

LEARNING ABOUT MOMENTUM CONSERVATION

Jude W. Shavlik

Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801

ABSTRACT

A system has been designed in which *explanation-based learning* is applied to classical mechanics. An overview of the fully-implemented system and an example run by it are presented. From the example the system acquires a general formula describing how the momentum of a collection of objects changes over time. The derivation of this formula is prompted by the analysis of a human's solution to a simple two-body, one-dimensional collision. The human's solution is based on the principle of momentum conservation. The acquired formula serves to solve a three-dimensional collision in the presence of external forces, a situation where momentum is *not* conserved.

1 INTRODUCTION

Explanation-based learning [DeJong81] is a knowledge acquisition method that utilizes deep domain models. This type of learning involves generalizing a single problem solution into a form that can be later used to solve conceptually similar problems. The generalization process is driven by the *explanation* of why the solution worked. It is the deep knowledge about the domain that allows the explanation to be developed and then extended. This approach to learning has much in common with [Mitchell85], [Silver83], and [Winston83]. See [DeJong85] for a full discussion.

A system has been designed and implemented in which explanation-based learning is applied to classical mechanics. Initially this system is capable of performing many of the mathematical manipulations expected of a college freshman and is provided a representation of Newton's laws. Physical concepts taught in a first semester physics course are to be acquired; hence the name of the system, **Physics 101**. Newton's laws - which constitute the initial domain model - suffice to solve all problems in classical mechanics, but the general principles that are consequences of Newton's laws are interesting for their elegance as well as their ability to greatly simplify the solution process. Since the system's physical knowledge rests on the strong foundation of Newton's laws, only its mathematical sophistication will limit the physical concepts it can acquire.

System Overview

The operation of the system can be seen in figure 1. After a physical situation is described and a problem is posed, **Physics 101** attempts to solve the problem, reporting its solution if one can be found. When its knowledge is not adequate to solve a problem, the system requests a solution from the human user. It then verifies the correctness of the answer (relative to its current knowledge level). The human-provided solution need not fully specify the solution steps; some "reading between the lines" is possible. If the system cannot verify the human's solution, or if an erroneous step was detected, more details are requested. Once a satisfactory answer to the posed problem is obtained, the system determines why the

human's solution worked and attempts to extend the solution technique. Any resulting generalizations are recorded and can be used to solve future problems.

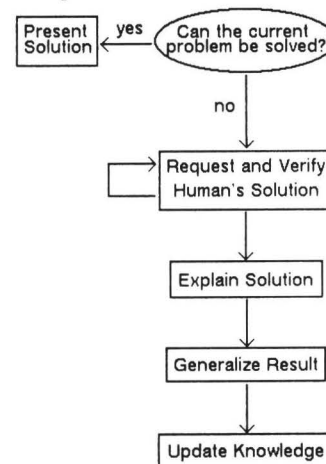


Figure 1. The Operation of Physics 101

Section II contains an example of this process. Conservation of momentum is used (by a human) to solve a one-dimensional, two-body collision problem that the system could not solve. The human's solution is verified and then generalized, during which the system discovers that momentum is only conserved in the absence of any net external force. A general equation describing how external forces affect momentum is derived by the system. A second problem, which involves three bodies under the influence of an external force, has been solved by **Physics 101** using this generalized result.

Explaining Human Solutions

The system classifies human solution steps into one of four categories:

- (1) A known formula could have been used; $force = mass \times acceleration$ is an example of this type.
- (2) New variables can be defined in order to shorten later expressions. A formula such as $momentum = mass \times velocity$ would fall into this class.
- (3) Equations can be algebraic variants of previous steps. The replacement of variables by their values also falls into this category.
- (4) The user can specify an equation that states a relationship among known variables, yet the system knows of no algebraically equivalent formula. These steps require full justification, which the system does by using its abilities to symbolically reason with calculus. Only the equations falling in this category are candidates for generalization.

Representing Physical Situations

Physical situations are represented as *worlds*, which comprise a number of *objects*. Objects have four measurable attributes; mass, position, velocity, and acceleration. All but mass are time-dependent, three-dimensional vectors. The relationships among these attributes, with respect to symbolic differentiation and integration, are known to the system.

