Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1 — Second Semester 2020/2021

Lesson 16 — 1 April 2021

Structured Knowledge Representations: Working Memory, Concepts and Relations

<u>Bibliography</u>:

1) W. Kintsch, *Comprehension: A paradigm for cognition*, Cambridge university press, 1998

2) G.A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological review, 63(2), p.81.1956

3) J. R. Anderson, A spreading activation theory of memory, Journal of Verbal Learning and Verbal Behavior, Volume 22, Issue 3, Pages 261-295, June 19835)

4) T. Berners-Lee, J. Hendler, J., and O. Lassila, (2001). The Semantic Web. *Scientific American*, 284(5), 28-37. 2001.

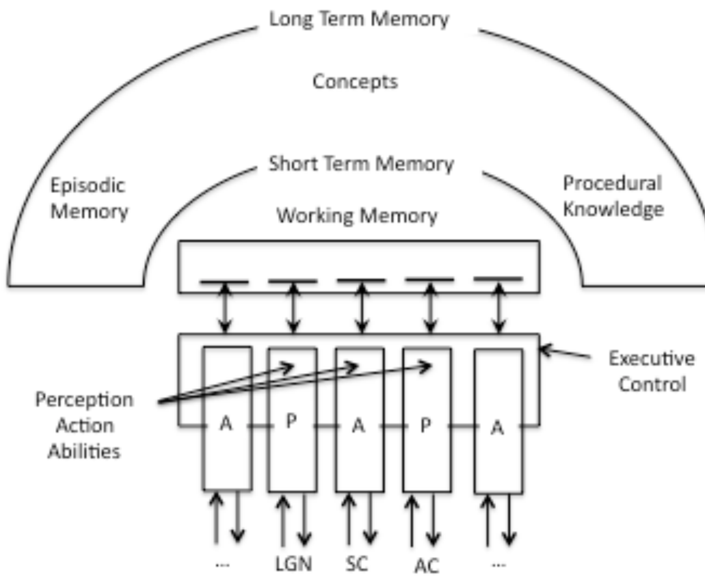
Background from Cognitive Science

Memory is the ability to encode, store and retrieve information. Memory serves to retain information over time for the purpose of influencing future action. Memory is fundamental to personal identity, social relationships, language, intelligence and survival.

Short Term and Long Term Memory

Memory is often understood as an informational processing system with explicit and implicit functions that is made up of a sensory processor, short-term memory, and long-term memory. This is an overly simplified model that, none-the less, makes it possible to explain and predict many aspects of human intelligence.

Short-term memory (also known as "primary" or "active memory") describes the ability to hold a small amount of information in a state that is easily recalled. For example, short-term memory can be used to remember a phone number that has just been recited.



Long-term memory (LTM) provides several different cognitive abilities:

- Episodic Memories: recordings of significant sensory experiences
- Semantic Memory: Abstract representations for sensory experiences
- Procedural Memory: Sequences of operations to accomplish goals
- Spatial memory (e.g. network of places as in the hippocampus)
- Muscle (Motor) Memory: A form of procedural memory for automatic sequences of muscle activations.
- Perceptual memory (short and long term memory of perceived patterns)

This list is not exclusive

Recall

Recall is the ability to retrieve information from long-term memory. There are three main types of recall: free recall, cued recall and serial recall.

Free recall describes the process in which a person is given a list of items to remember and then is tested by being asked to recall them in any order. Free recall often displays evidence of primacy and recency effects. Primacy describes the ability to recall items at the beginning of a list. Recency describes the ability to recall items at the end of a list.

Cued recall refers to the ability to recall associations of pairs of phenomena (usually words). An experimenter gives a subject pairs of word to recall, and then tests the ability to recall one of the words, given the other. The word presentation can be visual or auditory.

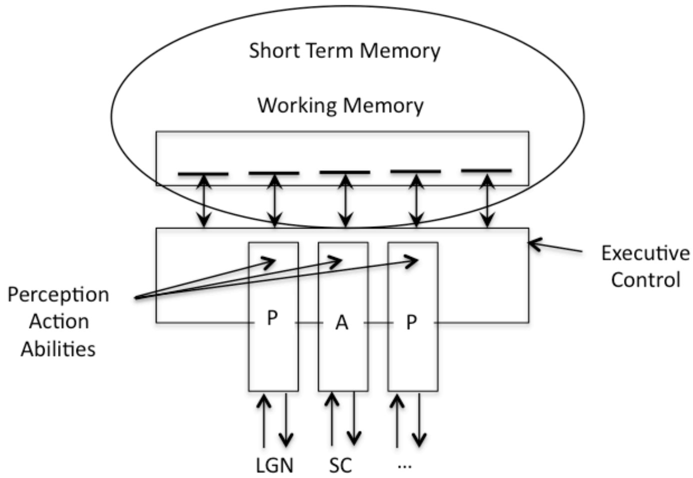
Serial recall is the ability to recall items or events in the order in which they occurred. The ability of humans to store and recall items in order is important for the use of language, and many other human abilities.

