

# Introduction: Cognitive Autonomy in Machine Discovery

JAN M. ZYTKOW

ZYTKOW@WISE.CS.TWSU.EDU

*Department of Computer Science, Wichita State University, Wichita, KS 67208*

## 1. How is discovery different from learning?

This special issue of *Machine Learning* on discovery raises the question: How is discovery different from learning? In machine learning we typically distinguish between (1) learning as acquiring new knowledge in the form of concepts, taxonomies, regularities, and the like, and (2) learning as performance improvement and skill acquisition. “Discovery” applies to things that exist, such as the moons of Jupiter or the laws of nature, so in this section we confront discovery with learning in the first sense, of acquisition of objective knowledge. In the second section we discuss the distinction between discovery and invention. The third section reviews the ways in which the articles in this issue contribute to the growing autonomy and integration of machine discoverers, and it mentions recent research in machine discovery not represented in this issue.

### *1.1. A discover must be autonomous*

There is no difference between discovery and learning based on the type of knowledge to be acquired. Each piece of knowledge available to a learner must have been discovered earlier, and everything that has been discovered can become a subject of learning. The distinction becomes clear when we focus on the source of knowledge. A learner depends upon a teacher or a knowledge source, while a discovery can be granted only if it has been made without help of anybody who already knows.

Learning is easier than discovery, because a human teacher or any agent who knows a given piece of knowledge can guide the learner’s cognitive process in many ways. Consider concept learning from examples. The teacher selects a useful concept and prepares a collection of examples and counterexamples. The teacher can focus the attention of the learner on a particular task, for instance, on learning a conjunctive concept, a general boolean concept, or a recursive concept. The teacher can provide the evaluation criteria or even select the best alternative among solutions produced by the learner. Computer systems in all areas of machine learning receive help in the input preparation, in the learning process, and in the evaluation of results. Much of that help is typically called a learning bias.

Human discoverers throughout history did not rely on external authority because there was none at the time of discovery, or even worse, because the discovery contradicted what the existing authorities believed. Whether an individual person, mankind as a collective discoverer, or a computer system, a discoverer must be equipped with its own autonomously

applicable repertoire of techniques and values. Take concept learning as an example. An unbounded number of predicates can be defined by the primitives of any language, but only some make useful concepts. While a dependable teacher guarantees that the concept to be learned is useful, a discoverer must select concepts using its own judgment. A teacher who understands a concept can prepare a collection of examples and counterexamples and describe them by suitable attributes, while a discoverer must use its own strategies for data collection, and its own judgment about relevant attributes. The notion of autonomy requires few comments. Kepler, for instance, discovered his laws from data collected by Tycho de Brahe, so neither the data collection strategies nor the attributes were his. However, his discovery was autonomous in many ways. First, Brahe's data came without a guarantee that their exploration will lead to any discovery. Kepler picked them without assurance that he would be successful. Second, Kepler used his ideas of patterns in data, generating and evaluating many patterns before he made his discoveries. Third, autonomy best applies to the whole historical process of discovery, less to a small episode. Great discoveries were usually made possible by the contributions of many people over a long time. Uncountable observations, many attributes, and many hypotheses on planetary motions were considered before Kepler, but when the history of science puts a spotlight on the most spectacular events, we tend to disregard a huge number of previous efforts. When we consider the scientific community as a collective discoverer rather than the culminating moments of individual discoveries, we can better understand the process and see the price paid for knowledge of relevant attributes and relevant data, which allowed Kepler to make his discoveries.

### *1.2. Towards greater autonomy*

Two complementary notions of autonomy are useful for us. First, the more an agent can do in the external world, the more autonomous it is. To increase this type of autonomy, we can give the agent more means, for instance, more sensors or better manipulators. Second, within the same means, an agent is more autonomous if it can make more choices, uphold more values, and explore a richer space of goals. An agent can have meager means to influence the external world, yet be able to set its cognitive goals in response to external situations, to internal values, and possessed knowledge, so that it can make its own decisions and understand as much of the world as its limited means permit. Philosophers call this an existential concept of autonomy.

