Thoroughly Modern Meno

by

Clark Glymour, Kevin T. Kelly

August 1988

Report CMU-PHIL-10



Philosophy
Methodology
Logic

Pittsburgh, Pennsylvania 15213-3890

Thoroughly Modern Meno

Clark Glymour

Carnegie Mellon University
and
University of Pittsburgh

and

Kevin Kelly
Carnegie Mellon University

1. Introduction

The *Meno* presents, and then rejects, an argument against the possibility of knowledge. The argument is given by Meno in response to Socrates' proposal to search for what it is that is virtue:

> Meno: How will you look for it, Socrates, when you do not know at all what it is? How will you aim to search for something you do not know at all? If you should meet with it, how will you know that this is the thing that you did not know?[1]

Many commentators, including Aristotle in the *Posterior Analytics*, take Meno's point to concern the *recognition of an object*, and if that is the point there is a direct response: one can recognize an object without knowing all about it. But the passage can also be understood straightforwardly as a request for a discernible *mark* of truth, and as a cryptic argument that, without such a mark, it is impossible to acquire knowledge from the instances that experience provides. We will try to show that the second reading is of particular interest.

If there is no mark of truth, nothing that can be generally discerned that true and only true propositions bear, Meno's remarks represent a cryptic argument that knowledge is impossible. We will give an interpretation that makes the argument valid; under that interpretation, Meno's argument demonstrates the impossibility of a certain kind of knowledge. In what follows we will consider Meno's argument in more detail, and we will try to show that similar arguments are available for many other conceptions of knowledge. The modern Meno arguments reveal a diverse and intricate structure in the theories of

---

[1]G. Grube, trans., *Five Dialogues*, Hackett, 1981. All quotations are taken from this translation.

knowledge and of inquiry, a structure whose exploration has just begun. While we will attempt to show that our reading of the argument fits reasonably well with Plato's text, we do not aim to argue about Plato's intent. It is enough that the traditional text can be elaborated into a systematic and challenging subject of contemporary interest.[2]

2. The Meno

In one passage in the Meno, to acquire knowledge is to acquire a truth that can be given a special *logical form*. To acquire knowledge of virtue is to come to know an appropriate truth that states a condition, or conjunction of conditions, necessary and sufficient for any instance of virtue. Plato's Socrates will not accept lists, or disjunctive characterizations.

> Socrates: I seem to be in great luck, Meno; while I am looking for one virtue, I have found you to have a whole swarm of them. But, Meno, to follow up the image of swarms, if I were asking you what is the nature of bees, and you said that they are many and of all kinds, what would you answer if I asked you: "Do you mean that they are many and varied and different from one another in so far as they are bees? Or are they no different in that regard, but in some other respect, int heir beauty, for example, or their size or in some other such way?" Tell me, what would you answer if thus questioned?
>
> Meno: I would say that they do not differ from one another in being bees.
>
> Socrates: Suppose I went on to say: "Tell me, what is this very thing, Meno, in which they are all the same and do not differ from one another?" Would you be able to tell me?
>
> Meno: I would.
>
> Socrates. The same is true in the case of the virtues. Even if they are many and various, all of them have one and the same form which makes them virtues, and it is right to look to this when one is asked to make clear what virtue is. Or do you not understand what I mean?

There is something peculiarly modern about the Meno. The same rejection of disjunctive characterizations can be found in several comtemporary accounts of explanation[3] We might say that Socrates requires that Meno produce an appropriate and true universal biconditional sentence, in which a

---

[2]A view of inquiry related to that developed here has been championed by A. Goldman in his *Epistemology and Cognition*, Harvard, 1984. Recent commentaries on the Meno by historians of philosophy have varied in the importance they give to Meno's paradox, but none of them give it the reconstruction considered here. Compare T. Irwin, *Plato's Moral Theory*, Oxford, 1977; J. Maravcsik, "Learning as Recollection" in G. Vlastos, ed., *Plato I, Metaphysics and Epistemology*, 1971; N. White, *Plato on Knowledge and Reality*, 1978. In what follows, a variety of technical claims are asserted without proof. In all cases the claims are simple applications of results shown in D. Osherson, M. Stob, and S. Weinstein, *Systems That Learn*, M.I.T. Press, 1985; C. Glymour "Inductive Inference in the Limit" *Erkenntnis*, 21, 1984; D. Osherson and S. Weinstein, "Identification in the Limit of First-Order Structures" *Journal of Philosophical Logic*, 15, 1986; D. Osherson and S.Weinstein, *Identifiable Collections of Countable Structures*, *Philosophy of Science*, forthcoming; and, especially, K. Kelly and C. Glymour, "Converging to the Truth and Nothing But the Truth," *Philosophy of Science*, forthcoming; and D. Osherson and S. Weinstein, "Paradigms of Truth Detection," manuscript, 1988.

[3]For example, in H. Field, "Mental Representation" in N. Block, ed. *Philosophy of Psychology*; and in R. Stalnaker, *Inquiry*.

predicate signifying 'is virtuous' flanks one side of the biconditional, and a conjunction of appropriate predicates occurs on the other side of the biconditional. Let us so say. Nothing is lost by the anachronism and, as we shall see, much is gained.

Statements of evidence also have a logical form in *The Meno*. Whether the topic is bees, or virtue, or geometry, the evidence Socrates considers consists of instances and non-instances of virtue, of geometric properties, or whatever the topic may be. Evidence is stated in the singular.