Newton's second and third laws are also known to the system. (Newton's first law is a special case of his second law.) His second law says that an object's acceleration is determined by the net force acting on the object, divided by its mass. The net force can be decomposed into two components; the external forces (due to sources outside the physical system), and the internal forces (due to each of the other objects within the system). These inter-object forces are constrained by Newton's third law, which says that every action has an equal and opposite reaction.

When solving physics problems it often proves useful to represent the instantaneous state of a world at two separate times (e.g., the initial and final states). For this reason all variables in the world can have known values at two specific times, as well as a value (which may also be unknown) for all time. A constraint system [Shavlik84] maintains the known algebraic relationships among these values. Several of these values - for object attributes or any of the forces - are set by the user to characterize a physical situation.

II ILLUSTRATIVE EXAMPLE

In the one-dimensional problem shown in figure 2, there are two objects moving in free space, without the influence of any external forces. (Nothing is known about the forces between the two objects. Besides their mutual gravitational attraction, there could be a long-range electrical interaction and a very complicated interaction during the collision.) In the initial state (**state A**) the first object is moving toward the second, which is stationary. Some time later (**state B**) the first object is recoiling from the resulting collision. The task is to determine the velocity of the second object after the collision.

As there is insufficient information for the constraint subsystem to determine this velocity, **Physics 101** attempts to determine an expression for object two's velocity as an explicit function of time. The resulting expression could then be evaluated in state B. Progress past the last line in figure 3 can not be made though, as the time-dependence of the inter-object force is not known. (All of the calculus in figures 3 and 5 is generated by **Physics 101**. Several minor intermediate expressions have been excised, for the sake of brevity.)

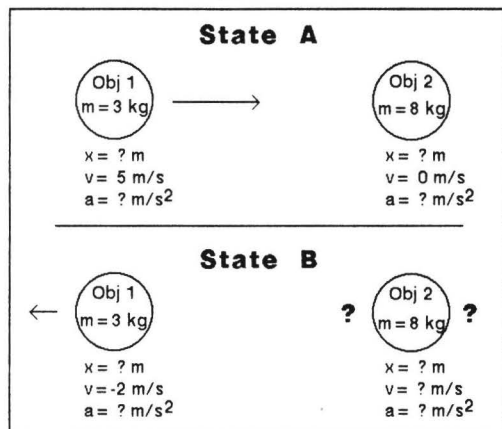


Figure 2. A Two-Body Collision Problem

$$\begin{aligned}
 \text{velocity}_{\text{obj2},x}(t) &= \int \text{acceleration}_{\text{obj2},x}(t) dt \\
 &= \int (\text{force}_{\text{net,obj2},x}(t) / \text{mass}_{\text{obj2}}) dt \\
 &= 1 / \text{mass}_{\text{obj2}} \int \text{force}_{\text{net,obj2},x}(t) dt \\
 &= 1 / \text{mass}_{\text{obj2}} \int (\text{force}_{\text{external,obj2},x}(t) + \text{force}_{\text{internal,obj2},x}(t)) dt \\
 &= 1 / \text{mass}_{\text{obj2}} \int (\text{force}_{\text{external,obj2},x}(t) + \text{force}_{\text{obj1,obj2},x}(t)) dt \\
 &= 1 / \text{mass}_{\text{obj2}} \int (\text{force}_{\text{external,obj2},x}(t) - \text{force}_{\text{obj2,obj1},x}(t)) dt \\
 &= -1 / \text{mass}_{\text{obj2}} \int \text{force}_{\text{obj2,obj1},x}(t) dt
 \end{aligned}$$

Figure 3. The Computer's Failed Solution Attempt

At this point the system requests a solution from the user. The solution provided can be seen in figure 4. Without explicitly stating it, the user took advantage of the principle of conservation of momentum, as the momentum of the world at two different times was equated. After that, various algebraic manipulations lead to the answer. In order to accept the solution, the system has to verify each of the steps in this solution.

