Positive Feedback Processes Underlying the Formation of Expertise

Brian R. Gaines
Knowledge Science Institute
University of Calgary
Alberta, Canada T2N 1N4

Abstract

Experts may be modeled as managing the inductive dynamics of knowledge acquisition in the knowledge processes of society. Who becomes an expert may be modeled as a random process under the influence of strong positive feedback loops in the social mechanisms giving access to knowledge. These models have implications for the design of expert systems.

1 The Nature of Expertise

Expert systems add a new element to knowledge acquisition and transfer processes in human society. However, the nature and formation of expertise is not well understood. In taking part of the social process of knowledge transfer and embedding it in technology, it is desirable to have a more overt understanding of what an expert is and does. Use of the technology also raises questions about the possibly changing role of experts—what will experts do once their expertise is captured—are they accountable for the advice given by expert systems based on their expertise?

Hawkins (1983) has abstracted from industrial experience in developing mineral exploration expert systems and proposed a model of human expertise relevant to expert systems. His model is summarized in figure 1:

- The expert first elicits data about the problem from the client;
- He or she develops a minimal model that accounts for the data provided;
- He or she generates advice based on the model and feeds this back to the client;
- The client may accept the advice, or query it and, possibly, the model;
- The queries lead to further data elicitation, and repeat of the modeling/advice/query cycle.

Thus, in Hawkins' model, the client plays an active role in further developing the model by providing more data until he or she is satisfied with the model and consequent advice. Expert advice giving and taking is part of a cycle of negotiation around a process of model formation.

This model of the expert-client interaction as the negotiation of a mutually acceptable model may be given formal foundations in mathematical modeling theory. Figure 2 shows a general hierarchical modeling process (Gaines 1987) based on Klir's (1985) analysis of model formation:

- At level one, *constructs* are those distinctions that the particular modeling system makes, a language for describing of the world;
- At level two, *data* are descriptions of actual case histories in terms of the constructs, an account of experiences of the world;
- At level three, *hypotheses* are the means of regenerating particular case histories from generalized accounts, rationalizations of the world (often called *models*);
- At level four, *analogies* are similarities between differently generated generalized accounts, correspondences between models;

- At level five, *abstractions* are accounts of a wide range of models, underlying foundations of analogies;
- At level six, *principles* are systemic foundations for abstraction, accounts of abstractions.