The task of acquiring knowledge thus assumes the following form. One is presented with, or finds, in whatever way, a series of examples and non-examples of the feature about which one is inquiring, and from these examples a true, universal biconditional without disjunctions is to be produced. In *The Meno* that is not enough for knowledge to have been acquired. To acquire knowledge it is insufficient to produce a truth of the required form; one must also *know* that one has produced a truth. What can this requirement mean?

Socrates and Meno agree in distinguishing knowledge from mere true opinion, and they agree that knowledge requires at least true opinion. Meno thinks the difference between knowledge and true opinion lies in the greater *reliability* of knowledge, but Socrates insists that true opinion could, by accident as it were, be as reliable as knowledge:

> Meno:..But the man who has knowledge will always succeed, whereas he who has true opinion will only succeed at times.
>
> Socrates: How do you mean? Will he who has the right opinion not always succeed, as long as his opinion is right?
>
> Meno: That appears to be so of necessity, and it makes me wonder, Socrates, this being the case, why knowledge is prized far more highly than right opinion, and why they are different.

Socrates answers each question, after a fashion. The difference between knowledge and true opinion is in the special *tie*, the binding connection, between what the proposition is about and the fact of its belief. And opinions that are tied in this special way are not only reliable, they are liable to stay, and it is that which makes them especially prized:

> Socrates: To acquire an untied work of Daedalus is not worth much, like acquiring a runaway slave, for it does not remain, but it is worth much if tied down, for his works are very beautiful. What am I thinking of when I say this? True opinions. For true opinions, as long as they remain, are a fine thing and all they do is good, but they are not willing to remain long, and they escape from a man's mind, so that they are not worth much until one ties them down by an account of the reason why. And that, Meno my friend, is recollection, as we previously agree. After they are tied down, in the first place they become knowledge, and then they remain in place. That is why knowledge is prized higher than correct opinion, and knowledge differs from

correct opinion in being tied down.

Plato is chiefly concerned with the difference between knowledge and true opinion, and our contemporaries have followed this interest. The recent focus of epistemology has been the special intentional and causal structure required for knowing. But Meno's argument does not depend on the details of this analysis; it depends, instead, on the capacity for true opinion that the capacity to acquire knowledge implies. That is the capacity to find the truth of a question, to recognize it when found, to stick with it after it is found, and to do so *whatever* the truth may be.

Suppose that Socrates could meet Meno's rhetorical challenge and recognize the truth when he met it: what is it he would then be able to do? Something like the following. In each of many different *imaginable* (we do not say *possible* save in a logical sense) circumstances, in which distinct claims about virtue (or whatever) are true, upon receiving enough evidence, and considering enough hypotheses, Socrates would hit upon the right hypothesis about virtue for that possible circumstance, and would then (and only then) announce that the correct hypothesis is indeed correct. Never mind just *how* Socrates would be able to do this, but agree that, if he is in the actual circumstance capable of coming to know, then that capacity implies the capacity just stated. Knowledge requires the ability to come to believe the truth, to recognize when one believes the truth (and so to be able to continue to believe the truth), and to do so whatever the true state of affairs may be.

So understood, Meno's argument is valid, or at least its premises can be plausibly extended to form a valid argument for the impossibility of knowledge. The language of possible worlds is convenient for stating the argument. Fix some list of predicates V, P1,..., Pn, and consider all possible worlds (with countable domains) that assign extensions to the predicates. In some of these worlds there will be true universal biconditional sentences with V on one side and conjunctions of some of the Pi or their negations on the other side. Take pieces of evidence available from any one of these structures to be increasing conjunctions atomic or negated atomic formulas simultaneously satisfiable in the structure. Let Socrates receive an unbounded sequence of singular sentences in this vocabulary, so that the sequence, if continued, will eventually include every atomic or negated atomic formula (in the vocabulary) that is satisfiable in the structure. Let ω range over worlds. With Meno, as we have read him, say that Socrates can come to know a sentence, S, of the appropriate form, true in world ω, only if

(i) for every possible sequence of presentation of evidence from world ω Socrates eventually announces

that S is true, and

(ii) in every world, and for every sequence from that world, if there is a sentence of the appropriate form true in that world, then Socrates can eventually consider some true sentence of the appropriate form in that world, can announce that it is true in that world (while never making such an announcement of a sentence that is not true is that world), and

(iii) in every world, and for every sequence from that world, if no sentence of the appropriate form is true in the world then Socrates refrains from announcing of any sentence of that form that it is true.

Meno's argument is now a piece of mathematics, and it is straightforward to prove that he is correct: no matter what powers we imagine Socrates to have, he cannot acquire knowledge, provided "knowledge" is understood to entail these requirements. No hypotheses about the causal conditions for knowledge defeat the argument unless they defeat the premises. Skepticism need not rest on empirical reflections about the weaknesses of the human mind. The impossibility of knowledge can be demonstrated *a priori.* Whatever sequence of evidence Socrates may receive that agrees with a hypothesis of the required form, there is some structure in which that evidence is true but the hypothesis is false; so that if at any point Socrates announces his conclusion, there is some imaginable circumstance in which he will be wrong.

We should note, however, that in those circumstances in which there is no truth of of the required form, Socrates can eventually come to know that there is no such truth, provided he has an initial, finite list of all of the predicates that may occur in a definition. He can announce with perfect reliability the absence of any purely universal conjunctive characterizations of virtue if he has received a counterexample to every hypothesis--and if the number of predicates are finite the number of hypotheses will be finite, and if no hypothesis of the required form is true, the counterexamples will eventually occur. If the relevant list of predicates or properties were not provided to Socrates initially, then he could not know that there is no knowledge of a subject to be had.

