JOURNAL > SITE 1: LOGIC GATE, THE POLITICS OF THE ARTIFACTUAL MIND | 2017

What Is It to Think?

Danielle Macbeth

What is it to think? Perhaps there is no one answer for everything we might reasonably characterize as thinking but only family resemblances among the various instances. Nevertheless, one sort of case is emblematic, at least in the Western intellectual tradition: reasoning that involves the manipulation of signs according to rules. Some instances of thinking in this sense are algorithmic; not only is each step licensed by a rule, rules also determine which (step-licensing) rule is to be applied at any given point. Doing an arithmetical calculation in the positional system of Arabic numeration is a paradigm of algorithmic thinking, which is why it is easy (relatively speaking) to build a machine to perform such calculations. More interesting cases of rule-governed thinking are not algorithmic. In these cases, each step is licensed by an antecedently specifiable rule but it is not determined in advance which rule ought to be applied at any given point. Chess playing is a familiar instance of non-algorithmic, rule-governed thinking. Although it is not so hard to build a machine, or for that matter, teach a child, to play chess according to the rules of the game, to make only legitimate moves, it is much harder to build a machine, or teach a child, to play chess well, to make only *good* moves, moves that will improve one's chances of winning. Finding proofs in mathematics, at least in interesting cases, is also like this. But although it is hard to build a machine to play chess well or to find a proof of an interesting mathematical theorem, it nevertheless has been done. The question, my question, is: are the machine and human reasoners both *thinking* in precisely the same sense when they engage in such rule-governed manipulations of signs? What I aim to show is that although we humans, we rational animals, can think mechanically, that is, in essentially the way a machine thinks, we humans can also think differently, even when the thinking involves the rule-governed manipulation of signs. The aim is not to show that physical systems cannot think. Clearly some can: we are physical systems and we can think. What is at issue is what precisely it is to think as we do.

Supercomputer Deep Blue at IBM's headquarters in Armonk, New York, on February 16, 1996. Photograph by Yvonne Hemsey/Getty Images.

Perhaps it will be objected that it is *obvious* that what the machine is doing in thinking is precisely the same as what we are doing in thinking because thinking (in the sense of concern here) just is the manipulation of signs according to rules and both we and the machine *are doing that*. Such differences as there are between the two cases—and, of course, there are differences—are simply irrelevant, and there is indeed a way of thinking to which this is obvious. Nevertheless, as our intellectual history over the past three millennia has again and again taught us, what is obvious given one conception of things can be revealed to be false in light of a subsequent conception. Consider, for example, the notion of a number. As the ancient Greeks thought of it, a number is a collection of units, from which it follows not only that zero is not a number, but also that there are no negative numbers, or fractions—although there are ratios of numbers. Indeed, as ancient Greek philosophers liked to point out, even one is not a number on this conception, but only the unit that provides the basis on which to form the collections that *are* numbers. The modern number concept is very different. A number on the modern conception is conceived computationally: a number is what stands in various well-known arithmetical relations to other numbers, in effect, a node in a web of arithmetical relations. And now it seems obvious that zero and one are numbers, and that there are negative and fractional numbers. Similarly, we need to see that although it is obvious that humans and machines think in precisely the same sense relative to the conception of logic and reasoning that is bequeathed to us by modernity—a conception that is, we will see, essentially Kantian—once one moves to a properly post-Kantian conception of logic and reasoning, it is not at all obvious that humans and machines think in precisely the same sense. Indeed, it is demonstrably false.

That both the practice of mathematics and the practice of physics were revolutionized in the 17th century is well known. Less well known is that both were again radically transformed a few centuries on, the practice of mathematics in the 19th century and that of fundamental physics in the 20th century. What is almost wholly unknown is that the practice of logic has followed suit. The received view is, in Quine's words, that "logic is an old subject, and since 1879 it has been a great one."[1] Logic, on this view, was first begun by Aristotle as syllogistic logic and then transformed into our standard polyadic predicate, quantificational logic by Frege in his 1879 *Begriffsschrift*. This is *not* what actually happened. In fact, logic was transformed already by Kant in the wake of Descartes' radically new work in mathematics and philosophy; it was *Kant* who gave us all the essentials of our standard logic, first, the division of terms into two logically different sorts of representations, and with that division the modern notion of a quantifier, and also the idea that logic is purely formal, without content or truth. What *Frege* did was to transform logic again, this time in the wake of 19th-century developments in mathematics due to Riemann and others.[2]

