MODULE 1: THE COGNITIVE REVOLUTION

MODULE OVERVIEW

Welcome to Phil/Psych 256, Introduction to Cognitive Science! In this first module, we will be getting better acquainted with the structure and content of the course, the disciplines that contribute to Cognitive Science as an interdisciplinary affair, and the central theory of mind that has informed much research on the mind since the 1950s. We will offer a historical overview of how the interdisciplinary study of the mind has developed since the 1950s. We will also be getting to know one another. Because Cognitive Science is the multi-disciplinary study of the mind, we will all be coming from our own disciplinary backgrounds with our own interests—interests that will often intersect in interesting and illuminating ways.

After some brief information about the course structure in Lesson 1, Lesson 2 will be a schematic overview of the disciplines that make up Cognitive Science, with a comment or two about how each one thinks of the mind and the role of the discipline within cognitive science. Lesson 3 will look in more detail at the Cognitive Revolution, and Lesson 4 will introduce a dominant paradigm within which Cognitive Science has been done: the Computational and Representational Theory of Mind (CRTM). This theory of mind will be critically explored as we progress throughout the ensuing modules.

LEARNING OUTCOMES

By the end of this unit, you should be able to:

a) Recount the major research developments after World War II that led to the Cognitive Revolution—unifying work in Philosophy, Psychology, Linguistics, and Computer Programming;
b) Explain why so many people have found the mind/computer analogy so powerful in their own research;
c) Discuss the difference between mental activity (computation) and mental objects (representations);
d) List the different goals of the different disciplinary studies of the mind.
e) Distinguish normative questions about how we *ought* to think from factual and explanatory accounts of how we *actually* think.

MATERIALS

A. READINGS

Miller, George A (March, 2003). "The Cognitive Revolution: a historical perspective". *Trends in Cognitive Science*, Vol. 7, No. 3, pp. 141-144.
Thagard, Paul (2009). "Why Cognitive Science Needs Philosophy and Vice Versa". *Topics in Cognitive Science* 1, pp. 237-254.

B. WORKS CITED

Chomsky, Noam (1959). "Review of B.F. Skinner's *Verbal Behaviour*" in Language, 35, No. 1 (1959), 26-58.
Moran, Joe (2002). *Interdisciplinarity: the New Critical Idiom*. London: Routledge.
Skinner, B.F. (1957). *Verbal Behaviour*. Acton, MA: Copley Publishin Group.

LESSONS

A. Road Map for the Course
B. The Disciplines that Study the Mind
C. Before the Cognitive Revolution
D. The Cognitive Revolution: Computation and Representation

STUDY QUESTIONS

1. Why does Thagard think Cognitive Science needs Philosophy?
2. Miller and others have argued that there was a Cognitive Revolution in the 1950s and 60s. What research programmes did this revolution make obsolete?
3. What was Chomsky's criticism of Skinner's behavioural linguistics?
4. Many in Cognitive Science believe the mind is like a computer. What kind of computer are they talking about?
5. Do you think there are kinds of mental processes that aren't adequately captured by the computational/representational theory? If so, which ones?

KEY TERMS

Algorithm; Anthropology (Cognitive Anthropology; Linguistic Anthropology); Argument form; Artificial Intelligence; Artificial Neural Network; Behaviourism (B.F. Skinner); Cognition; Cognitive Revolution; Computational-Representational Theory of Mind (CRTM); Computability; Computer Science; Conceptual Analysis; Connectionist Theories of Mind; Corpuscles; Critical Period (of language acquisition); Deductive Validity (deductively valid); Descriptive Study of the Mind; Generality; Functionalism; Information Theory; Intentional Content; Interdisicplinary (Interdisiplinarity); Lambda Calculus; Language Faculty; Lexicon; Linguistics; Markov Processes; Multiply-Realized (Multiple Realization); Multi-disciplinarity; Neuron; Neuroscience; Normativity; Operant Conditioning; Philosophy; Psychology; Psychophysics; Pragmatics; Reinforcement (Positive and Negative); Semantics; Shannon & Weaver Mathematical Model of Communication; SPAUN; Stimulus and Response; Syllogism; Synapse; Syntax; Turing Machine; Turing Test (Imitation Game); Universal Grammar; Weber-Fechner's Law.

LESSON A: ROAD MAP FOR THE COURSE

INTRODUCTION

Welcome to PHIL/PSYCH 256: Introduction to Cognitive Science. The course is divided into 12 modules, which are designed so that you can complete a single module each week. Information in each module will build on previous discussions; connections will be made with material covered earlier in the course, and the significance of each module for future modules will be flagged (when useful). The modules should, therefore, be completed sequentially.

Each module is broken down into lessons. Usually, a module will have an introductory lesson and two or three other lessons that get into detail regarding the overarching topics for the module. Do not think of the lessons as mere overview of the readings, though the readings will be discussed. The primary purpose of the lessons is to offer historical context, remarks about the state of the art when it comes to the relevant scientific ideas, and to help you think through some of the main concepts at play in the readings (and occasional videos or other media).

Finally, each module includes a list of key terms. These are often important bits of technical terminology. Because Cognitive Science is interdisciplinary, and each discipline has its own methods and theoretical vocabulary, part of your work in the course will involve keeping track of the different ways we speak about the mind. If you are a philosopher, you likely have a decent grasp of some of philosophy's concepts (and the terminology used to pick them out); if you are a psychologist, then you likely feel more at home with the way psychologists talk about the mind. Getting acquainted with how we use language, the similarities and differences, is often the first step to interdisciplinary research. (For an overview of the disciplines that make up Cognitive Science, and some remarks about interdisciplinarity, see Lesson 2 of this Module.) Each term will be clearly defined, and its meaning made clear through use, as it arises in the lessons.

COURSE ORGANIZATION

The course is loosely based around Four Units, each with its own theme:

1. Conceptual and Historical Background (Modules 1 and 2);
2. Minds: Computation and Representation (Modules 3-5);
3. Concepts, Consciousness, and Mental Content (Modules 6-9);
4. Problems for Computational/Representational Theories of Mind (Modules 10-12)

Here is a brief breakdown on what each of the four units focuses upon:

UNIT 1, Conceptual and Historical Background (Modules 1 and 2)

Module 1: Following the introductory comments you're now reading, this module moves on to examine the historical development of Cognitive Science starting in the 1950s and 60s. Cognitive Science is primarily a response to technological developments in computer science, but also new theoretical ideas in Linguistics, Cognitive Psychology, and Philosophy (especially logic, and the philosophy of mind and language).

Module 2: This module will look at the *prehistory* of Cognitive Science. We will establish the origins of the concept of mind, starting with the philosophical ideas of the ancient Greek philosophers Plato and Aristotle, and then moving on to the early modern scientific work of René Descartes and his 17th century contemporaries. We will see that the broad outline of our current conception of the mind, as a thing that engages in certain kinds of processes over internal representations, is part and parcel of the Scientific Revolution of the 17th century. In fact, while we use the phrase "the Cognitive Revolution" to label large-scale transformations in the study of mind starting in the 1950s, we will explore why Noam Chomsky thinks the first Cognitive Revolution actually starts in the early 17th century.

