DISCOVERING THE WAVE THEORY OF SOUND: INDUCTIVE INFERENCE IN THE CONTEXT OF PROBLEM SOLVING

Paul Thagard and Keith Holyoak
Program in Cognitive Science, University of Michigan
Perry Building, 330 Packard, Ann Arbor, MI 48104

ABSTRACT

This paper describes a program which simulates the discovery of the wave theory of sound, using several kinds of inductive inference that are triggered in the context of problem solving. The most novel of these is conceptual combination, which produces new concepts by combining existing concepts, represented as frame-like clusters of production rules. Combined concepts are not a linear amalgam of existing ones, since the conflicting expectations of the rules in the donor concepts must be resolved by a set of top-down and bottom-up procedures. The theoretical concept of a sound wave is produced by conceptual combination. The rule that sound consists of waves is produced by applications of other kinds of inductive inference: generalization and abduction.

I THE DISCOVERY OF SCIENTIFIC THEORIES

Artificial intelligence research on scientific discovery has primarily been concerned with discovery of *laws.* The programs in the BACON series have used heuristics to "discover" Kepler's third law of planetary motion, Snell's law of refraction, and other important laws from the history of science (Langley et al. 1983). But there is more to scientific knowledge than empirical laws: the greatest scientific achievements are *theories,* which unite empirical laws systematically, usually by postulating non-observable entities. A standard example is Newton's theory of gravity, which postulated a non-observable force in order to explain a wide variety of phenomena including Kepler's laws.

This paper describes a program which has been used to simulate the discovery of a simple but important theory, the wave theory of sound. This theory has both central features which distinguish theories from laws: it postulates the existence of non-observable entities, sound waves; and it explains different classes of empirical phenomena, the propagation and reflection of sound. The discovering program is called PI, for "processes of induction". It solves problems by a process of rule firing and spreading activation; and in the course of problem solving various kinds of inductive inference are executed, including generalization, abduction, and conceptual combination. The wave theory of sound is generated in the course of attempts to explain why sound propagates and reflects.

n PROBLEM SOLVING

Space does not permit description here of the architecture and problem solving process of PI (see Holland, Holyoak, Nisbett, and Thagard, forthcoming). The most important data structures in PI are rules, similar to the production rules of Newell and Simon (1972), and concepts, similar to the frames of Minsky (1975). Rules can have any number of conditions and actions, each of which is represented using a kind of predicate calculus notation to allow for n-ary relations.

At each timestep, the only rules considered for firing are those attached to active concepts. (Compare the architecture of Wallace 1983.) The spreading activation of concepts in PI is directed by its problem solving activity, working forward with concepts activated because they occur in the actions of fired rules, and backward by concepts activated through sub-goal formation. Unlike the ACT system of Anderson (1976), spreading activation is not automatic: concepts become active because rule firing or sub-goaling suggests their relevance to the current situation. Unused concepts suffer a gradual decline of degree of activation until they drop below a threshold and cease to be active.

Parallelism is simulated by the firing of more than one rule and the activation of more than one concept at a timestep. At each timestep, the current state of the system is monitored and various kinds of inductive inference are triggered. Any number of such inferences can be made at a given timestep, allowing for the convergence of evidence in support of a conclusion.

in INDUCTION IN PI

In current operation, PI triggers four kinds of induction: two kinds of generalization, instance-based and condition-based; abduction; and conceptual combination. (We use "induction" in the general sense of "forms of inference that expand knowledge in the face of uncertainty." Cf. Rescher 1980.) These four by no means exhaust the list of varieties of induction, and we are in the process of implementing additional mechanisms for rule learning and concept formation. Our description of generalization and abduction will be brief, for we want to concentrate on the more novel operation of conceptual combination which creates the theoretical concept of a sound wave.

