Supplementary notes for lecture 8:
Computational modeling of cognitive development

Slide 1

Why computational modeling is important for studying cognitive development.

Let's think about how to study the mind and brain. Why would taking a computational approach be useful? To show you why, I'm going to give an example not about the brain but about this artifact here.

Slides 2-3

Here is an artifact that looks like a simple candle, but it has some additional properties. There are some markings on the back, there's a metal base, and there's a nail attached to it. Here is a drawing that makes it easier to see all the parts.

Imagine that for some reason, you want to be able to predict how this artifact behaves with very high accuracy.

Slide 4

One way you can solve this problem is by taking a bottom-up approach. The idea is that if you're trying to understand something that is very complex, you can break it down into its most basic parts. Once you understand the basic components, and how they interact with each other, you can make a model of the system and predict its future states.

Slide 5

With this artifact, you might begin by studying the basic properties of wax. Then you can move on to studying the properties of the flame. You can study the material the nail is made of, and you can look at its precise shape. You can make very precise measurements of the markings on the back and you can study everything about the metal plate on the bottom.

Once you understand all of this, you can predict what will happen to the whole system. If you find one of these artifacts, you can change a few parameters in your model, run a simulation, and predict what the artifact will be like at any given time. Again, this type of approach is called bottom-up.

Slides 6-11

The top-down approach takes the opposite perspective. The idea in the top-down approach is that if you're trying to understand a system that serves some function, you can begin by figuring out what that function is. Once you know that, you can then begin to understand what each component does in relation to its functional purpose.

If we take a top-down approach with our artifact we would begin by trying to figure out what its function is. If you did this, you would notice that these artifacts are only used in the nighttime. You would see that every night a person comes up, looks at the markings in the back, puts a nail in the candle, lights the candle, and goes to bed. As the night passes the flame keeps burning until the wax holding the nail melts. The nail falls down and hits the metal base making a loud noise. This sound wakes the person up and then the candle is extinguished.

Notice the difference of what you understand using each of the two approaches. Someone who took a bottom-up approach might be able to predict the future states of the artifact, without ever knowing that the artifact is an antique alarm clock. In contrast, someone who took the top-down approach might not know anything about the specific properties of wax, or about the exact shape of the nail, but knowing that the object is an alarm clock allows her to predict what will happen in the future. If you know that the artifact is an alarm clock, you can look where the nail is placed relative to the markings in the back and you'll be able to predict when the artifact will make a loud noise, without any need to simulate the underlying complex physical process that led to it.

Slide 12

In this example, the top-down approach lets us realize that the candle artifact serves the same function as mechanical or an electrical alarm clocks. You can understand that they are all performing the same function using different physical implementations (wax and nails, springs and gears, and capacitors and transistors), and different algorithms (through burning wax, having springs oscillate, or by charging capacitors). And this type of understanding can be much more difficult to reach through a bottom-up approach.

These three levels of understanding that you could have -The function or computation, the algorithm, and the implementation- are known as Marr's three levels of analysis.

A full theory of the mind and brain (or of alarm clocks) must be able to explain what is happening on all three levels, but where we should begin is still an open question, by studying the implementation (bottom-up), or the computation (top-down).

The top-down approach is convenient from an artificial intelligence perspective. If we can understand how intelligence works at the computational level, then we may be able to implement it using different algorithms in different technologies.

Slides 13-14

So far we've only been talking about how to study alarm clocks. What about the brain? Intuitively, there's something off with the analogy. Alarm clocks were designed to serve some function, is this true for the brain as well? This is a difficult question, but many people believe this is the case. The brain evolved to perform certain computations, just like other organs. We can understand other organs purely at the functional level: Hearts pump blood, lungs transport oxygen into the bloodstream, etc. Much in the same way, we can assume the brain is serving some function, take a top-down approach, and see how far this takes us.

Slides 15-16

To recapitulate, we want to find what we gain by studying the mind and brain at the computational level of analysis, where we want to understand the logic of computations, not the specific algorithm of implementation.

There are many ways we can do this. Every approach has its strengths and weaknesses, but they all have the same underlying philosophy. Here we'll focus on Bayesian models of cognition because they've been pretty successful in helping us understand many aspects of higher-level cognition in recent years.