UNIT 2, Minds: Computation and Representation (Modules 3-5)

Module 3: This module will look at the important role of computer science and embryonic research into artificial intelligence, starting with Alan Turing's ideas regarding the possibility of machine intelligence. Computer science provides us with the notion of a program: a finite set of rules for manipulating symbolic representations. We will see how this inspired a whole new theory of mind, based on an analogy between mind and computer: the so-called "Computational and Representational Theory of Mind". Though this idea is briefly introduced toward the end of the current Module, Module 3 will delve into it in much more detail. We will also explore Functionalism: a philosophical theory of the mind that is inculcated in artificial intelligence research and in much cognitive psychology.

Module 4: This module will examine logical and other rule-based attempts to make sense of how the mind works. Can we adequately model the mind if we think about it as a truth-preserving system of inferential rules? These rules allow us to start with some representational content, and, if that content is true, ensure that we reason properly to other truths. While this is a powerful idea in philosophy and mathematics, we will see some of its shortcomings, and how other rule-based systems of cognition try to address some of these problems. While the mind might not infer in accordance with the rules of first-order predicate logic (i.e. the kind of logic you learn in an introductory philosophical logic class, or computer science class), we are still captured by the idea that the mind might just be the application of formal rules—whatever they turn out to be.

Module 5: This module will examine how the study of non-human animal minds might inform our theories of cognition, mental content, and communication. Our intuitions about animal minds pull us in two opposing directions. The first intuition is that animals, especially animals that are like us (e.g. non-human primates), have minds. The second intuition is that mindedness is associated with the capacity for linguistic communication and with related capacities for rational thought (certain forms, at least) that animals would seem to lack. We will see what anthropologists, cognitive and linguistic ethologists, and psychologists who focus on animal cognition, have to say about these issues. Some cognitive scientists think that animal cognition calls into question our pre-theoretical, and even our theoretical and scientific, conceptions of the mind.

UNIT 3, Concepts, Mental Content, and Consciousness

Module 6: This module will discuss the centrality of *concepts* in cognitive science research. Both philosophical and psychological work done on concepts will be addressed. We will also explore classical theories of concepts as definitions and as images before the mind.

Module 7: We will start to examine psychological theories of concepts that reject previous classical accounts, whether definitional or imagistic. We will look at Prototype and Exemplar theories, which together are often referred to as "resemblance theories" of concepts. What theoretical considerations and empirical evidence speak in favour of these new theories? Does one or the other better model the evidence of contemporary cognitive psychology? What are some of the pitfalls of resemblance theories of concepts when it comes to thinking about not only our psychology, but how we might think about machine learning?

Module 8: This module will examine one further theory of concepts that has been developed since the 1980s: the Theory theory of concepts. Theory theory tells us that concepts just are the explanatory frameworks and predictive strategies we employ as we negotiate our environment, so they function like miniature scientific theories. However, the bulk of this module will be devoted to a recent discussion, initiated by philosopher and cognitive scientist Edouard Machery, regarding whether or not philosophers and psychologists use the term "concept" in the same way, and whether or not the philosophical notion of a concept is useful for scientific research into the mind.

Module 9: Much of our mentality straightforwardly fulfills some functional role. Beliefs and desires are causally related to our action; perception plays a role in providing the cognitive system with knowledge about the external world that is causing mental representations; etc. However, philosophers and psychologists have had a hard time explaining the subjective nature of experience—the "what it is like to be" of subjectivity. This aspect of our mental lives is much less amenable to functional (causal) description. We will explore whether or not consciousness poses a "hard problem" for the science of mind, and what this might mean for our prospects of developing a science of consciousness.

UNIT 4, Problems for Computational/Representational Theories of Mind

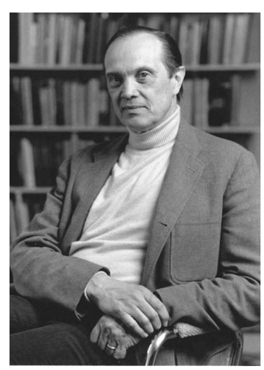
Module 10: Is the mind really like a computer program? The dominant theory of mind in Cognitive Science, the Computational/Representational theory, takes the analogy between minds and programs quite seriously. But programs are largely independent of the physical systems that realize them. Your new video game app doesn't really care whether it is running on an android phone, an iPhone, or your home PC. In this module, we will explore connectionist theories of the mind. We will see that there has been a lot of research in the past 30 years regarding the influence of the underlying neural architecture of cognition. What happens if we start thinking about the mind as if it *just is* the brain, as opposed to thinking of it as a software program that only *uses* the brain as "hardware"? How does our conception of cognition change when we see the mind as a systematically organized network of neurons connected to one another, and obeying certain rules for how these neurons communicate information to one another?

Module 11: We will explore the role of emotion in cognition. Until very recently, emotions were thought to be "non-cognitive": they only interfered with reliable belief formation and cogent reasoning patterns. Much recent neuroscience of emotion calls the sharp distinction between reason and emotion into question, and asserts the necessity of emotions in effective reasoning. If we cannot explain how we reason or how we act without recourse to emotion, does this pose a problem for the Computational/ Representational theory of mind?

Module 12: Computational/Representational theory of mind supposes that the mind is, in a certain sense, *cut off* from external reality. The mind engages in computation over the syntax of representations it has been given through sensation or by its own faculty of imagination (or by some external programmer). This module will explore work on extended, embodied, enactive, and social minds. Perhaps the mind does not end at the skull's boundary. Perhaps cognition isn't always a private and individual affair. We will explore these theoretical possibilities, and see what new evidence supports them.

LESSON B: THE DISCIPLINES THAT STUDY THE MIND

INTRODUCTION



As you will see when you read the George A. Miller piece for this module, Cognitive Science has been, since its inception, an interdisciplinary affair. (George A. Miller is pictured to the right.) We study the mind from a variety of disciplinary perspectives in the hopes that we can learn something about it that wouldn't be possible if we all remained in our disciplinary silos. Here is what Miller says:

"… at least six disciplines were involved: psychology, linguistics, neuroscience, computer science, anthropology and philosophy. I saw psychology, linguistics and computer science as central, the other three as peripheral. These fields represented, and still represent, an institutionally convenient but intellectually awkward division. Each, by historical accident, had inherited a particular way of looking at cognition and each had progressed far enough to recognize that the solution to some of its problems depended crucially on the solution of problems traditionally allocated to other disciplines." (Miller 2003, pg. 143)

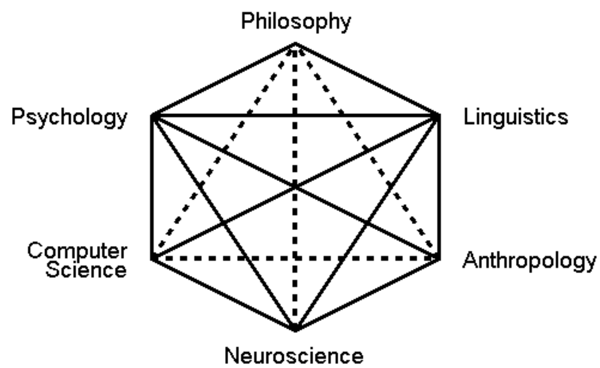
But what does "**interdisciplinary**" mean? It is one of those words that gets thrown around a lot, especially by university administrators. We want to make sure we have a working definition of the term, and that we understand the place of those various disciplines that make up Cognitive Science, and why we think interdisciplinary study is necessary when it comes to the mind.