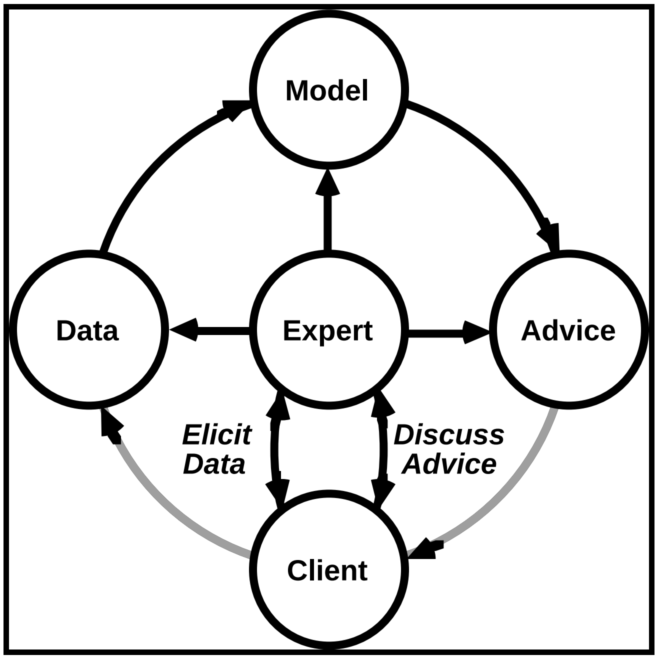

Figure 1 The negotiation cycle in expert consultation

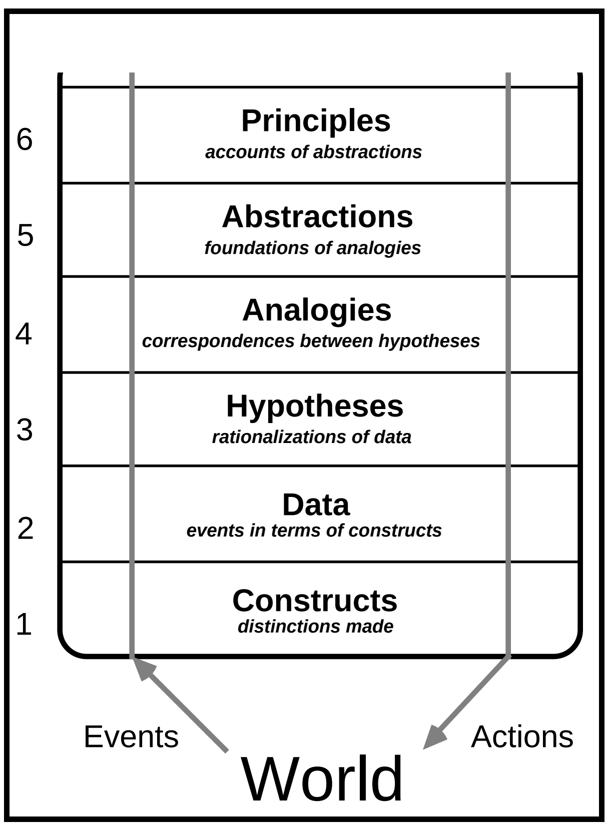

Figure 2 General hierarchical modeling process

Rasmussen (1985) has noted the importance, and analyzed the roles, of such hierarchical knowledge representation in individual and organizational decision making.

Modeling may be evaluated by the degree to which a level accounts for the information flowing through it. This may be measured as the *surprise* at an actual event in relation to one predicted, and the expected surprise is a form of entropy (Gaines 1977). The formal theory of modeling is one in which selections are made at each level down the hierarchy to minimize the rate at which surprise is passing up the hierarchy. The criteria for selection independent of the data are generally thought of as being ones of simplicity/complexity: of two models which fit the data equally well choose the simplest. The simplicity/complexity ordering is arbitrary and in its most general form is just one of *preference*. Hence the general modeling schema is one in which surprise flows up the hierarchy and preference flows down.

It should be noted that the modeling is an active process also—constructs are not just static partitions of experience. They may be operations: actions in psychological terms; processes in computational terms. Whether a system finds a distinction in the world, imposes it passively as a view of the world, or imposes it actively as a change in the world, is irrelevant to the basic modeling theory. The model of figure 2 encompasses both the expert's model-formation and model-application processes. There is neurological and behavioral evidence of the existence within the brain of the two channels of communication shown in figure 2 (Tucker & Williamson 1984). The arousal system passes surprise up to the cortex from the limbic region as unexpected events occur. The activation system passes preferences down from the cortex to the motor regions. The contextual interpretation of surprise according to its implications also provides a model for a wide range of human emotions (McCoy 1981, Gaines & Shaw 1984).

Hawkins' studies exemplify the roles of each of these levels in expert behavior and in the negotiation between expert and seeker of advice. In particular, he illustrates the significance of shared constructs between expert and user, and the potential for 'data' from falsely constructed events to invalidate expertise. The basic modeling theory, however, does not account for these types of social interaction, the relation between experts and users, the assumptions that an expert makes about the problem situation, the relation between experts and society. Indeed, it does not account for the concept of an 'expert' except as a particularly well-validated modeling system, leaving open the question as to how one system becomes better validated than another.

The following section extends the basic modeling theory to encompass social processes and the formation of expertise.

2 Social Processes in the Formation of Expertise

The weakness of the analysis of the previous section is that it accounts for only a minor part of the knowledge acquisition processes of an individual. Most of our knowledge does not come from forming models of events in the world, but rather from the social transfer of others' knowledge by a variety of cultural processes, some formalized in education but many very informal. However, the account can be extended to encompass such cultural knowledge acquisition if one notes that human civilization as a whole is representable as a general modeling system having no sources of knowledge other than its combined, cumulative experience of the world (Gaines 1987).

