Cognitive Science: An Introduction to the Science of the Mind - 2010

Tom Rochette <tom.rochette@coreteks.org>

July 24, 2025 — daae079c

0.1 Context

0.2 Learned in this study

0.3 Things to explore

1 Overview

2 Notes

- (p21) The human perceptual systems are information channels with built-in limits of about 7 items, or around 3 bits
- (p21) Chunking can be used to work around this limitation
- (p21) Natural language is the ultimate chunking tool
- (p22) The phenomenon of selective attention occurs in every sense modality
- (p22) Broadbent intepreted the dichotic listening experiments as showing that we can only attend to a single information channel at a time (assuming that each ear is a separate channel) - and that the selection between information channels is based purely on physical characteristics of the signal
- (p24) Information is everywhere, but in order to use it organisms need to represent it
- (p25) (the idea that) Information processing is done by dedicated and specialized systems
- (p29) Cognition can be understood as information processing and information processing can be understood as an algorithmic process
- (p29) One can try to understand how particular cognitive systems work by breaking down the cognitive tasks that they perform into more specific and determinate tasks
- (p34) SHRDLU consists of twelve different systems. Winograd himself divides these into three groups. Each group carries out a specific job
  - Syntactic analysis
  - Semantic analysis
  - Integrating the information acquired with the information the system already possesses
- (p34) We can identify distinct components for each of these jobs - the syntactic system, the semantic system, and the cognitive-deductive system
- (p34) What makes this possible for these systems to call upon each other is that their different forms of knowledge are all represented in a similar way: they are all represented in terms of procedures
- (p47) Marr distinguishes three different levels for analyzing cognitive systems
  - The top level is the computational level (computational theory)
    * to translate a general description of the cognitive system into a specific account of the particular information-processing problem that the system is configured to solve
    * to identify the constraints that hold upon any solution to that information-processing task
  - The middle level down is the algorithmic level (representation and algorithm)
    * tells us how the cognitive system actually solves the specific information-processing task identified at the computational level

- ∗ tells us how the input information is transformed into the output information
    - – The bottom level is the implementational level (hardware implementation)
        - ∗ find a physical realization for the algorithm - that is to say, to identify physical structures that will realize the representational states over which the algorithm is defined and to find mechanisms at the neural level that can properly be described as computing the algorithm in question
- (p49) Marr drew two conclusions about how the visual system functions from Warrington's neuropsychological observations
    - – Information about the shape of an object must be processed separately from information about what the object is for and what it is called
    - – The visual system can deliver a specification of the shape of an object even when that object is not in any sense recognized
- (p50) Marr's theory of vision uses 3 different types of representations
    - – primal sketch: distribution of light intensity across the retinal image, basic geometry of the field of view
    - – 2.5D sketch: depth and orientation of visible surfaces from the viewer's perspective
    - – 3D sketch: viewer-independent representation
- (p53) What are the starting-points for the information processing that will yield as its output an accurate representation of the layout of surfaces in the distal environment?
    - – The visual system needs to start with discontinuities in light intensity, because these are a good guide to boundaries between objects and other physically relevant properties
    - – Representational primitives: zero-crossings (registers of sudden changes in light intensity), blobs, edges, segments, and boundaries
- (p62) Anatomists distinguish three different parts of the mammalian brain
    - – the forebrain
    - – the midbrain
    - – the hindbrain
- (p129) The more informationally encapsulated an information system is, the less significant the frame problem will be
- (p130) In what format does a particular cognitive system carry information?
- (p130) How does that cognitive system transform information?
- (p132) How is the mind organized so that it can function as an information processor?
- (p142) Newell and Simon's characterization of physical symbol systems
    - – Symbols are physical patterns
    - – These symbols can be combined to form complex symbol structures
    - – The physical symbol system contains processes for manipulating complex symbol structures
    - – The processes for generating and transforming complex symbol structures can themselves be represented by symbols and symbol structures within the system
- (p148) Means-end analysis
    1. Evaluate the difference between the current state and the goal state
    2. Identify a transformation that reduces the difference between current state and goal state
    3. Check that the transformation in (2) can be applied to the current state
        - – If it can, then apply it and go back to step (1)
        - – If it can't then return to step (2)
- (p149) The trick in writing the GPS program, is building into it search strategies and sub-routines that will ensure that it reaches the goal state as efficiently as possible
- (p155) Fodor's computer model of the mind
    1. Causation through content is ultimately a matter of causal interactions between physical states
    2. These physical states have the structure of sentences and their sentence-like structure determines how they are made up and how they interact with each other
    3. Causal transitions between sentences in the language of thought respect the rational relations between the contents of those sentences in the language of thought
- (p164) (In reply to Searle Chinese room argument) Using an English dictionary to look up words up is not entirely straightforward, and what Searle is envisaging is more complex by many orders of difficulty.