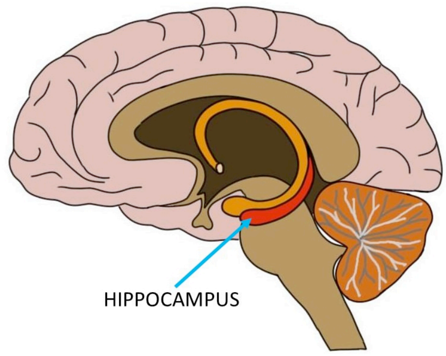
Working memory

For humans and other mammals, perception, action and cognition are controlled and coordinated in a structure known as working memory. Most models of human cognition include the working memory as the location where sensory information (phenomena) are associated with real or synthetic memories (episodes), actions, procedures (procedural knowledge), concept and other perceptions.

Working memory is used to process information from action and perception. It is easily demonstrated that humans can retain from 5 to 9 elements of information in working memory at one time. This can be demonstrated by asking people to recall a series of random letters or numbers. The average person can recall no more than 7 letters in working memory. Some individuals can hold 8, or even 9. Others are limited to 6 or 5. Without rehearsal working memory decays in a few tens of seconds. This can be demonstrated by asking someone to retain a letter while distracting them with a different task. Unless the person refreshes working memory by internal repetition (rehearsal), elements in working memory tend to decay within a minute.

There is evidence that that working memory is located in an organ at the center of the brain known as the hippocampus, although this controversial, with some authors arguing that working memory is distributed. The hippocampus is known to be involved in relational reasoning, particularly with relations of space and time. For example, the hippocampus of rats can be shown to hold a network of places, and there are demonstrations in which researchers drive a rat through a maze by stimulating different regions of the hippocampus with electrodes.





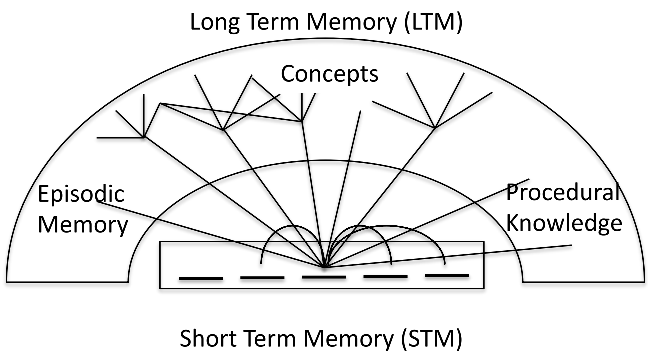
Perception is Active, Action is Perceptive

Information from the various senses is processed by specialized regions of the brain and integrated in working memory. In most cases, sensory information takes the form of a spatio-temporal map that is processed and combined by a series of neural layers to produce interpretations.

Recognition is an active process, with hypotheses in working memory driving a generative process that generates a synthetic image that is compared (top-down) to the sensor image. When the synthetic image and the perceived image match, the concept is reinforced. Errors in matching reduce the activation and cause the process to examine other hypotheses. Similar active processes are involved in perception from all of the senses including hearing (auditory perception), touch (tactile perception), Taste (gustatory perception) and smell (olfactory perception). A similar phenomena occurs with action, as the results of muscle movements are felt through a phenomena called "proprioception". Internal organs are sensed through the somatic system.

Spreading Activation

Spreading activation is a mechanism for associating items in working memory with long-term memory. Most cognitive theories assume a form of Hebbian memory, where associations are performed by some form of "spreading activation" (Anderson 83) in which activation energy propagates through a network of cognitive "units". Units that receive energy from several other units can become "activated" and can replace one of the 7±2 active units in working memory.



For example, W. Kintsch explains that when we read a word, we use frequency of occurrence of word sequences (other words that frequently co-occur with that word) to activate many possible interpretations. As additional words are perceived, the associations are correlated and the most activated association provides meaning.

Activation energy spreads from working memory to other elements of working memory and to long-term memory including concept memory, episodic memory, procedural knowledge, etc. Activated units then spread their energy to other units where it can arrive from multiple paths and accumulate. At the same time the energy decays with time. Theories differ in describing how activation energy propagates and how this propagation can be controlled by emotions and physiological state.