The modeling schema of figure 2 applies as much to groups of people, companies and societies as it does to the roles of a person. The modeling processes of an individual may be viewed as a cross-section of those of the civilization of which they are part. The modeling processes of society may be viewed as carried out by a distributed system of sub-processes in individuals. Cultural processes of knowledge transfer are then just various modes of communication between such sub-processes. The distributed system analogy nicely accounts for a variety of psychological and social phenomena: one person may assume many psychological roles (process switching), whereas a group of people working together may act as a single modeling entity and hence behave as one process (distributed processing).

These concepts may be formalized by considering the inductive inference process underlying knowledge acquisition. Deduction guarantees to take us from valid data to valid inferences, but the inferences are thereby part of the data—no new knowledge is generated. Induction takes us from valid data to models of that data that go beyond it—by predicting data we have not yet observed, and by giving explanations of the data in terms of concepts that are unobservable. Induction generates new knowledge but, as Hume (1739) pointed out over 200 years ago, the process is not deductively valid and it is a circular argument to claim that it is inductively valid. Goodman (1973) proposed that we accept the circularity but note that it involves a dynamic equilibrium between data and inference rules:

> *"A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend."*

Rawls (1971) in his theory of formal justice as induction from informal moral behavior terms this a *reflective equilibrium*.

Stich and Nisbett (1984) noted flaws in Goodman's argument in that individuals do not necessarily behave rationally in their own behavior relative to known good practice. They repaired them by proposing that the equilibrium is social not individual:

> *"a rule of inference is justified if it captures the reflective practice not of the person using it but of the appropriate experts in our society."*

This gives an operational definition of the role of experts within an overall knowledge acquisition system as referential sub-systems managing a domain of the knowledge acquisition process. It also leads to an concept of the role of expert systems in society, as making the referential process more overt and widely available. It is noteworthy that professional development books concerned with good practice emphasize the importance of balancing experts' authoritarian roles as referential sub-systems with their entrepreneurial roles as knowledge acquirers (Schön 1983).

Thus the pragmatic model of expertise developed by Hawkins can be brought into the formal framework of modeling theory by taking society to be a distributed modeling system in which experts are those with responsibility for managing the inductive process in particular domains. However, even this extension does not specifically account for the formation of expertise—how does a particular individual become an expert. The following section shows that positive feedback processes in the management of the various levels of knowledge transmission driven by random processes in society is adequate to account for expertise formation.

3 Expertise Formation by Positive Feedback Processes

An expert in Stich and Nisbett's account is a modeling system that has particular authority to decide on changes when events violate rules—in terms of figure 2, are new constructs needed, should the data be rejected, should the hypothesis be changed, and so on. However, this authority is a social construct, not usually one of right in present society, and the ascription of such authority to an individual is part of the larger modeling system's processes in determining which of its sub-processes is effective. Now, to become effective in modeling is, from figure 2, dependent on many factors: having access to knowledge at the upper levels such as principles, abstractions, analogies and classes of hypothesis; having access to knowledge at the lower level such as adequate construct systems and relevant data from case histories. Such access is itself a social resource, not generally available to all, and there are social mechanisms for allocating such access to individuals. It is the interaction between the criteria for access and the acquisition of knowledge that leads to the formation of expertise.

Sociologists have noted the very strong positive feedback processes in the dynamics of the scientific community (Hagstrom 1965). Merton (1973) coined the term the "Matthew effect" for those features of the reward system in research that were biased towards allocating greater credit for the same discovery to those with an already established reputation. The qualitative effect of such positive feedback processes is to amplify random differences within a population, creating strong distinctions which have no other basis than the feedback process itself. That is, even if all scientists were created equal in their capabilities, some would become experts relative to others because random differences in early performance generated credit that affected resource allocation and gave them greater opportunities to acquire knowledge.