The social historian Joe Morgan, who studies the historical development of academic disciplines and their interaction, defines **interdisiplinarity** as "any form of dialogue or interaction between two or more disciplines." (Morgan 2002, pg. 16) This is admittedly quite vague, but Moran is proposing a broad definition on purpose. This is because there are many types and levels of such dialogue or interaction, and we don't want our definition to favour some and preclude others. But notice that interdisciplinarity requires *dialogue* and *interaction*. This is different from a related, though ultimately distinct notion: **multi-disiplinary research**. Imagine we were to bring together a number of people who all study some common subject matter, but from different disciplinary perspectives, e.g. a work of art. We might invite an art historian, an art critic, a restoration specialist, other artists, etc. We might not be concerned with how they talk to one another; maybe they don't talk to one another. The restorationist might have little to say to the art critic, for example. We certainly need not be concerned with having them work on research programs together. We just want everyone to line up and say what they think about some piece of art. This is multi-disciplinarily. Interdisciplinarity, by contrast, requires that people speak with one another, come to understand other disciplinary outlooks (even if they ultimately disagree, in detail or in outline, with methods or findings), and that research happens in such a way that it exhibits influence of one discipline upon another—a reciprocity between research programmes.

In this way, Cognitive Science has always been more than joint record-keeping. When we teach courses in Cognitive Science, for example, we don't just spend a couple of weeks talking about each discipline. We see how philosophy, psychology, linguistics, anthropology, neuroscience, and computer programming/engineering/AI inter-penetrate one another. This course will be an example of how interdisciplinary studies can be exciting and fruitful. It will also examine, from time to time, some of the pitfalls of interdisciplinarity. But before we can see how interdisciplinary research is done, we'll need to look at the disciplines which make up Cognitive Science, and what they take to be their central contributions to the joint study of the mind.

HISTORY OF DISCIPLINARY CONGRESS

In 1978, Miller received Sloan Foundation funding to bring together an interdisciplinary committee to talk about the state of Cognitive Science. Even at this point, with Cognitive Science in its infancy, there was a great deal of interdisciplinary research. The image on the bottom of Miller 2003 (pg. 143) shows the state of interdisciplinarity in Cognitive Science circa 1978. Each line in this heptagram represent existing research programmes that utilized expertise in multiple disciplines. In Thagard (2009, pg. 243) an updated version of the heptagram from 1985 shows new connections between philosophy, computer science, and neuroscience (amongst others). These connections are even stronger in the present day, where much work in the philosophy of mind and philosophy of psychology is deeply informed by neuroscientific findings, and often uses computer modeling. In the other direction, AI research and neuroscience have learned from philosophers about the conceptual relationships that exist between all of these disciplines.



And let us remember that there are "at least" six disciplines that make up Cognitive Science. In actuality, that number is probably larger, and growing year-by-year. There are still other connections to disciplines not normally associated with the Cogntive Science heptagram (though perhaps they should be!): mathematics, evolutionary biology, cognitive and linguistic ethology, rhetoric and communication studies, etc. For now, we will list and briefly describe the six disciplines that Miller and Thagard talk about, though even more connections will be made as we progress throughout the course.

BRIEF OVERVIEW OF COGNITIVE SCIENCE DISCIPLINES

Philosophy

Philosophy is the most ancient study of mind. As we will see in Module 2, philosophers have been talking about the mind since at least the time of Plato and Aristotle, about two and a half thousand years ago.
The philosophical conception of the mind has changed much in that time, mostly keeping in line with the various sciences of mind, but some very old questions still hold the philosophical imagination. Philosophers are concerned with knowledge, its relationship to perception, and how perception, cognition, and action are conceptually related to one another. Using the tools of logic, formalized argument, and conceptual analysis, philosophers pursue questions that are usually abstract and general regarding the nature of the mind.

Philosophy brings to Cognitive Science this focus on general and abstract issues, but also, as Thagard (2009) tells us, a focus on **normativity**. Whereas the other disciplines that make up Cognitive Science are descriptive (informing us about how we actually *do* think), philosophy often examines questions regarding how we *ought* to think. Philosophers have long been concerned with logic as a result: isolating and applying those inference rules that allow us to preserve truth from thought to thought. The following is a **syllogism** (i.e. an argument with two premises that each share a term with the conclusion):

> P1. All dolphins wear tuxedos.
> P2. Anything in a tuxedo is good at poker.
> C. Dolphins are good at poker.

Now, it is highly unlikely that any mind has actually entertained the particular contents of the premises and conclusion of the above syllogism—since they're just silly statements that no one actually believes—so there isn't much for the psychologist to say about such contents. However, the philosopher has much to say. In Philosophy, we abstract an **argument form**, irrespective of content, and examine the logical properties it possesses. This argument is an instantiation of the following form:

> P1. All As are Bs
> P2. All Bs are Cs
> C. All As are Cs

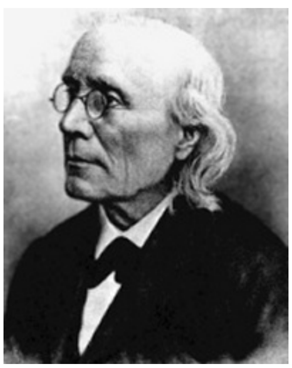
This argument form, regardless of the content we plug in for A, B, and C, is a **deductively valid form.** This means that any actual argument that is of this form counts as a deductively valid argument. It might surprise you, then, to find that the above syllogism about tuxedo-wearing, poker-playing dolphins, counts as a deductively valid argument. But it does! We'll come back to this surprising result in Module 4, where we will have much more to say about logic and other rule-governed theories of cognition, but we can already see how philosophers, with their emphasis on valid forms of inference, might have something to offer computer programmers, for example, when it comes to developing AI that engages in truth-preserving inference rules.