3. Weakening Knowledge

Skepticism has an ellipsis. The content of the doubt that knowledge is possible depends on the requisites for knowledge, and that is a matter over which philosophers dispute. Rather than supposing there is one true account of knowledge to be given, if only philosophers could find it, our disposition is to inquire about the possibilities. Our notion of knowing is surely vague in ways, and there is room for more than one

interesting doxastic state.

About the conception of knowledge we have extracted from Meno there is no doubt as to the rightness of scepticism. No one can have *that* sort of knowledge. Perhaps there are other sorts that can be had. We could weaken the set of possibilities that must be considered, eliminating most of the possible worlds, and make requirements (i), (ii) and (iii) apply only to the reduced set of possibilities. We would then have a revised conception of knowledge that requires only a reduced *scope*, as we shall call the range of structures over which Socrates, or you or we, must succeed in order to be counted as a knower. This is a recourse to which we will have eventually to come, but let us put it aside for now, and consider instead what might otherwise be done about weakening conditions (i), (ii) and (iii).

Plato's Socrates emphasizes this difference between knowledge and mere true opinion: knowledge stays with the knower, but mere opinion, even true opinion, may flee and be replaced by falsehood or want of opinion. The evident thing to consider is that for Socrates to come to know the truth in a certain world, Socrates be required to be able to find the truth in each possible world, and never to abandon it, but not be required to announce that the truth has been found when it is found. Whatever the relations of cause and intention that knowledge requires, surely Meno requires too much. He requires, as we have reconstructed his argument, that we come to believe through a reliable procedure, a procedure or capacity that would, were the world different, lead to appropriately different conclusions in that circumstance. But Meno also requires that we know when the procedure has succeeded, and that seems much like demanding that we know that we know when we know. Knowing that we know is an attractive proposition, but it does not seem a prerequisite for knowledge, or if it is, then by the previous argument, knowledge is impossible. In either case, the properties of a weaker conception of knowledge deserve our study.

The idea is that Socrates comes eventually to embrace the truth and to stick with it in every case, although he does not know at what point he has succeeded--he is never sure that he will not, in the future, have to change his hypothesis. In this conception of knowledge, there is no mark of success. We must then think of Socrates as conjecturing the truth forever. Since Socrates did not live forever, nor shall we, it is better to think of Socrates as having a *procedure* that could be applied indefinitely, even without the living Socrates. The procedure has mathematical properties that Socrates does not.

For Socrates to know that S in world ω in which S is true now implies that Socrates behavior accords with

a procedure with the following properties:

(i*) for every possible sequence of evidence from world ω, after a finite segment is presented, the procedure conjectures S ever after, and

(ii*) for every possible sequence of evidence from any possible world, if a sentence of the appropriate form is true in that world, then after a finite segment of the evidence is presented the procedure conjectures a true sentence of the appropriate form ever after.

These conditions certainly are not sufficient for any doxastic state very close to our ordinary notion of knowledge, since Socrates' behavior may in the actual world accord with a procedure satisfying i* and ii* even while Socrates lacks the *disposition* to act in accord with the procedure in other circumstances. For knowledge, Socrates must have such a disposition. But he can only have such a disposition if there exists a procedure meeting conditions i* and ii*. Is there? If the logical form of what is to be known is restricted to universal biconditionals of the sort Plato required, then there is indeed such a procedure. If Socrates is unable to acquire this sort of knowledge, then it is because psychology or sociology or biology, not in virtue of mathematical impossibilities. Scepticism about this sort of knowledge cannot be *a priori*. There is no general argument of Meno's kind against the possibility of acquiring this sort of knowledge.

The weakening of knowledge may be unPlatonic, but it is not unphilosophical. Francis Bacon's *Novum Organum* describes a procedure that works for this case, and his conception of knowledge seems roughly to accord with it. John Stuart Mill's canons of method are, of course, simply pirated from Bacon's method. Hans Reichenbach used nearly the same conception of knowledge in his "pragmatic vindication" of induction, although he assumed a very different logical form for hypotheses, namely that they are conjectures about limits of relative frequencies of properties in infinite sequences.

So we have a conception of knowledge that, at least for some kinds of hypotheses, is not subject to Meno's paradox. But for which kinds of hypotheses is this so? We are not now captivated, if ever we were, by the notion that all knowledge is definitional in form. Perhaps even Plato himself was not, for the slave boy learns the theorem of Pythagoras, which has a more complicated logical form. We are interested in other forms of hypotheses: positive tests for diseases, and tests for their absence; collections of tests one of which will reveal a condition if it is present. Nor are our interests confined to single hypotheses considered individually. If the property of being a squamous cancer cell has some

connections with other properties amenable to observation we want to know *all about* those connections. We want to discover the whole theory about the subject matter, or as much as we can of it. What we may wish to determine, then, is what classes of *theories* can come to be known according to our weaker conception of knowledge. Here, as I use the notion of theory, it means the set of all true claims in some fragment of language. Wanting to know the truth about a particular question is then a special case, since the question can be formulated as a claim and its denial, and the pair form a fragment of language whose true claims are to be decided. What we wish to determine is whether all of what is true and can be stated in some fragment of language can be known.

Either the possibility of knowledge depends on the fragment of language considered or it does not. If it does, then many distinct fragments of language might be of the sort that permit knowledge of what can be said in them, and the classification of fragments that do, and that do not, permit such knowledge becomes an interesting task. For which fragments of language, if any, are there valid arguments of Meno's sort against the possibility of knowledge, and for which fragments are there not? These are straightforward mathematical questions, and their answers, or some of their answers, are as follows:

Consider any first order language (without identity) in which all predicates are monadic, and there are no symbols taken to represent functions. Then any true theory in such a language can be learned, or at least there are no valid Menoan arguments against such knowledge.