In many areas of expertise where expert systems are expected to play a major role these positive feedback processes are exceptionally strong. In medicine, for example, the key learning resource is access to medical problems, and the 'owner' of such a problem has a keen personal interest in only allowing someone of very good reputation to handle it. The system, including considerations of legal liability, strongly funnels problems to those who are regarded as experts. It is, however, precisely these problems from which new knowledge is generated. Hence the data at level two in Figure 2 is made available primarily to those with already established reputations for having capabilities at level three and above to deal with it. Similar considerations apply to the award of scholarships, invitations to scientific congresses, and so on (Blume 1974). They also apply not only to individuals but also to social units such as a company subject to government procurement procedures that are heavily biassed to contractors with 'prior experience' and with whom the government agency has 'prior experience.'

It is obvious and expected that there should be many processes in society that seek out expertise to use or develop it. What is less obvious, however, is that the positive feedback involved in these processes is completely adequate in its own right to create 'experts' within a pool of initially undifferentiated people.

Simple simulation experiments of a competetive environment for two experts can illustrate the strength of these phenomena. For example, let the rules of a basic phenomenological simulation each problem requires certain knowledge; if the expert does not have the knowledge necessary it guesses with a probability of success, learning if it succeeds; the community chooses the expert for a problem with equal probability initially, gradually biasing the choice according to success or failure; there is no communication of knowledge between experts. Figure 3 plots: the probability that one expert will be always preferred and shows that this rapidly approaches 1.0—a best expert is determined; the expected knowledge of the better expert and that of the rival—one goes to 100% rapidly and the other is asymptotic to about 36%—there is objective evidence of the superior ability of the chosen expert. Which of the two individuals becomes the 'best expert' is, of course, completely chance.

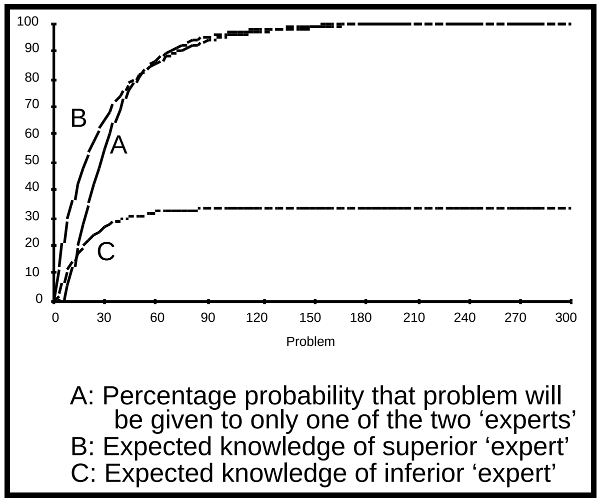

Figure 3 Simulation of the effects of positive feedback on the formation of expertise

The simulation shown is not Monte Carlo but based on the calculation of the exact probability distributions involved. It can easily be adjusted to take into account differences between the experts: that one starts with greater knowledge; that one learns faster; that one is favored initially (the prima facie credibility or 'well-dressed consultant' phenomenon). Similar simulations have been made of different positive feedback mechanisms; for example, if both experts are given the same problem but the problem difficulty is adjusted upwards if either gets it right—the situation of keeping up in the scientific 'rat race.' Effects have been introduced of the loss of knowledge through inadequate opportunities for its application, the growth of scientific knowledge so that there is always more to be acquired, and differential access to cultural knowledge transfer. All the simulations bear out the expected qualitative result, that a variety of different positive feedback mechanisms in society are adequate to account for differential expertise in a sample with initially equal knowledge and capabilities.

These results are not intended to suggest that individual differences are insignificant, or be a basis for egalitarian ethics. They do demonstrate, however, that the process of expertise formation is a natural systemic phenomenon in a learning society. It is functional because it distributes the processors available to different tasks through global mechanisms with no localized control. The random mechanism is efficient because it does not matter to the overall system which processors are allocated to particular tasks, and it will tend to pick out individual differences in capabilities if they exist. It puts experts into the role of traders who must offer their services to solve problems in order to gain access to those problems in order to increase their knowledge.

In practice, the professional and scientific communities also have strong mechanisms set up to reduce the effects of positive feedback in exaggerating differences, such as the ethical pressure to disseminate knowledge as widely as possible, and mechanisms for doing this through publications and seminars (Whitley 1984). However, much knowledge underlying skilled performance cannot be shared in this way (Nisbett & Wilson 1977) and the positive feedback mechanisms will always exert a strong effect.

