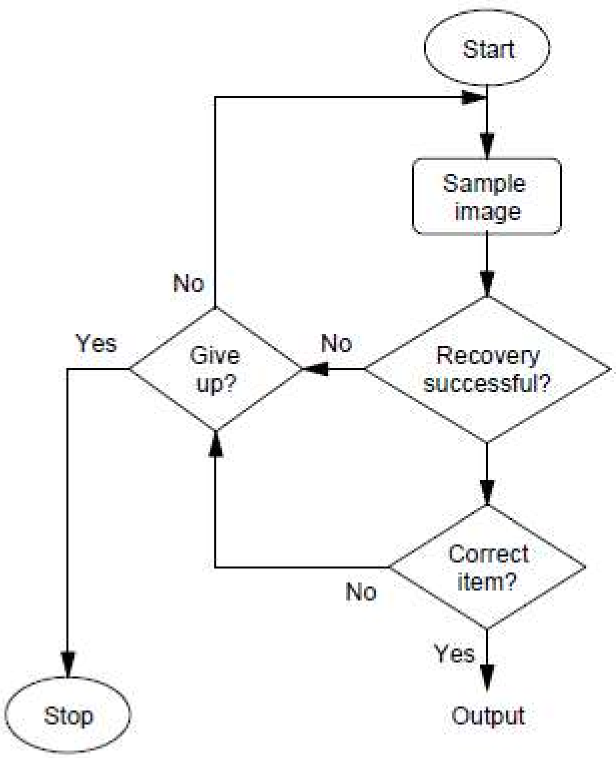


Figure 3: Flowchart representing the SAM retrieval process in a cued recall task.

SAM implementation

- Assume the retrieval probe contains multiple cues Q_1, Q_2, … Q_m
- Combined activation of image I_i is a multiplicative combination of individual image-cue associations

$$A(i) = \prod_{j=1}^{m} S(Q_j, I_i)$$

- Probability of sampling an image is a simple Luce choice rule

$$P_S(I_i) = \frac{A(i)}{\sum A(k)}$$

  – Not the same as the probability of successful memory retrieval
  – An image may or may not activate recovery of the correct target of retrieval

SAM implementation

- Probability of successful recovery

$$P_R(I_i) = 1 - \exp\left[- \sum_{j=1}^{m} S(Q_j, I_i) \right]$$

- Probability of recall success

$$P_{recall}(I_i) = \left[1 - \left(1 - P_s(I_i) \right)^{L_{\max}} \right] P_R(I_i)$$

- If attempt succeeds, association between probe cues and sampled image strengthens
  - SAM had basic learning built into it

SAM recognition algorithm

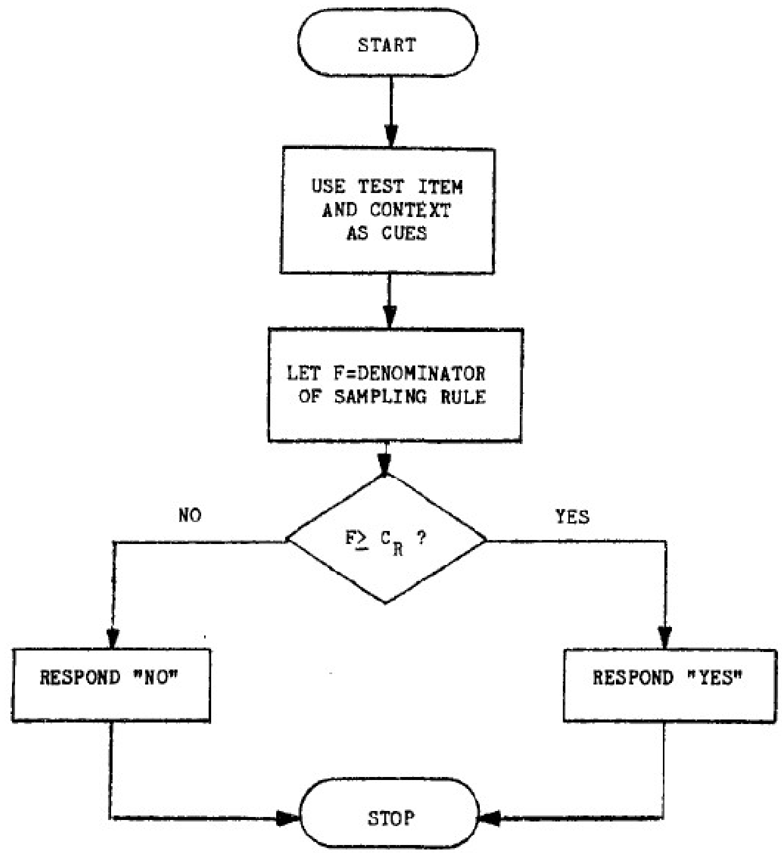


Figure 5. A flow chart for the basic SAM recognition process.

