SCIENTIFIC DISCOVERY AND CREATIVE REASONING WITH DIAGRAMS

Peter C-H. Cheng
ESRC Centre for Research on Development, Instruction and Training
Department of Psychology
University of Nottingham
University Park, Nottingham NG7 2RD
United Kingdom
email: pcc@psychology.nottingham.ac.uk

Herbert A. Simon
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213, USA.
email: has@a.gp.cs.cmu.edu

(Sept. 1993)

To appear in
S. Smith, T. Ward, & R. Finke (Eds.),
The Creative Cognition Approach
Cambridge,MA: MIT Press. (forthcoming).

1 INTRODUCTION

Scientific discovery is a highly creative human endeavor well worth studying as an example of creative cognition. There is substantial research interest in the nature of the processes of scientific discovery in cognitive science, many empirical investigations have been undertaken, and numerous computational models have been constructed. Langley *et al.* (1987) have argued that scientific discovery can be viewed as problem solving by heuristic search and in a series of simulation programs have shown how laws can be discovered inductively under this conception. Cheng (1992a) reviews the many different computational discovery systems that now exist, and Shrager and Langley (1990) and Zytkow (1992) contain papers describing a broad selection of the recent and current computational research. The empirical work has studied how subjects make discoveries in different simulated discovery environments; for example, Klahr & Dunbar (1988), Qin & Simon (1990) and Schunn and Klahr (1992). Gorman (1992) considers the empirical studies from the perspective of falsification in discovery.

Although much has been learned about the processes of scientific discovery there are aspects that still remain to be examined. One of these is the role of multiple representations and diagrammatic representations in discovery. Shepard (1978, 1988) describes many examples of creative imagery from the history of science and considers some of the developmental factors that may be associated with the power of imagery. Existing computational models and empirical work have tended to focus on discoveries made with a single non-diagrammatic knowledge representation despite the ubiquity of diagrams in scientific reasoning and discovery and the wide acknowledgement of the need to examine the role of diagrammatic representations (e.g., Langley at al., 1987; Shrager & Langley, 1990; Cheng, 1992a). Research on multiple representations and diagrams in discovery is still rather novel, even though there has been significant progress on reasoning with diagrammatic representations more generally (e.g., Novak, 1977; Larkin & Simon, 1987; Larkin, 1989; Koedinger & Anderson, 1990; Tabachneck, 1992; Qin 1992; Narayanan, 1992). With respect to discovery, Shrager (1990) has produced a computer program that uses diagrammatic and propositional representations as two different *modalities* that are grounded on sensory experience, to demonstrate his theory of *common sense perception*. Theory formation in this view employs processes that work within and between the different modalities to compare and combine information in the different representations.

Our own work has investigated the role of discovery with diagrams in early physics. Cheng (1992b) demonstrates the computational benefits that Galileo achieved in his kinematic discoveries with diagrams over a more conventional approach using algebraic equations. Galileo made many important discoveries but we noted that he never arrived at the principle of conservation of momentum. We were curious as to whether this was because, dealing only with gravitational force on a single body, he never had to introduce the concept of mass; or because the algebraic and diagrammatic methods he employed didn't lead readily to the statement of momentum conservation. To satisfy our curiosity, we examined the reasoning methods of Huygens and his contemporaries, and found that their diagrammatic methods were quite similar to Galileo's, but also quite conducive to inferring the law of conservation of momentum. Cheng and Simon (1992) show how it was easier for those early physicists to discover the conservation of momentum using diagrams than to induce the law directly from numerical data.

This chapter describes our continuing investigation of the role of discovery with diagrams in early physics by presenting a system that performs law induction using one-dimensional diagrams. The system is called HUYGENS and whilst we do not like to claim that HUYGENS gives a specific description of the thought methods of the scientist, Huygens, it does at least at a qualitative level, give some indication of the way in which diagrammatic representations shaped thought early in the 17th Century.

The chapter will first give some examples of the creative use of diagrams in scientific discovery and reasoning from the history of science. Then as a further step towards understanding the role of diagrams in discovery, the basis for one-dimensional diagrammatic law induction is considered. The regularity spotters, operators and heuristics used by the HUYGENS diagrammatic law induction program are considered and one of its simulations of a discovery is described. The limitations of one-dimensional diagrammatic law induction are discussed and the possibilities of diagrammatic discovery are considered more generally.