A. *Instance-Based Generalization*

In instance-based generalization we infer that all A are B on the basis of instances of A that are also instances of B. In PI, such inductions are triggered when the short-term memory of the problem solving system contains the information that two or more objects are instances of the same two highly active concepts. It is not necessary to select from the huge space of all the generalizations that the system *might* be able to make, since only a small subset will be triggered by the current state of the system. Once generalization is triggered, the information stored with concepts makes possible evaluation of whether generalization is warranted, taking into account both number of instances and the variability of the relevant kinds (Thagard and Nisbett 1982). For example, you will require fewer instances to make a generalization about the combustion properties of a kind of metal than you would to make a generalization about the color of a kind of bird, since birds are more variable with respect to color than metals are with respect to combustion properties. Variability judgments are made possible by the hierarchical organization of the frame-like concepts.

B. *Condition-based Generalization*

Condition-based generalization is performed by taking the intersection of the conditions of two rules, using two or more categories to create a new, broader category that contains the original categories as special cases (Cf. Hayes-Roth and McDermott, 1978). For example, from the rules:

If x is a MacDonald's hamburger, it is greasy.
If x is a Burger King hamburger, it is greasy.

We can infer:

If x is a hamburger, it is greasy.

This kind of generalization is triggered in PI by noticing active rules that have the same actions and have conditions differing in only one place.

C. *Abduction*

Abduction involves the generation of hypotheses to provide potential explanations of puzzling phenomena (Peirce 1931-1958; cf. Buchanan 1983). Explanation can be thought of as a kind of problem solving. In PI, requests for explanations are recast as different kinds of problems to be solved. If what is to be explained is a fact such as Fred's being late, then that fact is set as a goal to be reached by the system. If, on the other hand, what is to be explained is a general rule such as that people caught in traffic tend to be late, then a problem is set up with starting conditions consisting of a description of an arbitrary person *x* caught in traffic and a goal to be derived stating that *x* is late. An explanation is found when the corresponding problem is solved. Abduction is triggered only when the problem to be solved is one involving explanation, and search for explanatory hypotheses is constrained by the current state of activation. If the fact to be explained is why Reggie is ill-behaved, then the availability of the rule that children with learning disabilities are ill-behaved might give rise to the abduction that Reggie has a learning disability.

D. *Conceptual Combination*

Theoretical concepts are ones whose instances are not observable. Thagard (1984) described how theoretical concepts can be formed by *combining* existing concepts, through a process that accounts for several interesting psychological phenomena. That paper sketches how conceptual combination can be performed by reconciling slots in frames. PI uses essentially the same mechanisms of reconciliation, adapted to apply to rules rather than slots. The rules attached to concepts in PI have several properties found in frame slots, so the translation is straightforward. Slots in frames have *slot-names;* for example, the frame for SWAN has a slot named COLOR containing the value WHITE. Similarly, in PI the rule attached to the concept of swan that states, "If something is a swan, then it is white," is marked as concerning color. This property of a rule is called its *topic*

Slots in frames can contain different kinds of values, including default and range values (Winston and Horn 1981). Another useful kind is an *actual* value, which unlike a default value may not be overridden. Similarly, rules can be assigned a *status* as expressing either actual or default values. In the rule, "If *x* is a triangle, then *x* has three sides," the value of having three sides is actual rather than a default Most rules, however, will only express defaults.

To combine two concepts, PI proceeds as follows. First, the expectations produced by the rules attached to the two concepts are compared. If no topic with conflicting values is found, as in a mundane combination such as "red apple", no further processing is required. If a conflict is found, however, this triggers generation of a new concept to represent the combination. PI reconciles conflicts by considering a number of different cases. The most straightforward occurs when one rule has a default status while the other has an actual status: we clearly want to give the actual value priority over the default. Several such cases have been described by Osherson and Smith (1981), in arguments showing that conceptual combination is not well captured by mechanisms of fuzzy set theory. For example, our expectations about something characterized as a "striped apple" is no simple linear amalgam of STRIPED and APPLE. PI creates the concept of STRIPED-APPLE by noticing that the "striped" property attached to the concept STRIPED has actual status, whereas the competing rule attached to APPLE, "If *x* is an apple, then *x* is solid red," is only a default. The resulting rule, "If *x* is a striped-apple, then *x* is striped," explains why people consider an object that is a striped apple as more "typical" of the new concept than of either of the donor concepts.

