Serial Computer Architecture

- Many types and theories of memory in cognitive science coming from an information processing view, based on a metaphor to serial processing computer architectures.

Information Processing View

- The various types of computer memory stores map directly onto hypothesized human memory capabilities, as do the rehearsal, retrieval and loss processes.
Sensory Registers

- Different sensory registers exist for different sensory modalities, but the most studied are:
  - iconic register for visual sensation
  - echoic register for auditory sensation
- Two fundamental questions:
  - how much visual information can the iconic register store?
  - how long can the iconic register maintain this information?
Measuring Iconic Sensory Register Capacity

- A problem with measuring the capacity of the iconic register is that the information is lost before people can report its existence.
- Sperling (1960) developed an experimental approach that overcomes this problem.

<table>
<thead>
<tr>
<th>Display (50 msec)</th>
<th>Tone (after delay)</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>R G C L X N S B J</td>
<td>High Medium Low</td>
<td>If low tone was sounded “S, B, J”</td>
</tr>
</tbody>
</table>

Sperling Task Results

- Shows the number of letters able to be recalled as a function of the delay of the signalling tone.

Measuring Echoic Sensory Register Capacity

- The auditory equivalent of the Sperling task was conducted by Darwin, Turvey and Crowder (1972).
  - Presented a 3x3 ‘matrix’ of auditory information, by having stimuli read into left ear, right ear, or both.
  - Used a visual cue, in the form of a bar, to identify the required response.

<table>
<thead>
<tr>
<th>Left Ear</th>
<th>Both Ears</th>
<th>Right Ear</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>8</td>
<td>F</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>R</td>
</tr>
<tr>
<td>L</td>
<td>U</td>
<td>10</td>
</tr>
</tbody>
</table>
Sensory Register Capacities

- Even using the Sperling methodology, it is difficult to measure the information and time capacity of sensory registers, because of variations due to:
  - the individual
  - the nature of the sense-data
  - the context in which it is presented

Short Term Memory

- Short term memory holds in consciousness a small amount of the information in the sensory registers
  - e.g., remembering a phone number long enough to dial, but not other visual details of the phone book entry
  - typically information is held in short term memory for about 20 to 30 seconds
  - this can be extended by making deliberate efforts to repeat or rehearse the information

Capacity of Short Term Memory

- Ebbinghaus (1885) noted that STM seemed to have a limited capacity of about 7 items
- Digit span tasks are a common experimental approach for measuring STM capacity
- Miller (1956) argued that the capacity of STM is 7 ± 2 “chunks”
  - a “chunk” is a single meaningful unit of information
  - e.g., “14921914” is 8 chunks if you use the digits as basic units, but only two if it is parsed “1492” and “1914”
Expert and Novice Memory in Chess

- de Groot (1965, 1966) found that
  - Chess experts consider no more (perhaps fewer) moves than novices
  - Both experts and novices were equally bad in reconstructing random chess boards
  - Experts could reconstruct about 20 pieces of a realistic chess board, compared to 4 or 5 for novices
- This suggests expertise is based on representational chunking

Representational Chunking and Expertise

- Chase and Simon (1973) studied the nature of the reconstructed patterns for experts, and found they represented meaningful game configurations
  - Simon and Gilmar (1973) estimated that masters acquired of the order of 50,000 different chess patterns

Expert Memory

- Superior expert memory has been demonstrated in
  - the game of Go (Reitman, 1976)
  - bridge hands (Charness, 1979; Engle & Butskel, 1978)
  - electronic circuit diagrams (Egan & Schwartz, 1979)
  - computer programming (Shneiderman, 1976)
Working Memory

- Working memory is an alternate conceptualization of STM
  - arose from criticism of the serial nature of the information processing view, which assumes information cannot be stored in LTM without passing through STM
  - idea is we have different memory systems for different types of cognitive processing
  - working memory is for the temporary storage of information, and the active processing of information to achieve immediate goals
  - chunking provides evidence of the interaction between working memory and LTM