Ancient Aristotelian logic is a term logic, where a term is what things are called: Socrates, snub-nosed, human, sitting, and so on. Terms have, in other words, both referential and predicative aspects. In the subject position a term picks out what it is that is being talked about; in the predicate position a term serves to characterize that which is talked about. To judge, on the ancient Greek conception, is to *predicate of* one, some, or all things of some sort that they are or are not so. It follows that one can *judge of* only what exists, and that only what exists has a discoverable essence and can be defined. Aristotelian logic serves to determine the valid syllogisms involving two categorical sentences as premises and three terms; and these syllogisms are in turn of interest in light of their role in scientific demonstrations, where a demonstration is a syllogism in which the premises are true, primary, immediate, and better known than and prior to the conclusion, which is related to them as effect to cause.[3] The concern of Aristotle's system of logic is actual acts of inference in the context of scientific explanation of what is and must be so. Modern quantificational logic is very different. First, it draws a logical distinction among what Aristotle thinks of as terms between referring and predicative expressions. The distinction appears first in Kant. Kantian intuitions are purely referential; they belong to the receptivity of sensibility and give objects for thought. Kantian concepts are purely predicative; they belong to the spontaneity of the understanding and are that through which (given) objects are thought. This division is, furthermore, a successor of sorts to Descartes' division (grounded in his new mathematical practice of analytic geometry) of sensory experiences, which are confused mental effects of the impacts of external bodies on our sense organs, and clear and distinct ideas that are innate in us and fully meaningful independent of the existence of any objects that might correspond to those ideas. As Descartes explicitly notes, in what he thinks of as the "true logic", as contrasted with Aristotelian logic, it is false that only what exists has an essence; instead, essence precedes existence.[4] What is altogether original with Kant is, first, that intuitions (caused in us) and concepts (that are fully contentful independent of any relation to any object) are logically different as belonging to two essentially different cognitive faculties, and second, that *both* are constitutively involved in any cognition. In Kant's famous slogan: "[T]houghts without content are empty, intuitions without concepts are blind."[5] For Kant, though not for Descartes, all content lies in relation to an object, and all cognitive access to things is through concepts. It immediately follows that logic, insofar as it abstracts from all relation to any object, is strictly *formal*, without content or truth.[6] It also follows that some *new* mechanism of reference to objects is needed if judgment is to be possible. Because Kantian intuitions are blind without Kantian concepts (and hence can have no direct role to play in judgment as Kant conceives it), and Kantian concepts are not themselves in any way object involving, though *judgment* is, Kant thinks, constitutively object involving (because contentful, truth evaluable), we must learn to conceive judgment in a new way. We must learn to conceive judgment as an act of positing a relation of subject and predicate concepts *as* objectively valid, that is—subject and predicate concepts as combined in an object or objects. It is on behalf of just this conception of judgment that Kant introduces the universal and existential quantifiers as the mechanism of reference to objects in a judgment. The familiar Venn diagram, as it contrasts with an Euler diagram, provides a graphic illustration of judgment so conceived.