Philosophers are also concerned with the structure of concepts, and what counts as a scientific theory—and what counts as evidence for that theory. To these ends, the philosopher engages in **conceptual analysis**: understanding concepts by seeing how they are composed by more basic concepts, or understanding concepts by examining the theoretical and empirical presuppositions on which their normal use rests. So, when Thagard tells us that philosophy makes contributions regarding general questions and normative questions, he is right. But philosophy also has more specific contributions to make that are often under-appreciated.

Psychology

Psychology developed out of philosophy in the 19^{th} century. Psychologists, like philosophers, are concerned with the nature of concepts, how we form beliefs, and the relationship between mental contents. But, while some philosophers use empirical methods, many do not; psychologists, on the other hand, are expected to always ground their work in empirical evidence. The psychologist can also engage in conceptual analysis and the development of theories, but she is concerned primarily with what she can establish in the lab or in the field.

Experimental psychology dates back to the latter half of the 19th century. Wilhelm Wundt (1832-1920), a German physiologist and philosopher, opened the first laboratory for psychological research at the University of Leipzig. Wundt wanted to study those structural features of what he called the "intellect" to show that our mental representations of causation and concept were not themselves a product of sensory stimulation, but innate structures of cognition. Wundt's research is still a touchstone in psychology of perception, and in current debates in philosophy and psychology regarding the boundary between concepts and percepts, but we remember him more today for his part in establishing psychology as a discipline within the university.

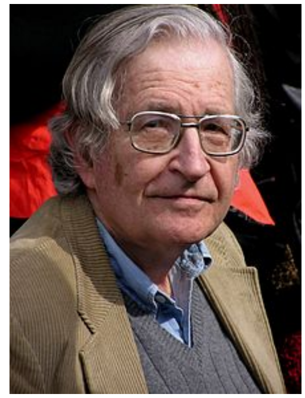


Before Wundt, psychologists didn't yet have labs to facilitate their work, but they nonetheless engaged in empirical research. Gustav Fechner (1801-1887) was the first to produce an experimental paper within psychology, hitherto an entirely theoretical discipline associated with philosophy. Fechner studied mind-body relationships under the label of "**psychophysics**", and was particularly interested in how we perceive light and colour. He demonstrated **Weber-Fechner's law**, which measures with mathematical precision the relationship between the intensity of physical stimulus (say, the brightness of a light) and our perception of that change. This was he first example of a psychological law that would allow for the precise co-ordination of psychological and physical phenomena, thus turning the mind into something that could be scientifically measured.

Psychologists' methods have expanded in the ensuing century and a half since Fechner's work, but the accurate description of mental phenomena is still the stock and trade of psychology. Psychologists are also concerned with offering mechanistic explanations regarding how we engage in thinking—when we get thinking right, and when we get it wrong. Psychologists have a wide-range of concerns, mapping onto various mental functions or faculties: perception, cognition, attention, action, emotion (or affect), intelligence (including intellectual development), personality, verbal and non-verbal behaviour, abnormal behaviour, psychopathy, etc. In this course, we will be concerned mostly with what psychologists have to say about perception and cognition, though we will touch on some other related concerns as well, e.g. the role of emotions in cognition.

Linguistics
Linguistics is the study of language(s) and linguistic ability. Linguists usually focus on the linguistic abilities of human beings in particular, but some linguistic ethologists study language (or at least proto-linguistic communication) in non-human animals. The primary areas of focus are **syntax** (language form), **semantics** (language meaning), and

pragmatics (the use of language in context). Linguistics also has roots in philosophy. Since at least the 1950s, however, it has been a scientific discipline pursued by folks trained primarily *as* linguists, but who often have backgrounds in multiple disciplines, such as anthropology, psychology, and philosophy—and sometimes mathematics.



Many, including George A. Miller, argue that linguistics is inaugurated as an autonomous scientific discipline with the work of Noam Chomsky (b. 1928, pictured to the right) starting in the 1950s. Chomsky argued that humans have an innate capacity, a language faculty, and that this shared faculty (which is part of our cognitive architecture) means that all human languages should share a common structure: what Chomsky called **Universal Grammar**. The syntax rules for forming recognized expressions in any given language, e.g. English, Swahili, or Urdu, must conform to the principles of Universal Grammar.

Even though the underlying cognitive structures required to explain linguistic competency and performance are "hidden", i.e. are not directly observable, linguistics is still a science. This is because we can show, via a structural analysis of languages, that there are common grammatical characteristics belonging to all of them—like **recursion** (the ability to embed meaningful expressions within other meaningful expressions) and **productivity** (our impressive ability to generate linguistic utterances that have never been uttered before, and yet easily understand them). Linguistics studies those hidden faculties that explain our productive abilities, and our linguistic comprehension. We can properly understand the workings of language only if we assume the mind can take a finite **lexicon** (or vocabulary) and a finite set of rules that allows for the production of indefinitely large and indefinitely many expressions which can be understood using those finite means. Studying those cognitive capacities that are necessary to explain our behaviour, linguistic and otherwise, has a long philosophical pedigree. Chomsky's programme in linguistics is one of the best examples of where it has been successful, getting a whole new scientific area up and running.

Anthropology

Anthropolgy is the study of human behaviour in past and present societies. In this course, we are especially concerned with two branches of anthropology: cognitive anthropology and linguistic anthropology. **Cognitive anthropology** is the study of human cognitive capacities based on human behaviour and cultural artifacts (especially tool technology). By looking at a society's use of tools, cognitive anthropologists hope to gain insight into how people engage in problem-solving within organized communities. This kind of "reverse-engineering" is very useful when we look at past societies, especially when those societies did not leave any *written* evidence about the way they organized themselves socially, or the way they solved problems—e.g., the problems of procuring shelter, engaging in hunting and gathering of food, or migrating with food sources (or migrating away from environmental and other threats).

The other branch of anthropology with which we are concerned is **linguistic anthropology,** which examines the differences between languages (or dialects of the same language) to see how language might influence social life, perception, and behaviour. Linguistic anthropology is also useful when we realize that differences in language can also track the geographical movement, inter-penetration, and isolation of different social groups. Linguistic anthropologists are primarily concerned with how language is used to communicate information, to keep information secret through in-group coding, and to inform the development of the groups themselves. Studying the language of people is also thought to give us insight into their culture, and the way they see the natural and social world they occupy. This is why it is such a massive loss when we lose languages (i.e. when people stop actively speaking a language), or when we lose collected knowledge about specific cultures and their languages. This sense of loss was acutely felt when flames engulfed Brazil's Museo Nacional in September 2018. The fire wiped out sound recordings and written records for several now-extinct Amazonian languages. We lost a key insight into those cultures, and those ways of seeing the world.