If we blame machine learners for their lack of autonomy, we must acknowledge that machine discoverers are not far ahead. No existing machine discoverer would reach much success in exploration of the real world if we did not provide help. The difference between machine discovery and learning is not so much in the current state of the art, but in the direction towards autonomy. The research in machine discovery should focus on increasing cognitive autonomy, reducing all types of need for external help. One way is to implement new components of the discovery process. Some components, such as instrument construction, may not directly produce discoveries, but by virtue of enabling them, they belong to discovery research. Further, we need to strengthen the integration and autonomous evaluation of results by machine discoverers. Greater autonomy means more discovery steps in succession performed without external intervention. When external intervention is replaced

by automated search and the overall search must stay within reasonable size, the accumulation of discovery steps is a big challenge. But it is also a big chance for asking the right research questions, and it gives the cognitive perspective necessary for the answers. A single step rarely permits a sound judgment about the results. A combination of steps provides a broader perspective on knowledge and more informed feedback on the reasons for acceptance.

When we eliminate, step after step, the need for external help, while maintaining or expanding the scope of knowledge that can be acquired by the discoverer, we get closer to the understanding of scientific values and the ways in which they support each other. Systems that perform a single activity, for instance, concept learning from examples, do not get that chance.

### *1.3. Concept learning and concept discovery*

In technical terms of logic, concepts are predicates that include free variables. Concepts by themselves have no truth value; they are neither true nor false. They name objects, properties, or patterns, but they do not make statement. Truth values can be assigned to statements, which use concepts and which have all variables bound by quantifiers. Statements are claims about the world. With the exception of tautologies, true and universally quantified statements are typically called laws and regularities.

A good, proven model of *concept discovery* comes from science. Initially, empirical concepts represented human sensors and human manipulations, for instance, the notions of warmth, acid taste, or labor needed to displace an object. Each of these concepts has helped to find regularities in nature, but as those regularities have been generalized, the initial concepts have been modified into concepts better suited to express knowledge. For instance, the acid taste helped to propose regularities in reactions of acids, but then patterns of reactions, initially discovered for acids recognized by their taste, were used to define acidity. Eventually the presence of hydrogen ions defined the theoretical notion of acidity, to be replaced by even more general concepts as chemical knowledge expanded to cover new substances and reactions. Many notions of acidity have been proposed in the historical process, and very few survived.

Concepts can be viewed as investments that produce payoff when they allow us to express regularities and laws. Better investments, that is, better concepts are recognized by more general knowledge that they help to express. Among an unlimited number of concepts that can be proposed, science uses a very limited number, choosing them based on the generality, utility, and accuracy of laws in which they occur. Concept discovery in science is not an isolated activity, because concepts are justified by feedback from knowledge. In machine discovery we can also use the same feedback.

In machine learning, concept learning from examples can be viewed as a very limited search for regularities. Membership in the target class is described by the target attribute, which indicates for each record whether it belongs to that class or not, that is, whether it is an example or a counterexample. The learner seeks the best definition of the target class in terms of other attributes. Such a definition has a truth value. If true, it shares many features of regularities, for instance, it can be used to predict class membership. The selection of the target class, however, is not autonomous. The target class is externally defined

and a learner searches only for a class definition. In contrast, a discoverer should be able to propose various target attributes and to search for regularities for each. A discoverer will see more value in the target concept when it occurs in many regularities or in more general regularities. While a learner may not understand the reasons why a concept has value, a discoverer should, because a broader scope of cognitive activities and additional evaluation criteria give it a better perspective for autonomous judgment of concepts.

#### *1.4. Conceptual clustering as limited discovery*

Conceptual clustering is a step towards autonomy in concept learning. Here the task is more open, aimed at autonomous creation of classes. Given a database of records, conceptual clustering seeks to divide all records into classes and to find a general description of each class. The concern for regularities in data was notably absent in early clustering systems, and resultant taxonomies have had little scientific value. A new generation of clustering systems guides the clustering process by predictivity of clusters (Fisher, 1987). The resultant cluster hierarchies demonstrate predictive power if regularities are present in the data. However, regularities are poorly represented by clusters, because the search for regularities is not the main task. For instance, a simple proportionality between attributes  $x$  and  $y$  must be approximated by many clusters, rather than by a simple pattern  $y = ax$ . This criticism applies equally to discrete numerical data, which can also be fit with equations, and to non-numerical data, where a functional dependency can be represented as a mapping between the sets of values of the attributes. In contradistinction to conceptual clustering, the main goal of many discovery systems is to find regularities in the data, while new concept construction has an instrumental role in the search for regularities.