SAM recall algorithm

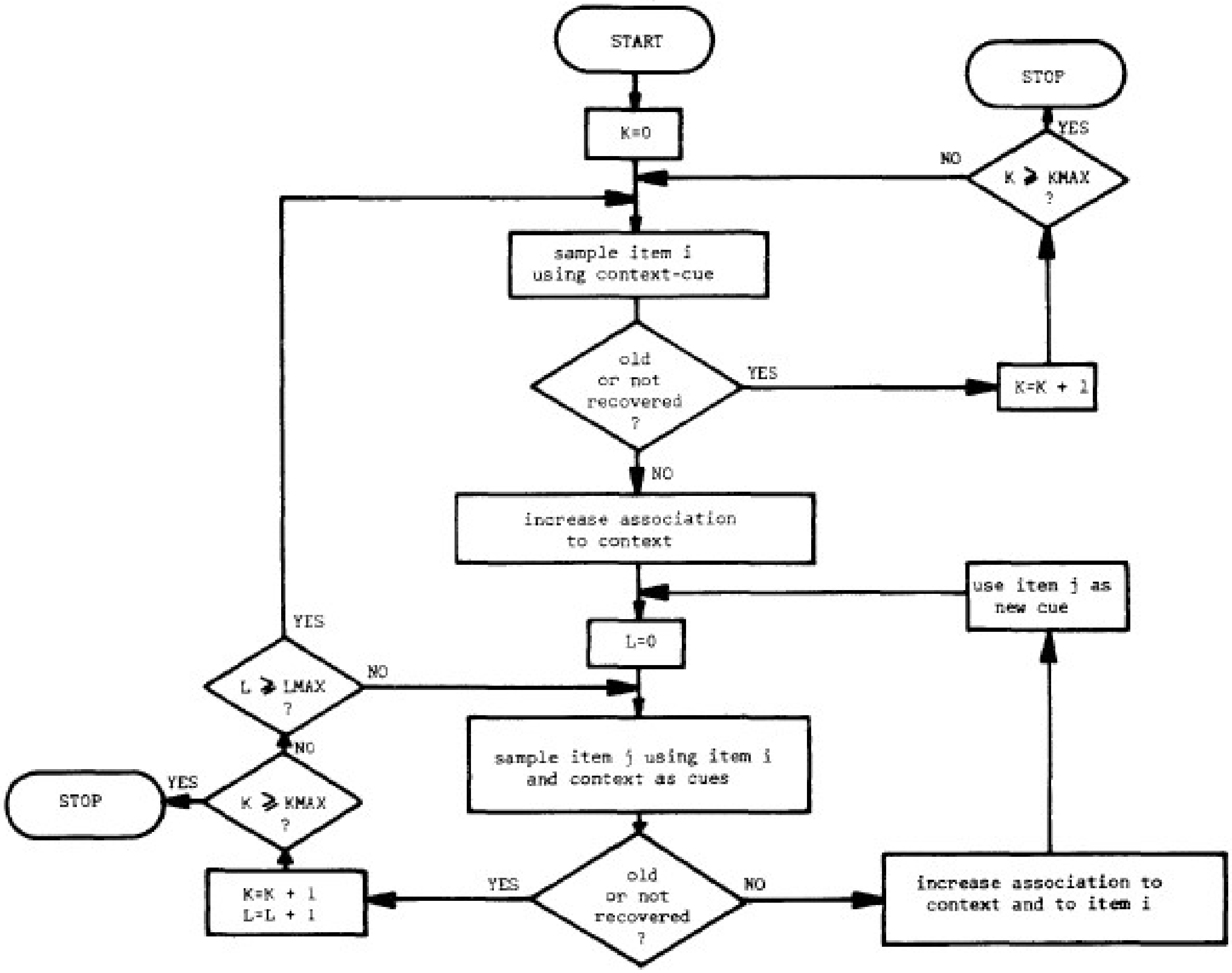


Figure 7. A flow chart of the SAM retrieval process for free-recall (taken from Raaijmakers & Shiffrin, 1980).

