ELSEVIER

Artificial Intelligence 82 (1996) 331-339

Artificial Intelligence

Research Note

A new theorem in particle physics enabled by machine discovery

Raúl E. Valdés-Pérez *

Computer Science Department and Center for Light Microscope Imaging and Biotechnology, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received February 1995; revised December 1995

Abstract

A widespread objection to research on scientific discovery is that there has been a noticeable dearth of significant novel findings in domain sciences contributed by machine discovery programs. The implication is that the essential parts of the discovery process are not captured by these programs. The aim of this note is to document for the AI audience a novel finding in particle physics that was enabled by the machine discovery program PAULI reported previously. This finding consists of a theorem that expresses the minimum number of conservation laws that are needed, mathematically speaking, to account for any consistent experimental data on particle reactions. This note also reports how a puzzle raised by the theorem—its conflict with physics practice—is resolved.

1. Introduction

A widespread objection to research on computational scientific discovery is that there has been a noticeable dearth of significant novel findings in domain sciences contributed by discovery programs. The implication is that the essential parts of the discovery process are not captured by these programs. If this implication is true, then the whole automated scientific discovery enterprise is so far of doubtful soundness. Of even wider consequence is that the theory of heuristic search is seriously incomplete, since it fails to account for a salient aspect of human reasoning: discovery in science. Therefore, it becomes critical to record instances of novel machine discovery in order to falsify the premise that underlies this serious implication, i.e., the premise that discovery programs have enabled no significant new discoveries.

* E-mail: valdes@cs.cmu.edu.

SSDI 0004-3702(95)00128-X

The aim of this note is to document for the AI audience a novel finding in particle physics that was enabled by the machine discovery program PAULI reported previously [4]. This finding consists of a theorem that expresses the minimum number of conservation laws that are needed, mathematically speaking, to account for any consistent experimental data on particle reactions.

I proceed by describing briefly the discovery task and summarizing prior work on it. Then I present the novel finding and the circumstances that led to it, and explore some of the implications of this finding within the particle physics domain. Since our research on scientific discovery emphasizes generality within science, I examine briefly the generic character of this task in science generally. This note closes by summarizing the lessons for machine discovery.

2. The particle physics task

The task of postulating conservation laws from observational data on particle reactions was apparently first mentioned in the scientific discovery literature by Langley, Simon, Bradshaw, and Zytkow [2]. I illustrate the task by a simple example. [1]

Let us suppose that experiments have shown that the following reaction among particles is observed to occur:

$$\overline{\pi} + p \rightarrow \pi^0 + n.$$

Furthermore, this second reaction

$$p \not\rightarrow \pi + \pi^0$$

has *never* been observed (this is implied by the symbol $\not\rightarrow$), despite much experimental effort, and despite the fact that the reaction is not ruled out by existing theory. These circumstances raise a quandary which calls out for resolution.

An adequate resolution lies in postulating a new conserved property that has the value of unity for the particles p and n and zero for the other particles. This property is conserved by the *observed* reaction and violated by the *unobserved* one, which explains why the unobserved reaction never occurs. Historically, particle physicists have faced similar puzzles involving more numerous reactions and have resolved them in such a partially data-driven manner by postulating conserved particle properties [3].

We can generalize the previous simple example into some general concepts of particle physics. First, a *conservation law* states that some aggregate quantity is conserved by a stated physical process. A *phenomenological conservation law* is one that is not discovered or justified on theoretical grounds, but instead serves as a rather ad-hoc explanation of observations; phenomenological reasoning in physics corresponds roughly to the data-driven reasoning of AI terminology. Conservation laws in particle physics are examples of *selection rules*: they select which hypothetical reactions cannot occur because they violate conservation; not all selection rules need be conservation laws.

[1] This example, and some other parts of this paper, are taken from previous publications in order to make the exposition self-contained.

Finally, the quantities that these conservation laws conserve are *quantum numbers*: simple, small numbers that characterize particles and which are not necessarily integers.

3. Prior work

Langley, Simon, Bradshaw, and Zytkow [2] were apparently the first in the AI literature to mention the discovery task of postulating phenomenological conservation laws. These authors pointed out the task's resemblance to the task carried out by their DALTON program in chemistry, implying that the former task was possibly amenable to the heuristic search methods that DALTON used, although no specific approach was given.

