Some Recent Human/Computer Discoveries in Science
and What Accounts for Them

Raúl E. Valdés-Pérez

Computer Science Department and
Center for Light Microscope Imaging and Biotechnology
Carnegie Mellon University
Pittsburgh, PA 15213 - USA
phone: (412) 268-7127
e-mail: valdes@cs.cmu.edu

Abbreviated title: Some Recent Human/Computer Discoveries in Science

Abstract
We have recently reported several human/computer discoveries in biology, chemistry and
physics that have appeared in domain science journals. One may ask what accounts for
these findings, e.g., whether they share a common pattern. My conclusion is that each
finding involves a new representation of the scientific task: the problem spaces searched
were unlike previous task problem spaces. Such new representations need not be wholly
new to the history of science; rather, they can draw on useful representational pieces from
elsewhere in natural or computer science. This account contrasts with earlier explanations
of machine discovery based on the expert-systems view. My analysis also suggests a broader
potential role for (AI) computer scientists in the practice of natural science.

Keywords: scientific discovery, representations, empirical AI
1. Introduction

My collaborators and I have recently reported several novel findings in the domains of biology, chemistry and physics that represent various mixes of human/machine effort. These results, published in domain science journals, can be fruitfully analyzed as a whole by asking what accounts for them, i.e., by seeking patterns within these results. The generality of any patterns that are found can be tested against other instances of human/machine discoveries. These patterns can also be tested against the conventional wisdom of the contributions to natural science that can be expected from computers and (AI) computer scientists.

This article carries out such a pattern-seeking study by making use of the concepts of problem space and representation. My conclusion will be that each of the three novel findings involves a new representation of the scientific task, in the sense that the problem spaces searched were unlike previous task problem spaces. I will contrast my explanations of these human/computer discoveries with an earlier explanation of machine discovery based on the PROSPECTOR expert system. I end with some conclusions about the potential role of (AI) computer scientists in the practice of natural science.

2. Patterns in AI

A frequent task within an empirical science is to seek patterns within the phenomena studied by that science. In AI as empirical science (Simon 1995), the phenomena involve AI tasks and systems, their detailed designs, and the processes followed by the AI scientists who build the systems. One can seek patterns within these phenomena, and this goal has been explicitly advocated by Simon (1979), in arguing that the data-driven character of AI discovery systems suggests that AI research should itself be more data-driven. As an early example of pattern-finding, the system-building projects of the early AI scientists of the 1950’s led, upon detailed analysis of these systems, to the fundamental concept of heuristic search in problem spaces (Newell and Simon, 1976).

Such data-driven theoretical work of inducing patterns from systems is not widespread in AI, but neither is it rare. I can cite two examples from our recent work: Valdes-Perez, Zytkow, and Simon (1993) analyzed six separately-developed discovery systems and observed that all the discovery tasks shared an implicit representation which they named “search in matrix spaces.” Elsewhere, I have proposed the concept of “generic task of scientific discovery” as a generalization of a broad array of work on computing in science (Valdes-Perez 1995a).
3. Representations

In a recent psychological article, Kaplan and Simon (1990) define a representation simply as a problem space, and a change of representation as a change in problem space (also see Newell and Simon, 1972). The invention of a new representation for a scientific task is, then, the invention of a new problem space in which to accomplish the task.

Let us recall that a problem space is defined as a start state, a collection of operators that generate new states by recursive application, and a test of whether a goal state has been reached. The decision of when to apply each operator is guided by heuristics. In real-life tasks, problem spaces are often too large to be searched unselectively; the selectivity is provided by the heuristics.

Simon (1992) has argued that a usual task of the scientist is not to invent completely unprecedented representations, because these are historically very scarce, e.g., a possible example is the invention of infinitesimals and their calculus, although even this representation has perhaps some roots in antiquity in the paradoxes of Zeno and others (Moore 1995). Rather, the usual challenge for the creative scientist who finds new representations for scientific problems is to adapt and mix elements of known representations that are our common scientific heritage, and to develop the operators and heuristics that are applicable in the new situation. In what follows, I will adopt this concept of representation and ask to what extent each of the discoveries analyzed here was enabled by a new task representation. That is, I will ask whether there are new operators or new types of states.