Successful predictions

- List length effects in recall
  - Probability of sampling an item decreases as the number of items increases
- Presentation time effects
  - Increased presentation time → stronger associations → better recall and recognition
- Primacy effects
  - In recall because of reinforcement of previous samples
  - In recognition because of the generation of new images corresponding to each successful recognition

Successful predictions

- Part set cuing
  - Cuing set reduces set of items that can be sampled from
  - This is what the browser experiment is testing
- Word frequency effects
  - Recall prefers high frequency items because they have more associations
  - Recognition prefers low frequency items because there is a lower variance in the associative strength distribution for LF items

SAM failures

- Long term recency effects
- Null list length effect in recognition
  - Empirical list length observations in recognition confounded with time to testing
  - Recognition memory for up to 10000 images documented (Standing, 1973; Brady et al., 2014)
- List strength effect
  - Strength of items = presentation time
  - Recognition of strong items in lists of purely strong items better than in mixed lists
  - Recognition of weak items in lists of purely weak items worse than recognition in mixed lists
- Mirror effect
  - Greater true positives and fewer false positives in recognition of low frequency items

Temporal context model

- SAM makes no assumptions about the effect of the environment on retrieval cues guiding the memory process
  - Accepted as inputs
  - Recent retrievals can become cues for subsequent retrievals
- The temporal context model (TCM) changes this
  - Assumes a linear drift of the *temporal* context cue that goes into every episodic memory encoding
  - Recommended reading: (Howard & Kahana, 2002)

TCM encoding

- Items are represented as feature vectors **f**
- Context is also represented as feature vectors **c** – on a different feature space
- Both item and feature vectors are time-indexed
- Construct an item-context mapping via an outer product

$$M^{FC} = \sum_{t}^{List} \mathbf{f}'_t \mathbf{c}_t$$

TCM retrieval

- Retrieval happens via spreading activation
- A state **c** on C will provide activation input $\mathbf{f}^{out} = M^{FC}\,\mathbf{c}$
- Similarity of this input to a given item **f** can be measured as a dot product
- This quantifies the retrieval pull the context exerts on each item

$$a_t = \mathbf{f}^{IN} \cdot \mathbf{f}_t \tag{1}$$

$$= M^{FC}\mathbf{c}_{t'} \cdot \mathbf{f}_t \tag{2}$$

$$= \sum_k (\mathbf{c}_k \cdot \mathbf{c}_{t'})\, \mathbf{f}_k \cdot \mathbf{f}_t \tag{3}$$

$$= \sum_k (\mathbf{c}_k \cdot \mathbf{c}_{t'}) \tag{4}$$

Follows from **f** orthonormality (assumed)