The person inside the room needs to be able to discriminate between different Chinese symbols - which is no easy matter, as anyone who has tried to learn Chinese well knows. They will also need to be able to find their way around the instruction manual (which at the very least requires knowing how the symbols are ordered) and then use it to output the correct symbols. The person inside the room is certainly displaying and exercising a number of sophisticated skills

- (p199) SHAKEY's five level of functionality
    1. Robot vehicle and connections to user programs: To navigate and interact physically with a realistic environment
    2. Low-level actions (LLAs): To give the basic physical capabilities of the robot
    3. Intermediate-level actions (ILAs): Packages of LLAs
    4. STRIPS: A planning mechanism constructing MACROPS (sequences of ILAs) to carry out specific tasks
    5. PLANEX: Executive program that calls up and monitors individual MACROPS
- SHAKEY's software packages are built around this basic idea that complex behaviors are hierarchically organized
- (p236) The backpropagation algorithm is not very biologically plausible. There is no evidence that error is propagated backwards in the brain. And nature rarely provides feedback as detailed as the algorithm requires
- (p236) However, there are other learning algorithms. Competitive networks using Hebbian learning do not require explicit feedback, and there is evidence for local learning in the brain
- (p243) Fodor thinks of the process of acquiring a language as a lengthy process of mastering the appropriate rules, starting with the simplest rules governing the meaning of everyday words, moving on to the simpler syntactic rules governing the formation of sentences, and then finally arriving at the complex rules such as those allowing sentences to be embedded within further sentences and the complex transformational rules discussed by Chomsky and other theoretical linguists
- (p245) Fodor argues that learning a language has to involve learning truth rules
- (p245) Learning a public language such as English, even if it is your first language, requires you to formulate, test, and revise hypotheses about the truth rules governing individual words. These hypotheses have to be formulated in some language. A truth rule is, after all, just a sentence. But which language are truth rules formulated in?
- (p245) Fodor thinks that it cannot be the language being learnt. You cannot use the language that you are learning to learn that language
- (p245) It can only be the language of thought
- (p247) There are robust data indicating that children go through three principal stages in learning how to use the past tense in English
- (p247) In the first stage you language learners employ a small number of very common words in the past tense. Most of these verbs aer irregular and the standard assumption is that children learn these past tenses by rote
- (p247) In the second stage children use a much greater number of verbs in the past tense, some of which are irregular but most of which employ the regular past tense ending of "-ed" added to the root of the verb. [. . .] Surprisingly, children at this stage take a step backwards. They make mistakes on the apst tense of the irregular verbs that they had previously given correctly. These errors are known as over-regularization errors
- (p247) In the third stage children cease to make these over-regularization errors and regain their earlier performance on the common irregular verbs while at the same time improving their command of regular verbs
- (p254) According to James, neonates inhabit a universe radically unlike our own, composed solely of sensations, with no sense of differentiation between self and objects or between self and other, and in which the infant is capable only of reflex actions. It takes a long time for this primitive form of existence to become the familiar world of people and objects and for reflexes to be replaced by proper motor behavior
- (p254) Researchers have developed techniques for exploring the expectations that infants have about how objects will behave. It is now widely held that even very young infants inhabit a highly structured and orderly perceptual universe. The most famous technique in this area is called the dishabituation