Kocabas [1] was the first to describe a program BR-3 capable of performing the task; his BR-3 used some algebraic manipulation to reduce the initial search space, followed by generating specific quantum numbers for particles, testing for contradictions, and backtracking. Given historical data on particle reactions, BR-3 was able to re-discover laws such as the conservation of baryon quantum number, lepton quantum number, and electron and muon numbers. However, Kocabas reported that BR-3 did not find the accepted strangeness quantum numbers when given the reactions that led historically to the discovery of strangeness.

Valdés-Pérez [4] then described the PAULI program which uses a combination of linear programming and backtrack search. PAULI's approach is based on re-representing the particle reactions as two sets of linear algebraic expressions,[2] making use of a well-known technique in chemistry that I employed in another discovery program MECHEM [5] for reasoning about multi-step reaction pathways. PAULI was able to re-discover the strangeness quantum numbers using the historical assumptions and data available to Murray Gell-Mann, the co-discoverer of strangeness [3].

3.1. Discrepancies between BR-3 and PAULI

Kocabas' paper showed how, given the reactions in our Table 1 (Kocabas' Table 3), BR-3 postulates the two accepted conservation laws of baryon and lepton number shown in Table 2. Every observed reaction in Table 1 conserves the sum of baryon numbers; that is, the summed baryon numbers of the reactants equals the corresponding products sum. The same conservation condition holds in the case of lepton numbers. On the other hand, each unobserved reaction in Table 1 violates at least one of the two laws. For example, the reaction $p \rightarrow \pi + \gamma$ violates baryon number conservation, but not lepton conservation.

On the same reactions data from Table 1, PAULI finds that one conservation law is enough to account for the observations; its quantum numbers also appear in Table 2. Only the three particles p, n, and K^0 receive unit quantum numbers, whereas the

[2] A reaction $A + B \rightarrow C + D$ that conserves some quantity Q implies the equation $a_Q + b_Q = c_Q + d_Q$ whose terms are variables rather than particles. Similarly, if the reaction fails to conserve Q, this implies the *inequation* $a_Q + b_Q \neq c_Q + d_Q$.

Table 1
Reactions giving rise to the baryon and lepton conservation laws

Observed reactions			Unobserved reactions		
$p+p$	$\rightarrow$	$p+p+\pi^0$	p	$\rightarrow$	$\overline{e}+\gamma$
$p+p$	$\rightarrow$	$p+\pi+n$	p	$\rightarrow$	$\pi+\pi^0$
$p+\pi$	$\rightarrow$	$\pi+p$	p	$\rightarrow$	$\pi+\gamma$
$\overline{\pi}+p$	$\rightarrow$	$\overline{\pi}+p$	p	$\rightarrow$	$\pi+\pi+\overline{\pi}+\pi^0+\pi^0$
$\overline{\pi}+p$	$\rightarrow$	π^0+n	$p+p$	$\rightarrow$	$\overline{\Lambda}+\overline{\Lambda}$
$\overline{\pi}+p$	$\rightarrow$	$p+\pi+\overline{\pi}+\overline{\pi}$			
$\gamma+e$	$\rightarrow$	$\gamma+e$			
$e+p$	$\rightarrow$	$e+p$			
π^0	$\rightarrow$	$\gamma+\gamma$			
$\overline{\pi}$	$\rightarrow$	$\mu+\overline{\nu}_\mu$			
π	$\rightarrow$	$\overline{\mu}+\nu_\mu$			
μ	$\rightarrow$	$e+\nu_\mu+\overline{\nu}_e$			
n	$\rightarrow$	$p+e+\overline{\nu}_e$			
$\overline{\pi}+p$	$\rightarrow$	$\Lambda+K^0$			

Table 2
Quantum numbers for the particles

Particle	Baryon number	Lepton number	PAULI's number	Particle	Baryon number	Lepton number	PAULI's number
p	1	0	1	$\overline{\nu}$	0	−1	0
n	1	0	1	μ	0	1	0
e	0	1	0	$\overline{\mu}$	0	−1	0
$\overline{e}$	0	−1	0	γ	0	0	0
Λ	1	0	0	π	0	0	0
$\overline{\Lambda}$	−1	0	0	$\overline{\pi}$	0	0	0
K^0	0	0	1	π_0	0	0	0
ν	0	1	0				

other particles are assigned zero. As is required of any solution, the observed reactions conserve PAULI's quantum number, whereas the unobserved ones violate it. Note that the program's quantum number is not a simple sum of the baryon and lepton numbers found by BR-3.