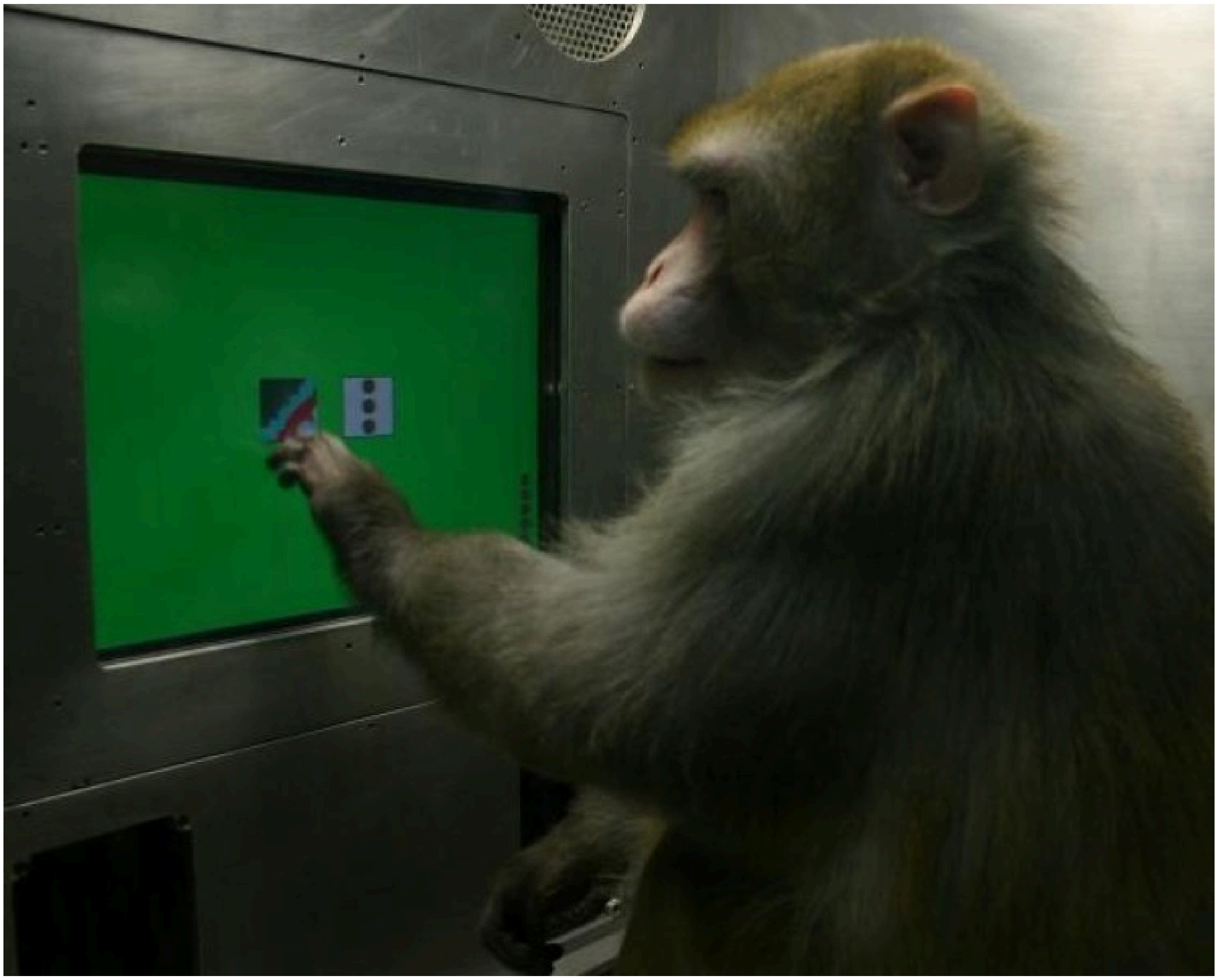
In a Venn diagram, one begins with two overlapping circles, one as the subject concept, the other as the predicate concept. But by contrast with what is found in an Euler diagram, the two overlapping circles of a Venn diagram do not thereby already express a judgment (the judgment that some S is P), as they would were that same drawing to be read as an Euler diagram. In a Venn diagram, there is in the two overlapping circles no reference to any objects or indeed any form of judgment at all, but only a space within which a judgment might be made. In order to make a judgment one must do something more to indicate that *objects* are thus and so in relation to the two concepts. Either one shades out some region showing thereby that *no* objects have some particular constellation of properties (the shading serving in effect as a universal quantifier), or one puts an 'X' in some region to indicate that *some* objects, at least one, have a particular constellation of properties (which is of course what the existential quantifier does). In neither case is one talking about any objects in particular; one is referring to objects in general by means of quantifiers. And this must be so because, again, concepts on Kant's view are *only* predicative and cannot give objects, but objects must yet somehow be given if judgment is to have any objectivity and truth. Kant's logic is, as our standard logic is, distinctively formal and quantificational.[7] Aristotle's logic is neither formal nor quantificational. And neither is Frege's.