In contrast to clustering, a regularity does not separate existing objects into classes, but it specifies a pattern obeyed by all objects, so that it distinguishes records that are possible from those that are impossible (statistical regularities distinguish statistically probable combinations of records from those that are improbable). Pattern extrapolation is poorly done by clustering. At least two classes must be produced, while to improve predictivity, the number of classes may be very large, even when all data follow one pattern. So even if clustering can be a limited form of discovery, the global regularities as such are overlooked, while their combinations are captured locally by clusters.

#### *1.5. Selection among regularities*

Consider fitting data with empirical equations. Candidate equations can be evaluated by their fit to data, but even when we limit the results to the simplest equations, several equations of comparable simplicity can typically fit the same data with a similar accuracy and without overfit. Each equation is an equally close approximation to the truth, at least according to the input data. A discoverer may not be able to make a choice in this situation, or it may not wish to make it, because its choice could be wrong.

Greater autonomy of a discoverer, however, gives a broader perspective on regularities, improving their evaluation. Various cognitive steps can help to choose among competing regularities. For instance, additional data can be collected in an area in which different

equations offer distinguishable predictions. Some equations are more susceptible to generalization, and some can be reduced to a known theory, while some others cannot. Also, when interpreted with respect to situations they describe, some equations permit a more plausible interpretation. For each of these reasons, the broader perspective contributed by further activities of the discoverer can disambiguate the choice.

### *1.6. Central role of discovery in learning*

Like everything in AI, machine learning turned out to be far more complex than believed 10 or 20 years ago, and applications are turning up slowly. Our concept learners have reached very high accuracy on individual tasks, but individual simple tasks have little practical appeal. Many learning tasks must be combined before a learning robot will be a reality. As long as we human teachers and experts on learning do not understand the interactions among different steps in the cognitive process, the chances are small that we can produce effective machine learners. We know how difficult and unrewarding it is to teach somebody who must be guided at each small step and who is not able to extend old methods to new situations so that we must explicitly teach all the links. At those times we realize how many steps may be required for effective learning. Once we realize how many steps are needed for even a simple learning task and how little guidance a good learner requires, we see that good learners are discoverers. A discoverer can take a piece of knowledge from a teacher and treat it as an element of the autonomous discovery process, that is, confront it with existing knowledge, and keep it if it can be merged. It is so much easier to teach a discoverer, because the discoverer will understand the links and autonomously fill in most of the task. Each of us is a discoverer, and we have been discoverers before we became learners, because as newborns we must have discovered, for instance, the meaning of the pointing gesture, the ideas of naming, truth, verification, and so forth.

Teachers are not perfect, so learners need their own independent judgment. As parents and teachers, we have bad moments. Fortunately our children and our students apply their own judgment. Our robots should have their autonomous system of values, too. We do not want them to learn everything somebody may try to teach them.

A discoverer does not have to do everything on its own for fear of losing autonomy. Search for knowledge is a collective process. Human discoverers inherit much, but they create their own synthesis that guides them in making their own discoveries. We can doubt whether any individual human discoverer has been completely autonomous, but collectively mankind has been. By a joint research effort in machine discovery, we are building such a collective discoverer. We are far from understanding what it will take to complete the task, but the pursuit of autonomy helps us to ask the right questions and to identify new research goals. Each component of the discovery method that we fail to anticipate will reveal itself in one or another shortcoming of our discoverers in comparison to human scientists.

Machine discovery can be equally helpful to understand knowledge acquisition. Knowledge acquisition systems should have enough autonomy so that they cannot be fooled or misled, so that they do not misinterpret advice, and so that they are able to seek additional knowledge.

## 2. How is discovery different from invention?

Machine discovery, like any other process in AI, has many features of a construction. A theory is gradually constructed as a combination of data structures. For many it does not sound quite right when we say that “a theory has been discovered.” Theories are often considered inventions, first because we construct them, and second because they are evaluated in pragmatic terms of simplicity, functionality, solvability of equations, and so forth. So while autonomy distinguishes discovery from learning, what is the difference between discovery and invention? The distinction seems obvious when we consider an example: hydrogen was discovered rather than invented by Cavendish, but Morse invented rather than discovered the telegraph. Hydrogen existed before Cavendish, while the telegraph is a complex spatio-temporal structure that did not exist before its invention. Natural laws exist eternally and independently of human actions, so they are also subject of discovery. All phenomena that occur naturally, even in very specific circumstances created by humans, can be discovered, not invented. It seems that truth in contrast to pragmatic values, “natural” in contrast to “conventional,” and “existence” in contrast to “construction” distinguish discovery from invention.