One can illustrate this notion of representation in science with a familiar example from theoretical computer science: the design of randomized algorithms, i.e., algorithms that make random choices using a stream of random bits. Karp (1990) states that “Isolated examples of randomized algorithms can be traced back to the very early days of computer science, but the central importance of the concept of a randomized algorithm became generally recognized only about fifteen years ago” and cites a 1976 article that drew attention to the general concept. He adds that, by now, “randomization is an extremely important tool for the construction of algorithms.” Hence, we see that there has emerged a new representation of the algorithm-design task, in the sense that algorithm designers now have an additional mental operator that incorporates coin-flipping into the algorithm being designed, whereas previously designers would only make use of deterministic approaches. The problem space for algorithm design has thus been broadened.

The concept of representation that I have borrowed is technical; thus, detecting changes in representation amounts to the reasonably well-structured task of deciding whether a different problem space is involved. I have borrowed this concept directly from the theory of heuristic search, while implying no conflict with other accounts of the representation concept. For
example, an earlier article in this magazine gave an account of (knowledge) representation in terms of the roles that it plays in reasoning, such as serving as a surrogate for the natural world (Davis, Shrobe, and Szolovits 1993). There is no substantial conflict between these roles and the current definition.

4. Recent Discoveries

In each of the three cases of novel findings below, I will proceed by giving some background on the task and the computational approach, documenting the finding, asking whether a new task representation is involved, and contrasting some of the previous or alternative task representations. I give only enough detail to appreciate the task and the role of the new representation; further details are available via the citations.

The three findings share the circumstance of involving more or less human inference, although they do not involve following the lead of domain scientists. Hence, my conclusions will address the issue of what a computer + (AI) computer scientist can accomplish in science, rather than what a computer by itself can accomplish. To address the latter issue, one would need to analyze in detail what fractional credit should accrue to human or machine, and this would take us far afield.

4.1. Catalytic Chemistry

For the last six years (Valdes-Perez 1994a, 1995b), I have worked to automate a task that has occupied the energies of experimental chemists since the beginning of this century, after it was first realized that many chemical reactions take place as a series of steps, rather than as a single elementary act. A recurrent chemical research problem has been to explain (elucidate) the multistep character, or pathway, of a particular chemical reaction. Typically this is an experimentalist’s problem, because chemistry still lacks a practical theory to predict what steps will occur based on first principles. To obtain insight that will complement background theory on what steps tend to occur, one carries out experiments with the reaction: measures the overall stoichiometry, detects reaction intermediates, infers precursor relations among the products, plots concentrations over time, and so on. Much reasoning particular to chemistry\(^1\) is involved, hence this task is much more theory-driven than many of the recent tasks addressed by data-driven discovery research (Langley and Zytkow 1989).

The approach to automating this discovery task began with a common AI question: what is the space of hypotheses, and how can one generate the elements (reaction pathways) of

\(^1\) Or more broadly, to generic pathway elucidation tasks, because pathways occur in branches of science other than chemistry.
this space sensibly and systematically? There have been obstacles to overcome, involving the usual problems of combinatorial explosion, and also some difficult algorithm designs; the solutions are described in the cited articles. For the present purpose, I focus on one of these obstacles, as follows.

A key problem in this and other scientific model-building tasks is that unseen entities have to be conjectured in order to formulate acceptable models. For example, in a chemical reaction there are usually intermediates that remain undetected because of the practical limitations of experimental technique. If a large base of chemical knowledge were needed to make such conjectures competently, then automating this task would be problematic, except perhaps on a narrow class of chemical problems.