The limited size of short-term memory appears to be a primary bottleneck for cognition. However, the limitation is actually an advantage, because it forces abstraction and formation of new concepts, in a process called "chunking".

Attention

Attention is the cognitive process of selectively concentrating on certain perceptual phenomena while ignoring other phenomena Attention is frequently described as as the allocation of limited cognitive processing resources.

Chunking

Chunking is a process of hierarchically grouping cognitive units into larger composite units. Chunking allows multiple cognitive units to be held as a single element in short term memory, overcoming the 7±2 limit. However, associations for a chunk in WM are with the chunk and not its individual elements.

To say more, we need to define what we mean by a cognitive "unit".

Conceptual Knowledge

Working memory elements are called "entities". Entities may be created for any perceivable phenomena (anything that can be reliably perceived). Entities are interpreted by associations with memories (recall), including long-term memories (episodic, procedural, etc) and association other entities in working memory (attention). The associations that direct recall and attention are formalized as relations, represented by predicates. Sets of relations are referred to as concepts.

Concepts

Concepts are mental constructs that organize sets of relations to direct recall and attention. Associating names with concepts also provides symbols for communications. Concepts are analogous to class definitions in object oriented programming. Indeed, early techniques for Object Oriented Programming was partly inspired by schema representations (such as Minsky's Frames) that were invented for knowledge representation in the 1970s and 80's.

Concepts are learned as generalizations from experience, from the transformation of existing concepts or by communication. Concepts can represent words, actions, perceived phenomena, experiences, feelings, etc.

A concept is instantiated (reified) by a collection of memories of actual or imagined phenomena. These memories may be images, sounds, image sequences, feelings, or any other perceived phenomena (e.g. taste, smell, etc).

In cognitive psychology, Concepts are studied as the units of human cognition. In philosophy, concepts are studied in the field of ontology, as part of the question of what entities may be said to exist and how such entities may be identified and associated. The study of ontologies has been carried over to informatics and semantic web where an ontology defines the elements of a domain.

Schema

Schema are declarative structures for representing concepts. The term Schema and was originally proposed by Emmanuel Kant in "Critique of Pure Reason (1781). Schema have been used in philosophy and cognition psychology since the 19th century to represent concepts for reasoning, perception, problem solving and natural language interaction.

A schema (plural schemata or schemas) describes a pattern of thought that organizes information. A key property of Schema is the association of concepts with procedures for perception, action and reasoning (association of entities with procedural knowledge and attention).

Schemas represent concepts as data structures with slots that define the properties of the concept and associate the concept with other concepts.

A typical Schema for a concept has
1) A name.
2) A definition.
3) Meanings: memories of examples of the concept
4) Roles: Operations or procedures that are enabled or prevented by the concept.
5) Relations to other concepts and other elements in LTM.
(In many schema systems, meaning and roles are part of the list of associations).

The definition of a concept can be Intentional or Extensional.

An intentional definition specifies a test (or set of tests) that determine if an entity is an instance of the concept. (The recognition function studied in the first half of this course provides intentional definition of concepts for perception).

An extensional definition provides a list of entities that can be identified as belonging to the concept.

A meaning typically denotes memories that serve as examples. Meanings can be from actual examples or can be imagined. Meanings can be visual, episodic, auditory, olfactory, emotional or examples of feelings.

Roles are operations or procedures (procedural knowledge) that are enabled or prevented by the concept. Roles can also refer to uses that the concept can have.

For Example: Consider the number 5.
The number 5 has
1) a name: (five in english, cinq in french etc).
2) a definition (the name of a set of all sets with 5 elements)
This is an intentional definition that may be implemented either by counting the elements (5 comes after 4) or by direct recognition.
3) Meanings: Experiences with examples of the concept 5. (visual patterns, sounds, mathematical operations, counting procedures, songs, ...).
4) Roles: Operations such as addition, subtraction, division, etc that are made possible. (Example 5 cannot be directly divided by 2).
5) Relations. Relations encode associations for concepts. They are the heart of knowledge representations, both in theory and in programming practice.

Multiple kinds of relations are possible:
ISA and AKO: Identifies the concept as a member of a larger class.
(AKO = A Kind Of). Examples: (5 ISA number) (5 ISA integer) (5 ISA odd)
Part-Of: Identifies the concept as a component of a larger concept
(5 is a part of the number 15, 5 is a part of the formula 15/3)
Order Relations: $(5 < 6)$, $(3 < 5)$, Time relations 5h is before 6h.

Relations

Relations organize organize perception, reasoning, recall and communication.