To explain its input data, PAULI prefers fewer conservation laws, ideally a single law (unless there are no unobserved reactions, in which case no "selection rules" are needed, since there is nothing to select *against*). This is the primary criterion of simplicity in the program. Given competing explanations of the same data, e.g., two alternative conservation laws, PAULI prefers the law that involves the smaller sum of the absolute values of quantum numbers; this preference expresses the program's secondary simplicity criterion.

4. Some puzzles, and an explanation

On the same input data, PAULI consistently found simpler (one-conservation-law) solutions than did BR-3, which was puzzling, since BR-3's achievements were rediscoveries of accepted results in particle physics. Our previous explanation for this

puzzle [4] consisted of three alternatives:
(1) Physicists erred by proposing unnecessarily complex assignments of quantum numbers.
(2) Physicists used further constraints to postulate the new quantum numbers.
(3) PAULI's simplicity criteria (inductive bias) are different from (and inferior to) the criteria used by physicists.

The second of these alternatives was correct in the case of PAULI's surprising conclusions about strangeness. That is, on the same strangeness input data as were used historically, consisting of (1) observed particle reactions, and (2) unobserved particle reactions, PAULI's solution was at first simpler than the solution accepted in physics. However, Ne'eman and Kirsh [3] mention a further constraint (the nucleon and pion families possess zero strangeness) that was not mentioned in the textbook (Omnes, 1971) cited by Kocabas, but was assumed by Murray Gell-Mann for his discovery of the strangeness conservation law in 1953. When this constraint was incorporated into PAULI for the strangeness case, the program did find the accepted values of strangeness for the particles involved in the input reactions. This solution to the puzzle *in the case of strangeness* involved the second alternative above (i.e., physicists had used additional constraints), and suggested that perhaps the entire puzzle of why PAULI persisted in finding simpler solutions could be solved by recourse to the same explanation, of an omission of analogous constraints. Surprisingly, this suggestion turned out to be wrong in the general case.

To gain further insight into this mystery, Kocabas and this author each ran his program on other inputs besides those in Kocabas' original paper; PAULI invariably found that the simplest solution involved only one conservation law, despite the fact that BR-3 would find multiple ones, and despite the fact that particle physicists had also postulated several conservation laws. From these observations of PAULI's invariant behavior, I conjectured that, on *any* consistent[3] reaction data *whatsoever*, one conservation law was provably sufficient to rule out the unobserved particle reactions and rule in the observed reactions. I turned to a colleague (Michael Erdmann) for help in proving this theorem, which he carried out by building on the matrix algebraic representations used to design PAULI.

Theorem. *For any set O of observed reactions and set U of unobserved reactions, at most one quantum number conservation law suffices to rule out U and rule in O. That is, there exists a numerical assignment to each particle appearing in the reactions such that every reaction in O conserves this number via summation, and every reaction in U fails to conserve it.*

We then reported these results to a physics audience [9] as three contributions: (1) an automation of the discovery task based on simple principles; (2) a systematic derivation

[3] Consistency means that the observed and unobserved reaction sets can in principle be distinguished using selection rules of conservation. If some unobserved reaction U is linearly dependent on the observed reactions O, then the dataset is inconsistent, because if every reaction in O conserves a quantum number, then U must also.

of the strangeness quantum numbers using historically accurate data and assumptions; and (3) the cited theorem on the parsimony of phenomenological conservation laws (i.e., one law is enough for all conceivable, consistent experimental data).

Section 5 resumes the discussion from the viewpoint of research on scientific discovery. I now proceed to explore some of the implications of this theorem, which is not without interest for those concerned with the philosophy and practice of induction in scientific inference.

4.1. *One law versus multiple laws*

For any given input data, the single all-encompassing conservation law found by PAULI is not in general logically equivalent to the alternative multiple conservation laws. That is, although both theories explain the reactions data, they make discrepant predictions about *unseen* data. In general, each theory prohibits some reactions that the other theory allows, although one expects that the multiple conservation laws will be more restrictive, since each law serves as an independent constraint on the possible reactions.