We have seen that from Kant's perspective Aristotle's notion of a term conflates the logically distinct functions of referring and predicating. So, from Frege's perspective, Kant's distinction of intuitions, through which (alone) objects are given, and concepts, through which (alone) objects are thought, conflates two essentially different distinctions: that of object and (Fregean) concept with that of *Bedeutung* (designation) and *Sinn* (sense). Whereas for Kant all cognitive significance—all being for a thinker—is through concepts, and all objectivity lies in relation to an object or objects, Frege teaches us to distinguish, on the one hand, between cognitive significance—that is, Fregean sense, *Sinn*, and concepts, which are laws of correlation that are the *Bedeutung* of concept words—and on the other, between objective significance, *Bedeutung*, and objects. Kant, Frege teaches us, was after all mistaken in thinking that all objective significance lies in relation to an object, mistaken in holding that all cognitive significance is predicative.

(This is of course not a criticism of Kant any more than it is a criticism of Aristotle that he did not recognize Kant's distinction. In the unfolding of the science of logic, each stage is necessary as the ground on which further distinctions can be made.)

Because, on Frege's account, concept words designate concepts conceived as functions, laws of correlation arguments to truth-values, there can be relations directly among concepts, relations unmediated by any features of objects. Frege, unlike Kant, has no need of quantifiers. What he needs instead are various second-level properties and relations, among them the property of being universally applicable. Such a second-level property holds, for example, of the first-level concept being a mammal if a cat since everything that is or would be a cat is or would be also a mammal.[8] But concept words, as well as object names, also express senses in Frege's system of logic. Where Kantian logic is founded on a dichotomy of logical form and empirical content, in Frege's logic, all expressions are contentful, and contentful in the same way: all express Fregean sense, *Sinn*. Because Frege also requires that one can ask after the designation or signification, the *Bedeutung*, of an expression only given a context of use, it is possible in Frege's logical language, though not in a Kantian one, to exhibit the contents of concepts in a way enabling rigorous reasoning on the basis of that content. Frege's is a language within which to reason from the contents of explicitly defined mathematical concepts. Like Aristotle's logic, and unlike Kant's, Frege's logic is designed to provide a language within which to reason in the exact sciences.

In standard, Kantian logic inferences are good, if they are, in virtue of their form; content is wholly irrelevant to the goodness of an inference in this system. Indeed, it is not even inference, the act of inferring, that is the focus in Kantian logic but instead the relation of logical consequence. To prove something in this logic is, as Wittgenstein notes in his *Tractatus* (6.1262), nothing more than a "mechanical expedient"; it is merely a mechanical means by which to show, make explicit, that the information in the conclusion is contained already in the premises. And *this* is something that a machine can do as well as any human being. Whether I manipulate the signs mechanically according to rules or build a machine so to manipulate the signs, the activity and the achievement is, in the two cases, precisely the same. Thus, it comes to seem that computers can think in just the sense we do, and correlatively that we human thinkers are in essence nothing more than biological computers.



A rhesus monkey chooses between images on a touch-screen computer monitor. Photograph by Herbert Terrace, Columbia University.

As conceived in Frege's system of logic, to say that an inference is good in virtue of its form is to say that it is an instance of something inherently general, something applicable in other cases as well. Any actual inference is, on this account, an application of a rule that applies also in other cases. Because in this system logical form does not contrast with content but is itself contentful, so-called formal inferences (that is, strictly logical inferences) are in a way *material* insofar as they depend on the *meanings* of the signs of logic. Again, the signs of Frege's logic express senses just as non-logical signs do, and because they do, logical truths really are *truths* on this account. They are thoughts we can know to be true, and about which we can find ourselves to have been mistaken. They are not mere forms. And to infer is, or at least can be, actually to do something, to make a move from one judgment or judgments to another, different judgment. Inferring is, in such a case, not merely a matter of making what is implicit in the premises explicit in the conclusion but is instead to make what is potential in the premises actual in the conclusion. In a mathematical proof, as Frege conceives it, the premises contain everything that is needed in order to realize the desired conclusion but the conclusion is not already there in the premises, even implicitly. One must go through the steps of reasoning in order to bring about the conclusion; and here Frege's thought is, significantly enough, very like Kant's as regards constructions in mathematics. It is Frege's conception of language in terms of *Sinn* and *Bedeutung* that enables us to understand precisely how this is to work.