There is an element of invention in each discovery, because *discovery must be expressed in one form or another*. A distinction between “laws of nature” and “laws of science” can be used to express this intuition: the former exist in nature and await discovery, like the moons of Jupiter, while the latter are constructed to represent the former. To the degree that the laws of science capture the laws of nature and make true claims about the world, they are discoveries. The formalism into which laws are put can be viewed as invention. Theories include a component of invention, because they are artifacts generated by humans and are subject to many pragmatic requirements. For instance, different formalizations of classical mechanics by Newton, Hamilton, and Lagrange are inventions that express the same laws of motion. Because theories include both invention and discovery we often prefer to say that theories are introduced or developed.

Discoveries are not limited to “natural” phenomena. Anything that exists can be a subject of discovery. We can discover knowledge about artifacts. Such artifacts as computer programs are a legitimate object of scientific inquiry and discovery. For instance, we can use a discovery system to discover the computational complexity of an algorithm. As another example, consider heuristics. They are inventions, neither true nor false, but useful when they improve the behavior of a program. But we can say true and false things *about* heuristics, and we can discover those things. For instance, regularities in how heuristics affect search belong to the category of discoveries.

When we invent a particular class of heuristic search spaces, we actually *create a domain of study*. After a domain has been defined, it makes sense to distinguish between true and false claims and to discover things about that domain. If cost is one of the dimensions in a domain, the least-cost heuristic might be discovered, while various other heuristics could be invented.

Acquiring new knowledge and improving performance are two basic directions in learning. Machine discovery concentrates on autonomous knowledge acquisition, but to be practical it must use inventions such as more efficient search mechanisms.

### 3. Putting together a machine discoverer

The number of discovery systems has been growing considerably in the last decade. The articles in this issue form a small but significant part of this trend. The majority of work on machine discovery focused on the reconstruction of the *scientific method*. A picture of scientific craftsmanship is emerging from our joint research, confirming the old thesis of Newell, Shaw, and Simon (1962) that discovery is problem solving and that it can be carried out by computer programs. Although discoveries are abundant in everyday life, the focus on scientific discovery has proved successful. Everyday concepts notoriously elude formalization, while scientific formalism has historically proved to lead to spectacular results of scientific reasoning and knowledge representation. A typical scientific phenomenon can be studied in a compact domain that is limited to a few concepts yet rich in knowledge representable in a formal way. The continued concentration on the scientific method should lead to a practical success in the future, and the widespread use of computers and data-acquisition equipment by scientists permits rapid installation of a successful discovery system. Robot-discoverers are another application target, since scientific domains are similar to robotic applications.

#### 3.1. Empirical context of discovery

Recent discovery systems have expanded in numerous directions their treatment of the empirical context of knowledge, that is, the representation of objects, states, and processes, which were very primitive in early systems. In this issue, the articles by Shen, Scott and Markovitch, Rajamoney, and Nordhausen and Langley contribute to that subject. Other recent contributions include BLAGDEN (Sleeman, Stacey, Edwards, & Gray, 1989), GALILEO (Żytkow, 1990), and the work on analogy (Falkenhainer, 1987; Falkenhainer & Rajamoney, 1988), which used a mapping between elements of theories and elements of empirical situations that was developed in qualitative physics (Forbus, 1984).

**Interaction with the environment** is very important for autonomy of a machine discoverer. Environmental feedback can reduce drastically the number of observations needed to develop a theory, by making them where they seem to produce the greatest effect on the nascent theory. In this issue, discovery systems DIDO (Scott & Markovitch) and LIVE (Shen) exploit feedback from the environment. Selection of new experiments in response to discoveries has also been central in KEKADA (Kulkarni & Simon, 1987) and FAHRENHEIT (Żytkow & Zhu, 1991).

DIDO monitors the uncertainty of its own knowledge to select new experiments in the least known area in the space of possible experiments, and it keeps experimenting in that area long enough to enable significant improvement of knowledge. DIDO's mechanism of dynamic goal selection can be used by other systems whenever their knowledge can be interpreted probabilistically. While DIDO delays theory revision until many experiments have accumulated, LIVE employs deterministic rules and revises them after it finds an inconsistency with even a single experiment.

Scientists deal with competing theories by **designing crucial experiments**. Science inevitably reaches such conflicts, which are the focus of Rajamoney's DEED (this issue).

DEED takes a number of competing theories that make the same predictions in a given empirical situation. Knowing the correspondence between components of a theory and components of experimental situation, DEED uses this correspondence to plan to change the situation so that competing theories make conflicting predictions.