2 REASONING AND DISCOVERY WITH DIAGRAMS

There are several reasons why diagrams are useful to the scientist trying to make discoveries. At a general level, multiple representation gives the scientist the possibility of switching to an alternate representation when an impasse is reached. Diagrams have certain properties that often give them advantages over other representations. Diagrams can facilitate problem solving by reducing the amount of computation required to search for relevant information and by reducing the effort required for recognizing appropriate operators or inference rules (Larkin & Simon, 1987). Further, diagrams permit perceptual inferences to be made that are easier than can be done with their more difficult logical counterparts (Larkin, 1989).

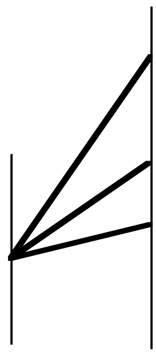


Figure 1 Quickest Descent Problem - Proposition 30

Consider Galileo's kinematic discoveries described in his *Two New Sciences* (1974). All the propositions described here are from the third section of that book. The 30th proposition considers the following situation: Given inclined planes, ramps, running between two parallel horizontal lines, Figure 1, what is the inclination of the plane that will give the quickest time of descent? This will be called the "quickest descent" problem. The acceleration of the ball increases with greater inclination. However, the total distance to be travelled also increases. The problem is to find when the two effects combined to produce a minimum. Galileo considered this problem after he had discovered his law of free fall, the second proposition, but he did not approach the problem by trying to apply the law directly to the given situation. Rather, he used a diagrammatic approach involving the sixth proposition, known as "Galileo's theorem" (Drake, 1978).

Galileo's theorem is concerned with the times of descent along inclined planes within a vertical circle. Figure 2 shows inclined planes running from points on the circumference of a circle to the lowest point in the circle, and planes running to the circumference from the highest point. The times of descent on all such inclined planes are equal, which Galileo proves from the law of free fall and the geometry of circles (by relating the law's square relation between distance and time to the mathematical description of circles as the sum of two squares). The way Galileo solved the quickest descent time problem was to combine Figures 1 and 2, giving

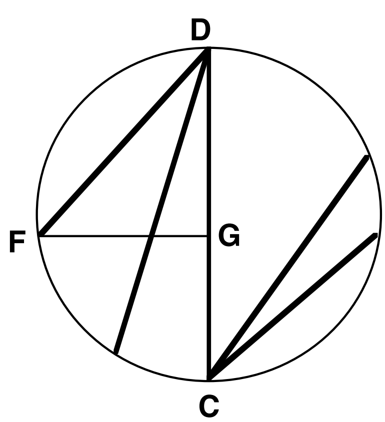


Figure 2 Galileo's Theorem - Proposition 6

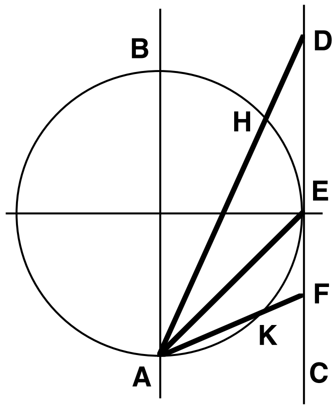


Figure 3 Solution to Quickest Descent Problem

Figure 3. The circle has been drawn with a radius equal to the distance between the parallel lines with its bottom at the point of intersection of all the inclined planes. From Galileo's theorem it immediately follows that that the times of descent along the inclined planes within the circle are equal; $t_{EA} = t_{HA} = t_{KA}$. However, as the inclined planes *DA* and *FA* are longer than *HA* and *KA*, their times of descent must be greater than times for *HA* and *KA*. Therefore, the descent time between the verticals along the incline *EA* is the shortest; the inclination is 45 degrees.