The context drift assumption

- Assume a linear drift in context

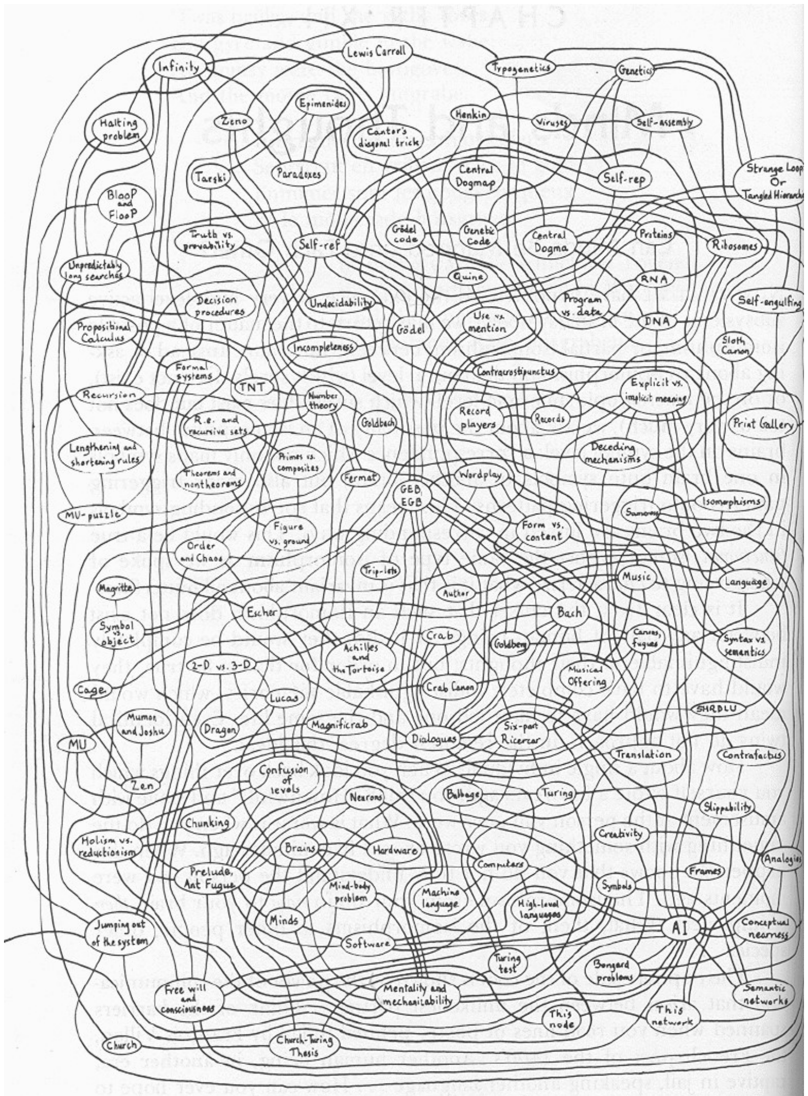
$$\mathbf{c}_t = \rho\ \mathbf{c}_{t-1} + \mathbf{c}_t^{in}$$

- A little bit like a recurrent network
- Naturally makes contexts at closer times more similar than contexts at farther times from the probe point
- Yields long-term recency predictions

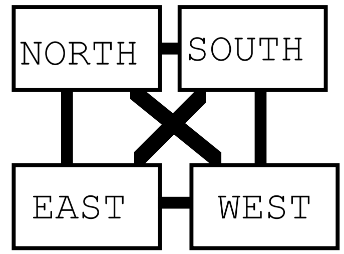
Search of associative memory

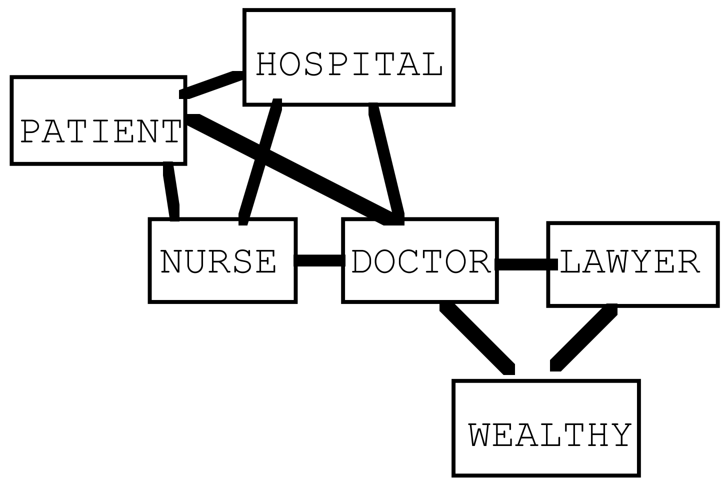
- If we can assume that we know the associative strength between all possible targets and cues
  - We can predict various experimental outcomes in memory experiments
- But how do associative strengths get to be the way they are in your head?

Semantic Networks



Hofstadter. Godel, Escher, Bach.