The last three steps are easily verified, as they are simple algebraic manipulations. The hard part is determining a justification for the first equation in figure 4. **Physics 101** must ascertain that this equation is consistent with its physical knowledge. Since the two sides of the equation only differ as to the state in which they are evaluated, an attempt is made to determine a time-dependent expression describing the general form of one side of the equation. Figure 5 presents the calculus involved in determining how the general expression explicitly depends on time. This calculation can go to completion because the unknown force that was an obstacle in figure 3 is now cancelled by the other inter-object force. Even though the values of these two forces are unknown, Newton's third law constrains them so that they cancel. Since the expression is constant, its value in any two states can be equated. The initial step has been validated.

Next, the system attempts to generalize the initial equation in the human's solution. In the initial equation of figure 4, four variables are used to determine the value of object two's velocity. The system ascertains that the first object's velocity produces the inter-object force that cancels the force that had been an obstacle in figure 3. **Physics 101** also determines that the two mass terms are needed for this additive cancel to occur. Now that the system has explained why the human's attempt worked, it can generalize the technique to situations with more than two objects. Any object in a world may exert a force on any other object. So to cancel the total inter-object force in the general case, a $\text{mass} \times \text{velocity}$ contribution must come from every object in the world. **Physics 101** produces the left-hand side of equation 1, and then uses its knowledge of physics and calculus to arrive at the right-

$$\begin{aligned}
 &\text{mass}_{\text{obj1,stateA}} \text{velocity}_{\text{obj1,stateA},x} \\
 &\quad + \text{mass}_{\text{obj2,stateA}} \text{velocity}_{\text{obj2,stateA},x} \\
 &= \text{mass}_{\text{obj1,stateB}} \text{velocity}_{\text{obj1,stateB},x} \\
 &\quad + \text{mass}_{\text{obj2,stateB}} \text{velocity}_{\text{obj2,stateB},x} \\
 3 \text{ kg } 5 \text{ m/s} &= 3 \text{ kg } -2 \text{ m/s} + 8 \text{ kg } \text{velocity}_{\text{obj2,stateB},x} \\
 15 \text{ kg m/s} &= -6 \text{ kg m/s} + 8 \text{ kg } \text{velocity}_{\text{obj2,stateB},x} \\
 \text{velocity}_{\text{obj2,stateB},x} &= 2.63 \text{ m/s}
 \end{aligned}$$

Figure 4. The Human's Solution

$$\begin{aligned}
& \text{mass}_{\text{obj1}} \text{velocity}_{\text{obj1},x}(t) + \text{mass}_{\text{obj2}} \text{velocity}_{\text{obj2},x}(t) \\
&= \text{mass}_{\text{obj1}} \int \text{acceleration}_{\text{obj1},x}(t) dt + \text{mass}_{\text{obj2}} \int \text{acceleration}_{\text{obj2},x}(t) dt \\
&= \text{mass}_{\text{obj1}} \int (\text{force}_{\text{net,obj1},x}(t) / \text{mass}_{\text{obj1}}) dt + \text{mass}_{\text{obj2}} \int (\text{force}_{\text{net,obj2},x}(t) / \text{mass}_{\text{obj2}}) dt \\
&= \int \text{force}_{\text{net,obj1},x}(t) dt + \int \text{force}_{\text{net,obj2},x}(t) dt \\
&= \int (\text{force}_{\text{external,obj1},x}(t) + \text{force}_{\text{internal,obj1},x}(t)) dt + \int (\text{force}_{\text{external,obj2},x}(t) + \text{force}_{\text{internal,obj2},x}(t)) dt \\
&= \int (\text{force}_{\text{external,obj1},x}(t) + \text{force}_{\text{obj2,obj1},x}(t)) dt + \int (\text{force}_{\text{external,obj2},x}(t) + \text{force}_{\text{obj1,obj2},x}(t)) dt \\
&= \int (\text{force}_{\text{external,obj1},x}(t) - \text{force}_{\text{obj1,obj2},x}(t)) dt + \int (\text{force}_{\text{external,obj2},x}(t) + \text{force}_{\text{obj1,obj2},x}(t)) dt \\
&= \int (\text{force}_{\text{external,obj1},x}(t) + \text{force}_{\text{external,obj2},x}(t) + \text{force}_{\text{obj1,obj2},x}(t) - \text{force}_{\text{obj1,obj2},x}(t)) dt \\
&= \int (\text{force}_{\text{external,obj1},x}(t) + \text{force}_{\text{external,obj2},x}(t)) dt \\
&= \int (0 \text{ kg m} / \text{s}^2 + 0 \text{ kg m} / \text{s}^2) dt \\
&= \text{constant}_1 \text{ kg m} / \text{s}
\end{aligned}$$