paradigm
- (p255) The basic idea behind the dishabituation paradigm is that infants look longer at events that they find surprising
- (p256) Baillargeon's drawbridge experiments, together with other experiments using the same paradigm, have been taken to show that even very young infants have the beginnings of what is sometimes called folk physics (or naïve physics) - that is to say, an understanding of some of the basic principles governing how physical objects behave and how they interact
- (p257) Elizabeth Spelke four principles:
    - Principle of cohesion: surfaces belong to a single individual if and only if they are in contact
    - Principle of contact: only surfaces that are in contact can move together
    - Continuity constraint: an object cannot be present and then suddenly absent (peekaboo)
    - Solidity constraint: it is impossible for more than one object to be in a single place at one time
- (p262) According to the neural networks approach to object permanence, the expectations that infants have about how objects will behave reflect the persistence of patterns of neural activation - patterns that vary in strength as a function of the number of neurons firing, the strength and number of the connections between them, and the relations between their individual firing rates
- (p271) Fodor and Pylyshyn's argument against neural networks being a serious competitor of the physical symbol system hypothesis (aka Fodor-Pylyshyn dilemma)
    1. Either artificial neural networks contain representations with separable and recombinable components, or they do not.
    2. If they do contain such representations, then they are simply implementations of physical symbol systems.
    3. If they do not contain such representations, then they cannot plausibly be described as algorithmic information processors.
    4. Either way, therefore, artificial neural networks are not serious competitors to the physical symbol system hypothesis.
- (p281) An agent is a system that perceives its environment through sensory systems of some type and acts upon that environment through effector systems
- (p286) Horizontal faculty psychology: Domain-general
- (p286) Vertical faculty psychology: Domain-specific
- (p287) Domain-specific faculties are informationally encapsulated, they can only call upon a very limited range of information
- (p287) Each vertical cognitive faculty has its own database of information relevant to the task it is performing, and it can use only information in this database (Fodor call these cognitive faculties cognitive modules)
- (p288) Modular processes have the following four characteristics:
    - Domain-specificity: Modules are highly specialized mechanisms that carry out very specific and circumscribed information-processing tasks
    - Informational encapsulation: Modular processing remains unaffected by what is going on elsewhere in the mind. Modular systems cannot be "infiltrated" by background knowledge and expectations, or by information in the databases associated with different modules
    - Mandatory application: Cognitive modules respond automatically to stimuli of the appropriate kind, rather than being under any executive control. It is evidence that certain types of visual processing are modular that we cannot help but perceive visual illusions, even when we know them to be illusions
    - Speed: Modular processing transforms input (e.g. patterns of intensity values picked up by photoreceptors in the retina) into output (e.g. representations of three-dimensional objects) quickly and efficiently
- (p289) Two further features that sometimes characterize modular processes:
    - Fixed neural architecture: It is sometimes possible to identify determinate regions of the brain associated with particular types of modular processing
    - Specific breakdown patterns: Modular processing can fail in highly determinate ways. These breakdowns can provide clues as to the form and structure of that processing
- (p289) Here are some mechanisms that Fodor thinks are likely candidates for cognitive modules:

- – Color perception
- – Shape analysis
- – Analysis of three-dimensional spatial relations
- – Visual guidance of bodily motions
- – Face recognition
- – Grammatical analysis of heard utterances
- – Detecting melodic or rhytmic structure of acoustic arrays
- – Recognizing the voices of conspecifics
- (p298) It is a sad fact that organisms tend to learn by getting things wrong. Learning requires feedback and negative feedback is often easier to come by than positive feedback. But how do we know when we have got things wrong, and so be able to work out that we need to try something different? In some cases there are obvious error signals - pain, hunger, for example.
- (p298) Domain-general cognitive mechanisms could not have been selected by natural selection because they would have made too many mistakes - whatever criteria of success and failure they had built into them would have worked in some cases, but failed in many more
- (p303) Fodor's argument against the massive modularity thesis: Modular systems take only a limited range of inputs, how is input filtering implemented? Filtering needs a broader range of inputs than the module for which it is doing the filtering. But, on the other hand, since the filtering process is modular, it must have a limited range of inputs. The process repeats itself until we eventually arrive at a pool of potential inputs that includes everything. The filtering here involves processing so domain-general that it cannot be described as modular at all
- (p306) The ACT-R/PM architecture: two layers: a perceptual-motor layer and a cognitive layer
    - – The modules within each layer are generally able to communicate directly with each other
    - – Communication between modules on different layers, on the other hand, only takes place via a number of buffers
- (p307) The cognition layer is built upon a basic distinction between two types of knowledge - declarative (knowledge-that) and procedural (knowledge-how)
    - – The first type of knowledge involves the storage and recall of a very specific piece of information
    - – The second is a much more general skill, one that is manifested in many different ways and in many different types of situations
- (p307) Declarative and procedural knowledge are both represented symbolically, but in different ways
    - – Declarative knowledge is organized in terms of "chunks". A chunk is an organized set of elements. These elements may be derived from the perceptual systems, or they may be further chunks
    - – Procedural knowledge is represented in terms of production rules. Production rules are also known as Condition-Action Rules
- (p308) What makes ACT-R/PM a hybrid architecture is that this symbolic, modular architecture is run on a subsymbolic base
- (p308) The process of selection (of which production rule to execute) takes place subsymbolically. The job of selecting which production rule is to be active at a given moment is performed by the pattern-matching module. This module controls which production rule gains access to the buffer. It does this by working out which production rule has the highest utility at the moment of selection
- (p309) The utility of a particular production rule is determined by two things. The first is how likely the system is to achieve its current goal if the production rule is activated. The second is the cost of activating the production rule (this idea is similar to Schmidhuber Gödel machine that will only replace their executing program if its expected utility is higher than the currently running program and replacing it)
- (p309) There are two basic components determining a chunk's overall activation level.
    - – The first component has to do with how useful the chunk has been in the past
    - – The second component has to do with how relevant the chunk is to the current situation and context
- (p311) Two lessons learned from ACT-R/PM
    - – Thinking properly about the modular organization of the mind requires thinking about how the different modules might execute their information-processing tasks
    - – Different parts of a mental architecture might exploit different models of information processing.

