Spring 2017

Knowing How: A Computational Approach

Joseph A. Roman '17, Gettysburg College

Knowing How: A Computational Approach

Abstract
With advances in Artificial Intelligences being achieved through the use of Artificial Neural Networks, we are now at the point where computers are able to do tasks that were previously only able to be accomplished by humans. These advancements must cause us to reconsider our previous understanding of how people come to know how to do a particular task. In order to unpack this question, I will first look to an account of knowing how presented by Jason Stanley in his book Know How. I will then look towards criticisms of this view before using evidence presented by the existence of Artificial Neural Networks to present a new view that addresses the problems present in Stanley's work. Finally, I will argue that knowing how to do something is a matter of heuristics, or knowing certain shortcuts which approximate a solution to the task one is trying to accomplish.

Keywords
Artificial Intelligence, Philosophy, Epistemology, Knowledge

Disciplines
Artificial Intelligence and Robotics | Philosophy of Mind | Philosophy of Science

Comments
Capstone Research produced as part of the Philosophy Senior Seminar.

Joseph Roman

Prof. DeNicola

PHIL 400

March 9th, 2017

Knowing How: A Computational Approach

Certain aspects of human knowledge have always been of great interest to epistemologists. There are certain things that humans almost certainly know and can only be called knowing, but which seems to defy all traditional forms of defining knowledge. One such form of knowledge is the way in which humans appear to know how to do complex physical tasks with little to no cognitive thought. For the sake of speaking on a field which I myself understand, consider a fencer.

When an experienced fencer is participating in a bout, while the fencer may be thinking in terms of general strategy and deciding when and where to attack, the attacks and defenses are almost automatic. Compare this to a new fencer, whose movements will be stiff and awkward as they are forced to think through their actions and their natural reaction to being attacked may actually be detrimental to them. The experienced fencer seems to possess a much greater degree of knowledge of how to fence, but this knowledge is not apparent by looking simply to cognitive action. This demonstrates the difference (or at least apparent difference) between two different forms of knowledge. Those forms are knowing that, and knowing how.

Knowing that refers to knowing something in a propositional or true way. This is often defined as justified, true belief, and takes the form *X* knows *p*. These are factual statements about the world that can be proven true, or false, and tell us information about the world and what a

single person knows. Though this definition has been up for debate and has some clear flaws, it will not be the focus of this paper.

Knowing how is a far more complex problem. Knowing how refers not to some proposition that expresses truth about the world, but rather knowledge about how to take some action. This action can be as simple as ringing a doorbell, or as complex as playing a sport at a professional level. At first glance, this seems to be very different from knowing that, but there is debate to this point. Intellectualists argue that while knowing how appears to be different, knowing how can be accurately and completely explained in terms of knowing that.

For this paper, I will be assessing the Intellectualist argument and then looking at counter-arguments to that position. I will then be using my background in Computer Science, as well as looking at advances in Artificial Intelligence (or AI) to show how, with some slight definitional changes, artificial neural networks help demonstrate that knowing how can be reduced to knowing that through the use of heuristics.

We will first approach one of the stronger arguments for the Intellectualists as made by Jason Stanley in his book *Know How*. The book covers in great detail the various aspects of what knowing how exactly entails. In the 5th chapter, Stanley outlines exactly what knowing how entails.

> "Knowing how to do something is a kind of knowledge-wh, a relation to a question. It is also a first-person mental state… It is time to add the remaining elements of of the analysis of the practical knowledge. These two remaining elements are the *modality* associated with infinitival embedded questions, and the distinctive *mention-some* reading of infinitival questions" (Stanley 111).

We will look at each of these aspects of knowing how individually to understand how Stanley wants to construct knowing how and how this construction supports the Intellectualist approach to knowing how.

The first aspect of knowing how, according to Stanley, is knowledge-wh. This has to do with the idea that statements of fact usually take the form *X* knows that *p*. However, there are statements that exist that do not fit into this model. For example, take some of the example statements from the book such as, “Hannah knows how to vote” and “Hannah knows who Obama is” (Stanley 36). These are statements which imply questions which an agent knows the answer to. In other words, these are statements of fact that can most simply be stated as knowing the answer to a given question. This is a fairly simple, but essential part of knowing how statements.

The next aspect we must cover is the idea that knowing how involves a first-person mind state. This aspect of knowing how actually involves several different aspects to be covered in turn, starting with the idea that a statement that expresses knowing how to do something will contain an infinitive clause, rather than a finite clause. This means that these statements take the following form. “For every s and F, s knows how to F iff for some way of F-ing, s knows that w is a way to F” (Stanley 71). Note here that this way of formulating knowing how demonstrates a way of understanding knowing how that can be formulated into a propositional statement. This is actually the most important point in this entire theory for the arguments I will be making later.