If the language is monadic but with identity, or if the language contains a predicate that is not monadic, then neither the fragment that consists only of universally quantified formulas, nor the fragment that consists only of existentially quantified formulas, nor any part of the language containing either of these fragments, are such that every true theory in these fragments can be known.

In each of the latter cases an argument of Meno's kind can be constructed to show that knowledge is impossible.

4. Times for All Things

The weakened conception of knowledge is still very strong in at least one respect. It requires for the possibility of knowledge of an infinite wealth of claims that there be a time at which all of them are known; that is, a single time after which all and only the truths in a fragment of language are conjectured. We might instead usefully consider the following circumstance: When investigating hypotheses in a fragment

of language, Socrates is able, for each truth, eventually to conjecture it and never subsequently to give it up; and Socrates is also able, for each falsehood, eventually *not* to conjecture it and never after to put it forward. Plato's Socrates illustrates that the slave boy can "recollect" the Pythagorean theorem from examples and appropriate question, and presumably in Plato's view the slave boy could be made to recollect any other truth of geometry by a similar process. But neither the illustration nor the view requires that the slave boy, or anyone else, eventually be able to recollect the whole of geometry. There may be no time at which Socrates knows all of what is true and can be stated in a given fragment of language. Yet the disposition to follow a procedure that will eventually find every truth and eventually avoid every falsehood is surely of fundamental interest to the theory of knowledge. Call a procedure that has the capacity to converge to the whole truth at some moment, as in the discussion of the previous section, an EA learning procedure, and call an AE learner a procedure that for each truth has the capacity to converge to that truth by some moment, and for each falsehood avoids it ever after some moment. Every EA learner is an AE learner, but is the converse true? Or more to the point, are there fragments of language for which there are AE procedures but no EA procedures?

There are indeed. Consider the set of all universal sentences, with identity, and with any number of predicates of any arity and any number of function symbols of any arity. By the negative result stated previously, there is no EA procedure for that fragment of language, no procedure that, for every (countable) structure, and every way of presenting the singular facts in the structure, will eventually conjecture the theory (in the language fragment) true in that structure. But there is an AE procedure for this fragment. If, for knowledge about a matter, Socrates is required only to have a disposition to follow an AE procedure for the language of the topic, then no Menoan argument shows that Socrates cannot acquire knowledge, even if Socrates does not know the relevant predicates or properties beforehand.

The improvement does not last. If we consider the fragment of language that allows up to one alternation of quantifiers, whether from universal to existential or existential to universal, it again becomes impossible to acquire knowledge; there are no AE procedures for this fragment that are immune from arguments of Meno's kind.

5. Discovery and Scope

Whether we consider EA discovery or AE discovery, we soon find that arguments of Meno's kind succeed. The same sort of results obtain if we further weaken the requirements for knowledge. We might, for example, abandon Plato's suggestion that when a truth is known it is not subsequently forgotten or rejected. We might then consider the requirement that Socrates be disposed to behave in accordance with a procedure that as it considers more and more evidence about a question, is wrong in its conjectures only finitely often, correct infinitely often, but may also suspend judgement infinitely often. Osherson and Weinstein have shown that with even this remarkably weak conception there are questions that cannot, in senses parallel to those above, be known. Or we might allow various sorts of approximate truth; for many of them, arguments parallel to Meno's are available.

The conceptions of knowledge we have discussed place great emphasis on *reliability*. They demand that we not come to our true beliefs by chance, but in accordance with procedures that would find the truth no matter what it might be, so long as the procedures could be carried out. What the Meno arguments show is that in the various senses considered, for most of the issues that might invite discovery, procedures so reliable do not exist. The anti-skeptical response ought to be principled retreat. In the face of valid arguments against the possibility of procedures so reliable, and hence against the possibility of corresponding sorts of knowledge, let us consider procedures that are not so reliable, and regard the doxastic state that is obtained by acting in accord with them as at least something better and more interesting than accidental true belief.

For each of the requirements on knowledge considered previously, and for others, we can ask the following kind of question: For each fragment of language, what are the classes of possible worlds for each of which there exists a procedure that will discover the truths of that fragment for any world in the class? The question may be too hard to parse. Let us define it in pieces. Let a discovery problem be any (recursive) fragment F of a formal language, together with a class K of countable relational structures for that fragment. One such class K is the class of all countable structures for the language fragment, but any subsets of this class may also be considered. A discovery procedure for the discovery problem is any procedure that for every k in K and every presentation of evidence from k, "converges" to all of the sentences in F that are true in k. "Convergence" may be in the EA sense, the AE sense, or some other sense altogether (such as the weak convergence criterion considered two paragraphs previously).

What the results I have described tell us is that for many fragments F, if K is the set of all countable

structures for F, then there are no discovery procedures for pairs <F,K>. That does not imply that there are no discovery procedures for pairs <F, K'> where K' is some proper subset of K. Must it be that for knowledge, true belief have been acquired in accordance with a procedure that would lead to the truth in *every* imaginable sequence?

Suppose we think of inquiry as posing discovery problems, a question or questions and a class of possible worlds or circumstances that determine various answers to the question. Depending on which world or circumstance is ours, different answers will be true. Successful inquiry, which leads to some kind of knowledge, accords with a procedure that will converge to the truth of the matter, whatever it may be, in each of these possible circumstances. It is possible for procedures to have the capacity to find the truth in each of a class of circumstances without having the capacity to find the truth in every imaginable circumstance.