Some tasks lend themselves to a symbolic approach. Others to a subsymbolic approach
- (p315) How do the individual cognitive sub-systems work?
- (p315) How are the individual sub-systems connected up with each other?
- (p319) Brodmann's basic insight was that different regions in the cerebral cortex can be distinguished in terms of the types of cell that they contain and how densely those cells occur
- (p319) By using the Nissl stain to examine the distribution of different types of neuron across the cerebral cortex, Brodmann identified over fifty different cortical regions
- (p319) Principle of segregation: The idea that the cerebral cortex is divided into segregated areas with distinct neuronal populations
- (p321) Tract tracing: Injecting a chemical that works as a marker into a particular brain region. Typical markers are radioactive amni acids or chemicals such as horseradish peroxidase (HRP)
- (p325) Principle of integration: The idea that cognitive functioning involves the coordinated activity of networks of different brain areas, with different types of task recruiting different networks of brain areas
- (p327) The reason that EEGs (electroencephalography) are so useful for studying ERPs (event-related potentials) is that EEGs have a very fine temporal resolution
- (p328) Magnetoencephalography (MEG) measures the same electrical currents that are measured by EEGs, however they are measured through the magnetic fields that they produce. This allows a finer spatial resolution than possible with EEGs. It is also much less susceptible to distortion due to the skull than EEG. However, it brings with it all sorts of technical issues, such has requiring to be carried in a room specifically constructed to block all alien magnetic influences, including the earth's magnetic field
- (p329) Both PET and fMRI have high spatial resolution and relatively poor temporal resolution
- (p331) Broadbent thinks of attention as occurring at the early stages of perceptual processing. His model is what is known as an early selection model
- (p331) The locus of selection problem is the problem of determining whether attention is an early selection phenomenon or a late selection phenomenon
- (p337) There are many different types of selective attention. Attention operates in all the sensory modalities. We can attend to sounds, smells, and tactile surfaces, as well as things that we see
- (p340) There are two dominant hypotheses about how visuospatial attention works
    - Visuospatial attention exploits certain memory mechanisms. In this case, we would expect brain networks associated with spatial working memory to be active during tasks that involve attention
    - Attention is linked to preparatory motor signals. The prediction generated by this hypothesis is that brain areas associated with motor planning will be active in tasks that exploit visuospatial attention
- (p344) To make meaningful comparison across different subjects, the data need to be normalized - that is, the data from each subject need to be reinterpreted on a brain atlas that uses a common coordinate system, or what is known as a stereotactic
- (p354) The development of pretend play in infancy appears to follow a fairly standard trajectory. The most basic type is essentially self-directed - with the infant pretending to carry out some familiar activity
- (p355) The next stage is other-directed, with the infant pretending that some object has properties it doesn't have
- (p355) A more sophisticated form of pretense comes with what is sometimes called object substitution. This is when the infant pretends that some object is a different object and acts accordingly
- (p355) Leslie's model of infant pretense starts off from three basic observations:
    - Pretend play in the infant depends crucially on how the infant represents the world (and hence on his primary representations)
    - We cannot explain what is going on in pretend play simply with reference to the infant's primary representations
    - The pretend representations must preserve their ordinary meanings in pretend play
- (p356) Pretend representations are somehow "quarantined" from ordinary primary representations. The key problem is to explain how this quarantining takes place.
- (p356) Leslie's explanation of how primary representations are quarantined exploits a very basic parallel between how representations function in pretend play and how they function when we are representing other people's mental states in mindreading. When we represent what other people believe or desire,