Stanley then goes into great detail describing two linguistic elements referred to as PRO, which is an unpronounced pronoun, and *de se* attitudes, or attitudes about oneself. These two elements are part of infinitival statements and therefore, are parts of any statements about knowing how. These two aspects of infinitival statements, for Stanley, point towards the idea

that statements about knowing how are in reference to opinions about oneself. In other words, when we we know how to do something, we know that we know a way to do something. This excludes knowing of ways that other people may do that act. When we talk of knowing how to do something, we are talking about knowing a way in which we do something.

The last pieces of the puzzle that is knowing how are both covered in the fifth chapter of *Know How*. The first piece is the presence of modalities in infinitival statements. What this is referring to is the idea that knowing how statements often have modalities embedded within the sentence that are not explicitly stated. An example from the book is the two statements, “John knows how to solve the problem” and “John knows how he could solve the problem” (Stanley 113). These two statements are basically equivalent and show that when a person is expressing a know how statement, they are expressing some modality as well. The model above is the “ability” model and it deals with those statements which have an embedded “can”. However, it is possible for knowing how statements to also have an embedded “ought” within them and take a deontic form. This tends to be the exception rather than the rule, and is dependent on the context of the statement.

The last important aspect of knowing how statements is that they seem to have, rather than a mention-all reading, a mention-some reading. To simplify this, this means that a knowing how statement is saying that a person simply knows a way of how to do something, rather than all ways of doing something. However, this approach is slightly problematic. You have to know how to separate those mention-all statements, and those mention-some statements. The answer is that mention-some statements have a limited domain, or number of cases where a judgement can be made. This are based on context surrounding a situation. For example, when we talk about a child knowing how to write, we hold them to a lower standard than a college educated adult.

How we are to construe these domains is complicated and will be discussed in greater detail later.

With these aspects of knowing how statements, we can start to unpack an idea of how to reduce these knowing how statements to knowing that statements. First, we must combine these various aspects of knowing how into one definition. One is helpfully given directly by Stanley.

> "Relative to a context in which the domain is the set of normal ways of of doing things with which John is acquainted, a sentence like 'John knows how to swim' is true at world w if and only if John knows at w a proposition true in a world w′ if and only if there is some way in which John could swim in w′ with which John is acquainted and that way is also a way John is acquainted with by means of which John could swim in w" (Stanley 122).

Stated more simply, this is the idea that John knows a way to swim that is both consistent with the world that exists, and the world that John is aware of existing. By this definition, John has propositional knowledge of a way to swim, which while not explicitly stated, is necessarily propositional knowledge. This explains exactly how Stanley builds his system to support his assertion that knowing how can be reduced to knowing that. However, this theory is not without its own weaknesses.

The first weakness of this theory is brought up in some form in a review of the work done by Kent Bach of San Francisco State that it does not give a robust explanation as to the difference between knowing how to do something, and being able to do that act. In the following passage, he explains how Stanley would reply to this objection.

> "Presumably a necessary condition for this success is that the agent have the resources (physical wherewithal, tools, materials, etc.), opportunity, and motivation (I'll call this the

> "ROM" for short), but knowing-how does not require actually being able to succeed in performing the action, since one might not actually have the ROM. So, we might say, knowing how to do something is the cognitive component of the ability to do it. However this knowledge gets put to work, it would help explain one's success if one were to try to perform the action when the ROM condition is met. But it is not clear that the counterfactual success condition is necessary, since one can know how to do something one has the ROM to do even though one may not succeed, as when shooting a free throw, or is even likely to fail, as when taking a three-point shot" (Bach 12).

He goes on to further discuss this point, but I raise my own objections to this defense. It makes some sense that a person may know how to do something in one context, but with slightly different factors surrounding the situation, a person may be unable to perform a given action. This however, seems to introduce the potential issue that knowing how to do something applies only to a given situation with specific parameters, which it can be argued would not be sufficient knowledge to accomplish that same task in the future. If a person only knows how to perform a task in a very specific scenario, can that be called knowledge how in a meaningful way. I would argue that without a robust explanation of why we can fail to accomplish something we know how to do, the explanation is insufficient. The explanation given, while introducing the idea of context, does not give enough information as to what that context is and what context is important.