Rehearsal in Working or Short Term Memory

- Without any deliberate action, new material tends to replace the old in STM
- Rehearsal of information in STM may be used to preserve older information at the expense of the new:
  - **maintenance rehearsal**: literally and passively repeating the information in its original form
  - **elaborative rehearsal**: actively thinking about the information while rehearsing

Retrieval from STM

- Sternberg’s (1966) reaction time experiment for studying STM retrieval
  - presented a stimulus set in the form of a short list of numbers, faces, words, etc.
  - gave participants time to study the stimulus set, until it was remembered
  - measured reaction time for correct answers

<table>
<thead>
<tr>
<th>Memory Set</th>
<th>Q1</th>
<th>Q2</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 3 6 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was 6 in the set?</td>
<td>Was 9 in the set?</td>
<td></td>
</tr>
</tbody>
</table>
Basic Findings from the Sternberg Task

- There is always a 400 msec time associated with producing the response behavior.
- Each additional list item contributes about another 40 msecs.
- There is no difference between items in the list, and those not in the list—implies search is exhaustive (or parallel).

Functionalism and the Sternberg task

- Two regularities: increased time with stimuli and independence of answer.
- Two models: serial: time handled naturally, but need exhaustive search; parallel: independence handled naturally, but need to explain increase in time.

Long Term Memory

- Long term memory is a very long term, sometimes lifelong, information store with a large capacity.
- Information held in LTM must be retrieved into STM (or working memory) before it can be used.
- STM is: brief, limited in capacity, easy to access.
- LTM is: enduring, virtually limitless in capacity, harder to access.
Forgetting from LTM

- Much research has measured the rate at which information is lost from LTM (Rubin et al. 1996)
  - Known as the curve of forgetting, or the retention curve
  - Generally, these curves show rapid initial loss of information, followed by a more gradual decline
- This is one area of memory research that has examined both humans and various animals (e.g., pigeons, rats)

![Graph showing the curve of forgetting](image)

Forgetting Mechanisms

- There are different (competing?) accounts of the way in which information is forgotten, based on
  - decay: information decays from memory, and is eventually lost unless it is periodically used
  - emphasis on time
  - interference: information is lost because other information interferes
  - emphasis on information
  - context: information is not remembered because of a difference between the context in which it was encoded, and the one in which it is trying to be retrieved

Time Capacity of LTM

- The time capacity of LTM was studied by Bahrick, Bahrick and Wittlinger (1975)
  - tested 392 US high school graduates on their memory for the names and faces of classmates, using yearbooks
  - tested at 9 retention intervals ranging from 3 months to almost 48 years
- Basic findings:
  - face recognition level was about 90% over 34 years
  - name recognition and matching started at similar levels, but declined after 15 years
  - free recall performance was less impressive
Information Capacity of LTM

- The information capacity of LTM was studied by Shepard (1967):
  - presented participants with 612 magazine pictures
  - after various delays, tested by showing pairs of pictures, one new and one old
  - participants were able to identify the old picture almost 100% of the time, even after several days
- Nickerson (1965, 1986) and Standing (1973) have reported similar results using 10,000 pictures

Interference

- A distinction is made between two types of interference
  - retroactive interference: when the presence of new information interferes with the retrieval of old information
    - trying to remember an old cell number for a friend, but failing because you’ve learned their new number
  - proactive interference: when previously stored information interferes with the retrieval of new information
    - trying to learn to ski, but being worse than someone new to the snow, because you already know how to snowboard, and it interferes with you learning to ski

Evidence for Interference

- Paired-associate learning provides one source of evidence for the role of interference in forgetting
- A typical cued recall study involves an experimental group and a control group, with test performance being worse for the experimental group because of interference
  - The experimental design below tests for retroactive interference

<table>
<thead>
<tr>
<th>Experimental Group</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learn A-B</td>
<td>Learn A-B</td>
</tr>
<tr>
<td>Learn A-D</td>
<td>Learn C-D</td>
</tr>
<tr>
<td>Test A-B</td>
<td>Test A-B</td>
</tr>
</tbody>
</table>
Context and Forgetting

- There is evidence that changes in context affect whether information is remembered.
- Smith (1979) had subjects learn a list of 80 words, then do a surprise recall test the next day under various conditions:
  - Same room (18.0)
  - Different room, no reinstatement (12.0)
  - Photos of study room (18.8)
  - Think about study room (17.2)
  - At home, thinking about home (9.6)
- The more context is reinstated, the better the performance in recalling words.