we do so with representations that are also quarantined from the rest of our thinking about the world
- (p358) Leslie thinks that we need to supplement our account of how primary representations function with two extra components:
  - The first component is a way of marking the fact that a primary representation has been decoupled and is now being used for pretend play
  - The second is a way of representing the relation between agents and decoupled representations
- (p358) Leslie proposes that the first of these is achieved by a form of quotation device (Sarah said: "The world is flat.")
- (p363) False belief test: Test whether a child is able to understand that others may hold different beliefs than their own and that these beliefs may be incorrect/different than their own
- (p364) Baron-Cohen et al.: "Our results strongly support the hypothesis that autistic children as a group fail to employ a theory of mind. We wish to explain this failure as an inability to represent mental states. As a result of this the autistic subjects are unable to impute beliefs to others and are thus at a grave disadvantage when having to predict the behavior of other people."
- (p366) Pretend play emerges during the second year of life. But children do not typically pass the false belief test until they are nearly 4. There is a very clear sense, therefore, in which the BELIEVES operation must be much harder to acquire than the PRETENDS operation
- (p367) Onishi and Baillargeon used a violation of expectations paradigm that measured looking time. They hypothesized that the length of time that the infants looked at each of the scenarios would be a guide to their implicit understanding of false belief. They predicted the infants would look significantly longer when the actor did not behave as expected. The robust effect that they discovered is that infants looked significantly longer when the actor searched in the yellow box (unexpected) than when the actor searched in the green box (expected). So, they conclude, infants have an understanding of false belief much earlier than suggested by the traditional false belief task
- (p368) The Onishi and Baillargeon experiments identify an implicit understanding of false belief, whereas the standard false belief tasks are testing for an explicit understanding of false belief
- (p370) Baron-Cohen model of the mindreading system:
  - 0-9 months
    * Intentionality detector (ID): Allows the infant to distinguish the animate, goal-driven entities from the other objects it encounters
    * The emotion detector (TED): Allows infant to understand not just that agents make movements towards particular goals, but also why those movements are being made and what sort of movements they are
    * Eye direction detector (EDD): Help the infant identify the goals of the mouvement (A good way to find out the apparent goal of a purposeful movement is to check where the agent is looking - since agents tend to keep their eyes on the target)
  - 9-14 months
    * Shared attention mechanism (SAM): Joint visual attention between the infant and an agent
  - 14 months
    * The empathy system (TESS)
  - 18-48 months
    * Theory of mind mechanism (TOMM)
- (p372) For normal social development it is not enough simply to be able to identify other people's emotional states and moods. The developing child needs to learn to respond appropriately to those emotional states and moods. This is where empathy comes in
- (p372) Psychopaths have profound social problems, but these problems are very different from those suffered by autistic people. Psychopaths are typically very good at working out what is going on in other people's head. The problem is that they tend not to care about what they find there - and in fact they use their understanding to manipulate other people in ways that a normal person would find unacceptable
- (p374) Leslie and his collaborators have a subtle solution to the problem of explaining the long time lag between when they think that the capacity for metarepresentation first emerges (during the second year) and when children generally pass the false belief test (towards the end of the fouth year)
  - Leslie thinks that there are two very different abilities here