When attention is restricted to a discovery problem that contains a restricted class of possible worlds or circumstances, that restriction constitutes a kind of *background knowledge* brought to inquiry. The background knowledge says that the actual circumstance is one of a restricted class of circumstances or possible worlds. The theory of recollection, Plato's solution to Meno's paradox, claims that inquiry is conducted with a special sort of background knowledge, stamped in the soul before birth. Two different reconstructions of Plato's solution fit the story, and we offer them both without choosing between them.

In the first account, the correct *definitions* are stored in the soul and need only be brought to mind. The presentation of examples and the process of recollection eventually brings forth the truth, and provides knowledge, not because the process using that same background knowledge would succeed no matter how the world (or rather the forms) might imaginably be, but because there is a guarantee that the world (or, rather again, the forms) accords with knowledge the soul possesses. The background knowledge is so complete that no *inference* from examples is required; examples only ease access to knowledge we already have.

In the second account a complete list of *definiens*, each characterizing a distinct form, is stored in the soul. An inquiry into the nature of virtue must then match instances of the usage of "virtue" with the appropriate definiens in the list. In this case the process of recollections involves an inductive inference from particular examples to a universal biconditional connecting a definiens in the list with a term denoting the subject of inquiry. On the assumptions that no two forms are such that the same individuals

participate in both, and that there are only finitely many forms, Socrates can eventually conjecture the form of virtue, know that his conjecture is correct, and can do so no matter which definiens in the list happens to represent the form of virtue.

On either reconstruction, Plato's reply to Meno's paradox has two aspects, and the slave boy's rediscovery of the theorem of Pythagoras illustrates each of them. First, knowledge may be had by means other than the means of inquiry. It may be inherited, innate, stamped on the soul, and not acquired by generalization from examples given in this life. Second, given such prior knowledge, the task of discovery or the acquisition of knowledge is reconceived and becomes feasible, for the inquirer need not be able to fix upon the truth in every imaginable circumstance, but only in those circumstances consistent with prior knowledge.[4]

Plato has little to say in *The Meno* about what souls do that gives them the knowledge we recollect in successful inquiry. We (or our souls) have background knowledge through a causal process that is not itself inquiry. We could instead entertain the thought that we acquire background knowledge through inquiry conducted in our past lives. The second alternative raises a number of interesting questions.

When we inquire into a question, the discovery problem we address depends upon our knowledge. The class of alternative circumstances, and thus alternative answers, that need be considered is bounded by our prior knowledge. If we know nothing, it is the class of all imaginable circumstances; if we know a great deal, the class of alternative circumstances may be quite small. Suppose as we go through life (or through a sequence of lives) we form conjectures about the answers to various questions, and while we reserve the right to change these conjectures upon further evidence, in the meanwhile we use them as though they were background knowlege for still other questions. Should evidence later arrive that causes us to abandon our conjectures, we will also have to reconceive the discovery problems in which we had taken those conjectures as backgound knowledge.[5].

Since we are not only uncertain what discovery problems we shall face, but more profoundly, we may be wrong in our construal of the discovery problems we presently face, it would seem only prudent to rely on

---

[4]Chomsky gave a strictly parallel answer to questions about how children learn their native language. "Cartesian linquistics" ought perhaps to have been called "Platonic linguistics."

[5]The result is a picture of inquiry something like that proposed by Issac Levi. Compare Levi's *The Enterprise of Knowledge*, M.I.T. Press.

learning procedures that have the widest possible scope. We know from what has gone before that Meno's argument, and derivatives of it, show that there is no procedure adequate for all discovery problems, but some procedures may do better than others. We can characterize a dominance relation between discovery procedures: Procedure A dominates procedure B provided A solves (in whatever sense may be specified) every discovery problem B solves, but not vice versa. A procedure is then *maximal* if no procedure dominates it. We might then take prudence to require that our manner of inquiry accord with a maximal procedure. Some second thoughts are called for. In the well studied case in which what is to be learned is not a theory but a language, it is known that every maximal procedure solves the discovery problem that consists of learning any finite language on a fixed vocabulary, but no procedure solves any larger problem, posed by any larger class of languages on that same vocabulary. There is no maximal procedure that identifies even one infinite language. For problems that concern the learning of theories, one should expect something analogous: the maximal procedures will be very sparse, and will fail to solve discovery problems that are readily solved by other methods.

Since, in all likelihood, we cannot fix beforehand on maximal methods, prudence can only recommend something more modest. When we recognize that one discovery procedure dominates another then, *ceteris paribus*, it is prudent to use the dominant procedure rather than the dominated procedure. Whether that is a sensible or feasible recommendation depends on the dominance structure of discovery procedures. If, for example, there is a readily described infinite chain of procedures, later members of the sequence dominating all earlier members, then the recommendation would give us a task worthy of Sisyphus. We would ever be changing one procedure for another, without rest and without end. Sometimes, much as the Existentialists say, the best thing to do is to stop preparing to make inquiries and make them.

6. Hypermodern Meno

Methodology amounts to recommendations restricting procedures of inquiry. Any such restriction can be thought of as determining a class of procedures, those that satisfy it. Besides methodology, psychology is another source of restrictions on procedures, and computation theory still another. For example, we might nowadays suppose that the discovery procedures available to us, even with the aid of machines, must be *computable* procedures, and invoking Church's thesis, restrict our attention to the class of Turing computable procedures for inquiry.