The obstacle of conjecturing unseen entities is overcome in MECHEM by a simple and seemingly naïve method: conjecture “wild cards” such as $X$, $Y$, $Z$, etc., use these wild cards together with the seen entities to formulate hypotheses, and then use the domain laws of a science to constrain these variables sufficiently (within the context of a specific hypothesis) to entail a small set of possible identities for the variables. For example, using the conservation constraint of reaction balance, the unknown $X$ in the following single-step hypothesis

$$CH_3 + MCH_2OOH \rightarrow X + CH_3OH$$

is inferred to consist of one $M$, one carbon, two hydrogens, and one oxygen atoms, i.e., its chemical formula is $CH_2OM$. In more complex cases involving multiple steps and unknowns, and possibly more than one unknown per step, a generalized linear equation solver is used to infer the chemical formulas of the wild cards. The introduction of wild cards is guided by simplicity: $N$ “wild cards” are introduced when $N-1$ wild cards prove insufficient to account for the constraints, i.e., when no adequate reaction pathway can be built. The program begins with $N$ equal to zero.
Now consider the nature of the problem space that reaction paths are navigated through. A state in the problem space that is reached by applying operators that propose under-determined chemical steps involving wild cards, or variables; each additional step that uses a variable places more constraint on the variable, resulting eventually in a completely determined step. This type of operator has not been used so far in computer applications to chemistry, which instead generally employ operators that result immediately in specific chemical steps (i.e., not including variables); further elaborations of the reaction pathway do not then constrain the steps further. Wipke, Ouchi, and Krishna (1978) have described

$$
\begin{align*}
1. & \quad H_2 + M_2 \rightarrow 2(HM) \\
2. & \quad M_2 + \text{ETHANE} \rightarrow HM + CH_3CH_2M \\
3. & \quad M_2 + CH_3CH_2M \rightarrow CH_2MCH_2M + HM \\
4. & \quad M + CH_2MCH_2M \rightarrow HM + CH_2MCHM \\
5. & \quad 2(CH_2MCHM) \rightarrow CH_2MCH_2M + CHMCHM \\
6. & \quad CHMCHM \rightarrow 2(CHM) \\
7. & \quad HM + CHM \rightarrow M + CH_3M \\
8. & \quad 2(CHM) \rightarrow CHM + CH_3M \\
9. & \quad HM + CH_3M \rightarrow M_2 + \text{METHANE}
\end{align*}
$$

Figure 2: MECHEM's Reaction Pathway for Ethane + H₂ → 2(Methane)

However, modern chemists reason at the level of molecular structures (i.e., graph-like connectivity between atoms), not just chemical formulas. But, it is also possible to infer the molecular structure of any wild cards, given their already-inferred formulas, and given the overall multi-step context in which they appear. For example, figure 1 shows a nine-step pathway in which there appear the six wild cards $U, V, W, X, Y,$ and $Z$ ($M$ and $M_2$ are not wild cards, but catalyst reaction sites). Figure 2 shows the same pathway but with the wild cards replaced by the molecular structures that MECHEM inferred for them. MECHEM found this pathway, which Valdes-Perez (1994b) then proposed as an alternative explanation for the catalytic reaction ethane + H₂ → 2(methane), which has been studied for over 20 years and whose pathway was considered largely solved. The above documents the first example of a novel finding: a simple, plausible alternative reaction pathway for a long-studied reaction.

I now consider the nature of the problem space that MECHEM searches.² Figure 1 shows a state in the problem space that is reached by applying operators that propose under-determined chemical steps involving wild cards, or variables; each additional step that uses a variable places more constraint on the variable, resulting eventually in a completely determined step. This type of operator has not been used so far in computer applications to chemistry,³ which instead generally employ operators that result immediately in specific chemical steps (i.e., not including variables); further elaborations of the reaction pathway do not then constrain the steps further. Wipke, Ouchi, and Krishna (1978) have described

²Actually MECHEM searches in more than one space, because, for example, the space of molecular structures is searched to provide structural information on the conjectured intermediates. However, the analysis focuses on the top-level problem space of possible pathways, i.e., steps of the form reactants → products.

³Much less in human practice: e.g., a recent book on writing reaction pathways (Miller 1992) contains nothing like this operator.
these typical operators for an AI audience, although in the context of chemical synthesis, not elucidation, but the principle is the same. Another group at Munich, Germany (Ugi and Wochner 1988) has over the years developed an alternative, more mathematical representation of chemical inference, which again differs from MECHEM's representation.

MECHEM's problem space involves new types of state and new operators, and therefore implies that MECHEM uses a new representation of the chemical task of elucidating reaction pathways. Of course, the notion of using variables and solving for their identities by inference rules is hardly unprecedented, even though the application to molecular structure here involves quite complex inference. However, as Simon (1992) has argued, the ordinary task of the scientist who proposes new task representations (even the scientist who carries out Kuhnian "revolutionary" science, rather than "normal" science) is to stitch together the representation from familiar elements; the key issue is whether the resulting cloth is new to its particular dress.