7

* The first is the ability to attribute true beliefs to someone else
        * The second is the ability to attribute false beliefs
    - The default setting of the theory of mind mechanism is to attribute true beliefs
    - Success on the false belief task only comes when young children learn to "switch off," or inhibit, the default setting
    - According to Leslie, this requires the development of a new mechanism. He calls this mechanism the selection processor
* (p377) Perner's thinking about mindreading is very much informed by influential theories in philosophy about the nature of belief and other mental states that philosophers collectively label propositional attitudes
* (p377) Belief is called a propositional attitude because it involves a thinker taking an attitude (the attitude of belief) towards a proposition
* (p378) The child is certainly representing another person as being in a psychological state. But they can do that without engaging in metarepresentation. Since the content of the psychological state tracks what the child considers to be the state of the world, the child does not need to deploy any resources over and above the resources that she herself uses to make sense of the world directly
* (p381) According to simulation theory, the core of the mind-reading system does indeed exploit a specialized cognitive system, but this cognitive system is not actually dedicated to information processing about beliefs, desires, and other propositional attitudes
* (p381) [The idea of simulation theory is] that we explain and predict the behavior of other agents by projecting ourselves into the situation of the person whose behavior is to be explained/predicted and then using our own mind as a model of theirs
* (p382) According to standard simulationism, the process od simulation has to start with the mindreader explicitly (although not necessarily consciously) attributing beliefs and desires to the person being simulated
* (p382) For Goldman we identify other people's beliefs and desires by analogy with our own beliefs and desires. We know which beliefs we tend to form in response to particular situations. And so we assume that others will form the same beliefs, unless we have specific evidence to the contrary
* (p383) Standard simulationists are typically committed to the following two basic principles:
    - We understand the psychological states of others by analogy with our own psychological states
    - We have a special self-monitoring mechanism for keeping track of our own psychological states
* (p384) The intuitive idea behind radical simulationism is that, instead of coming explicitly to the view that the person whose behavior I am trying to predict has a certain belief (say, the belief that p), what I need to do is imagine how the world would appear from her point of view
* (p384) For radical simulationist, children who fail the false belief test lack imaginative capacities
* (p390) One of the basic claims of simulation theorists is that mindreading is carried out by what they call co-opted mechanisms. These are information-processing systems that normally serve another function and that are then recruited to help make sense of the social world
* (p391) Paired deficits: problems with experiencing the relevant emotion and in identifying it in others
* (p391) There is evidence of paired deficits for several different emotional states
    - Fear: Amygdala
    - Anger: Dopamine levels
    - Disgust: Insula
* (p392) Mirror neurons were first discovered in macaque monkeys by an Italian research group led by Giacomo Riozzolatti in the mid-1990s. Rizzolatti and his colleagues were recording the responses of neurons that showed selective activation when the monkey made certain hand movements when they noticed completely by chance that the same neurons fired when the monkey saw an experimenter making the same movement
* (p392) The mirror neuron system could serve as a neural substrate both for TED (the emotion detector system) and TESS (the empathy system)
* (p392) The mirror neuron system is part of what makes imitation possible
* (p405) A dynamical system is any system that evolves over time in a law-governed way
* (p405) One of the basic theoretical ideas in dynamical systems modeling is the idea of a state space. A state space has as many different dimensions as it has quantities that vary independently of each other -

as many different dimensions as there are degrees of freedom in the system

- (p407) Van Gelder's basic point is that cognitive scientists are essentially engaged in reverse engineering the mind. Cognitive scientists have tended to tackle this reverse engineering problem in a particular way - by assuming that the mind is an information-processing machine. But what Van Gelder tries to show is that this approach is neither the only way nor the best way.
- (p409) Four very important features of the Watt governor:
  - Dynamical system
  - Time-sensitivity
  - Coupling
  - Attractor dynamics
- (p412) Dynamic system theorists think not only that the mind is a dynamical system, but also that when we look at the relation between the organism and the environment what we see is a coupled system
- (p420) The dynamical systems approach adds a powerful tool to the cognitive scientist's toolkit, but it is unlikely ever to be the only tool
- (p421) The principal objection that situated cognition theorists make to traditional cognitive science is that it has never really come to terms with the real problems and challenges in understanding cognition
- (p422) The basic complaint is that SHRDLU only works because its artificial micro-world environment has been stripped of all complexity and challenge
- (p423) For situated cognition theorists, SHAKEY is not really a situated agent, even though it propels itself around a physical environment. The point for them is that the real work has already been done in writing SHAKEY's program. SHAKEY's world is already defined for it in terms of a small number of basic concepts
- (p423) SHAKEY already has the building blocks for the solution. But working out what the building blocks are is perhaps the most difficult part of real-world problem-solving
- (p425) One of the basic design principles stressed by situated cognition scientists is that there are direct links between perception and action
- (p427) Webb's robot crickets nicely illustrate one of the basic themes of biorobotics and situated cognition. Input sensors are directly linked to output effectors via clever engineering solutions that make complicated information processing unnecessary
- (p427) Morphological computation is a research program for designing robots in which as much computation as possible is done for free
- (p430) Subsumption architectures are organized very differently from modular architectures. Their basic components are activity-producing sub-systems. Subsumption architectures are made up of layers. The layers are autonomous and work in parallel. The higher layers subsume the lower layers, but they do not replace or override them
- (p430) This makes it easier to design creatures with subsumption architectures. The different layers can be grafted on one by one. Each layer can be exhaustively debugged before another layer is added. And the fact that the layers are autonomous means that there is much less chance that adding a higher layer will introduce unsuspected problems into the lower layers
- (p434) We can identify three basic features of subsumption architectures, as developed by Brooks and other AI researchers:
  - Incremental design
  - Semi-autonomous sub-systems
  - Direction perception-action links
- (p434) The problem is that subsumption architectures don't seem to have any decision-making processes built into them. Conflict resolution is purely mechanical
- (p435) SSS:
  - Servo-based layer that controls the robot's effectors and processes raw sensory data
  - Subsumption layer that reacts to processed sensory input by configuring the servo-based layer
  - Symbolic layer that maintains complex maps of the environment and is capable of formulating plans; the symbolic layer configures the subsumption layer
- (p436) Behavior-based architectures incorporate some of the basic design features of subsumption architectures. But they have two additional features that separate them from subsumption architectures