For example, Table 2 above showed two theories for the reactions of Table 1: the baryon/lepton numbers accepted by physicists and PAULI's numbers. As an experiment, I formed all 1575 possible reactions of the form $A \rightarrow B + C$ that involve the given particles. Of these reactions, the baryon/lepton theory prohibits 1325 reactions and accepts 250, while PAULI's theory prohibits 675 and allows 900 (as expected, the dual-conservation-law theory is more stringent). There are 168 reactions that are allowed by both theories, 593 reactions that are prohibited by both, and 814 reactions (more than half) on which the two theories disagree. Two examples of these 814 discordant reactions are: $p \rightarrow e + n$ is prohibited by the baryon/lepton theory, but not by PAULI's theory, and $K^0 \rightarrow \gamma + \pi$ is prohibited by PAULI, but not by baryon/lepton.

In brief, one law and multiple law theories can make conflicting predictions on unseen data; they are not generally equivalent.

4.2. *A second puzzle and its resolution*

The beginning of Section 4 mentioned one puzzle: PAULI consistently yielded single-conservation-law solutions where BR-3 and particle physics practice yielded solutions based on multiple conservation laws. This puzzle was resolved by discovering that single law solutions were to be expected *mathematically*, so that PAULI, which is based on a systematic, mathematical formulation of the search space, should not find anything other than single laws.

However, this mathematical resolution of the puzzle immediately raised a second puzzle which the closing statement in our physics article [9] expressed thus: "It might be worthwhile to reconcile this theorem with the multiplicity of phenomenological quantum properties". In other words, what criteria could lead, under a purely data-driven regime, to a *justification* of the multiple conservation laws found in particle physics phenomenology?

This latter question motivated a follow-up article [7] that examined and rejected two such criteria: *optimism* (all unseen reactions can occur) and *pessimism* (all unseen reactions are impossible) because they are easily shown to fail to provide the needed justification. A third, *minimax* criterion, according to which one seeks to *minimize* the *maximum* quantum number, does lead to a theoretical justification, since a single conservation law might involve, for example, a maximum quantum number of 3, whereas multiple laws might lead to a maximum quantum number of 2. However, even though minimax is a powerful concept in optimization, a justification of physics laws based on it seems somewhat esoteric and *ad hoc*.

The fourth and final justification examined in the cited follow-up article leads to the seemingly most satisfactory resolution, and is based on the following preliminary observation: If the number of observed, linearly independent reactions equals or exceeds the number of particles, then elementary matrix algebra indicates that *no* conservation law can exist, that is, no law can rule in the observed reactions and rule out the unobserved ones. [4] In such a case, one must either seek alternative selection rules not based on conservation of a quantum number, or perform *divide-and-conquer*.

The divide-and-conquer approach implies dividing all the observed reactions into two or more groups which need not be disjoint, such that within each group, the number of observed reactions is less than the number of particles. One then finds a conservation law for each group separately. Since no single law will cover all the reactions, the result is that every group's conservation law will find exceptions among some reactions outside that group. This is precisely the situation in particle physics practice, in which some reactions fail to conserve one quantum number while conserving all the others.

The resolution of this second puzzle, then, is that one can show, via our machine discovery aided theorem and some further analysis, that the multiple conservation laws of particle physics phenomenology are *mathematically necessary* whenever the observed reactions become numerous relative to the number of particles. Curiously, this constitutes a top-down justification of the state reached in a partly bottom-up manner by physics phenomenology.

5. Generic views of the discovery task

Most research on scientific discovery has been concerned with the relation of discovery processes to *general* problem solving [2]. Our recent research, in contrast, has been overtly preoccupied with the relation of a given discovery task to problem solving *in science*. That is, we seek generalization *not* within general problem solving, *but within science* [8]. We proceed thus not from lack of ambition, but in the belief that a body of computationally oriented theory about science will yield different, and perhaps crisper, results than will an analogous theory about much broader phenomena. These research goals are explicated elsewhere under the organizing concept of *generic task*

[4] Because a matrix R of observed reactions will be nonsingular and invertible, so that the only solution to the equation $RP = 0$, where P is the matrix of particle quantum numbers, will be $P = 0$, which means that no unobserved reaction can fail to conserve the quantum number.

of scientific discovery [6]. Previously, Valdés-Pérez, Zytkow, and Simon [10] showed that the current discovery task lies within a formally defined generic category that they called *scientific model building as search in matrix spaces*. Here I will analyze briefly some broader, but less formal, generic aspects of the discovery task.