For any restriction on discovery procedures, the preceding discussion should suggest the following sort of question: What arguments of Meno's sort can be made against all procedures of this class? More exactly, for any restriction on discovery procedures, does the restriction also limit the class of discovery problems that can be solved? For both the EA and AE conceptions of successful inquiry, the requirement that procedures be computable limits the class of discovery problems that can be solved. There are discovery problems that can be solved by EA procedures but not by any computable EA procedures, and there are discovery problems that can be solved by AE procedures but not by any computable AE procedures. Methodological principles that are often regarded as benign also limit discovery when they are imposed in combination with the requirement of computability. A consistency principle applies to procedures that always conjecture theories consistent with the evidence; a conservative principles applies to procedures that never change a current conjecture unitl new evidence contradicts it. Either of these requirements, in combination with the requirement of computability, restricts the class of discovery problems that can be solved. It is easy to see that reverse is not true. That is, for every conservative, consistent, computable procedure, there is an inconsistent or unconservative (or both) procedure whose scope includes all discovery problems that can be solved by the first procedure.

When we investigate the restrictions on discovery that are implicit in methodological restrictions, we are entertaining recommendations to hop from one procedure to another. The picture of inquiry sketched in the previous section suggests the same thing for different reasons: as we reconceive the discovery problems with which we are faced, we may change our minds about which methods are appropriate. In that spirit, some philosophers have recommended methodological principles on empirical grounds: procedures that accord with the principles have worked in the past.[6]

The effect of hopping from one procedure to another can only be itself some procedure for discovery that mimics other procedures when given various pieces of evidence. From the inside, a hopping procedure may feel different than a procedure that does not hop, but behaviorally, the disposition to hop from procedure to procedure as evidence accumulates simply is a procedure, located somewhere in the vast ordering of possible discovery procedures. Recommendations about when and how to change procedures as evidence accumulates thus amount to restrictions on acceptable procedures, and form part (thus far an uninvestigated part) of methodology as we have just construed that subject. Despite these

---

[6]Compare C. Glymour, *Theory and Evidence*, Princeton University Press, 1980, and R. Boyd, "Realism, Underdetermination and a Causal Theory of Evidence," *Nous* 7 (1973).

caveats, if we are familiar with only a small set of methods, as seems to be the case, hopping among them can constitute a better procedure.

Recommendations about preferences among procedures may also come from the study of the scope of procedures, but that study cannot be algorithmic. There is no computable function that will tell us, for all ordered pairs of indices of discovery procedures, whether the first member of the pair dominates the second. We are instead landed somewhere within the analytical hierarchy of recursion theory, and just where it is that we have landed is an open question.

The general notion of hopping among procedures suggests an apparent paradox: Can an effective procedure that hops among procedures hop from itself to some other procedure? Can it hop back to itself? In a sense it can. If we think of a hopping procedure as a program that simulates other programs, then (by the recursion theorem) it can at various stages pursue a simulation of itself, or cease to simulate itself, and thus accept or reject itself as a method. Of course, no procedure can *behave* differently than it does.

7. Real Learning

Some people may think that results and questions such as those we have derived from the Meno paradox are remote from real concerns about the acquisition of knowledge. One might complain that these are all *formal* results, and because of that, for some reason mysterious to us, of no bearing on real science and its philosophical study. The study of the connection between logical form and the possibility of successful inquiry, in various senses, strikes us as both theoretically interesting and profoundly practical. For every question that has a logical form, or at least a tolerable variety of possible logical forms among which we may be undecided, these studies address the prospects for coming to know the answer.

Problems of a similar kind abound in the sciences, and questions (whose answers are in many cases unknown) about the existence of Menoan arguments against the acquisition of knowledge affect very practical issues about procedures of inquiry. We will give a few illustrations.

7.1. Language Learning

Consider a child learning its first language. Somehow, within a few years, the child comes to be able to produce and to recognize grammatical sentences in the native language, and to distinguish such sentences from ungrammatical strings. Grammatical sentences of any possible language can be regarded as concatenations of symbols from some finite vocabulary. If we fix the finite vocabulary then the number of distinct sets of strings built from that vocabulary is of course infinite, and in fact uncountably infinite. Suppose, however, we make the reasonable assumption that if a collection of strings is the collection of grammatical strings of some possible human language, then the collection is recursively enumerable. That is, for any set of strings of this kind there is a computable function such that, if a string is in the collection, the computable function will determine that it is. So, restricting attention to the languages that can be built on some particular vocabulary, the collection of possible natural languages is restricted to the recursively enumerable sets of strings made from that vocabulary. For each recursively enumerable set there is a program, actually an infinity of different programs, that when given an arbitrary string will compute "yes" if and only if the string is in the set (and will not return anything otherwise). The recursively enumerable sets can be effectively indexed in many different ways, so we can imagine each possible language to have a name that no other possible language has, and in fact we can imagine the name just to *be* a program of the kind just mentioned.

One way to think of the child's problem is this: on the basis of whatever evidence the environment provides, the child forms a sequence of programs that recognize a sequence of languages, until, eventually, the child settles on a program that recognizes the actual natural language in the child's environment. Psychological investigation suggests that children use *positive* evidence almost exclusively. That is, the evidence consists of strings from the langauge to be learned, but does not include evidence as to which strings are not in the language.

With this setting, due essentially to E. Mark Gold,[7] an important aspect of human development is made formal enough to permit mathematical investigations to bear on issues such as the characterization of the collection of possible human languages. For a language to be possible for humans, humans must be capable of learning it. Assuming that any possible human language could have been learned by any one human, it follows that the collection of possible human languages must be identifiable, or learnable, in the

[7]Aspects of the framework were developed independently and slightly earlier by Hilary Putnam.

sense that for every language in the collection a human child, if given appropriate positive evidence, can form a program that recognizes that language. There are surprising results as to which collections of languages are, and are not, learnable. Gold himself proved that any collection containing all finite languages and at least one infinite language cannot be identified. Imposing psychologically motivated constraints on the learner, Osherson and Weinstein have argued that any learnable collection of languages is finite. A wealth of technical results are now available about language learning.