4 Some Practical Consequences

The popular stereotype of an expert is a fount of knowledge in some domain. If that fount can be captured in an expert system then the role of the person who previously embodied is called into account. Those outside the expert systems industry often ask the question, “why should experts put themselves out of their jobs by allowing their knowledge to be transferred to an expert system.” The reply is usually that experts complain about the number of phone calls they have to answer to give advice on trivial problems. They would prefer to have the time to read journals, attend professional meetings, and think more about what they are doing. The model of expertise put forward in this paper gives theoretical foundations for this informal empirical result. The expert is not a static fount of knowledge but a dynamic process for continuously maintaining that fount by managing the acquisition of new knowledge.

The model of the expert’s relation to those seeking advice as one of trading it for access to problem-solving, and hence learning, opportunities is also consistent with the actual behavior patterns of experts in relation to expert systems. Divesting the unprofitable trading to focus on the profitable is a sensible strategy, that is, passing to an expert system those standard problems where little learning opportunity is expected. However, the uncertainties about the classification of any problem situation suggest there may be dangers in this approach which could be avoided by adequate design and application of expert systems. The advice giving system should contain rules which indicate problem-solving situations that are beyond the boundaries of its capabilities and refer these to a human expert (in the same way that human experts refer difficult problems to others). It should log actual case histories of advice giving for later review by human experts so that these rules can be validated. It should make provision for feedback as to the results of advice so that learning opportunities are not lost.

The continuity between the modeling role adopted by the expert in Hawkin’s model of expert-client negotiation, and the expert’s wider modeling role in managing Goodman’s reflective equilibrium, suggests that the performance and knowledge acquisition phases of expert systems should not be separated as they are currently. It would be better to think in terms of *knowledge support systems* that integrate knowledge acquisition and advisory roles (Gaines & Shaw 1987). In particular, the model of system-user relations as mutual negotiation rather than authoritarian advice giving is a good model for expert system.

All these considerations are also relevant to questions about the accountability of expert, expert system, and user for the advice. However, that is a side-effect of the social pressures on the inductive process—accountability, and the resultant rewards and punishments, are cultural mechanisms to encourage learning and the appropriate use of expertise. Well designed expert systems will make provision for the clear presentation of the quality of advice, reference to human experts, audit trails, logging, and case history review, because these are necessary to an effective technology—the possible legal consequences of not doing this are just a reminder of what is an effective technology in the domain of knowledge acquisition and transfer.

There are implications for the detailed technical design of expert systems also. As an example consider the form of rule often induced from case data which states that if the value of an attribute is unknown it is reasonable to fill in it with a default value based on the value of other attributes; for example, if the induction algorithm relates the value of attribute C to that of attribute A:

If (A Equal B) And (C Equal Unknown) Then (Set C Equal D)

This rule uses known data (facts) and induced relations (hypotheses) to make suppositions (expected consequences). It is also reasonable, but unusual, to add a meta-rule based on the same relation:

If (A Equal B) And (C Not-Equal D) Then (Set Surprise Equal True)

This meta-rule uses known data and induced relations to highlight anomalies (unexpected facts). The new attribute 'Surprise' picks up that the value of C is anomalous compared with previous data used in inductive inference. It can be used to trigger query or alarm reports, and the normal explanation mechanism based on back-tracking will trace the surprise to its source. It can be used to indicate that the problem under consideration is unusual and should be referred to a human expert or, at least, be recorded and flagged of interest.

In this example, the surprise variable corresponds exactly to the upward flow of modeling information in figure 2, and the conflict between the rule and the data is exactly the reflective equilibrium managed by experts discussed in the second section. It is was noted in the first section that human emotion can be modeled as contextual interpretation of surprise, and it is possible to regard meta-rules based on surprise as incorporating some of the attention-directing aspects of human emotion. The meta-attributes generated through anomaly-detection rules may be used as a basis for machine learning, as Lenat (Davis & Lenat 1982) has done in AM (terming surprise 'interesting'). However, they are also a basis for effective integration between expert and expert system, with the human expert continuing to have the role of managing the inductive process.