Most rule conflicts will not be so easily reconciled. If working memory describes instances of the donor concepts which suggest a resolution, then a new rule can be formed in a fairly bottom-up manner. For example, suppose you are forming the concept of a Canadian violinist, with conflicting expectations (derived from bad movies) that Canadians are rugged, lumberjack types, whereas violinists are considerably more refined. In such a situation, PI consults instances of concepts with properties of the relevant type. If it has some examples of Canadian violinists, it uses their properties to resolve the conflict. Thu6 if the instances of Canadian violinists are all refined, then the new rule attached to the concept CANADIAN-VIOLINIST will express this expectation.

A more top-down mechanism can proceed without any instances. In combining the concepts of feminist and bank teller, PI encounters conflicting expectations about how political a feminist bank teller will be. One natural way of reconciling the conflict is to notice that feminists are much less *variable* in their political views than are bank tellers, so that we would expect feminist bank tellers to have the political attitudes of feminists rather than those expected much less reliably of bank tellers. (Cf. Tversky and Kahneman 1983). To assess variability, PI uses much the same mechanism that plays a crucial role in instance-based generalization.

The simple mechanisms so far described will not always suffice to reconcile conflicting rules. A more sophisticated version would consider whether either of the new possible rules would make possible the solution of any problems.

Conceptual combination is akin to the biologically inspired *crossover* operation of Holland (1984), although it combines more complex structures. It is thus more like the schema composition of DeJong (1982). Contrast the heuristics of Lenat (1983) which produce new concepts by mutating a single concept, not by combining two concepts.

IV SIMULATION OF THE DISCOVERY OF THE WAVE THEORY OF SOUND

Using these kinds of induction, the program PI simulates the discovery of the wave theory of sound. The first systematic discussion of the wave theory of sound is due to the Roman architect Vitruvius (1926 pp. 138f.), writing around the first century A.D. We have no detailed information about how Vitruvius or his predecessors actually discovered the wave theory of sound, but the text suggests that discovery occurred in the context of trying to explain why sound propagates and reflects, and that it depended on noticing a crucial analogy with water waves. Similarly, in one simulation, PI is set the problem of explaining why sound propagates and reflects. In the latter case, this amounts to solving the problem of starting with an arbitrary example of sound *x* and explaining why *x* reflects.

To get things going, the concepts SOUND and REFLECTS are activated. Activation of the latter concept leads to activation of several rules concerning reflection, including the information that water waves reflect and rope waves reflect. The presence of these rules on the active list then triggers a condition-based generalization that all waves reflect. This new rule then provides the basis for the abduction that *x* is a wave, since that in conjunction with the rule that all waves reflect would explain why *x* reflects. But now we have active messages that *x* is a wave and that *x* is sound, triggering both the generalization that all sounds are waves and the conceptual combination of sound-wave. Generalization succeeds because *x* was chosen arbitrarily, so that consideration of number of instances and variability is irrelevant

In PI conceptual combination only produces a new permanent concept when combination requires the resolution of some conflict, as in the striped apple and feminist bank teller examples. The text of Vitruvius shows that the combination of the concepts of sound and wave did require resolution of conflicting expectations: he remarks that sound spreads spherically whereas water waves spread out in a single plane. PI resolves this conflict by supposing that sound waves will inherit the more specific property of sound. The result is the newly-stored theoretical concept of a sound wave, containing the information that a sound wave is a kind of wave as well as numerous new rules formed from those attached to the concepts of sound and wave. Having formed the rule that sound consists of waves, PI is able to deduce why the sound *x* reflects, so the problem of explaining why sound reflects is solved.

This is clearly not the only way in which the fortunate confluence of ideas concerning sound and waves might have occurred. Another simulation takes into account the Greeks' strong interest in stringed instruments. PI is set the problems of explaining why sound reflects and propagates. It reaches a solution by rule firing and spreading activation, via associations from sound to music to stringed instruments to vibrations to waves.