9

- Distributed representations
- Real-time functioning
- (p436) Behavior (as defined by Mataric) is a control law that satisfies a set of constraints to achieve and maintain a particular goal
- (p450) Priming experiments. Subjects are exposed very briefly to some stimulus - an image on a screen, perhaps, or a sound. The time of exposure is short enough that the subjects do not consciously register the stimulus. Nonetheless, the exposure to the stimulus affects their performance on subsequent tasks
- (p458) The neuropsychologists David Milner and Melvyn Goodale have developed a sophisticated theory of vision that is built around this idea that one of the roles of consciousness is to permit voluntary and deliberate action
- (p460) Conscious awareness is restricted to the ventral pathway while the dorsal stream governs the visual control of movement non-consciously
- (p463) The retention of information is very impaired in the absence of consciousness
- (p463) Greenwald, Draine, and Abrams experiment suggests another hypothesis about the function of conscious awareness - namely, that consciousness allows information to be explicitly retained and maintained. According to this hypothesis, information that is picked up non-consciously can indeed be deployed in relatively sophisticated tasks, but it can be used only within a very limited time horizon. Conscious information, in contrast, is more transferable and flexible
- (p464) Two related ideas emerged about the function of consciousness
  - Conscious awareness seems extremely important for planning and initiating action (as opposed to the online control of behavior, which can be carried out through non-conscious information processing)
  - Concious information persists longer than non-conscious information
- (p464) Phenomenal consciousness (P-consciousness): P-consciousness is experience... We have P-conscious states when we see, hear, smell, taste, and have pains. P-conscious properties include the experiential properties of sensations, feelings, and perceptions, but I would also include thoughts, wants, and emotions
- (p464) Access consciousness (A-consciousness): A state is A-conscious if it is poised for direct control of thought and action. To add more detail, a representation is A-conscious if it is poised for free use in reasoning and for direct "rational" control of action and speech
- (p470) Dehaene and Naccache global workspace theory focuses on three different things that they believe consciousness makes possible. These are:
  - the intentional control of action
  - durable and explicit information maintenance
  - the ability to plan new tasks through combining mental operations in novel ways
- (p471) Dehaene and Naccache consider these three hypothesized functions of consciousness within a framework set by two basic theoretical postulates about mental architecture and the large-scale organization of the mind
- (p471) The first theoretical postulate is a version of the modularity theory. Consciousness is restricted to information within the global workspace
- (p471) The second theoretical postulate has to do with how information becomes available to the global workspace. Attention is the key mechanism here. It functions as a gate-keeper, allowing the results of modular information processing to enter the global workspace

3 See also

4 References

- Bermúdez, José Luis. Cognitive Science: An Introduction to the Science of the Mind. Cambridge: Cambridge University Press, 2010.
- https://www.amazon.com/Cognitive-Science-Introduction-Mind/dp/0521708370