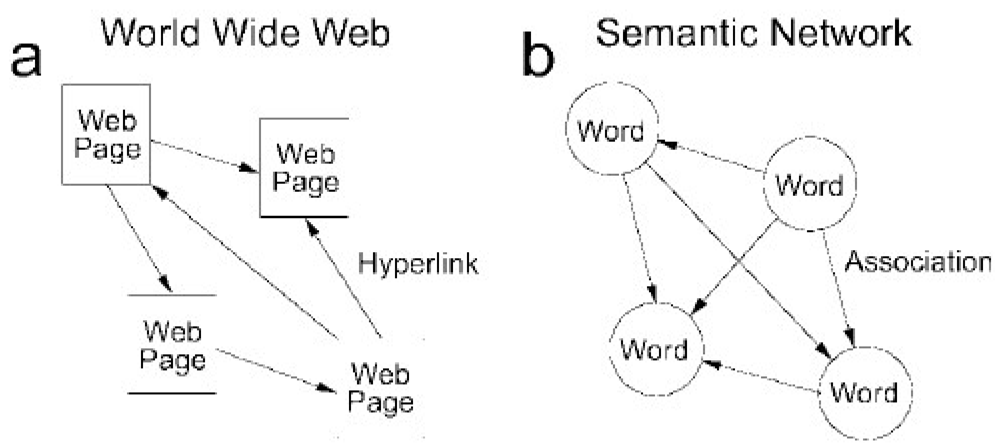




Explains everything and predicts nothing

Can we be more precise in dealing with semantic networks?

- Griffiths and Steyvers made a great observation
  - Search in the semantic network has the same information-theoretic constraints as relevance search on the internet



Recommended reading: Griffiths, Steyvers & Firl. Google in the mind.

Basic idea

- Use PageRank to predict word completing task performance
- PageRank:
  - Consider the adjacency matrix of all web pages L
  - Important websites are linked to by other important websites
  - Consider each link to be transporting some of each website's importance to the outbound connection
  - Solve for importance; list websites containing search term in order of importance

Memory hypothesis

- All brain areas stimulated by a particular retrieval cue constitute nodes in a graph
- Consider the adjacency matrix of this graph, measure in terms of synaptic connectivity
- <u>Consider accessibility of a memory engram as the equivalent of website ‘importance’</u>
- We have a correspondence between web and memory search

Word completion task

- Given a letter, come up with as many words as you can that start with that letter

Beginning letter						
A	B	C	D	P	S	T
Human responses						
Apple (25)	Boy (11)	Cat (26)	Dog (19)	People (5)	Snake (11)	Tea (5)
Alphabet (7)	Bat (6)	Car (8)	Dad (16)	Penguin (3)	Stop (4)	Television (5)
Ant (6)	Banana (5)	Cool (3)	Door (5)	Pizza (3)	Saw (2)	Time (4)
Aardvark (3)	Balloon (4)	Card (2)	Down (4)	Play (3)	Sea (2)	Tree (4)
Ace (2)	Book (4)	Class (2)	Dark (3)	Pop (3)	Sex (2)	Table (3)
Ambulance (2)	Baby (3)	Coke (2)	Dumb (3)	Puppy (3)	Silly (2)	Tall (3)
Animal (2)	Ball (2)	Cookie (2)	Day (2)	Piano (2)	Sister (2)	Tank (3)
Absence (1)	Barn (2)	Crack (2)	Devil (2)	Pie (2)	Sit (2)	Telephone (3)
Acrobat (1)	Bear (2)	Cross (2)	Dinosaur (2)	Pig (2)	Slither (2)	Town (3)
Act (1)	Beef (2)	Cut (2)	Do (2)	Power (2)	South (2)	Train (3)
PageRank						
Animal (2)	Big (0)	Cold (0)	Dog (19)	Pretty (0)	Small (1)	Time (4)
Away (0)	Bad (1)	Car (8)	Dark (3)	People (5)	Sad (1)	Tall (3)
Air (0)	Boy (11)	Cat (26)	Drink (1)	Paper (0)	School (0)	Talk (1)
Alone (0)	Black (0)	Color (0)	Down (4)	Pain (0)	Sun (2)	Tree (4)
Apple (25)	Beautiful (0)	Clothes (0)	Death (1)	Puppy (3)	Smile (0)	Tired (0)
Arm (0)	Blue (2)	Child (1)	Door (5)	Person (1)	Stop (4)	Tiny (0)
Ache (0)	Book (4)	Cute (0)	Day (2)	Play (3)	Soft (1)	Thin (0)
Answer (1)	Body (0)	Clean (0)	Dirty (0)	Place (1)	Sex (2)	Top (1)
Apartment (0)	Bright (0)	Close (0)	Dirt (0)	Party (0)	Sky (0)	Together (0)
Alcohol (0)	Baby (3)	Cry (0)	Dead (0)	Pen (0)	Sleep (0)	Train (3)

How to model this?