The wave theory of sound was not the only one that the Greeks constructed. Democritus recommended a particle theory of sound, which PI discovers using rules about the behavior of balls. However, like most of the Greeks, PI rejects this theory because it cannot explain how sounds can pass through each other with little interference.

Thus the discovery of the concept of sound waves and of the wave theory of sound can be simulated using conceptual combination and other methods of induction occurring in the context of problem solving.

ACKNOWLEDGMENTS

We are grateful to Pat Langley for comments on an earlier draft, and to John Holland and Richard Nisbett for many conversations.

REFERENCES

Anderson, J.R. (1976), *Language, Memory and Thought,* Hillsdale, New Jersey: Erlbaum Associates.

Buchanan B.G. (1983), "Mechanizing Hypothesis Formation." In *PSA 1982,* Vol 2. Edited by P. Asquith and T. Nickles. East Lansing, Philosophy of Science Association, 129-146.

DeJong, G. (1982), "Automatic Schema Acquisition in a Natural Language Environment," *Proceedings of the National Conference on Artificial Intelligence,* American Association for Artificial Intelligence, 410-413.

Hayes-Roth, R., and McDermott, J. (1978), "An Interference Matching Technique for Inducing Abstractions," *Communications of the ACM 21:* 401-410.

Holland, J. (1984). "Escaping Brittleness: The Possibilities of General Purpose Learning Algorithms Applied to Parallel Rule-Based Systems." University of Michigan Cognitive Science Technical Report No. 63. Forthcoming in *Machine Learning II,* edited by R. Michalski, et al.

Holland, J., Holyoak, K, Nisbett, R., and Thagard, P. (forthcoming), *Induction: Processes of Inference, Learning, and Discovery,* unpublished manuscript, University of Michigan.

Langley, P., Bradshaw, G., and Simon, H., (1983), "Rediscovering Chemistry with the BACON System," in R. Michalski, J. Carbonell, and T. Mitchell, eds., *Machine Learning: An Artificial Intelligence Approach,* Palo Alto: Tioga, 307-330.

Lenat, D., (1983), "The Role of Heuristics in Learning by Discovery: Three Case Studies," in R. Michalski, J. Carbonell, and T. Mitchell, eds., *Machine Learning: An Artificial Intelligence Approach,* Palo Alto: Tioga, 243-306.

Minsky, M. (1975), "A Framework for Representing Knowledge," in P.H. Winston (ed.), *The Psychology of Computer Vision,* New York: McGraw Hill, 211-277.

Newell, A. and Simon, H. (1972), *Human Problem Solving,* Englewood Cliffs, N.J.: Prentice-Hall.

Osherson, D. and Smith, E. (1981), "On the Adequacy of Prototype Theory as a Theory of Concepts," *Cognition 9:* 35-58.

Peirce, C.S. (1931-1958), *Collected Papers,* 8 vols. Edited by C. Hartshorne, P. Weiss, and A. Burks. Cambridge: Harvard University Press.

Rescher, N. (1980), *Induction.* Oxford: Basil Blackwell.

Thagard, P. (1984), "Conceptual Combination and Scientific Discovery," in P. Asquith and P. Kitcher, eds., *PSA 1984,* vol. 1., East Lansing: Philosophy of Science Association, 3-12.

Thagard, P. and Nisbett, R. (1982), "Variability and Confirmation," *Philosophical Studies 42:* 379-394.

Tversky, A. and Kahneman, D. (1983), "Extensional vs. Intuitive Reasoning: The Conjunction. Fallacy in Probability Judgment," *Psychological Review 90:* 293-315.

Vitruvius, (1926), *The Ten Books on Architecture,* trans. M. H. Morgan, Cambridge: Harvard University Press.

Wallace, J., (1983), "Motives and Emotions in a General Learning System," *IJCAU83, Proceedings,* Los Altos: William Kaufmann: 84-86

Winston, P. and Horn, B. (1981), *USP,* Reading, Mass.: Addison-Wesley.