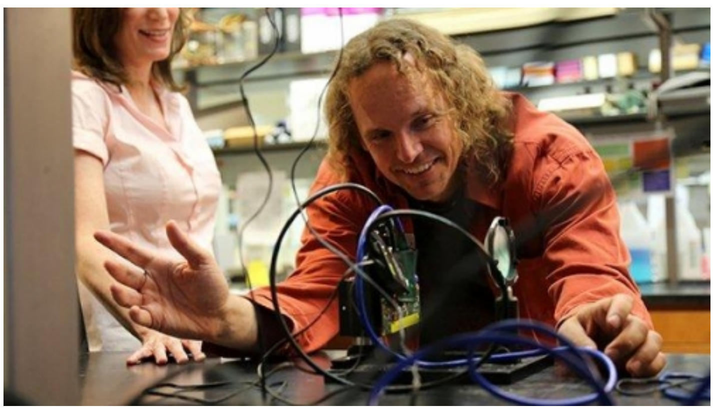
Computer Engineering/Programming/Artificial Intelligence Research
Computer Science (i.e. research focussed on computer engineering, computer programming, and artificial intelligence) is the practical study of digital (and perhaps non-digital) computers. Computer Scientists theorize and build complex computing machinery that is able to take a set of "problems" that can be "solved" algorithmically. An **algorithm** is a bit of formal mathematics that lays out a rule (or set of rules) that, if followed, will produce a solution as output from a set of inputs.

But computer science also provides us with a powerful analogy for the mind, as is shown by the "computational representational theory of the mind", which will be discussed on Lessons 3 and 4 of this module. Computer science also provides the other disciplines within Cognitive Science with useful tools, such as computer models/simulations. These allow us to see what happens within an interconnected system, given certain initial conditions we input to the computer program that abstractly simulates the system. This can be extremely useful for our thinking about actual systems that are thought to share some structural affinities with the computer model. We do this all the time in population studies, and when we track weather systems. But in this class, such modeling techniques are useful because they can teach us about how people learn languages, how languages change over time, and how we form beliefs in light of certain environmental stimuli—important for theoretical work in anthropology, linguistics, psychology, and philosophy.

Neuroscience

Neuroscience is the study of the nervous system, usually with a focus on some specific part of the nervous system, and usually some part of the brain. Neuroscientists use animal testing, computer models, and imaging technology (such as functional magnetic resonance imaging, or fMRI) to see what is happening in the brain at the level of neural activation and inter-neuron communication. **Neurons** are nerve cells that carry electrical signals within a network, where neurons "upstream" from the signal either activate or do not activate based on what their neighbours are doing. "Neighbours" are determined by synaptic connections, so that two neurons are neighbours if they are connected by a **synapse**. Understanding patterns of activation in a neural network allows us to see how the brain and nervous system are functioning when we engage in cognition, e.g. when we parse linguistic utterances of fellow speakers, solve puzzles, empathize with a friend, or remember someone's address.

We know that we engage in certain cognitive processes; this requires no nuanced brain science. But learning the underlying neural mechanisms that make cognition possible offers us not just an explanatory tool for how thinking gets done. It also tells us how we might intervene in the brain to cure cognitive deficits, or how we might train a brain to re-learn some ability after damage to those parts of the brain normally associated with that ability. This can be very important for teaching people how to speak again after a stroke, or for understanding how cognitive function can be impaired by neural degeneration, say in Alzheimer's patients. As we shall see in Module 10, **connectionist theories of the mind** have expanded our understanding of machine minds, programmed at the level of artificial neurons that form an **artificial neural network**. This might be extremely significant for AI, especially if we want to design AI with cognitive capacities like our own. The Module 10 reading from Chris Eliasmith, a philosopher and cognitive scientist working out of the University of Waterloo's Centre for Theoretical Neuroscience, will show us how we might build artificial brains. Eliasmith has already done this with his **SPAUN** (Semantic Pointer Architecture Unified Network), a functional brain that can recognize and remember numbers and patterns much the same way a human brain does.



LESSON C: BEFORE THE COGNITIVE REVOLUTION

INTRODUCTION—TWO QUESTIONS

This course will focus on two main questions: *What does the mind do?*—or as philosophers and psychologists like to say, what is mental function? The second question is related to, and sometimes confused with, the first: *What sorts of things are before the mind?*—or, as we say nowadays, what mental entities must we posit to explain how the mind functions? As you can see, the two questions are related. The first question asks us to think about minds by studying mental activity or mental processes: decision-making, attention, perception, understanding, imagination, memory, etc. The second question asks us to think about minds in terms of mental content: representations of external states of affairs, concepts, mental images, thoughts, feelings, etc. To even begin to get a full picture of what the mind is, we need answers to both of these questions.

For example, think about how you might work out a fairly complicated math problem. You might use reasoned inference, following the rules of arithmetic (or geometry, or group theory, etc.) to arrive at the answer. You might use your imagination at various points to think through a difficult lemma. You might (and should!) write some of these ideas down, keeping track of the inferences in a way that is unlikely to be achieved just by using your memory alone—though you will certainly be using your memory. In order to solve the problem, or even to try but fail, requires the following mental activities: remembering, understanding, planning, reasoning, and imagining. And all of these mental activities have to be coordinated with perception and action. Call all of these activities, and very many more to be explored in the weeks to come, "cognition". **Cognition** just is the process of thinking: of sensing, perceiving, and acquiring knowledge. Cognition, or as some people used to say, "thinking", is what minds *do*.

But we must also remember that we are concerned with mental content, not just mental activity. We also want answers regarding what is being thought *about*. In the case of solving a complicated math problem, you are thinking about a great many things: numbers, geometrical shapes, formulas, inference rules, what sorts of criteria a good proof should meet, etc. Sometimes, the problem will be phrased in such a way that it is about physical objects: the number of apples in a basket, or when two trains, leaving two stations, travelling at different speeds, might pass each other on the tracks. But some of the things we think about are less concrete. They are not things that I can obviously judge to be such-and-such using just my five senses, like I can with apples and trains, tables and chairs, and the other concrete things with which I might have extensive experience. I can see a table, and how it is extended in space and time. I can get much of the same information from my sense of touch, though with touch I also receive information about how the table feels against my skin—if it is rough or smooth, if its edges are pointy, etc. I could rap the table with my knuckle, and hear the sound it makes—maybe even coming to form the belief that the table is mostly hollow with a veneer (like a table from Ikea) as opposed to it being made of solid wood. Your sense of sound lets you "see inside" of things without destroying them. (Your sense of hearing means that you can romantically eavesdrop on the beating of a loved one's heart without the need for advanced and dangerous surgery!) You can also taste and smell the table, though this is rarely recommended.

But abstract objects, like numbers, aren't very much like tables. They have no obvious spatio-temporal location. They aren't *in my kitchen*, at least not in the same way as my kitchen table. And it's not just numbers! I have several uncles. Some of them live in Canada, but many live in the United States. Some of them have died. But whether in Toronto, Fresno, or the grave, it makes sense to think of them as having some spatio-temporal location. But what about the property they all share? They all possess—through birth, marriage, or adoption—the property of *being my uncle*. But where does this property reside? Where and when is the concept MY UNCLE? (The question doesn't even make sense.)