Imagine a person who knows how to count things but not any mathematical facts, even the most basic. Imagine further that this person wishes to know what, say, the sum of seven and five is. A good mechanical expedient for finding the answer would be to count out seven things, then count out five more things, then put all the things together and count the resulting collection, something we can do with, for example, marks on a page:

/////// /////

The collection on the left is a collection of seven strokes and that on the right is a collection of five strokes. By adding five more as I did after making the first seven strokes, I made, whether I knew it or not, a collection of twelve things, as I can learn by counting the whole. That there are twelve strokes is *already there*, contained implicitly in the display as soon as the five more strokes are added. What my subsequent counting does is only to make that fact explicit. A valid proof in standard, Kantian logic is essentially the same. One writes all the premises and thereby has, whether one knows it or not, the information in the conclusion. The steps of the proof, like the subsequent counting in our little example with the strokes, serve only to make that fact explicit.

Kant clearly does not think of the example of seven and five in the way just outlined.[9] According to Kant, the truth that seven plus five is twelve is not analytic but instead synthetic a priori. That is, according to Kant, being twelve is *not* contained already in the starting point, in the given numbers seven and five. Rather, Kant thinks, the number twelve can be *constructed* given the starting point, the given numbers seven and five. We can see how this is to work if we conceive the strokes on the page not as mere *things*—strokes on a page that perhaps also stand in for other things—but instead as primitive signs of an exceedingly simple mathematical language, as signs that express Fregean senses. We then can use these signs to construct complex signs designating numbers. On this reading, the collection of seven strokes is not merely that—a complex sign that pictures a collection by itself being a collection (as, on Wittgenstein's Tractarian view, a proposition pictures a fact by itself being a fact)—but is instead a complex sign designating *one* thing, *the* number seven, through a sense that reveals that number as a certain multiplicity. Similarly, the collection of five strokes read as Frege would have us read is a complex sign for the (one) number five, a sign that designates *the* number five, again through a Fregean sense. On this reading, the individual strokes in the complex sign do not designate; only the whole complex sign, the whole collection of strokes designates. But once we have the *two* collections, the complex sign for the number seven *and* the complex sign for the number five, we can *construct* a sign for the number twelve by combining all the primitive signs into one. This is a mathematical operation, not a merely mechanical one, insofar as, first, we are using the strokes to display the arithmetical contents of the numbers seven and five, *what it is to be* such numbers (namely, certain multiplicities). Subsequently, on the basis of that displayed content, we manipulate the primitive signs in a rigorous, rule-governed way in order to show something about the numbers involved, that the sum of (the numbers) seven and five is (the number) twelve, a properly mathematical result.

The same distinction of mechanical and mathematical reasoning can be made for the case of a computation in Arabic numeration because, again, it is possible to read a numeral in that notation in either of two essentially different ways. The first, merely mechanical way is to read each of the ten digits as a numeral for a number no matter what the context of use. On this (mechanical) reading, '3' designates the number three, '7' the number seven, and so on for all the digits, no matter the context. Complex signs in the positional system of Arabic numeration are then read additively, with the positions of the numerals serving to mark *what* is being counted: '375', for example, pictures a collection of three hundreds and seven tens and five units. Such collections can then be manipulated mechanically, using, for example, an adding machine. But we can also read the sign '375' differently, as exhibiting the computational content of *the* number three hundred and seventy-five, its content as it matters to arithmetical computations. On this second (non-mechanical, properly mathematical) reading, the primitive signs only express Fregean senses independent of a context of use; they designate only in and as complexes. Moreover, the same point applies to the case of reasoning from the contents of inferentially articulated mathematical concepts. Although one *can* read the signs mechanically, as recording, or picturing, information (as in standard logic), one can also read them as Frege teaches us to read, as expressing Fregean senses that contain modes of presentation of the designated concepts. As I show in *Realizing Reason*, even strictly deductive reasoning—conceived as Frege conceives it—can be ampliative, a real extension of our knowledge; and it can be because there is in such cases a kind of *construction* of the conclusion on the basis of the premises. By deductively reasoning on the basis of the contents of the concepts with which one begins, as given in their definitions, one proves theorems about those very concepts, about their relations one to another. The reasoning is not mechanical; it is mathematical reasoning of the sort mathematicians actually engage in.[10]