7.2. Statistical Inference

One of the prinicipal statistical tasks is to infer a feature of a population from features of samples drawn at random from that population. One can view an ideal statistician as drawing ever larger samples, and using the statistical estimator to guess the value of the quantity of interest in the population. Some of the usual desiderata for statistical estimators are founded on this picture. For example, it is desired that an estimator be *consistent*, meaning that whatever value the quantity has in the population, for any positive epsilon the probability that the estimate of the quantity differs from the true value by more than epsilon approaches zero as the sample size increases without bound. This is clearly a convergence criterion; it implicitly considers a family of possible worlds, in each of which the quantity of interest has a distinct value. When the quantity is continuous there will be a continuum of such possibilities. A consistent estimator must, given increasing samples from any one of these possible worlds, converge with probability one to a characterization of the value the quantity has in the world from which the data is obtained.

7.3. Curve-Fitting

Every quantitative empirical science is faced with tasks that require inferring a functional dependency from data points. Kepler's task was to determine from observations of planetary positions the function giving the orbits of planets. Boyle's task was to infer the functional dependency of pressure and volume from measures on gas samples. These sorts of challenges can usefully be viewed as discovery problems. Data is generated by a process that satisfies an unknown functional dependency, but the function is known (or assumed) to belong to some restricted class of functions. In principle, more data points can be obtained without bound or limit, although in practice we may lose interest after a while. In real cases, the data are subject to some error, but something may be known about the error--its bounds, for example, or its probability distribution. The scientist's task is to guess the function from finite samples of data points. The conjecture can be revised as more evidence accumulates.

Many procedures have been proposed for this sort of discovery problem. Harold Jeffreys[8] for example, proposed a procedure that uses Bayesian techniques together with an enumeration of the polynomial functions. Nineteenth century computational designs, such as Babbage's, used differencing techniques for computing polynomials, techniques that could (in the absence of error) be turned round into discovery procedures. More recently Langley, et al.[9]

have tried doing exactly that, and have described a number of other procedures for inferring functional dependencies from sample data.

For any of these procedures, and for others, the foremost questions concern reliability. For any procedure we can and should ask under what conditions the conjectures will converge to an appropriate function. We can ask such questions for many different senses of convergence, and for many different accounts of what makes a function (other than the correct one) appropriate, but we should certainly try to formulate the issues and answer them. Very little work of this kind has been done; neither Jeffreys, nor Langley and his collaborators characterize exactly when their procedures will succeed, although in both cases it is easy enough to find many classes of functions (e.g., classes including logarithmic, exponential and similar transcendental functions) for which the procedure will fail in the long run. A more systematic study has been done for a related class of problems in which the data are finite pieces of the graph of a recursive function, and the discovery task is to identify the function by guessing a program that computes it.[10]

7.4. Generating Functions

One of the characteristic kinds of discovery tasks, at least in the physical sciences, is the discovery of generating functions. The idea is easiest to understand through an example. When monatomic gases are heated they emit light, but only light of certain definite frequencies. For example, when atomic hydrogen emits light, the spectrum contains a series of lines following a line whose wavelength is 6563 Angstroms. In addition, the spectrum of hydrogen contains a number of other series of lines. The spectral likes of other elements, notably the alkaline earth and alkali metal elements, can also be arranged in various series. Here is a kind of discovery problem: given that one can obtain the spectrum of such a gas, and

---

[8]See his *Scientific Inference*, Cambridge University Press, 1973.

[9]See P. Langley, H. Simon, G. Bradshaw and J. Zytkow, *Scientific Discovery*, M.I.T. Press, 1986.

[10]See Angluin and Smith, *A Survey of Inductive Inference Methods*, Technical Report 250, Yale University Department of Computer Science, 1982.

can identify lines as lines of a common series, what is the function that determines the frequencies (or wave-lengths) of the lines in the series? For the principal hydrogen series, Balmer solved this problem in 1885. Balmer's formula is

$$1/\lambda = R(1/4 - 1/n^2)$$

where n is an integer greater than or equal to 3, λ is the wavelength, and R is a constant (the Rydberg constant). Balmer generalized his formula to give a parametric family

$$1/\lambda = R(1/m^2 - 1/n^2)$$

for which series for m =1, 3, 4 and 5 have been found.

Balmer's formulas give a collection of discrete values for a continuous quantity, in this case the wave number, and they specify that collection by giving a (partial) function of the positive integers.

There are other famous discoveries in the natural sciences that seem to have an analogous structure. The central question in chemistry in the 19th century was the reliable determination of the relative weights of atoms. Alternative methods yielded conflicting results until in 1859 Cannizzaro noted that the relative vapors densities of compounds form series; for example, all compounds of hydrogen form a series, as do all compounds of oxygen, and so forth for any element. Of the continuum of possible values for compounds on hydrogen, only a discrete set of values is founded, and Cannizzaro discovered that the vapor density of any hydrogen compound is divisible by half the vapor density of hydrogen gas. Analogous results held for compounds of other elements. Cannizzaro's discovery was of crucial importance in putting the atomic theory on a sound basis; Balmer's discovery formed the crucial evidence for the early quantum theory of matter.