**Hypotheses about a physical situation** generated by Karp's HYPGENE (this issue) exploit similar links. HYPGENE responds to an incorrect prediction of a relatively complex theory applied to a hypothetical situation by searching for a better hypothesis. To do that, HYPGENE uses the links between reasoning steps and hypothetical elements of the input situation. Understanding how changes in the situation influence a theory's predictions, together with a goal-driven mechanism, reduces a large search space to a small number of alternatives that can explain the mismatch.

### *3.2. Feedback between concepts and regularities*

Scientific knowledge is shaped by an interplay between various discoveries, predominantly regularities and the concepts used to express them. In the current issue, Nordhausen and Langley's IDS and Shen's LIVE confront this problem.

Shen's LIVE couples its concept-generation mechanism with feedback about regularities to guide concept selection. For instance, faced with seemingly non-deterministic behavior, LIVE can find concepts instrumental in discovery of deterministic regularities in that behavior.

IDS creates clusters of states linked by similarities, finds qualitative regularities among those states, and finally finds quantitative laws that describe change of qualitative states. The system is a significant accomplishment in integration, but the missing feedback between regularity discovery and concept formation reduces its autonomy. Integration within IDS is such that clusters of states cannot be changed in view of discovered regularities. Another one-way link in IDS leads from qualitative to quantitative laws, whereas in science the opposite direction is no less important: quantitative regularities often lead to discovery of qualitative change. Without the feedback, the system may require a tutor/teacher who knows how to substitute for the missing feedback and feeds data in the right order.

### *3.3. Interaction among theory evaluation criteria*

At the heart of discovery is the detection of patterns in data. Pattern detection involves tradeoffs among the complexity of a pattern, accuracy of fit, and breadth of scope. It is difficult to find a justified tradeoff among these three values. Milosavljevic and Jurka (this issue), working on evolution of patterns in a particular class of DNA substrings, found a natural and convincing linkage between all three values through the minimum length encoding principle. In their domain, divergence from patterns occurs by mutations that are rare, so that every bit literally counts. The number of bits to reconstruct data from the patterns equals the number of bits to express each pattern or its divergence from another pattern, to link data to patterns, and to express divergence between data and patterns. The same basic unit of a bit makes all three evaluation criteria directly compatible.



Every equation-finding system employs a particular combination of evaluation criteria. Schaffer (this issue) introduces a method for external evaluation of scientific equation finders, which can throw light on their internal evaluation mechanisms. His results, showing low correlation between the equations preferred by scientists and the findings of BACON-1 (Langley et al., 1987) on the same data, offer an interesting challenge to BACON-1 but also raise the following question: What are the valid conditions to compare the results of automated equation finding with those of a scientist? A scientist has a far broader perspective than an equation finder when both consider the same data. For instance, the scientist may know measurement error, may believe that data must match a specific theory, may see a plausible empirical interpretation of terms in a particular equation, and so forth. To put a scientist and a discovery system on equal terms, we should give the scientist's broader perspective to the discovery system or we should confront human subjects with data free of context. Experiments by Qin and Simon (1990) demonstrated that several students who were shown Kepler's data, but were not informed of the data's origin, reached conclusions similar to those of BACON-1 and of Kepler.

### *3.4. Invention of new heuristics*

Speeding up  $A^*$  search with more efficient but admissible heuristics is an interesting challenge in invention. Prieditis's Absolver 2 responds to this challenge by autonomous application of two searches. Search in the space of abstractions of the initial problem is combined with search for a speedup of the solution to the abstracted problem. A new heuristic is returned when both searches are jointly successful. Since each new heuristic leads to a new  $A^*$  search, the speedup process can recur. Whether Absolver 2 *discovers* or whether it *invents* heuristics may depend on the preferred notion of existence, that is, on the ontology of mathematics.

### *3.5. Other directions in research on discovery*

Several research directions active in the last few years are not represented in this issue:

**Knowledge discovery in databases** has attracted considerable interest in recent years, leading to many automated methods of mining databases for useful knowledge. Three collections of papers, edited by Piatetsky-Shapiro and Frawley (1991), Piatetsky-Shapiro (1991), and Żytkow (1992) overview the state of the art.