- PageRank-like associativity is the outcome
  - What is the process?
- One possibility
  - Activation spreads from node to node along associative links
  - Assume each node spreads its activation equally over all nodes it is connected to
  - New activation = old activation – decay + inputs

Modeling formally

- Assume the vector **x** is activation for all nodes

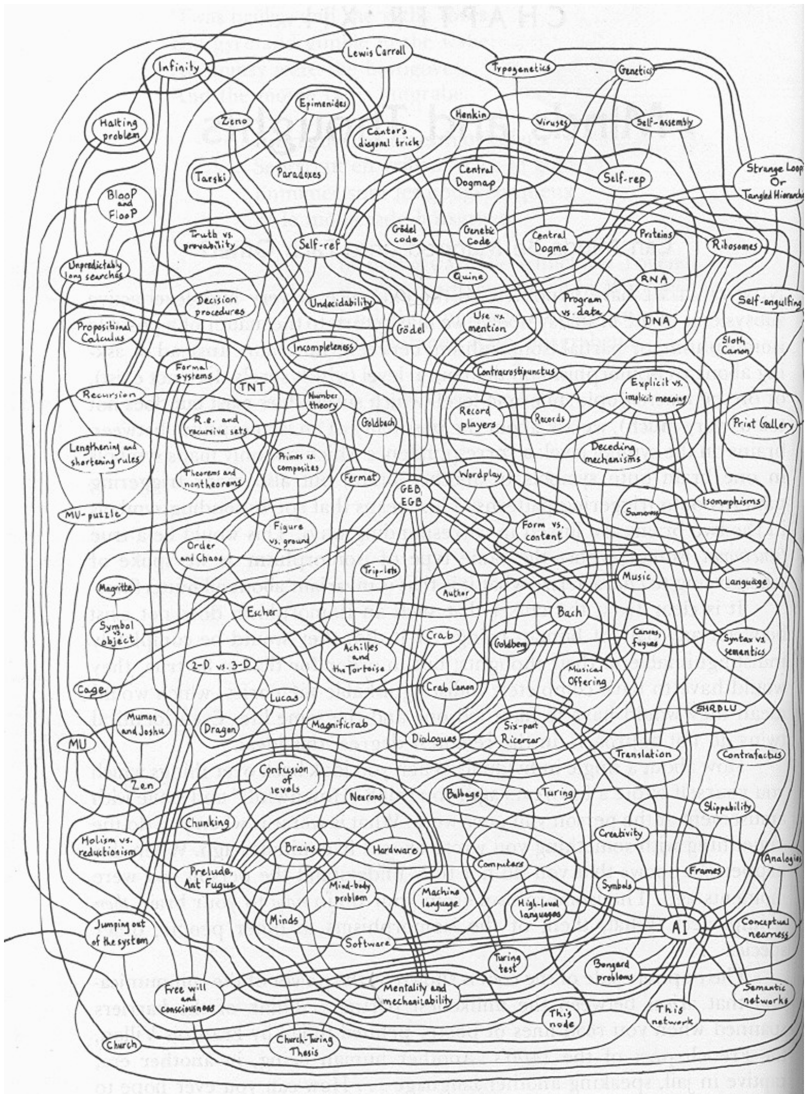
$$\mathbf{x}^{(t+1)} = \alpha \mathbf{x}^{(t)} + (1-\alpha)\mathbf{M}\mathbf{x}^{(t)}$$

- Here **M** is a matrix whose entries are

$$\mathbf{M}_{ij} = \frac{\mathbf{L}_{ij}}{\sum_{k=1}^{n} \mathbf{L}_{kj}}$$

- **L** are binary outbound links in the graph

Semantic Networks



Hofstadter. Godel, Escher, Bach.