Although it is quite feasible to do, there is no record that Galileo attempted to solve the problem by the direct application of his various laws of motion using a more conventional mathematical approach. However, it is interesting to consider what is involved under that approach as the contrast shows the ingenuity of Galileo's diagrammatic solution. A general inclined plane is represented by the triangle *DAC*, in Figure 4. The time of descent down the plane is given by the fifth proposition on uniform velocity motion from the *Two New Sciences*;

$$t_{DA} = d_{DA} / V_{DA} , \qquad \ldots (1)$$

where t, d, and V are time, distance and mean speed, respectively. The mean speed is equal to the maximum speed at the end of the descent, $Vmax$, divided by two, thus:

$$t_{DA} = 2 \, . \, d_{DA} / Vmax_{DA} \, . \qquad \ldots (2)$$

Galileo's law of free fall relates the time, t, for descents down an inclined plane to the

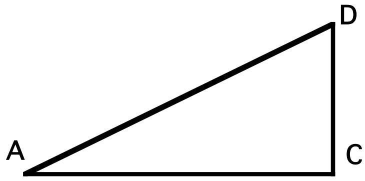


Figure 4 Simple Inclined Plane

vertical distance, h, travelled;

$$h_{ST} / h_{SY} = t_{ST}^2 / t_{SY}^2 , \qquad \ldots (3)$$

where the subscripts refer to different parts of the descent from rest. The terminal velocity at the end of an inclined plane is thus proportionate to the square root of the height of the plane; so:

$$t_{DA} = 2 \, . \, d_{DA} / \sqrt{h_{DC}} , \qquad \ldots (4)$$

where h is the height. From Pythagoras's theorem, the length of the inclined plane can be replaced by its height and the horizontal component of its length, giving:

$$t_{DA} = 2 \, . \, \sqrt{(d_{AC}^2 + h_{DC}^2)} / \sqrt{h_{DC}} . \qquad \ldots (5)$$

The final step is to find the relation between h_{DC} and d_{AC} that yields the minimum value of t_{DA}. There are various ways to do this but none are straightforward and simple; for example, it is possible to reasoning directly about the values of numerator and denominator of Equation 5 as h_{DC} varies. The solution is when the angle of the plane is 45 degrees, that is $h_{DC} = d_{AC}$.

The conventional mathematical approach is more complex than the diagrammatic approach, because the bulk of the reasoning is centered around the abstract equations expressing kinematic laws. Four laws were combined to find Equation 5, which requires further difficult reasoning to determine the minimum time. Under the diagrammatic approach, the minimum time was found by spotting the line that did not extend beyond the circumference of a circle; a simple piece of perceptual reasoning. Cheng (1992b) has modelled various Galilean kinematic discoveries under the diagrammatic and the conventional mathematical approaches, including this example, and has shown that the diagrammatic approach often requires less computation than the conventional approach.

In the following sections the discovery of more complex laws using diagrams are considered. First, we consider the diagrams that the early physicists may have used to discover the conservation of momentum.

3 CONSERVATION LAWS AS ONE-DIMENSIONAL DIAGRAMS

The momentum of a body is the product of its mass and velocity. For two bodies colliding in one-dimension the momentum conservation law is usually written as an equation:

$$m_1.U_1 + m_2.U_2 = m_1.V_1 + m_2.V_2 , \qquad \ldots (6)$$

where, m_1 and m_2 are the masses of the two bodies, U_1 and U_2 are their velocities before collision, and V_1 and V_2 their velocities after collision. If energy is also conserved in the collisions then the following energy conservation equation holds simultaneously;