Examples include temporal relations, spatial relations, Part-whole relations, family relations, social relations, administrative organizations, military hierarchies, and class hierarchies.

Kinds of Relations
A non-exhaustive list of relations between concepts includes:

1) Class membership (ISA, AKO) relations
2) Structural (Part-of) Relations
3) Ordinal relations (bigger-than, smaller than)
4) Spatial Relations (right-of, left-of, above, below, in-front-of, behind, etc)
5) Temporal relations (Allen's 13 relations between intervals).
6) Organizational relations (team member, leader, etc)
7) Family (parents, brothers, sisters, etc)
8) Causal (action A caused phenomena P)

Class membership (is-a) are useful as part of the definition of a schema.
This list is NON-EXHAUSTIVE. Relations can be defined as needed by a domain.

Predicates

Relations are formalized as Predicates (Truth functions).

A predicates is function that associates concepts. Traditionally, predicates are assumed to be Boolean functions, but probabilistic predicates are increasingly used to represent relations.

A predicates is a function that tells whether or not a relation is valid for a set of entities. Classically, predicates are treated as Boolean functions that can only return a value of TRUE or FALSE. As we have seen, in probabilistic reasoning, predicates represent the likelihood that the relation holds, with a value between 0 and 1.

Relations as N-Ary Predicates

The "Arity" of a relation is the number of arguments. Arity represents the number of entities associated by the relation. Relations may have an arity of 0 or more arguments.

The valence or Arity of a relation is the number of entities that it associates.

Nullary:	Friday()	;; a statement. An assertion. An axiom.
Unary:	Man(Bob)	;; Bob is a human of of gender male
Binary	Brother(Bob, Chris)	;; Bob and Chris are brothers
Ternary	Triangle(A, B, C)	;; a geometric relation associating points or lines

In some systems it is possible to have functions with variable arity. These are called polyadic functions or variadic functions. Set(A, B, C, D)

The philosopher CS Pierce (1871) demonstrated that any system of relations could be reduced to a hierarchy of binary relations by chunking. This is used in the widely used notation: (Subject Relation Object).

Implicit vs Explicit representations for Relations
Relations can be represented "implicitly" or "explicitly".

Implicit representation
Most programming systems for structured knowledge representation provide data structures that represent cognitive units as "objects" or instances of a class.

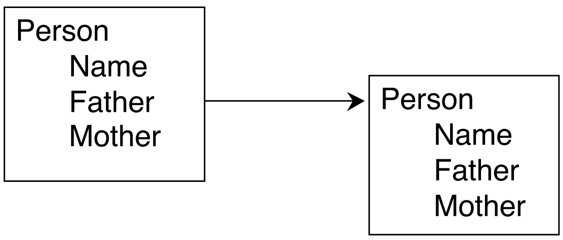
In such a system, each object is a list of named slots. Slots can have types, default values, possible ranges and other defining information. Slots can contain values for properties, contain pointers to other units or pointers to code that can be executed. Cognitive units are organised in class hierarchies, providing inheritance of structures and procedures.

Such structures are typically accompanied by a number of procedures (methods), and are implemented as a form of object-oriented programming system.

In such systems, the cognitive unit associates properties, code and other units "implicitly". With an implicit representation, the relation is represented as a pointer in a slot. For example, the following is a Class for family relations:

(defclass PERSON (slot NAME) (slot FATHER) (slot MOTHER)).
John <- (make-instance PERSON (NAME John) (FATHER Joe) (MOTHER Jane))

The slot FATHER contains a pointer to an object of the class PERSON that represents the father of the person. The pointer is the object address.



Implicit representations are simple and more efficient in computing and memory.

However, with an implicit (slot based) representation for a concept, the set of relations is fixed and cannot change dynamically.

An implicit (slot based) representation for relations is not easily completed with meta information. Some forms of reasoning are much easier with an explicit representation.

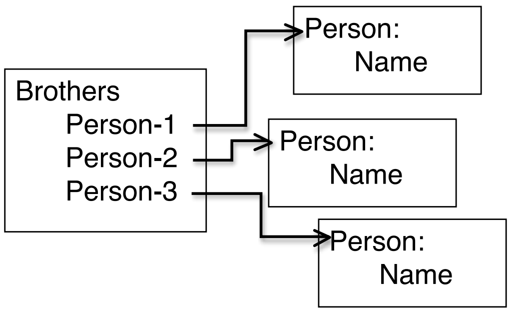
Explicit representation

With an explicit representation, relations are represented by a schema whose arguments are concepts.

Explicit Representations for relations can be changed dynamically without changing the underlying concept.