4.2. Particle Physics

Previously, Langley, Simon, Bradshaw, and Zytkow (1987) pointed out resemblances between the task addressed by their DALTON program and the following task from particle physics: Given observational data on reactions among elementary particles, one notices the absence of certain reactions that do not violate any known conservation laws, such as conservation of electric charge or baryon number. Because it is expected that any reaction that violates no conservation law will occur with some probability, these persistent absences are surprising and require an explanation. One form of explanation that has occurred is to invent new quantum numbers and assign them to the known particles such that the absent reactions fail to conserve the quantum numbers, but all the observed reactions successfully conserve the quantity. Devising such explanations for observational data (or phenomena) is what goes on in particle physics phenomenology; for example, the baryon and strangeness quantum numbers historically were postulated in such a data-driven way (Ne'eman and Kirsh 1986).

To illustrate the discovery task more concretely, let us consider the following simple problem that involves a single observed and a single never-observed reaction:

\[
\text{observed} : \quad \pi^+ + p \rightarrow \pi^0 + n \\
\text{unobserved} : \quad p \rightarrow \pi + \pi^0
\]

To exclude the unobserved reaction while admitting the observed one, it suffices to postulate a new conserved quantum number having the value of unity for \( p \) and \( n \) and zero for the other particles. This quantum number is conserved by the observed reaction and disconserved by the unobserved one.

Kocabas (1991) followed up on the observation by Langley et al. and wrote a Prolog program
BR-3 capable of postulating quantum numbers by a generate-and-test method that proposed specific numbers for each particle and backtracked when necessary. BR-3 was used on historical data to rediscover some accepted quantum numbers. In some cases, the program would find two new conserved properties to account for the observations, as has been the historical case.

In turn, I followed up on BR-3 after realizing that the logic of this discovery task could be formulated systematically, with guarantees about the simplicity of the resulting conservation laws. This realization was based partly on a representation that had been used in MECHEM and which is familiar to mathematical chemists: a reaction subject to conservation conditions can be represented as an algebraic equation. For example, the reaction $A + B \rightarrow C$, in which $A$, $B$, and $C$ are substances, particles, or whatever, implies the equation $a + b = c$ over a numeric domain such as the number of hydrogens present in the respective substances; the equality “=” expresses the conservation condition.

For the particle physics task, I formulated the relevant constraints explicitly and algebraically, and then implemented the PAULI program, which is a search/linear-programming hybrid based on an algebraic formulation of constraints (Valdes-Perez 1994c). PAULI is guaranteed to find the simplest assignment of quantum numbers, first in terms of the number of new conserved quantum properties (fewer is simpler), and second in terms of the magnitudes of the quantum numbers (smaller is better).

PAULI was applied to historical data that led to the independent discoveries of strangeness by Gell-Mann and Nishijima (Ne’eman and Kirsh 1986), and the program verified their solution consisting of the strangeness quantum numbers. However, further experiments with PAULI (consisting of trying various inputs) led to an invariant pattern: PAULI always found that one new quantum property or conservation law was enough even when BR-3 (and physicists) found several. This led to my conjecture and a proof (by M. Erdmann) of the theorem that on any input data consisting of observed and nonobserved reactions among elementary particles, at most one conservation law suffices to explain the observations, if they can be explained at all in terms of conserved quantum numbers. We have reported this theorem, together with a description of PAULI, to the physics community (Valdes-Perez and Erdmann 1994). This theorem constitutes the second example of a novel finding; this case involved more human inference than the first case, because (simple) human reasoning was involved in the conjecture of the theorem, but prompted by the invariant behavior of the discovery program.