Slide 17

We need to make some initial assumptions to get off the ground. We don't know how the mind represents knowledge, but we are going to assume that our belief about anything can be quantified and that we can assign it a real number between 0 and 1, where 0 is something like "definitely not true," 1 is "definitely true," and 0.5 is complete uncertainty.

We can use this scale to describe how certain we are that this animal is a cow, that this baby is happy, or that the yellow blocks will be the first to fall if this table were bumped.

Slides 18-22

Once we've made this assumption, we can use Bayes' rule to describe how our beliefs change as we interact with the world. Intuitively, Bayes' rule tells us that the belief that a hypothesis is true given the evidence we observe is proportional to the belief we had in the hypothesis before observing the data (your prior belief) times the likelihood that you would have observed that data if the hypothesis were true.

Slides 23-30

To get a better idea of how Bayes' rule works in cognition, let's look at a toy example in world learning. Imagine you're trying to figure out what the word "Dier" means. You know it means "Animal", "Mammal", or "Giraffe", but you don't which one it is.

These are our hypotheses and since we have no idea which one is true we assign a prior probability of 1/3 to each.

One day, you're walking with your friend (who doesn't speak English) and you come across four animals: An alligator, a giraffe, a tiger, and an elephant. You friend points to the giraffe and calls it a Dier. Now you can use Bayesian inference to update your belief on what Dier means. If Dier meant "animal" the chance of your friend pointing at the giraffe is 1/4 (there are four animals). If Dier meant "mammal" then the likelihood of your friend pointing at the giraffe is 1/3. And if Dier meant "giraffe" then the likelihood of your friend pointing at the giraffe is 1. By multiplying your prior belief in each hypothesis by how likely it is that your friend would've called the giraffe a Dier under each of them, we can get the posterior belief. After normalizing the distribution, the updated belief that "Dier" means "animal", "mammal", or "giraffe" is .16, .21, and .63 respectively.

Slides 31-42

As our friend keeps pointing at other animals and calling them "Dier" we can continue using Bayesian inference to update our beliefs, giving us a model of how we can learn the meaning of a word in a very simple setting.

Slide 43

So far we've just outlined a normative model that tells us how to integrate information into a probabilistic framework and make rational inferences. But the last 10 or so years have found that these models predict very accurately how humans reason in many domains ranging from theory of mind, to intuitive physical reasoning, to pragmatic inferences in language.

So we know that adults do something like Bayesian inference. Although there is some evidence that our brain actually is doing Bayesian inference in many tasks, all we need is the weaker claim that the brain is doing something that is computationally equivalent to Bayesian inference. A way to think about this is like trying to model an alarm clock made of electrical circuits and showing that it is functionally equivalent to a mechanical alarm clock, without making any claims as to whether the mechanical alarm clock is using the same algorithms.

Slides 44-46

We've talked about why it's a good idea to take a top-down approach to study cognition, we have a general framework to build computational models, and we have evidence that adults reason in a way that's predicted by these Bayesian inference models. What do we gain from using the same approach to study development?

There are many reasons why these models would be helpful for studying development, but the main point that cognitive scientists have made is that the most difficult problems in cognition, the problems we are the furthest from solving, are the problems that every five-year-old has solved. If we could formalize what the core components of each domain of knowledge are, and how children learn, then we'll have a computational solution to all of these problems.

Slides 47-48

If you combine Bayesian inference with the Theory Theory – the idea that knowledge is structured in the form of causal theories – we can use Bayesian inference to try to find the best theory that explains the data, design an experiment where we can give infants or toddlers the same data that we give to our model and then compare how well the model's predictions match infant/toddler behavior.

Slide 49

In past lectures we've already read some papers that have taken this approach, where the authors were able to get a better idea of how infants make inferences on some tasks by comparing their behavior to normative Bayesian models.

Slides 50-52

For example, in Gweon et. al. 2010, the authors measured how much infants expected the new balls to squeak by looking at how long the infant tried to make a it squeak. Using this measure they were able to compare how certain the infants were that the balls squeaked to how much confidence a normative Bayesian model would have had under different sampling assumptions. This way, the authors were able to figure out what sampling assumptions the infants were making by showing how well the model predicted the behavioral data.

Slide 53

Conclusion of slides

MIT OpenCourseWare
http://ocw.mit.edu

9.85 Infant and Early Childhood Cognition
Fall 2012

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.