So if we want to learn about minds, we have to study them through their activity, and through their contents. But immediately, a methodological problem arises: a problem that goes back at least as far as Descartes (pictured, right), and maybe even to the ancient Greek philosophers we will be reading next week.

The problem is this: There is a certain sense in which the mind seems *closed off* to the study of its activities and its contents. When I see you trying to solve a problem in geometry, I might recognize some of the telltale signs of cognitive activity. You might scratch your head from time to time, or scrunch up your face. There are, in fact, many ways you might evince your frustration. And, when you solve the problem, your eyes might widen and brighten. A smile might take up residence on your previously sullen mug. But these external cues are *only* cues: they do not tell us much about the underlying cognitive states. And they tell us little, if anything, about those peculiar abstract objects about which you were thinking. In fact, if we take Descartes seriously, they might just be a bit of theatre: For all that anybody else can tell by looking at you, you might not even have a mind! You might just be some kind of well-designed **automaton**: a machine that looks as if it is thinking, and produces the right kind of output behaviour, but which has no mind at all.

Even if you are not willing to take Descartes seriously regarding the possibility of automatons, everyone working in the sciences of the mind has to take seriously a more practical methodological problem. Even with fMRI and other imaging technologies, we still have only very limited abilities to see into the brain, and neural activity might give us very little information about what is happening at the level of cognition. And we still don't have any account of how there can be things like numbers before the mind—abstract things that can be thought about by many minds, or by the same mind at many different times.

So how do you study something that, for both principled and practical reasons, seems so closed off, so opaque to the usual forms of scientific inquiry?

BEFORE THE COGNITIVE REVOLUTION

The 1950s and 60s saw a sea-change in the study of the mind. Because these changes were so rapid, and they constituted such a fundamental shift in the way we research the mind, they are often referred to collectively as "The Cognitive Revolution". (See George A. Miller's short piece on the history of these shifts in our readings for this module.) Cognitive Science was born out of revolutionary work in computer science, cognitive neuroscience, and linguistics starting in the 1950s and 60s. While the Cognitive Revolution can be characterised in a number of different ways, we usually make sense of it as a methodological reaction to some rather restrictive research programmes in psychology, linguistics, and the mathematical study of communication. In particular, psychological **behaviourism**, sometimes of a quite radical sort, had become quite popular, especially within the American academy. (Miller, 2003; see also the previous lesson in this module.)

We will be examining a much longer history in Module 2, but a bit of the immediate history of the Cognitive Revolution is necessary if we want to make sense of why these changes in the study of mind were so profound for so many academic disciplines.

In psychology and linguistics, researchers like B.F. Skinner (1904-1990, pictured to the right) tried to limit their theories to talk only about the stimulation of an organism's sensory apparatus, the physical response of the organism to this stimulus, and various aspects of an organism's environment. This meant psychology could be done without making reference to the internal mental states of the organism (if these states even exist). One could study minds by reducing them to behaviour. Even complex behaviour could be explored in this way. For example, language learning for human beings could be explained by examining how **positive reinforcements** are correlated with linguistic output that matches our expectations for syntax: the speaker is rewarded for making syntactically well-formed utterances. When the speaker produces a grammatically flawed utterance, then the speaker will meet with some kind of negative **reinforcement**: censure and correction. Over time, the organism will speak in accordance with recognized grammatical rules due to **operant conditioning**. Behaviourism as a research programme could reduce psychology to the study of behaviour.

The Cognitive Revolution challenged the behaviourist research programme in psychology, and variants of it in philosophy as well. It did so not because there is some fundamental error in behaviourist theory. Behavioural psychology can explain a lot of what is going on in the mind. But, as a unified and exclusive explanatory framework, there is much about the mind that cannot be explained using only the theoretical tools available to the behaviourist: **stimulus and response measurement**, and the associations between behavioural output and reward/punishment supposed in operant conditioning theories of learning.

Here is an example to help show the problem. Many people have heard of the Spanish explorer and conquistador Juan Ponce de León (1474-1521). While he was an important figure in the Spanish military, and was thrice appointed to the position of Governor of Puerto Rico, most people know little of the actual Ponce de León. If you are from Florida, you might know that he was the first Spaniard to attempt to establish a permanent colony there. But likely, you don't even know this. The only "fact" most people know about Ponce de León is that he was searching for the *Fountain of Youth*. It is almost certainly apocryphal that Ponce de León was driven by this search, as opposed to, say, imperial conquest, but let's assume that the legend is true—that Ponce de León really was looking for the Fountain of Youth. If you are a behaviourist, how do you make sense of his behaviour under this hypothesis? How do you even form the hypothesis?

The problem isn't just that the Fountain of Youth is absent. A behaviourist can explain how a rat will negotiate a maze in search of an expected reward, say some bit of food at the end, even if the food is not actually there. That is because the experimenter knows the history of conditioning this rat has gone through. In the past, it has several times negotiated this maze for a piece of fruit. Experimenters have enough experimental data about rat behaviour that they can categorize it as "reward seeking behaviour" even when there is no reward at the end of the maze. Pavlov's dog will salivate when the bell is rung, even if there is no tasty treat! All that we need is a history of bell-ringing followed by food as a reward.

But the case of Ponce de León is different. The Fountain of Youth is not just absent; it is mythical! There is no Fountain of Youth, and there never has been. It has played no part in a history of operant conditioning, and has never been (in fact, cannot be!) a physical part of Ponce de León's environment. This means that no experimenter, no matter how careful she may be, knows what "Fountain of Youth seeking behaviour" looks like, for she has never seen any test subject search for the Fountain of Youth. The Fountain of Youth is not a stimulus, is not a response behaviour, nor is it an objective feature of an organism's environment, or part of its behavioural history. If a scientist wants to study

Ponce de León's yearning and search for the Fountain of Youth, she has to speak about what Ponce de León *has in mind*, and this means re-introducing mental terms into her study, over and above behavioural terms. Ponce de León is certainly *thinking* about an object—just not one that actually exists in space and time. Maybe it is a figment of his *imagination*, but this means we have to talk about the imagination—precisely the sort of closed-off mental faculty that behaviourism wanted to avoid in its explanations of behaviour.

We will come back to talk about mental objects in terms of their **intentional content** later in the course, especially Modules 6 through 8. For now, let us just note that complete explanations of agent behaviour require our theories of cognition and action to posit more than what the strict behaviourist wants to allow. This realization will have important consequences, especially in linguistics. We will explore some of those consequences in the next Lesson, and also (in more detail) in Module 4 when we discuss the influence of Noam Chomsky work in linguistics.