Figure 5. Verifying the First Step in the Human Solution

hand side. In the general case, all the internal forces cancel but the external ones remain.

$$\sum_{i=1}^N \text{mass}_i \text{velocity}_i = \int \sum_{i=1}^N \text{force}_{\text{external},i} dt \quad (1)$$

This formula says: *The total momentum of the objects in a world is determined by the integral of the sum of the external forces on those objects.*

III CONCLUSION

By analyzing a worked example, **Physics 101** is able to derive a formula describing the temporal evolution of the momentum of any arbitrary collection of objects. This formula can be used to solve a collection of complicated collision problems.

Others [Forbus83], [Langley81], [Larkin81] have investigated the learning of physical concepts. Larkin's **ABLE** system learns how to algebraically apply principles of physics. Langley's work involves the analysis of multiple data measurements and does not utilize deep knowledge about the physical world. Forbus and Gentner consider learning *qualitative* physics by analogical reasoning. Unlike any of these systems, **Physics 101** reasons with calculus to explore the implications of Newton's laws.

There are several possible long-term applications of this research. From the viewpoint of the development of expert problem solvers, the techniques outlined here could be used to "train" programs in domains that are based on the selection and manipulation of formulae. Their analogs of Newton's laws could be encoded and then sample solutions to various problems could be presented, much in the way technical textbooks are organized. This system should also prove applicable to cognitive modelling and intelligent computer aided instruction, as it is able to recognize incorrect applications of physical principles (e.g., assuming momentum is conserved in the presence of an external gravitational field).

ACKNOWLEDGEMENTS

This work greatly benefited from discussions with my advisor, Professor Gerald DeJong.

REFERENCES

- [DeJong81] G. DeJong, "Generalizations Based on Explanations," *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, Vancouver, B.C., Canada, 1981, 67-70.
- [DeJong85] G. DeJong, R. Mooney, P. O'Rourke, S. Rajamoney, A. Segre and J. Shavlik, "A Review of Explanation Based Learning," Technical Report, Coordinated Science Laboratory, University of Illinois, Urbana, IL, forthcoming in 1985.
- [Forbus83] K. Forbus and D. Gentner, "Learning Physical Domains: Towards a Theoretical Framework," *Proceedings of the 1983 International Machine Learning Workshop*, Urbana, IL, June 1983, 198-202.
- [Langley81] P. Langley, "Data-Driven Discovery of Physical Laws," *Cognitive Science* 5, 1 (1981), 31-54.
- [Larkin81] J. H. Larkin, "Enriching Formal Knowledge: A Model for Learning to Solve Textbook Physics Problems," in *Cognitive Skills and their Acquisition*, J. R. Anderson (ed.), Lawrence Erlbaum, Hillsdale, NJ, 1981, 311-334.
- [Mitchell85] T. Mitchell, S. Mahadevan and L. Steinberg, "LEAP: A Learning Apprentice for VSLI Design," LCSR-TR-64, Rutgers University, New Brunswick, NJ, January 1985.
- [Shavlik84] J. W. Shavlik, "Building and Processing Constraint Networks for Vector Algebra," Working Paper 61, Coordinated Science Laboratory, University of Illinois, Urbana, IL, December 1984.
- [Silver83] B. Silver, "Learning Equation Solving Methods from Worked Examples," *Proceedings of the 1983 International Machine Learning Workshop*, Urbana, IL, June 1983, 99-104.
- [Winston83] P. H. Winston, T. O. Binford and M. Lowry, "Learning Physical Descriptions from Functional Definitions, Examples, and Precedents," AI Memo No. 679, MIT, Cambridge, MA, January 1983.