- Can say something about how the semantic network comes about
- Spreading activation from node to node brings the graph into its present shape

Some predictions are possible

Exploration in the semantic network

- Exploration of memory is affected by the familiar exploration-exploitation tradeoff
  – But how? Search in memory is impossible?
  – By manipulation of cues
- What sort of effect can environmental factors or previous tasks have on memory exploration?

https://warwick.ac.uk/fac/sci/psych/people/thills/thills/2015hillstics.pdf

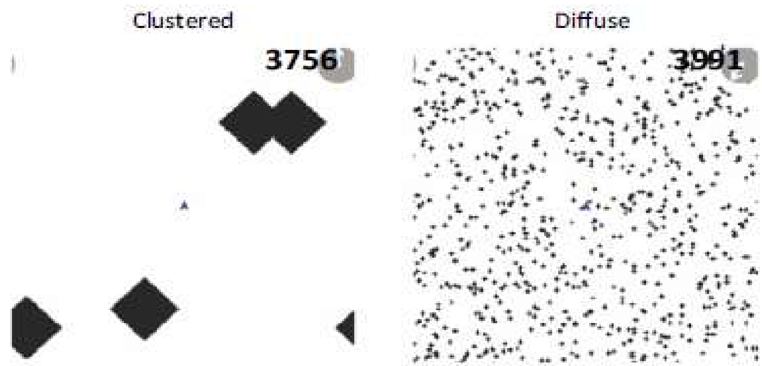
Thomas Hills' memory search

- Optimal foraging theory
  - Animal spends some time looking for food
  - When it finds a patch of food, the rate of food acquisition drops over time
  - Goal is to maximize rate of food accumulation
  - Optimal theory predicts: shift to a new patch when food acquisition rate drops below global mean

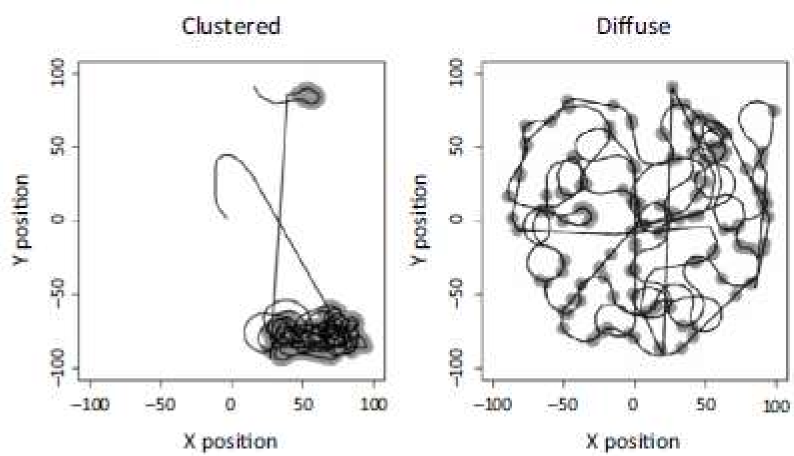
Animal foraging	Exploiting a known berry bush versus exploring for new bushes
Visual search	Analyzing one spot on a chest radiograph versus looking for the next spot to check
Information search	Searching within a document versus searching for new documents; deciding when to accept an item on a menu versus continuing to look for new items
Search in memory	Trying to remember more African animals versus switching to Australian animals
Search in problem solving	Focusing on solutions that have worked in the past versus seeking new solutions
Social (group) learning	Learning or copying existing knowledge versus using innovation to seek new knowledge