We can imagine a scientist faced with the following kind of problem: an infinite but discrete series of values of a continuous quantity is given by some unknown function of a power of the integers, I^n, or of the positive integers, but the function may belong to a known class of functions of this kind. The scientist can observe more and more members of the series, without bound, and can form a series of conjectures about the unknown function as the evidence increases. The properties of discovery problems of this sort have not been investigated either in the scientific or in the philosophical literature; and aside from the obvious procedure of looking for common divisors of values of a quantity, we know of no discovery procedures that have been proposed.

7.5. Theoretical Quantities and Functional Decompositions

If you have only a number of resistance free batteries, wires of varying but unknown resistances, and a device for measuring current through a circuit, you can discover Ohm's law, that voltage in a circuit equals the current in the circuit muliplied by the resistance in the circuit, even though you have no device to measure voltage or resistance, and even though at the beginning of the inquiry you have no belief that there are properties such as voltage and resistance. Pick a wire to serve as standard, and let the current through each circuit with each battery serve to a measure a property of each battery. Pick a battery to serve as standard, and let the current through each circuit with each wire and that battery serve to measure a property of each wire. You will then find, by simple curve fitting, that the relations between these two properties and the current is described by Ohm's law. Langley, Simon, *et al.* give a discovery procedure that solves this problem. But what is the general form of the problem?

Consider any real (or rational, or integer as the case may be) valued function of n-tuples of nominal variables. In the circuits considered previously, for example, current I is a function of each pair of values for the nominal pair (battery, wire). In general we have F(X1,..., Xn). Let F be equal to some composition of functions on subsets of the nominal variables. For example, I(battery, wire) = V(battery) * R(wire), where * is multiplication. A discovery problem consists of a set of functions on subsets of tuples of nominal variables, and for each tuple and set of functions, a function that is a composition of (i.e., some function of) that set. The learner's task is to infer the decomposition from values of the composite function.

Evidently a lot of clever science consists in solving instances of problems of functional decomposition, and thus discovering important but initially unmeasured properties. The properties of discovery problems of this kind, and of algorithms for solving them, are almost completely unstudied.

7.6. "Underdetermination," or Answerable and Unanswerable Questions

A scientist often has in mind a particular question to which an answer is wanted. The aim is not to find the whole truth about the world, but to find the answer to one particular question. There is a tradition in philosophy, in physics, and even in statistics of considering contexts in which particular questions cannot be answered. Philosophers talk about "underdetermination" in such contexts, whereas physicists tend to talk about similar issues in terms of "physical meaningfulness" and statisticians in terms of "identifiability." The examination of such issues is in structure very much like Gold's consideration of classes of

languages that cannot be identified. Arguments consider a collection of alternative structures of some kind, characterize the evidence generated from any structure, and establish that even "in the limit" some structures in the collection cannot be distinguished.

Consider a question about the shape of space: what is its global topology? In relativity, the evidence we can get at any time about that question is bounded by our past light cone; the discriminations we can make at any time are then determined by the data in that light cone and whatever general laws we possess. The general laws can be thought of as simply restricting the possible classes of space-time models. As time goes by, more and more of the actual universe is in the past of an imaginary, immortal observer. Are there collections of relativistic models for which such an observer can *never* determine the global topology of space? It turns out that there are, and some of them are not too difficult to picture. Imagine that space is a three-dimensional sphere, and that space-time is an infinite sequence of three dimensional spheres. Suppose the radius of the sphere expands as time goes on. At any moment the past light cone of an observer may include, at each past moment some but not all of the sphere of space at that past moment. If the radius of space expands fast enough, then at *no* moment will the past light cone include all of space. Now consider another space-time made mathematically from the first by identifying the antipodal points on the sphere of space at each moment. The shape of space will be different in the two space-times. The sphere is simply connected--any closed curve on the surface of a sphere, even a three dimensional sphere, can be contracted smoothly to a point. The projective space obtained by identifying antipodal points on the sphere is not simply connected. The two spaces have different topologies. Now imagine that space expands with sufficient rapidity that the past light cone of any point never reveals whether one is in the spherical sphace of the projective space. Many other classes of indistinguishable space-times have been described.[11]

7.7. Indistinguishability by a Class of Procedures

Issues of distinguishability also arise in settings that are remote from cosmology. In the social sciences, engineering, parts of biology and epidemiology, we often rely on statistical models of causal relations. Often an initial statistical model is thought to be in error, and a variety of algorithmic or quasi algorithmic techniques have been developed to find revisions. Factor analysis is one way ; procedures that modify an initial model by means of "fitting statistics" are another; procedures that try to match the empirical

---

[11]See C. Glymour, *Indistinguishable Space Times and the Fundamental Group*, and D. Malament, "Indistinguishable Space Times" both in J. Earman, ed., *Foundations of Space Time Theories*, University of Minnesota Press.

constraints entailed by a model with those found in the data are still a third.

For each of these kinds of procedures the discovery framework poses a relevant question: For what classes of models can the procedure succeed in identifying in the limit? What are the collections of models such that given data generated from any one model in the collection, as the size of the sample increases without bound the procedure will identify the model that actually generated the data?

Sometimes a variety of procedures share a feature; either they share a limit on the information they consider in forming a hypothesis, or they share a limit on the hypotheses they consider. In the latter case it is perfectly obvious that certain classes of models cannot be identified. In the former case, finding out what classes of models can and cannot be identified may take some work. The discovery paradigm empahsizes the importance of the work.

8. Conclusion

There is a lot of structure behind the words that translators have given to Plato's Meno and to Plato's Socrates. The structure is, we hope, plausibly attributed even though it is remarkably modern. That should be of no surprise to those who think philosophy really addresses enduring questions, and who think the questions of knowledge had the same force and urgency for the Ancients as for ourselves.