PAULI addresses the task of providing a theoretical basis to distinguish the possible reactions that particles undergo from the impossible reactions. In an abstract sense, this task is common throughout science: to find or postulate a feature that distinguishes two classes of objects. For example, a developmental biologist searches for developmental characteristics that distinguish a mutant organism from the wild-type (normal) organism. A psychiatrist looks for eye-movement patterns that discriminate between psychotics and normal subjects. A physiologist seeks features to classify diseased and healthy cells. All of these tasks are somewhat analogous to the concept learning task in machine learning. However, all these scientific tasks emphasize *finding* or *postulating* discriminatory features rather than selecting from a known list.

As in concept learning, a correct but uninteresting theoretical explanation of the observed and unobserved reactions consists of a trivial disjunction of the observed reactions, meaning that all other reactions are prohibited. A more interesting and useful partial explanation was attempted by the physicist Abraham Pais, who pointed out that, for the reactions that led to the discovery of strangeness, particles are generally produced in pairs and never as single particles [3]. This observational pattern of *associated production* could serve as a partial discriminating feature between possible and impossible reactions. However, particle .physicists eventually followed a more theoretical and complete approach by postulating unseen, conserved properties (quantum numbers) such that any hypothetical reaction that violated conservation was deemed impossible. This approach is, of course, the one followed by BR-3 and PAULI.

So, although the task of inventing features to discriminate between two classes is broadly generic throughout science, the detailed solution strategies may be highly particular, as in the approach followed by particle physics in this case.

6. Lessons

The lesson of this paper for artificial intelligence is that machine discovery based on heuristic search does lead to new findings in science, even in sciences of such celebrity and theory density as particle physics. In this instance, the new finding is a theorem which *was enabled by machine discovery* in the sense that observations of the invariant behavior of a machine discovery program directly led to the theorem's conjecture; its proof used representational techniques borrowed from the design of the program itself. Follow-up work then addressed—and resolved—the question of why physics practice seemingly conflicts with the theorem.

It is important to document such results for AI audiences, in order to falsify the premise of an otherwise powerful argument: a dearth of machine discoveries implies that machine discovery research is not addressing the essential parts of the discovery process in science.

Acknowledgements

The author thanks Sakir Kocabas for discussions about BR-3 and about the comparative outputs of BR-3 and PAULI. Yuval Ne'eman guided us through the "particle zoo" and offered encouragement. Michael Erdmann carried out the formal proof of the parsimony theorem. Clark Glymour, Peter Jansen, Aurora Pérez, and Herbert Simon commented valuably on drafts of this note, as did the anonymous reviewers.

This work was supported by a grant #IRI-9421656 from the (USA) National Science Foundation, by the Center for Light Microscope Imaging and Biotechnology, and by the W.M. Keck Center for Advanced Training in Computational Biology.

References

[1] S. Kocabas, Conflict resolution as discovery in particle physics, *Mach. Learn.* **6** (1991) 277-309.

[2] P. Langley, H.A. Simon, G. Bradshaw and J.M. Zytkow, *Scientific Discovery: Computational Explorations of the Creative Processes* (MIT Press, Cambridge, MA, 1987).

[3] Y. Ne'eman and Y. Kirsh, *The Particle Hunters* (Cambridge University Press, New York, 1986).

[4] R.E. Valdés-Pérez, Algebraic reasoning about reactions: discovery of conserved properties in particle physics, *Mach. Learn.* **17** (1994) 47-68.

[5] R.E. Valdés-Pérez, Conjecturing hidden entities via simplicity and conservation laws: machine discovery in chemistry, *Artif. Intell.* **65** (1994) 247-280.

[6] R.E. Valdés-Pérez, Generic tasks of scientific discovery, in: *Working Notes of the Spring Symposium on Systematic Methods of Scientific Discovery*, AAAI Technical Reports (1995).

[7] R.E. Valdés-Pérez, On the justification of multiple selection rules of conservation in particle physics phenomenology, *Comput. Phys. Commun.* **94** (1996) 25-30.

[8] R.E. Valdés-Pérez, Computer science research on scientific discovery, *Knowledge Eng. Rev.* (to appear).

[9] R.E. Valdés-Pérez and M. Erdmann, Systematic induction and parsimony of phenomenological conservation laws, *Comput. Phys. Commun.* **83** (1994) 171-180.

[10] R.E. Valdés-Pérez, J.M. Zytkow and H.A. Simon, Scientific model-building as search in matrix spaces, in: *Proceedings AAAI-93*, Washington, DC (1993) 472-478.