LESSON D: THE COGNITIVE REVOLUTION

INTRODUCTION

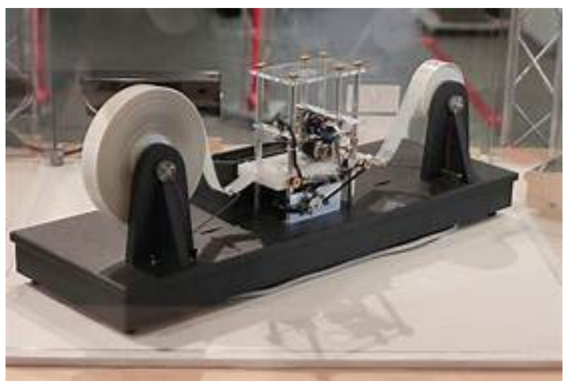
In the period following World War II, advances in communication and computing technology, as well as advances in brain science, gave rise to a new way of thinking about the mind and how it works. Scholars from a number of different disciplines began to think about the mind like a *computer*. The analogy has been a powerful one ever since. Computers are like minds in some important ways, after all. Like computers, we mind-havers take in information about the world. Like computers, we perform various operations on the incoming information. Like computers, we end up producing some output based on the information we have processed. But for this intriguing analogy to become a central tenet of interdisciplinary research, we needed a better understanding of what computers were, and how computation was supposed to work.

Alan Turing (1912-1954) was an English mathematician, cryptographer, philosopher and computer scientist. His work was influential on the development of digital computing (which is how most people remember him), but he was also responsible for more general work in mathematics regarding computability. **Computability** is the ability for a system to solve a given problem with the use of an algorithm**:** a set of rules stored in memory to be followed. Say the problem I want a solution for is arriving to my first class on time. If the class is on campus, then I want to arrive at the lecture hall at the right time when classes begin. The algorithm can be seen as a set of rules to be followed like steps in a recipe. Call this set of rules *S*. If *S* is mechanically followed, the solution will be achieved:

1. Log on to your student timetable on the University of Waterloo website.
2. Enter your WATIAM credentials.
3. Search "Class Timetable"

4. Find “Introduction to Cognitive Science, Phil/Psych 256”
5. Record the time and place of the lecture.
6. Arrive at the room recorded in step (5) at the time recorded in step (5).

Turing showed not only what kinds of problems can have an algorithmic solution, but also what a machine would look like that could perform such calculations. A machine that can arrive at some answer in a finite number of steps, using a finite number of rules, is called a “**Turing Machine**”. This idea was introduced in Turing’s 1936 paper, “On Computable Numbers, With an Application to the *Entscheidungsproblem*”. Turing just called it a “computing machine” in that paper, but we’ve since named it after him. Whatever a Turing Machine can do can also be done within the **lambda calculus**: a formal system developed by the mathematician and logician Alonzo Church (1903-1995, pictured to the right). We aren't concerned with a formal presentation of what a Turing Machine can do, or the expressive power of the lambda calculus. This isn’t a logic or math class. But we *are* concerned with the underlying notion that Turing conjured to talk about computability: *a digital computer*.



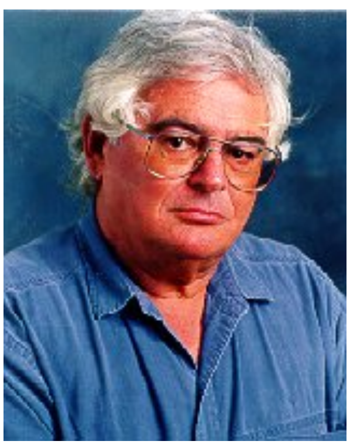
THE COMPUTER-MIND ANALOGY

When we say that the mind is like a computer, we mean “computer” in Turing’s formal sense. A computer has a finite set of rules. These rules tell the computer what to do when it is given a finite set of symbols as input. It computes (i.e. applies its rules) based on the syntax (or form) of the symbols, regardless of whether or not those symbols are taken to refer to things in the world. The computer doesn’t care about these reference relations—i.e., it has no capacity to tell whether or not there are things out there in the world to which the symbols refer. It is merely concerned with the form, or shape, of the symbols—say, the difference between a “1” and a “0” in binary code. The computer then follows the rules in a mechanical way to produce an output. This output might be just a new string of symbols that solves the original problem, or the computer might be hooked up to some kind of action mechanism—say, the unlocking of your phone, or the raising of the mechanical arm when you flash your parking pass in a parking garage.

On the face of it, some of our thought looks an awful lot like what computers do. When you do your taxes, solve a math problem, try to catch the next train, or figure out what courses will fit with your schedule, your mind is, for all intents and purposes, like a computer. But a very powerful theory of how the mind functions and what the mind is became popular, first in philosophy in the early 1960s, and then in cognitive psychology and AI. This new theory does not just say that your mind is *sometimes like* a computer; it says the mind just *is* an information processing system, so it *always* functions in essentially the same way a computer does. We will now introduce this theory. In some ways, the entire course is set up as a study of this theory of mind, including its triumphs and problems.

COMPUTATION AND REPRESENTATION: HOW THE MIND FUNCTIONS

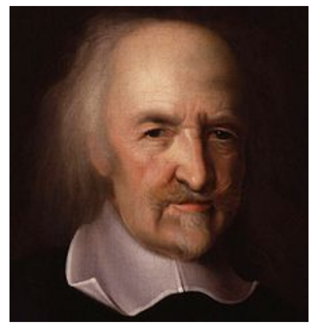
This way of thinking about the mind has been referred to as the Computational-Representational Theory of Mind. Some have noted that it is more than a theory. If former University of Waterloo philosopher and cognitive scientist Paul Thagard (b. 1950, pictured right) is correct, it is an "Understanding" of mind shared by many different theories. For our purposes, it makes no difference. You can call it the "Computational-Representational Theory of Mind" (CRTM), or "Computational-Representational Understanding of Mind" (CRUM). Some people just call it the "Computational Theory of Mind" (CRM) or just the Representational Theory of Mind (RTM). For our purposes, these labels will all be taken to mean pretty much the same thing.



The CRTM is not the only theory of mind in cognitive science. When we get to the end of the term, we will look at some others. But it has been the most common theory of mind in the context of cognitive science discussions, and it has driven research in cognitive psychology, computer programming, linguistics, and philosophy of mind like no other theory in the past half-century or so. Jerry Fodor (1935-2017, pictured left), the prominent philosopher and cognitive scientist, likes to say that CRTM "is the only game in town", meaning that it is central to all research in cognitive science. While this might be an overstatement, it is not far from being true, at least as a historical claim about the study of mind since the 1950s and 60s.

CRTM argues that minds (human minds, animal minds, machine minds, and even alien minds—should they exist) compute over the syntax of representations they receive. In the case of minds like our minds, we receive these representations via our sensory-motor system, and the output of processing is usually tied back to action when we move our bodies or produce linguistic behaviour. In the case of computers, the inputs come from a keyboard, mouse, or through its connection to a network.