(defclass PERSON (slot NAME))
(defclass Brothersp (slot person1), (slot person2), (slot person2))

A slot holds a pointer to the object that represents the relation. This object can then provide additional information about the object, such as what, where, why, when, who and how.



With an explicit representation for relations, it is possible to write a set of general procedures for acquiring (learning), reasoning, and explaining that apply to all relations.

With an implicit representation, such procedures would be specific to each class. Thus explicit relations support generalized methods acquiring (learning), reasoning, and explaining about relations.

For example, Allen's temporal reasoning is much easier to program using explicit models of relations because the set of relations between intervals changes dynamically.

We will see more about this in the following lectures.

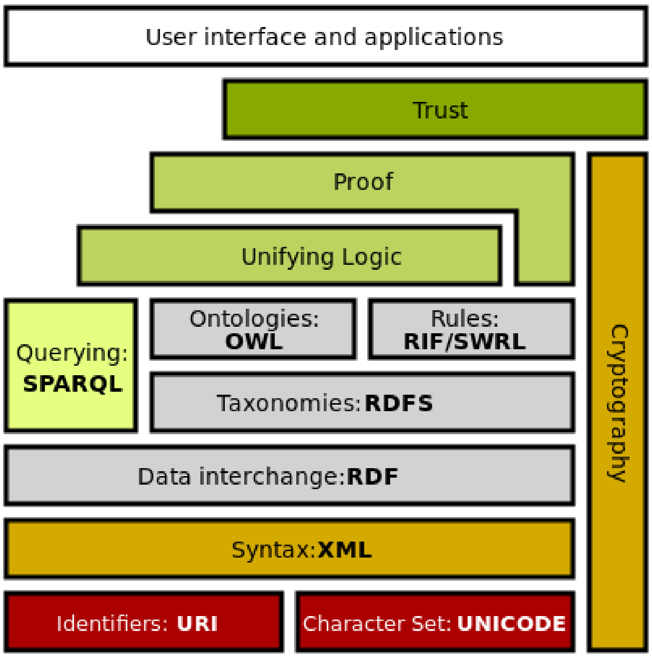
RDF and the Semantic Web

A modern use of Relations is provided by the Semantic web [Berners-Lee 2001]

Semantic web is an attempt to construct structured knowledge representations that organize and provide meaning for information on the world-wide-web. The Semantic Web is intended to provide a common framework that allows data to be shared and reused across application, enterprise, and community boundaries

The Semantic Web is an extension of the World Wide Web through standards that promote common data formats and exchange protocols. Standards for the Semantic Web are managed by the World Wide Web Consortium (W3C).

W3C has defined stack of standards for the Semantic Web that include XML, RDF, OWL and SPARQL. Meaning is provided by relations, expressed as semantic triples using RDF.



The elements of the Semantic Web stack are:

- URI are web resources (URL or internet addresses).
- XML provides a surface syntax for structured documents, but imposes no semantic constraints on the meaning of these documents;
- RDF is a data model for resources and the relations between them. RDF provides semantics for the data model;

- RDFS is a schema system for describing properties and classes of RDF resources, with a semantics for generalization hierarchies of properties and classes;
- OWL adds constructs for describing properties and classes.

XML (Extensible Markup Language) is a markup language that defines a set of rules for encoding documents in a both human-readable and machine-readable format. An XML document consists of a nested set of open and close tags, where each tag can have a number of attribute-value pairs. The vocabulary of the tags and their allowed combinations is not fixed, but can be defined as needed for an application. XML is used as a uniform data-exchange format that provides a common syntax for exchange of data.

Semantic triples are the atomic data entity in the Resource Description Framework (RDF).

RDF (Resource Description Framework) is a W3C standard used as a general framework for modeling information about web resources. The basic construction in RDF is the triple <subject, predicate, object>. The subject denotes a resource and the predicate expresses a relation between the subject and the object. The subject and object are URIs (Uniform Record Identifiers) that can be web addresses or URL's.

RDF Schema (RDFS) provide a vocabulary for RDF data. RDFS is an extension of RDF that provides basic elements for structuring RDF resources. RDFS allows definitions for Classes, Properties, Data-types and Hierarchies for both classes and properties.

OWL (Web Ontology Language) is a semantic markup language for defining, publishing and sharing ontologies. OWL can be used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms. This representation of terms and their relations is called an *ontology*

SPARQL is a query language for retrieving and manipulating data stored in RDF format. Most forms of SPARQL queries contain a set of triple patterns called a basic graph pattern in which the subject, predicate and object may be a variable (denoted by a question mark).

For example, the following SPARQL query retrieves pairs book together with its author author:

```
SELECT  ?book  ?author WHERE {?book :author ?author
```