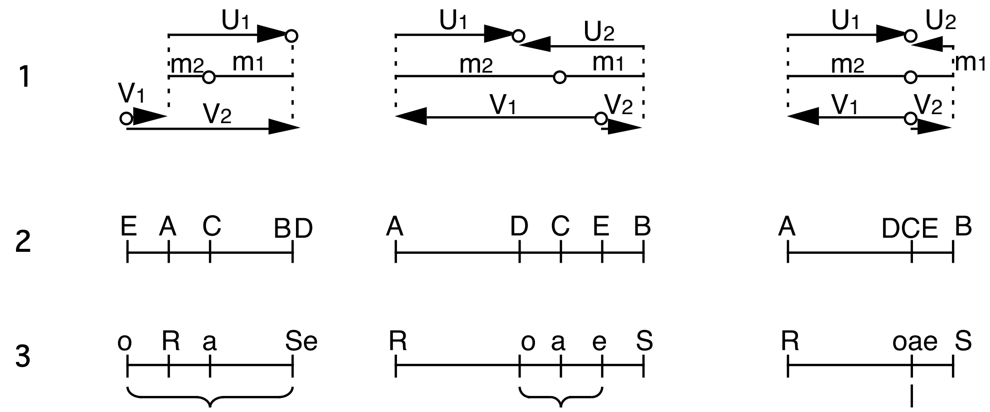


Figure 5 One-Dimensional Momentum Conservation Diagrams

$$m_1.U_1^2 + m_2.U_2^2 = m_1.V_1^2 + m_2.V_2^2 \ . \quad ..(7)$$

However, it is possible to encode the two laws in a diagrammatic form that embraces both equations; three examples are shown in the first row of Figure 5. The top line of the left diagram shows that body 1 comes in from the left and impacts body two, which is initially stationary. Body 1 is bigger than body 2, as shown by the middle line. The bottom line shows that both bodies travel off to the right. The speeds of the bodies are in proportion to the lengths of their respective lines. The center diagram shows a collision where the bodies approach from opposite directions with equal speeds, but depart with different speeds in opposite directions, because the masses have different magnitudes. The right diagram shows that when the ratio of the initial speeds, U_1/U_2 , is equal to the inverse of the ratio of their masses, m_2/m_1, then the final speeds for each body is the same as it was before collision, but the bodies' directions are reversed.

Notice that the total lengths for initial and final velocity lines are equal, that is:

$$U_1\text{-}U_2 = V_2\text{-}V_1 \quad (8)$$

This relation can be simply derived from Equations 6 and 7. The lines for the masses are drawn end to end with their total length equal to the length of the U_1-U_2 line. The structures of the laws are such that the ends of the lines, shown by the small circles, must always lie in a straight vertical or diagonal line. Huygens and Wren presented similar diagrams to the Royal Society of London when they first described the law of the conservation of momentum (Hutton, Shaw & Pearson, 1804). In Huygens's diagrams, row 2 in Figure 5, *A* and *B* are the two bodies, their velocities before impact are denoted by the lines *AD* and *BD*, the velocities after impact, by *EA* and *EB*; and the masses of *A* and *B* by *BC* and *AC* [*sic*], respectively. In the diagrams, the lengths of *DC* and *CE* are always equal. Wren's diagrams, row 3 in Figure 5, are essentially the same, except for differences in notation and the fact that Wren explicitly states that the diagrams

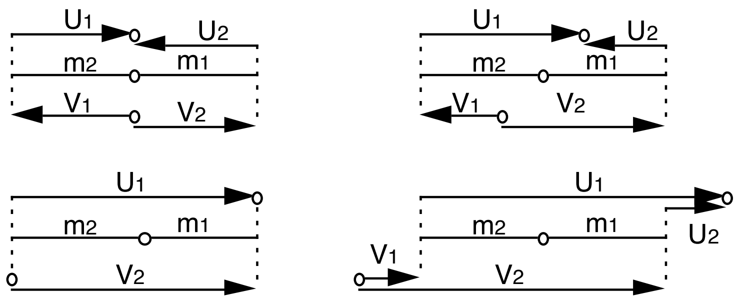


Figure 6 Configurations of Collisions

are reversible; that is, either *Ro* and *So* or *Re* and *Se* can indicate the initial velocities, with *eR* and *eS* or *oR* and *oS* the final velocities, respectively.