Prospective Memory

- Most memory studies have focused on retrospective memory for events that have happened in the past.
- It is also possible to study prospective memory for events that will happen in the future.
- Prospective memory places a heavy emphasis on time, since people need to remember what it is they have to do, when they have to do it.
- Prospective memory requires different forgetting characteristics than retrospective memory:
  - e.g., you want to remember a shopping list as well as possible, until you have done the shopping; then, you generally want to forget most of it.

Rational Forgetting?

- Decay, interference and context theories of forgetting sit at the algorithmic level in Marr’s hierarchy for analysis:
  - Focus is on inefficiencies in encoding or retrieval, assuming the goal is to be able to remember everything perfectly all the time.
- A computational level explanation for forgetting might argue it is an appropriate goal:
  - A perfect memory counteracts generalization and inductive reasoning.
  - It is efficient for our memory system to make recent and frequent memories more readily accessible (Anderson & Schooler, 1991).
- Is it desirable for an intelligent agent, in a changing world, to remember everything, with no forgetting?
  - There are clinical cases of patients much like this: Luria (1975) described S.V. Shereshevski, a Russian journalist with a seemingly unlimited memory.
Recognition, Recall, Reconstruction Tasks

- Different types of demand on memory lead to different retrieval processes, and different tasks for studying memory
  - Remembering or reconstruction is putting together previously encountered information in full
  - Recall is spontaneous conscious recall of individual pieces of information
  - Recognition is feeling of remembrance that something has previously been encountered
    - the tip-of-the-tongue phenomenon suggests that recall is harder than recognition

Remembering and Reconstruction

- Carmichael et al (1932) gave two groups of people the same images, but different groups got different verbal labels
  - Studied their reconstructions of the images from memory

Remembering and Reconstruction

- Shows (again) the top-down influence of cognition on perception
  - Also highlights the difficulty of quantitative measurement of reconstruction tasks
Recognition Memory Task

• Recognition memory tasks are probably the most straightforward experimental way to study memory
  − Lend themselves to easy and direct quantification of performance
• Words are probably the most commonly used stimuli, and allow for some independent variables to be manipulated easily
• A typical task involves two stages
  − A study phase where participants are presented with a list of words, and have to try to remember them
  − A test phase where participants are presented with a list of words, and have to decide whether each word is
    − On the study list (an “old” word, also known as a “target”)
    − Not on the study list (a “new” word, also known as a “distractor”)

Measuring Recognition Performance

• Accuracy is just a count of the number of correct decisions
  − e.g., measure that a subject made 30 correct decisions on the 40 test stage questions in a condition, and so their accuracy was 75%
  − a reasonable measure, but has limitations
• Compare the following two possibilities
  − Subject A responded “new” to every word in the test stage
  − i.e., they never recognized any of the 40 test words as being on their study list, even though 20 were
  − Subject B responded “old” to every word in the test stage
  − i.e., they thought they had seen all 40 test words on their study list, even though 20 were not
• These subjects seem to have very different memory recall properties, abilities, or performances
  − Accuracy would give them both the score of 50%

Signal Detection Measures

• A more sophisticated way to measure performance in the test stage is to think about the four possibilities for each decision
  − This approach is consistent with “Signal Detection Theory”
• Each word can either really be an old word or a new word, and the subject can either say it was an old word or a new word

<table>
<thead>
<tr>
<th>Subject’s Decision</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLD</td>
<td>Hit</td>
<td>Miss</td>
</tr>
<tr>
<td>NEW</td>
<td>False Alarm</td>
<td>Correct Rejection</td>
</tr>
</tbody>
</table>

“Truth”
**Hits, Misses, False Alarms, Correct Rejections**

- **Hit**: The word really was on the study list, and the subject recognized it.
- **Miss**: The word really was on the study list, but the subject did not recognize it.