As in my first example involving MECHEM, this result was enabled by a new task representation, which involves an algebraic representation of the relevant constraints, and which leads to guarantees about the simplicity of the solutions that the program will find. These constraints were reformulated through various linear-programming representational tricks into a working program. The operators in this new problem space add algebraic constraints
\[
\begin{align*}
(\pi_1^0 + n_1) - (\pi_1^1 + p_1) &= 0 \\
\pi_1 + \pi_1^0 - p_1 &\geq 1/100 \\
\pi_1 + \pi_1^0 - p_1 &\leq 99 \\
\pi_1 + \pi_1^0 - p_1 &\geq -99 \\
\pi_1 + \pi_1 &\leq 0 \\
\pi_1 &\leq 0 \\
\pi_1^0 &\leq E_{\pi_1^0} \\
-E_{\pi_1^0} &\leq \pi_1^0 \\
n_1 &\leq E_{n_1} \\
-E_{n_1} &\leq n_1 \\
\pi_1 &\leq E_{\pi_1} \\
-E_{\pi_1} &\leq \pi_1 \\
p_1 &\leq E_{p_1} \\
-E_{p_1} &\leq p_1 \\
\pi_1 &\leq E_{\pi_1} \\
-E_{\pi_1} &\leq \pi_1 
\end{align*}
\]

Figure 3: A State in PAULI’s Problem Space

and remove them, and the states consist of sets of algebraic constraints. To illustrate the problem space, the following two reactions:

- **observed:** \( \overline{\pi} + p \rightarrow \pi^0 + n \)
- **unobserved:** \( p \rightarrow \pi + \pi^0 \)

lead to the possible state in PAULI’s problem space shown in figure 3. The use of the subscript ‘1’ emphasizes that these are not particles, but variables over a numeric domain.

It was search within this representation, as carried out by PAULI, that created the phenomena (i.e., the program invariably finding single-conservation-law solutions) that led to the data-driven conjecture and Erdmann’s proof of the mentioned theorem. As in the case of MECHEM, this problem space is new to computer applications in particle physics, including AI applications. The new representation involved piecing together known elements of other representations with which the author was familiar, such as representing reactions as algebraic equations. Some technical problems had to be overcome, but the representational pieces were, nevertheless, not without precedent.

### 4.3. Cell Biology

A developmental biologist with whom I have collaborated has studied a developmental stage of early Drosophila embryos in which nuclei, or cells not yet having acquired a membrane,
Figure 4: Nuclear Divisions within Early Drosophila Embryos

1. If the nuclear divisions were truly random, i.e., randomly oriented between 0 and 180 degrees, then one would see occasional crowding due to chance.
2. Observations of time-lapse embryo images reveal little crowding; on the contrary, the nuclei maintain good spacing.
3. Therefore, the nuclear divisions are not randomly oriented, but must be somehow patterned.

Figure 5: Reasoning that Led to Finding of Patterned Behavior

undergo several successive rounds of division. These processes are illustrated in figure 4, which superimposes the original nucleus and the two daughter nuclei that result from each division. Earlier analysis of such data had concluded that nuclear divisions were disordered, or random, based on a histogram of division angles which appeared roughly uniform between 0 and 180 degrees (Minden et al. 1989).

Our analysis led to the conclusion that the divisions displayed patterned behavior (Valdes-Perez and Minden 1995), rather than the randomness cited earlier. This conclusion of patterning is my third and final example of a new finding in science. The detection of patterning was followed up by hypothesizing a model that would give rise to the observations, although this follow-up is not relevant to the analysis of this article.

Our finding of patterned behavior involved exclusively human reasoning, which after the event I was able, by retrospection, to reconstruct as shown in figure 5. This discovery can be credited to possessing or actively applying the heuristic rule which is the first premise of the deduction. As mentioned above, previous studies had considered a different problem space in the search for patterns, e.g., constructing a histogram of division angles, which revealed no simple pattern because the distribution was roughly uniform. Our finding involved, instead, reasoning in a different problem space which used the operators/inferences of figure 5.

In contrast to the first two cases from chemistry and physics, the current biological example
involved only human reasoning, not machine discovery. Nevertheless, afterwards we were able to devise a general, systematic method (PENCHANT) for finding subtle patterned behavior that is broadly applicable (Valdes-Perez and Perez 1994). This method, which is a complex instance of a permutation test (Good 1994), generalizes the three steps in figure 5 thus: nuclear division becomes the process under study, division orientation becomes a process parameter, and crowding becomes an arbitrary feature or quantity.