Some of the points made above about the potential advantages of diagrammatic representations can be seen with these diagrams. For example, different configurations of collisions may be distinguished when velocities of the balls have different signs, are equal, unequal, or are zero. How many different configurations exist? This is not a simple problem working directly from Equations 6 and 7, but using the diagrams it is simple to answer. For instance, when the masses are equal there are four configurations, see Figure 6. Huygens and Wren provided such series of diagrams in their expositions (Hutton, Shaw & Pearson, 1804). Further, it is possible to do simple quantitative reasoning and reasoning about extreme cases with these diagrams, using the fact that the small circles in diagrams must always lie in a straight line, vertically or diagonally (row 1, Figure 5). Figure 7A shows what happens when a stationary ball (m_1) is hit by another (m_2) as the ratio of their masses tends to infinity. The maximum speed (V_2) that the stationary ball can attain after the collision is just two times that of the first ball (U_1). Similarly, Figure 7B shows what happens when two balls approach with equal speeds from opposite directions as the mass of one tends toward infinity compared to the other. The maximum speed of the smaller ball after impact (V_2) is three times its initial speed (U_2).

There are several different ways in which the conservation of momentum may have been discovered. One possibility is that the law was discovered in a theoretical fashion by derivation from energy conservation and considerations of invariance of motion relative to Galilean transformation of coordinates (Barbour, 1989). A second possibility is that the law was

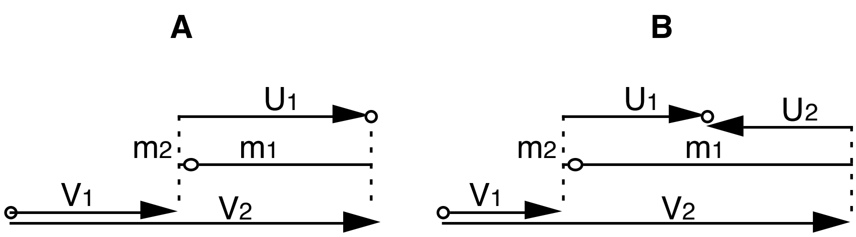


Figure 7 Collisions in which the Ratio of Mass Tends to Infinity

induced from sets of numerical data gathered in experiments. This method of discovery is quite feasible, as demonstrated by several systems that have modelled the discovery of the principle of momentum conservation: for example, BACON.5 (Langley et al., 1987) and ABACUS (Falkenhainer and Michalski, 1986). These systems find the law by searching a space of algebraic relations among the variables using the numerical data. A third possibility also considers that the law was induced from experimental data, but using diagrams rather than working directly with algebraic equations. Cheng & Simon (1992) have modelled the discovery in this manner and have shown that it would have been easier for Huygens and Wren to have made the discovery using diagrams than with the conventional mathematical methods available at the time. In the next section, general law induction with one-dimensional diagrams is considered.

4 THE HUYGENS SYSTEM

Cheng and Simon's (1992) system simulated the discovery of the conservation of momentum using one-dimensional diagrams, but some of its heuristics were specific to the structure of the conservation law problem: specifically, the heuristics to deal with the two independent variables (V_1 and V_2). Here a more general system for one-dimensional diagrammatic law induction is described. The system is called HUYGENS.

It considers one diagram for each *set* of experimental data obtained from a single experimental *test* (Cheng, 1991). Variables are represented as line segments on the number line. The length of a line segment, or *line* for short, is in proportion to the magnitude of the value of its variable and the orientation, positioning and relative sizes of lines encode different algebraic relations between the variables. HUYGENS operates by constructing diagrams from sets of data using diagrammatic operators. A *group* of diagrams is generated by applying the same sequence of operators to the sets of data. Relations that really exist are manifested as patterns common to all the diagrams in a group and it is the job of regularity spotters to find such patterns. When a pattern is found, an algebraic law is simply inferred from the regularity and the particular operators used to generate the diagrams. HUYGENS employs cycles of regularity spotting and operator application. The various operators, regularity spotters and heuristics that are used in the making of discoveries are described.

4.1 Operators

The job of the diagrammatic operators is to construct or modify diagrams so that they encode different relations between the variables. Tables 1 and 2 show the diagrammatic operators that are required in the cases of discovery modelled so far. Other operators may be needed in other cases of discovery. Various conventions are used when drawing the diagrams. The number line is assumed to increase towards the right. The lines in the diagrams are considered to lie on the number line, but for clarity they are drawn with vertical separation.