One startling consequence of adopting the CRTM is that it opens up a whole new field of research: artificial intelligence. If minds are just computers, and we know how to make computers, then (in principle) we know how to make minds! It might be the case that there are practical obstacles when it comes to making minds—especially human-like minds—but at least in principal scientists can reference CRTM as a framework that makes sense of their research programmes. Before Turing provided us with the mathematical model of the computer, lots of philosophers and psychologists flirted with the idea that the mind is just a machine. This idea actually goes all the way back to the Scientific Revolution of the 17th century. As we shall see in the next Module, some philosophers and scientists in the 17th century believed that the mind was independent of the body. For example, René Descartes argued that the mind and the body are two separable substances, which just happen to be combined together over the lifetime of a human being.



But some contemporaries of Descartes, such as the English philosopher Thomas Hobbes (1588-1679, pictured to the right), already disagreed with Descartes' distinction between mind and body as different kinds of substance. If mind was anything we could talk about sensibly, then we had to make sense of some underlying physical mechanism. For Hobbes, our ability to engage in rational deliberation, to store and access memories, to solve problems, etc. had to be explained by the often slight movements of the small bits of matter (which, at the time, were dubbed "**corpuscles**"), especially those in the brain. Unlike Turing, Hobbes never sussed out how the mechanism was supposed to work, and there is little chance Hobbes could have come up with the idea of a computer in the 17th century. Nevertheless, the idea that minds are just physical systems, in principle the sorts of things we could build, has a rich pedigree. Turing was just the first to show in detail what kind of a machine could be intelligent.

A second consequence of CRTM is related to the first. Because a mind is conceived of as a computer program, it could possibly be realized in a number of different ways. Just as there are different kinds of hardware that can run a favourite app, or video game, or digital video, so too is it possible that minds might be realized by brains like our own, or brains of very different architecture, or by some non-organic system. The language of "realization", and the idea that minds, or at least certain kinds of mental function, might be **multiply realized** by many different physical systems, is a philosophical idea known as **functionalism**. Minds aren't identified with the physical systems that realize them, but are identified, rather, in terms of their functions—by what they *do.* Anything that can reason, remember, problem-solve, imagine (simulate), attend to its own mental states, etc. is a mind. And these various functions are understood, in turn, in terms of their role in generating outputs on the basis of the inputs they process. Functionalism plays a key role in the development of artificial intelligence. Without it, we wouldn't be able to conceive of the possibility of machine minds. Prior to the development of functionalism, the most attractive theories of mind in psychology and philosophy were behaviourist or identity theories. If a mental process just *is* a set of behaviours, or just *is* a brain state, then entities that behave differently than us, or that have a different kind of brain, cannot undergo that mental process. Functionalism breaks the spell of behaviour-chauvinism and brain-chauvinism in the study of mind, since the same functions can (in theory at least) be executed by different kinds of physical system, just as a given software routine can be carried out on different hardware architectures.

The third consequence of CRTM is perhaps the most significant. CRTM assumes that the mind comes ready made with at least some "programming" in place. Recall that a computer just is this program implemented within a physical system. The program is the set of rules the computer has in place for manipulating symbolic inputs to produce symbolic outputs. If the human mind is just a computer, then it too must be programmed with rules—though their nature may as yet be unknown to us. Regardless of the details, though, we have some evidence (given our linguistic and conceptual capacities) to think that some aspects of the cognitive system are innate.

This became clear with Noam Chomsky's criticism of Skinnerian behaviourism as applied to the study of linguistic behaviour. In 1957, B.F. Skinner published *Verbal Behaviour*. In this book, the eminent American psychologist argued that we could explain how human children learn language by operant conditioning. (For an overview of Skinner's use of the notion of operant conditioning, see Lesson C of this module.) Human beings enrich their vocabulary, and learn grammar, by being positively or negatively reinforced at the time of a successful (i.e. grammatical) use of a phrase, or punished for an unsuccessful use. Over time, linguistic behaviour can be tailored in the same way that operant conditioning can tailor any other kind of behaviour.

While we will hold off until Module 4 for a complete discussion of Chomsky's contributions to linguistics, a brief on those ideas will help us to see what was so revolutionary about the Cognitive Revolution. In a review of Skinner's *Verbal Behaviour* (Chomsky, 1959) Chomsky argued that Skinner's account of our language acquisition made little sense. This was for three main reasons. First, children learn a language with such rapidity in an extremely small window of time, and there is evidence that children who are not stimulated by the speech of others during this window cannot be compensated for the deficit later in life. This means that language acquisition has to take place in a **critical period**, which makes little sense of it is just about operant conditioning. We should be able to learn by operant conditioning in a range of contexts, at any age, but evidence from linguistics and psychology shows this is not true for how children learn language. Second, the pace at which children acquire language during the critical period is astounding. They readily master very complex grammatical rules with only very occasional error. What's more, they seem to do so simply by being embedded in an environment with other language users. Operant conditioning is neither sufficient for explaining these rapid changes in ability, nor is it necessary. Children learn language without reward. They don't even require that much by way of oversight from adults. Being at the knee of their parents seems to suffice. Third, Chomsky argues that there are certain structural similarities between all human languages. The behaviourist has a hard time explaining why this is the case, since language rules ought to be different given different regimes of reinforcement and punishment.

Chomsky argues that we must have some **Language Faculty**. This is hardwired within the cognitive system, and it helps explain how we acquire language. We don't really *learn* language as Skinner says we do; rather, meagre environmental stimulus triggers an innate faculty—one that explains our surprising ability to, at a very young age, generate and understand sentences which we have never heard before. In fact, take any sentence in this lesson as a great example. Even though you have never read or heard most of these sentences, you can still easily understand them—at least the ones without the technical vocabulary! This is because all possible grammatical utterances follow certain rules of construction, and these rules are innate to the cognitive system. We have to posit them in accordance with our theories, and make reference to their existence, when explaining certain kinds of phenomena, such as language acquisition. Maybe there are other such theoretical posits. We will consider this possibility as we progress through the course.

Turing started this shift, but Chomsky completes it. According to the CRTM, there are certain internal states of the cognizer that are innate (e.g. the "rules" of the "program"). This is a dramatic turn from the strict behaviourist research programmes that dominated in psychology, linguistics, and philosophy from the 1930s to the 1950s. The shift is so dramatic that the use of the word "revolution" in "Cognitive Revolution" seems apt. It was once again acceptable to posit mental entities and mental processes in describing and explaining natural phenomena. The mind seemed less "hidden" than it did before. Now we understood that the mind was a functional entity, defined by what it *does*. We also understood how some of these cognitive processes were associated with the brain (thanks to advances in neuroscience), but not solely identifiable with the brain—since a computer could realize some of the same functions. Because of Turing's important contributions, we were no longer embarrassed to talk about the role of internal states when explaining how the mind worked. These were all important methodological advances.

Very rarely does the research community change its fundamental view of key subject matter so quickly and so thoroughly. Exploring the nature of this revolution, and the new research it made possible, will be a guiding theme for this course.