**Discovery of structure** has stimulated steady and significant interest. Several systems discover hidden components and their properties (Revolver: Rose 1989; GELL-MANN: Fischer & Żytkow, 1990); others discover hidden properties of observable objects (BR-3: Kocabas, 1991) or postulate hidden steps in chemical processes (MECHEM: Valdés-Pérez, 1993). Sleeman et al. (1989) suggested an interesting search in the space of qualitative models of a chemical system.

## **4. Summary**

We argued that a discoverer must be autonomous because discovery of X can be granted only if it has been made without help of anybody who already knows X. Autonomy is

a matter of degree, and it grows by acquiring more means, goals, and values. We have shown that combination of discovery steps leads to more autonomy and to informed feedback between acceptance of different elements of knowledge. We also reviewed the ways in which articles in this issue contribute to autonomy of the discoverer and to cognitive feedback between elements of the discovery process. We argued that discovery plays a central role in learning and that to build a practically useful learner we should aim at a capable discoverer.

## Acknowledgments

Discussions with Saul Amarel, Tom Dietterich, Sakir Kocabas, Jack Mostow, Pawel Stefanski, and Raúl Valdés-Pérez helped to clarify many issues.

## References

- Falkenhainer, B.C. (1987). Scientific theory formation through analogical inference. *Proceedings of Fourth International Workshop on Machine Learning* (pp. 218–229). Los Altos, CA: Morgan Kaufmann.
- Falkenhainer, B.C., & Rajamoney, S. (1988). The interdependencies of theory formation revision, and experimentation. *Proceedings of the Fifth International Conference on Machine Learning* (pp. 353–366). Los Altos, CA: Morgan Kaufmann.
- Fischer, P., & Żytkow, J. (1990). Discovering quarks and hidden structure. In Z. Ras, M. Zemankova, & M. Emrich (Eds.), *Methodologies for intelligent systems* (Vol. 5). New York: North-Holland, pp. 362–370.
- Fischer, D.H. (1987). Knowledge acquisition via incremental conceptual clustering. *Machine Learning*, 2, 139–172.
- Forbus, K.D. (1984). Qualitative process theory. *Artificial Intelligence*, 24, 175–200.
- Kocabas, S. (1991). Conflict resolution as discovery in particle physics. *Machine Learning*, 6, 227–309.
- Kulkarni, D., & Simon, H.A. (1987). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science*, 12, 139–175.
- Langley, P., Simon, H.A., Bradshaw, G.L., & Żytkow, J.M. (1987). *Scientific discovery: Computational explorations of the creative processes*. Cambridge, MA: MIT Press.
- Newell, A., Shaw, J., & Simon, H.A. (1962). The process of creative thinking. In H. Gruber, G. Terrell, & J. Wertheimer (Eds.) *Contemporary Approaches to Creative Thinking*. New York: Atherton.
- Piatetsky-Shapiro, G. (Ed.). (1991). *Proceedings of AAAI-91 Workshop on Knowledge Discovery in Databases*, Anaheim, CA, July 14–15.
- Piatetsky-Shapiro, G., & Frawley, W. (Eds.). (1991). *Knowledge discovery in databases*. Menlo Park, CA: AAAI Press.
- Qin, Y., & Simon, H.A. (1990). Laboratory replication of scientific discovery processes. *Cognitive Science*, 14, 281–312.
- Rose, D. (1989). Using domain knowledge to aid scientific theory revision. *Proceedings of the Sixth International Workshop on Machine Learning*. San Mateo, CA: Morgan Kaufmann.
- Sleeman, D.H., Stacey, M.K., Edwards, P., & Gray, N.A.B. (1989). An architecture for theory-driven scientific discovery. In K. Morik (Ed.), *Proceedings of the 4th European Working Session on Learning (EWSL-89)* (pp. 11–23). London: Pitman.
- Valdés-Pérez, R.E. (in press). Conjecturing hidden entities via simplicity and conservation laws: machine discovery in chemistry. *Artificial Intelligence*.
- Żytkow, J. (1990). Deriving basic laws by analysis of processes and equations. In P. Langley & J. Shrager (Eds.), *Computational models of scientific discovery and theory formation*. San Mateo, CA: Morgan Kaufmann.
- Żytkow, J.M., & Zhu, J. (1991). Automated empirical discovery in a numerical space. *Proceedings of the Third Annual Chinese Machine Learning Workshop* (pp. 1–11), July 15–19, 1991. Harbin Institute of Technology.
- Żytkow, J.M. (Ed.). (1992). *Proceedings of the ML-92 Workshop on Machine Discovery (MD-92)*. Wichita, KS: National Institute for Aviation Research.