When appropriate, the origin of a line is indicated by an 'o' and its point of interest (interest point) by a 'x'. An interest point is the end of a line; its position depends on the magnitude of the variable, and is determined by the data. Where appropriate, construction points are also marked by a '|'; a construction point being an intermediate point identified and used by operators in the construction of a line. The lines are labeled with symbols for their variables. HUYGENS's internal representations of the lines are in the form of triples of numbers for the origin, the interest point and the construction point, if applicable. These triples are stored together in lists corresponding to diagrams, and diagrams are themselves stored together in lists to form groups. The overall organization of the data is equivalent to the structure of the diagrams as if they were drawn on paper.

An internal representation that is even closer to real diagrams (marks on paper) could be employed. For example, the lines and markers for the different points could be held in a bit map. The operators could be modified to generate lines in such a format, and the regularity spotters modified to look for points on such lines. However, the operators and spotters would be functionally equivalent to those currently possessed by HUYGENS but would make the implementation more complex, without contributing anything further to the present analysis. HUYGENS does nevertheless use a simple routine to convert the sets of number triples into real lines for display in its output trace.

Values of various attributes for each variable or line are recorded by HUYGENS. The values are initially given as part of the input to the system and the attributes considered are: (i) the *type*, whether a line stands for an independent or dependent variable; (ii) the *property* of the line, for example, velocity or temperature; and, (iii) the *status* of the line, whether or not it is a new line just generated by an operator. The properties of new lines are inherited from the original lines in a manner appropriate to each operator employed.

Table 1 shows four elementary operators. The PLOT operator takes a variable, X, and draws a line with a length in proportion to the magnitude of X. All the origins of all the lines drawn by PLOT are at the same arbitrary position on the number line. The opposite end of a line is its interest point, because its location depends on the magnitude of X. The ADD operator finds the sum of two variables, X_1 and X_2, that have the same dimensionality by redrawing their lines end to end. The total length may be considered as a new variable or term, X', representing the sum of X_1 and X_2. The SUBTRACT operator finds the difference between the two variables, X_1 and X_2, having the same dimensionality, by redrawing the X_2 line with its interest point coinciding with the interest point of X_1. The length of the new line in between the origins, X', equals the difference between X_1 and X_2. The NEGATE operation makes the inverse of a variable by redrawing its line with the interest point on the opposite side of the origin. ADD, SUBTRACT and NEGATE assign to the new lines they generate the same property as the original lines.

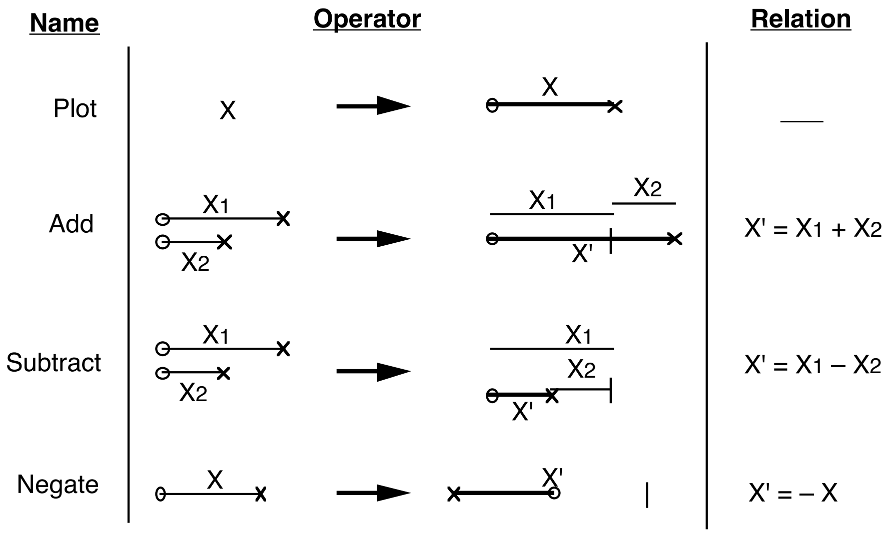


Table 1 Elementary Operators

The five normalization operators, Table 2, are used to incorporate variables standing for more than one type of property into a diagram. They all require a pair of variables, Y_1 and Y_2, for one type of property, P_y, and a single variable, X, for some other property, P_x. X is taken as the datum against which the other variables are standardized. NORMALIZE is the simplest and works by redrawing Y_1 with its length equal to X, and then redrawing Y_2 in proportion to Y_1. A new term, X', for property P_x is obtained with a value equal to X times Y_2/Y_1. The other