<table>
<thead>
<tr>
<th>Subject's Decision</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLD</td>
<td>Hit</td>
<td>Miss</td>
</tr>
<tr>
<td>&quot;Truth&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW</td>
<td>False Alarm</td>
<td>Correct Rejection</td>
</tr>
</tbody>
</table>

**False Alarm**: The word was not on the study list, but the subject thought they recognized it.

**Correct Rejection**: The word was not on the study list, and the subject correctly did not recognize it.

<table>
<thead>
<tr>
<th>Subject's Decision</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLD</td>
<td>Hit</td>
<td>Miss</td>
</tr>
<tr>
<td>&quot;Truth&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW</td>
<td>False Alarm</td>
<td>Correct Rejection</td>
</tr>
</tbody>
</table>

**Alternative Terminology**

- A common alternative set of names involves “true” and “false” “accept” and “reject” decisions.
  - This naming highlights which decisions are correct, and which are mistakes.

<table>
<thead>
<tr>
<th>Subject's Decision</th>
<th>OLD</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLD</td>
<td>True Accept</td>
<td>False Reject</td>
</tr>
<tr>
<td>&quot;Truth&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEW</td>
<td>False Accept</td>
<td>True Reject</td>
</tr>
</tbody>
</table>
Recall Tasks

- One source of evidence comes from recall tasks
  - In a first study phase, a list of items, usually words, are presented
    - e.g., “alcohol, cellar, soup, diary, scarf, tent, carry, sheet, bottle, glue, cloud, fan, shoe, pillow, grass, coffee, computer”
  - Participants then recall as many words as possible in a test phase
    - In a serial recall task, the words must be recalled in the order they are presented
    - In a free recall task, the order in which words are recalled is not important
    - In a cued recall task, the study is of pairs of items, and the test cues on the first item, requiring the participant to give its associate
    - If there is a long delay before recall starts, the (serial, free, cued) recall task is called a delayed recall task

Serial Position Effect

- The standard empirical findings in free recall are that:
  - There is a primacy effect, in which words near the beginning of the list are moderately well recalled
  - Words in the middle of the list are poorly recalled
  - There is a recency effect, in which words near the end of the list are well recalled

Murdock 1962 Data

- Murdock collected data under 6 conditions, varying the number of words, and the time between words
STM-LTM Explanation of Serial Recall

- Primacy and recency can be explained in terms of separate STM and LTM stores
  - early items are rehearsed more often, and so are more likely to be remembered from LTM
  - late items will be in STM at the end of stimulus presentation, and so are more likely to be recalled
- Introducing a distractor task supports this argument
  - e.g., once the list has been read, participants must count backwards from 158 in steps of 7
  - removes the recency effect, but not the primacy effect

Long Term Primacy and Recency

- Recall of the US presidents (from Washington to George HW Bush) shows primacy and recency effects

Application of Memory Theories to ADRD

- Methods and theories of memory are important in the applied problem of detecting, diagnosing and treating Alzheimer's and Related Disorders (ADRD)
  - Methods, like recognition and recall tasks
  - Theories, like signal detection and primacy and recall measures
- An example is diagnosing people into FAST Stages from memory tests
  - Stage 1: No functional decline
  - Stage 2: Personal awareness of some functional decline
  - Stage 3: Noticeable deficits in demanding job situations
  - Stage 4: Requires assistance in complicated tasks such as handling finances, planning parties, etc.
  - Stage 5: Requires assistance in choosing proper attire
  - Stage 6: Requires assistance dressing, bathing, and toileting
  - Stage 7: Speech ability declines to about a half-dozen intelligible words; progressive loss of abilities to walk, sit up, hold head up
Application of Recognition and Recall Tasks

- Two common behavioral protocols to screen for Alzheimers and Related Disorders are the MCIS and the ADAS-Cog
- The MCIS has six basic parts
  1. Free recall task: 3 repeats of same 10 words, presented in the same order
  2. A meta-memory estimate of delayed free recall ("how many out of 10 will you remember?")
  3. Triadic comparisons of animals
  4. Delayed free recall
  5. Recognition memory task with 10 words and 10 distractors, and confidence judgments
  6. Free recall on animals from triadic comparison
- The ADAS-Cog has a very similar recall protocol, except the words are re-trained in a random order