The second issue I take with this is that from a pragmatic viewpoint, this theory does not benefit us in understanding some of the underlying issues regarding know how. Stanley admits in his paper that it is difficult, if not impossible to know how to do something well without practicing the thing in question.

> “Stanley briefly discusses the relationship between knowledge, ability, and practice. Consider the difference between knowing how to perform an action fluidly and merely knowing how to perform it cumbersomely. Suppose, for example, that you know how to touch-type but want to do it with more speed and fewer errors. So you practice and get better. Stanley suggests that the benefit of practice is not to give you new knowledge but to enable you to apply your prior knowledge directly, without having "to engage in a distinct action of *consulting* the propositional knowledge" (p. 184). Practice yields so-called "muscle memory," as when you develop a "feel" for punching in your ATM number without having to think about how. You have the pattern of finger movements down pat but haven't learned anything new” (Bach 24).

This is important to the field of Computer Science, in which programming a computer to do these tasks that simply seem to require practice is incredibly difficult. This makes the definition given to us by Stanley, while well stated, not compelling enough to provoke any action or further understanding.

These two issues are the ones that I wish to solve by approaching this problem through the lense of modern computing, specifically Artificial Intelligence. Advances in Artificial Intelligence (or AI) have been somewhat common of late due to a couple of important advances and discoveries. The first way in which AI has been helped by modern computing is by the fact that our ability to create more powerful computers is growing at a seemingly exponential rate. While there are still some serious memory and power limitations, they are far less severe than they used to be. The second major advancement has to do with the creation of Artificial Neural Networks (or ANNs). ANNs have allowed AI to be programmed with learning algorithms that allow an AI to adapt its own code to better deal with a given problem. To explain how this helps

the Intellectualist argument in regard to knowing how, we must first look at how exactly a ANN is built and functions.

In order to keep this explanation fairly simply, I will be using an article written for Medium.com by an employee of the company Josh.ai who goes only by the name Aaron. This is a fairly simple article, but it will provide enough of an overview to suit my purposes. The most important thing to cover is describing how these ANNs work.

> The answer is that we need to essentially ask them a large amount of questions, and provide them with answers. This is a field called supervised learning. With enough examples of question-answer pairs, the calculations and values stored at each neuron and synapse are slowly adjusted (Aaron 6).

Essentially, an AI is given a question, attempts an answer, and then is judged on its answer. If it got the correct answer, or a preferable outcome, it adjusts accordingly, and a similar thing happens if it gets a question wrong. Basically, it takes the questions given and tunes itself to answer the questions given as accurately as possible in the future. This can be done either through direct observation where a human agent trains an AI's responses, or through a self-learning mechanism where an AI can run simulations and judge the outcomes entirely independent of a human overseer.

This approach to AI was created to mimic the manner in which humans learn in order to help create computers that could solve problems that are difficult to solve using normal algorithmic methods. This has to do with the way that neural networks operate as compared to usual problem solving methods.

When we make a program to play a game such as Chess, we normally solve this problem in an algorithmic way. In this way, the computer looks at the current board state, calculates as

many future board states as it can handle, and moves to create the board state that gives it the greatest chance for success based on the board states left available to the computer. In essence, the computer calculates all possible moves and picks the one that gives it the greatest chance of success considering all, or some, possible future moves. This is what I will call an algorithmic solution to the problem. This is incredibly useful in those circumstances where there are a limited number of configurations for a problem, or there is a problem that can be solved the same way regardless of the problems configuration.

ANNs work very differently. Instead of working off of one solution that is coded into the computer from the beginning, ANNs operate based on what are called heuristics. They do not always find the optimal solution to the problem, in fact they almost never do, but they instead learn a set of rules based on previous experience that allow the AI to solve most situations. This means that any solution the AI will decide on is not based on a set of concrete rules that will necessarily lead to a solution, but will take the actions it deems most likely to solve the problem. This is a simplification of the issue, but it is accurate to how a ANN behaves.

With this in mind, we can move to discuss how this advancement in AI affects our understanding of knowing how to do something. These ANNs can be said, in some sense to know how to do a particular act. For example, Google and IBM have created AIs based on these ANNs which can play the game Go, and make medical diagnoses respectively. The two AIs being referred to are Google's AlphaGo and IBM's Watson. AlphaGo recently played a set of 5 games of Go against one of the top ranked players and former world champion Lee Sedol, and won that series 4-1. This was unprecedented simply because Go was one of the last traditional board games that computers had yet to be able to compete at a high level in. In AlphaGo's first real public performance, it not only won the majority of its matches, but seemed unbeatable in

several of the games. An article on the website Quartz covered the matches, and specified some of the specific strange but brilliant moves made by AlphaGo. Here is one such move. “The fourth line from each edge of the board is known as ‘the line of influence,’ and it’s so important to the game that most boards mark it with dots. Young players are taught to play along the line of influence if they are after territory in the middle of the board. But AlphaGo, playing in black, played on the fifth line, which is generally thought to tilt the balance in favor of an opponent” (Wong and Sonnad). This move ended up helping AlphaGo secure the win. Even though this move is not a traditional move, based on the results of the game, I think it is fair to say that, if we do not take any arguments about mental states into account, that AlphaGo meets the requirements for knowing how to play Go.