D. W. Hagelbarger and Claude Shannon's "mind-reading machine", 1953.

I have argued, first, that from the perspective afforded by modern, Kantian logic there is and can be no essential difference between the thinking of the mathematician and that of a mere machine because, on that account—with its strict dichotomy of logical form and content—thinking is and can be nothing more than the rule-governed manipulation of signs the meanings of which are completely irrelevant. But I have also shown that from the perspective Frege provides, we can see how it is that the mathematician's thinking on the basis of the contents of mathematical concepts is essentially *different* from that of a mere machine. That the machine can *mimic* our thinking is unsurprising given that any notation within which to reason (such as our little stroke language, the system of Arabic numeration, and Frege's logical language) can also be treated mechanically. The signs themselves do not enforce a mathematical reading. Nevertheless, they *can* be read mathematically, at least by us human beings. That is, *they can be read* where to read is to extract the meaning expressed in a specially designed system of written marks. And because they can be read, they can also be misread. We (human) thinkers can make mistakes, and very often our mistakes are grounded in our failure adequately to grasp the contents of the concepts of concern to us. Consider again the example of the concept of number: we at first thought that negative and fractional numbers are not numbers because we understood numbers to be collections of units; we had the number concept, but did not yet fully understand the content of that concept. It is just as Frege says: "[O]ften it is only after immense intellectual effort, which may have continued over centuries, that humanity at last succeeds in achieving knowledge of a concept in its pure form."[11] The concept of *thinking* is another such concept. It is not simply given what *thinking* is. In the end, perhaps it is this very thing, the fact that we can think that we *know* in cases in which it is later revealed that we were mistaken, that makes most manifest the essential difference between a human thinker and a mere machine. *We* have second thoughts, and it is because we do that we can be said to have any thoughts at all.

Footnotes

1. W.V.O. Quine. *Methods of Logic*. London: Routledge and Kegan Paul, 1952. vii. Print.

2. These claims are developed and defended at length in my *Realizing Reason: A Narrative of Truth and Knowing*. Oxford: Oxford University Press, 2014. Print.

3. See Aristotle's *Posterior Analytics*. Book I, Chapter 2.

4. "We must never ask about the existence of anything until we first understand its essence." Replies to the First Set of Objections, in *The Philosophical Writings of Descartes*. Vol II. Trans. J. Cottingham, R. Stoothoff, and D. Murdoch. Cambridge: Cambridge University Press, 1984. 78. Print.

5. Immanuel Kant. *Critique of Pure Reason*. A51/B75. Trans. P. Guyer and A. W. Wood. Cambridge: Cambridge University Press, 1998. 193-4. Print.

6. This conception of logic as formal, lacking all content and truth, is made fully explicit in our model-theoretic conception of language. See, for example, Wilfrid Hodges. "Truth in a Structure." *Proceedings of the Aristotelian Society* 86 (1985): 135-51. Print.

7. Of course, Kant's logic is not precisely the same as our logic. In particular, it is not symbolic as our logic is, and it is monadic. Standard logic would be made symbolic in the second half of the 19th century, beginning with Boole, and extended to a full logic of relations at the turn of the 20th

century. This latter advance was made by Bertrand Russell and was, as he saw, a merely technical development. What Russell did not realize was that the essentials of modern logic were already in Kant; he thought they were due to Peano and, independently, Frege. See Bertrand Russell. *Our Knowledge of the External World*. London: George Allen and Unwin, 1914. 50. Print. Also, Jourdain's record of a conversation he had with Russell on 20 April 1909, in I. Grattan Guinness (Ed. and Trans.). *Dear Russell – Dear Jourdain*. New York: Columbia University Press, 1977. 114. Print.

8. See Chapter 2 of my *Frege's Logic*. Cambridge, Mass.: Harvard University Press, 2005. Print.

9. Kant discusses this case in the B Introduction of the first *Critique*. B15-16.

10. Again, the point is not that physical systems cannot think, but to highlight fundamental differences between the merely mechanical manipulation of signs according to rules and mathematical reasoning in a specially devised system of written marks, where the latter essentially involves what Frege has taught us to recognize as the *Sinn* of an expression as contrasted with its *Bedeutung*.

11. Gottlob Frege. *Foundations of Arithmetic*. Trans. J. L. Austin. Evanston, Ill.: Northwestern University Press, 1980. vii. Print.

Danielle Macbeth is T. Wistar Brown Professor of Philosophy at Haverford College in Pennsylvania, USA.