Real World Alzheimer's Recognition Data

- Each panel corresponds to a different FAST Stage
- Each cross corresponds to a patient, measured in terms of their recognition task performance, in what is often known as a Reciever Operating Characteristic (ROC) plot
  - False-alarms on the x-axis
  - Hits on the y-axis
- As disease severity increases, there are fewer hits and more false-alarms

Real World Data From the MCIS

- A sequence of free recall probabilities over the
  - three Immediate Free Recall (IFR1, IFR2, IFR3) tests
  - the Delayed Free Recall (DFR) test
- Note the
  - Learning over IFRs, but with serial position curve effect
  - Lack of recency for DFR
Real World Alzheimer’s Recall Data

A Taxonomy of Memory Systems

Declarative and Procedural Memory

- Numbers of researchers (e.g., Tulving 1972) have drawn distinctions between different types of information stored in LTM:
  - **declarative memory**: memory for facts and events, usually separated into two types
    - semantic (generic) memory: general or abstract knowledge
    - episodic memory: memories for particular events, in a way that is often autobiographical
  - **procedural memory**: memory for skills
Explicit and Implicit Memory

- There is also a distinction between different ways information is retrieved from LTM
  - **explicit memory** involves conscious retrieval, as in recall and recognition tasks
  - **implicit memory** is evident in skills, conditioned learning, and associations in memory that cannot explicitly be recalled or recognized
- Empirical evidence for implicit memory comes from
  - priming tasks: performance being affected by a perceptual or conceptual prime
  - associative learning: through classical conditioning (presenting a stimulus together with a reflex eliciting stimulus) or operant conditioning (reinforcement or punishment of behavior)
  - a very clear piece of evidence comes from laboratory studies on people’s sensitivity to underlying artificial grammars

Artificial Grammars

- Participants are given a set of sequences generated using a grammar, and are then able to classify new sequences as “grammatical” or “not grammatical”, without being aware of the existence of the grammar itself

<table>
<thead>
<tr>
<th>Grammatical sequences</th>
<th>Ungrammatical sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAZZN</td>
<td>AAAPN</td>
</tr>
<tr>
<td>AQ4</td>
<td>AAQN4</td>
</tr>
</tbody>
</table>

Associative Networks

- Collins and Loftus (1975) modeled semantic memory as an associative network
  - Links between semantically related words
  - A recent big study along these lines is at http://www.smallworldofwords.com/tests/
Associative Priming

- Associative priming experiments provide evidence for associate networks
- Meyer and Schvaneveldt (1971) asked participants to judge whether or not pairs of items were words, and recorded their response times

<table>
<thead>
<tr>
<th></th>
<th>Positive Pairs</th>
<th>Negative Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unrelated</td>
<td>Related</td>
</tr>
<tr>
<td>Nurse</td>
<td>940 ms</td>
<td>855 ms</td>
</tr>
<tr>
<td>Butter</td>
<td>Bread</td>
<td>Butter</td>
</tr>
</tbody>
</table>

Spreading Activation

- One explanation is that priming one part of a semantic network spreads ‘activation’ to related parts of the network, and speeds responses to the activated areas

Everyday memory

- Everyday memory research represents and attempt to study memory as it occurs in daily life
- Everyday memory is functional, in the sense that it used to store information with some adaptive significance to an individual
- Anderson and Pichert (1978) had participants read a passage about a house, under two different sets of experimental instructions
  - perspective of home-buyer
  - perspective of the burglar
- There was a difference in the recall of different items (e.g. “quality of plumbing” vs “how to get in and out”) relevant to each perspective
**Emotion and Everyday Memory**

- Memory for events often has an emotional component
  - Often emotional memories are the most vividly recalled ("flashbulb memories")
  - e.g., where were you when you saw or learned about
    - The moon landing
    - JFK being shot
    - John Lennon being shot
    - Shuttle disaster
    - Princess Di dying
    - 9/11
  - Very high levels of arousal, however, can interfere with storing information
  - Emotional arousal tends to lead to a narrowing of "cognitive focus"