How then does this help prove that knowledge how can be reduced to knowledge that? This argument is reliant on two points. First, the manner in which ANNs attempt to learn how to do an act is supposed to simulate the ways in which humans learn to act, and while the results are not the same as when we attempt to teach a human to do something and the scope of what ANNs can do is still relatively small, they do seem to be able to perform tasks normally thought to be only possible by humans sometimes better than any human agent. The second point I am relying on is the idea that no matter how complex a computer becomes, it is still simply code that can be copied and transferred to another system if we wish to do so. Essentially, as long as there is another computer that can handle the power requirements, any computer code can be copied and transferred to another computer and it would present no significant problems.

This leads me to my conclusion. Knowing how to do something can be reduced to a set of defined statements that take the form of heuristics rather than normal propositional statements. In

addition, I will argue that heuristics should be considered a form of knowing that because they do state a proposition that can be proved true or false.

The way in which a ANN operates means it is effectively using heuristics to get to its solution. It is not working with an algorithm that works every time, but looking at the context of a situation and picking the solution that is most likely to work. The same can be said of humans attempting to learn how to do a complex action, such as play a sport. We don't learn the best way to do a certain act, we learn a number of mental shortcuts which get us to our destination more often than not.

This sufficiently explains why we fail to accomplish our goals even if we seemingly know what to do. Heuristics operate on the idea that they probably get us to our goal. That means that inevitably, we will fail. When we do so, we revise our worldview and then continue on. This means that even if you do know how to do something, failing isn't inherently problematic for the Intellectualist position.

However, to make this argument, one would have to demonstrate that Heuristics can be seen as propositional statements. This is easy enough to do when you look at how these sentences can be broken down. Let's look at the statement, "John knows how to swim". What we need to do, if we take knowing how to be a heuristic, is break down what exactly "how to swim" entails. Stanley would break this statement down as "John knows a way to swim", however I would take this statement one step further. I would say that "John knows a way to swim that usually works". This statement can be broken down into sub-actions, but every single action within this set should carry some degree of uncertainty within it.

Note that while these statements do rely on probability, the statements themselves can be considered true or false. The uncertainty is only in the success of the action, and not in the truth

or falsity of the overall statement. This means that this statement can fit all the requirements necessary for propositional knowledge.

There is one weak point in this argument, however, which is regarding at what point we can consider a person to know how to do something. How often must a person succeed at a particular act before we can say that a person knows how to do something? If a person knows a way to do something that they are aware almost never works, can they be said to know how to do something? I would argue that a person needs to be able to do something beyond the degree of random attempts in order to be considered to have knowledge how. If a person cannot determine in what situations a method works, and that method works less than half the time, then a person cannot be said to know how to do something. However, any requirement I set can be attacked as arbitrary and I do not currently have the evidence to set a limit where I can refute such claims. As a result, this argument exists in a somewhat incomplete fashion.

However, I would still defend this idea as important in revealing something important about the nature of knowing how to do something. We are living in an age where the existence of computers can help to reveal a great many mysteries present in the world around us. The emergence of ANNs has given us substantial insight into the process of learning in general and we should not discount the information being presented to us. We are starting to know how learning takes place and we as humans should jump at the chance to demystify our understandings of ourselves. ANNs are helping not only to solve problems present in the world outside, but also to provide us better understanding of who we are as people.

Works Cited

Aaron. "Everything You Need to Know About Artificial Neural Networks – Technology, Invention, App, and More." *Medium*. Technology, Invention, App, and More, 28 Dec. 2015. Web. 01 Mar. 2017.

Bach, Kent. "Review, Jason Stanley, Know How." 2012. Web.

Stanley, Jason. *Know How*. Oxford: Oxford UP, 2011. Print.

Wong, Joon Ian, and Nikhil Sonnad. "Google's AI Won the Game Go by Defying Millennia of Basic Human Instinct." *Quartz*. Quartz, 25 Mar. 2016. Web. 01 Mar. 2017.