A spatial foraging task

- Experimental task: find hidden areas of high reward

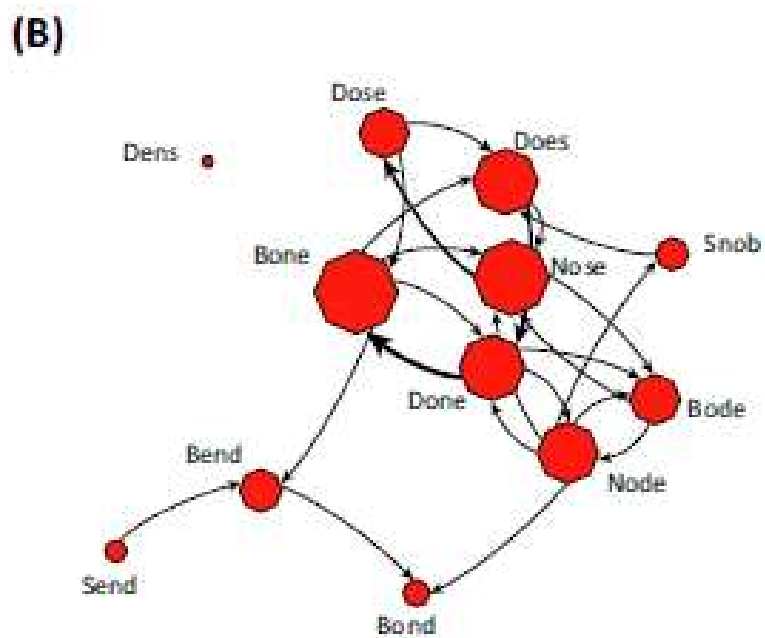


- Strategies differ

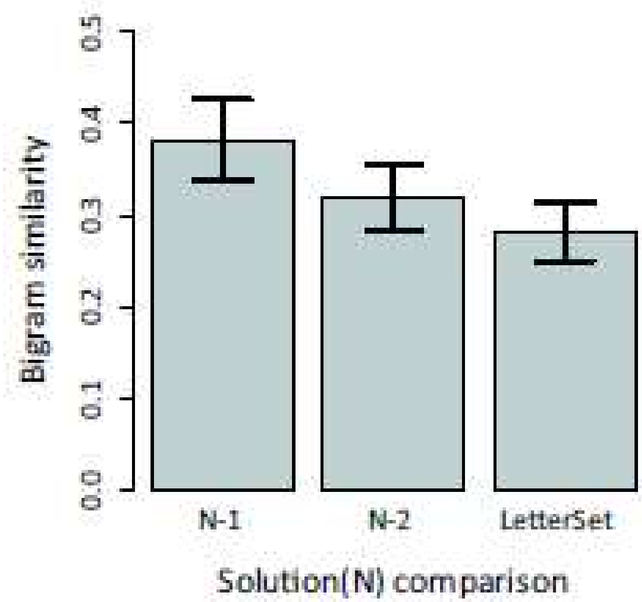


Same participants do a memory task

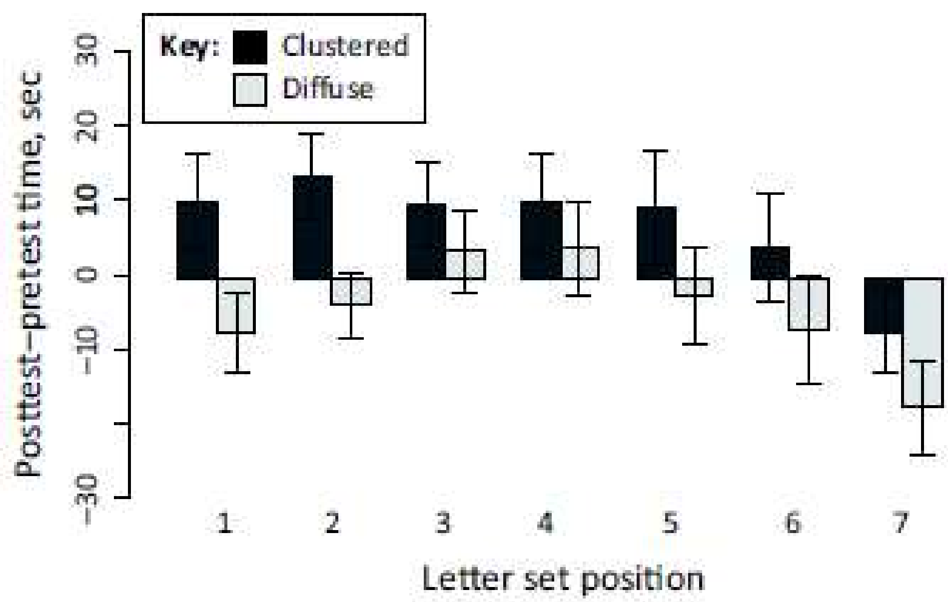
- Scrabble: find all words in the letter set NSBDOE



Word production shows sequential dependence



Previous task appears to control exploration propensity



Semantic network traversal is cognitively controllable

The mind-wandering mystery

- The literature in this sub-area is very crowded
- How does spontaneous retrieval from memory occur?
- What characterizes mind-wandering? (link)
- What is the relationship between memory retrieval and anxiety? (link)