5 Conclusions

The role of the expert in the knowledge processes of society has been analyzed and shown to be that of managing the inductive dynamics of knowledge acquisition. Experts trade the skills based on their knowledge for access to problems providing them with new material for the inductive process. The formation of expertise is functional in general because it leads to division of labor in the management of knowledge acquisition. The development of an individual expert is a random process brought about by strong positive feedback loops in the social process; for example, that a proto-expert with superior performance is brought more problems and hence has a greater opportunity to learn and improve that performance. A diversity of such positive feedback processes operate in the professions and sciences with little relation between them except their overall effect in promoting the formation of expertise. Sociological evidence for the existence of these effects has been discussed, and simulation studies illustrating the quantitative phenomena have been described.

This model of the nature of expertise provides foundations for knowledge acquisition and transfer processes in expert systems and associated knowledge acquisition tools. It places expert systems technology within a social context and provides a basis on which social and legal aspects of their use and its implications may be analyzed. It has implications for the design and application of expert systems.

Acknowledgements

Financial assistance for the studies on which it is based has been made available by the Natural Sciences and Engineering Research Council of Canada. I am grateful to the anonymous referees for detailed critical comments.

References

Blume, S.S. (1974). **Towards a Political Sociology of Science**. New York: Free Press.

Davis, R. & Lenat, D.B. (1982). **Knowledge-Based Systems in Artificial Intelligence**. New York: McGraw-Hill.

Gaines, B.R. (1977). System identification, approximation and complexity. **International Journal of General Systems**, **3**, 145-174.

Gaines, B.R. (1987). An overview of knowledge acquisition and transfer. **International Journal of Man-Machine Studies 26**(4), 453-472 (April).

Gaines, B.R. & Shaw, M.L.G. (1984). Emotion: a systemic development of the personal construct model. **Proceedings of Personal Construct Network Conference** (June). Cincinnati.

Gaines, B.R. & Shaw, M.L.G. (1987). Knowledge support systems. **ACM MCC-University Research Symposium**. Austin, Texas: MCC. pp.47-66.

Goodman, N. (1973). **Fact, Fiction and Forecast**. Indianapolis: Bobbs-Merrill.

Hagstrom, W.O. (1965). **The Scientific Community**. New York: Basic Books.

Hawkins, D. (1983). An analysis of expert thinking. **International Journal of Man-Machine Studies**, **18**(1), 1-47 (January).

Hume, D. (1739). **A Treatise of Human Nature**. London: John Noon.

Klir, G.J. (1985). **Architecture of Systems Problem Solving**, New York: Plenum Press.

McCoy, M.M. (1981). Positive and negative emotion: a personal construct theory interpretation. Bonarius, H., Holland, R. & Rosenberg, S., Eds. **Personal Construct Psychology: Recent Advances in Theory and Practice**. pp. 95-104. London: MacMillan.

Merton, R.K. (1973). **The Sociology of Science: Theoretical and Empirical Investigations**. Chicago: University Press.

Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we can know: verbal reports on mental processes. **Psychological Review 84**, 231-259.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decisionmaking and system management. **IEEE Transactions on Systems Man & Cybernetics, SMC-15** (2), 234-243 (March/April).

Rawls, J. (1971). **A Theory of Justice**. Cambridge, Massachusetts: Harvard University Press.

Schön, D.A. (1983). **The Reflective Practitioner**. New York: Basic Books.

Stich, S.P. & Nisbett, R.E. (1984). Expertise, justification and the psychology of inductive reasoning. Haskell, T.L., Ed. **The Authority of Experts**. pp. 226-241. Bloomington, Indiana: Indiana University Press.

Tucker, D.M. & Williamson, P.A. (1984). Asymmetric neural control systems in human self-regulation. **Psychological Review 91**(2) 185-215 (April).

Whitley, R. (1984). **The Intellectual and Social Organization of the Sciences**. Oxford: Clarendon Press.