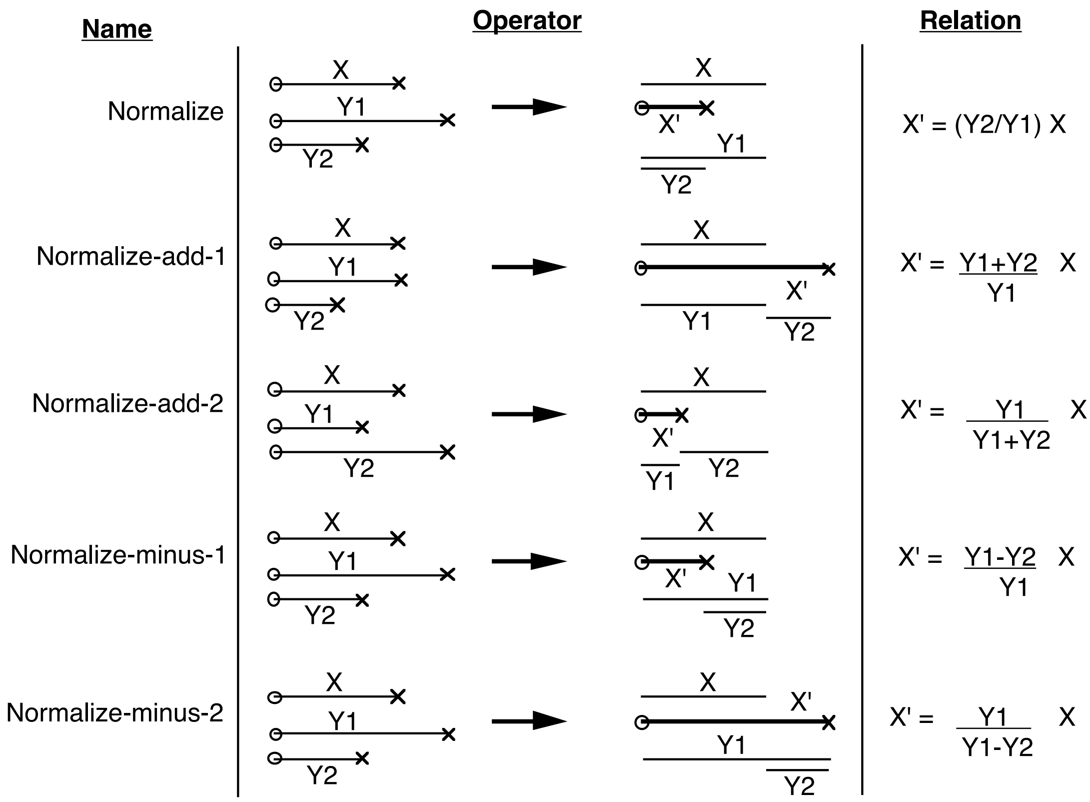


Table 2 Normalization Operators

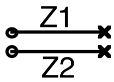
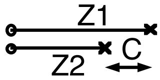
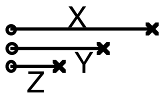
Name	Pattern	Relation
Equal		Z1 = Z2
Constant		C = Z1 - Z2
Mean		Y = (X + Z)/2

Table 3 Precise Regularity Spotters

normalization operators are similar to NORMALIZE but they take the sum (or difference) between Y_1 and Y_2 and use Y_1+Y_2 (Y_1-Y_2) or Y_1 alone as the basis for standardization of X; see Table 2.

The application of an operator usually reduces, and never increases, the number of lines in a diagram. Each operator generates a new line that stands for a variable or a term that expresses a potential relation among the variables for the old lines, as shown, for example, in the third column in Tables 1 and 2. Groups of diagrams are generated by different sequences of operators. A sequence of operators that successfully encodes a regularity in the data will produce the same pattern in every diagram in its group. Determining whether any relations really exist by looking for such patterns is the job of the regularity spotters, which are considered next.