PENCHANT reaches the same conclusion about patterned divisions that we had, which was not a trivial exercise, because the human and computer methods are similar only upon close analysis. We have since begun to apply the method to other problems both within biology and elsewhere. No significant machine discoveries have yet been made, but PENCHANT has uncovered unsuspected patterns of moderate interest to at least one experimentalist (Valdes-Perez and Perez 1994).

When PENCHANT makes an independent discovery of significance, rather than retrace our own steps, then a stronger case for the role of the computer could be made than is warranted in this, our final example of a human/computer finding in science. Again, the task representation or problem space involves familiar elements drawn from elsewhere; nothing in figure 5 is new to the history of science.

5. Discussion

I have argued that new task representations or problem spaces have been involved in the three documented discoveries in science that this article has analyzed. It is likely that the use of a new representation was a crucial cause of the new findings, rather than a mere coincidence.

The representation-based explanation differs from some earlier accounts of machine discovery that involved expert systems. For example, in a report of the novel recognition of a hidden mineral deposit by the PROSPECTOR program, Campbell et al. (1982) state that “Expert systems ... encode valuable knowledge and judgment for the purpose of making that expertise more broadly available than in the past.” They explained PROSPECTOR’s achievement as follows: “We believe that this result was achieved because PROSPECTOR programming adequately reflects limited but appropriate selections from the knowledge and judgment of a known porphyry molybdenum expert.” Thus, PROSPECTOR’s result is due to confronting a model of a human expert (embodied in PROSPECTOR) with new data unavailable to the original expert. This expert-systems view focuses on expanding the availability of expertise through the use of expert systems, rather than expanding the expertise itself. In contrast, my representation-change explanation of the three recent discoveries focuses on a change in task problem space, i.e., a change in expertise.
The new task representations that I claim were, of course, humanly designed or even searched (in the biology case) rather than invented by machine. However, much scientific work involves problem solving within a given representation, rather than inventing a new one. Hence, the computer’s role is not unlike the role of many an individual scientist who makes a publishable discovery within a given representation that served as a starting point.

Given the research interest in automating changes in representation (Benjamin 1990), one can also ask how drastic were the representation changes described here. I believe that these changes were not minor, e.g., they do not involve simple changes to the problem space such as deleting an operator, making an operator more specific by instantiating one of its parameters, and so on. Rather, the new problem spaces differ from the previous spaces in more complex ways. Moreover, these representation changes were largely theory-driven (rather than data-driven) in the sense that they involved analytical insights into the special task structure which were not obtained in any obvious way by the use of data. This theory-driven character makes it doubtful whether any attempt to automate these representation-change processes in a general way can succeed in practice, at least for scientific applications.

Besides the expert-systems viewpoint discussed above, other common views credit the achievements and/or potential of machine discovery to the computer’s greater speed and memory, or to capabilities for doing exhaustive searches within known representations. More broadly, the role for computer scientists in natural science is commonly perceived as limited to inventing or adapting faster algorithms and more sophisticated data structures, or devising means to deal with very large databases. However, the three discoveries above all involved modest computations (whether computer or mental), straightforward data structures, and modest amounts of input data.

These results suggest that new task representations, developed perhaps by a systematic analysis of the heuristic relation between data and discovery, will constitute a significant fraction of the machine-aided discoveries that will occur in the short and medium term. As illustrated above, these new representations need not involve elements wholly new to the history of science; rather, they can draw on a wide array of representational pieces that have proved useful elsewhere in natural or computer science. I surmise that a detailed analysis of other productive machine discovery programs (such as META-DENDRAL (Lindsay et al. 1993) and TETRAD (Spirtes, Glymour, and Scheines 1993) will confirm the pattern reported here.

Acknowledgments. This work was supported partly by a grant #IRI-9421656 from the National Science Foundation, by a Science and Technology Center grant #MCB-8920118 from the National Science Foundation, and by the W.M. Keck Center for Advanced Training in Computational Biology. I thank Paul Rosenbloom for asking, during a seminar at USC ISI, what accounted for these recent discoveries; this question led eventually to the writing of this article.
REFERENCES


Learning, 326–334. San Francisco, Calif.: Morgan Kaufmann.