4.2 Regularity Spotters

The regularity spotters look for patterns that are common to every diagram in a group. The spotters examine the interest points in the diagrams, because these reflect any patterns that are present in the data. Table 3 and Table 4 present regularity spotters required in the cases of discovery so far modelled.

The precise spotters, Table 3, look for *exact* patterns in diagrams. The EQUAL regularity spotter identifies when two lines are equal in length in every diagram in a group. The relation inferred is that Z_1 is equal to Z_2. The CONSTANT regularity spotter identifies when the difference between two variables is a constant for all the members of a group. The MEAN spotter identifies when a variable has a magnitude equal to the mean of the two others, for every diagram in a group. When a regularity is found, the relations can be written as an equation with appropriate symbols for the variables, as shown in column 3.

Table 4 presents two more regularity spotters. They are *relative* spotters that seek patterns based on the relative lengths of lines, rather than their exact lengths. The BETWEEN spotter identifies when a variable always has a magnitude between those of two others, in any

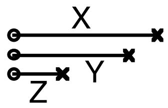
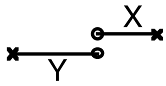
Name	Pattern	Relation
Between		X > Y > Z or X < Y < Z
Negative		X > 0 & Y < 0 or X < 0 & Y > 0

Table 4 Relative Regularity Spotters

order. The NEGATIVE spotter identifies when a pair of lines has interest points on opposite sides of a common origin. The main use of the relative spotters is in heuristics that suggest when different operators may be appropriate to consider (discussed in the next section).

The regularity spotters look for patterns across pairs or triplets of lines in each diagram in a group. How the pairs and triplets of lines are chosen for consideration and how the regularity spotters suggest the use of certain operators are considered.

4.3 Heuristics

HUYGENS needs heuristics to limit the size of the search space of diagrams. There are various places where appropriate heuristics are employed; they are considered in turn. Examples of these heuristics can be seen in the simulation of a discovery that is described below.

First, when HUYGENS is looking for a common pattern within all the diagrams from a particular group, pairs and triplets of lines have to be identified for the regularity spotters. Rather than consider all combinations of pairs and triplets, HUYGENS uses a strategy that takes into account the type, property and status attributes of the lines. Different combinations of the attributes are considered by the SELECTION-BY-ATTRIBUTE heuristic: (i) status and property; (ii) type and property; and (iii) property alone. For each combination HUYGENS finds the total number of pairs and triplets that can be generated when the lines have matching (non-nil) values for the specified attributes. For example, when the type and property combination is considered, only lines in a diagram standing for the same property and that are exclusively independent or dependent variables are taken as potential pairs or triplets. The set of pairs and triplets chosen is the one that has the greatest number of pairs plus triplets not exceeding the number of lines in the diagram.

All three attribute combinations consider the property of the lines, because it is not valid to compare quantities for different properties. The status and property combination is included because it is worthwhile looking for patterns among new lines that have just been generated.

The type and property combination is included because it is sensible to consider first the independent lines together, before complicating matters with the inclusion of the dependent lines too. The property attribute alone is included in case the other combinations do not yield any pairs or triplets.

Before applying the regularity spotters to the pairs and triplets of lines, HUYGENS attempts to focus the spotters on those pairs and triplets whose lines have the same property as the dependent line, using the DEPENDENT-PROPERTY-FOCUS heuristic. The simple rationale being that it seems sensible initially to seek relations for lines with the same property as the dependent line(s) before making things more complex by considering lines with other properties.

Regularity spotters are then applied to the selected pairs and triplets of lines to find any patterns that are common to all the diagrams in a group. The PREFER-PRECISE-SPOTTERS heuristic gives priority to the precise regularity spotters, because they identify exact relations between the lines. If on a particular cycle HUYGENS finds a group of diagrams with a precise regularity and another with a relative regularity, then only the precise group is considered in the next cycle. Depending on the regularity found and whether there are further lines to consider HUYGENS will apply different operators; the SPOTTER-OPERATOR-MATCH heuristic. When a precise regularity is found and there are no other lines left in the diagrams to be considered, HUYGENS states a law was found. However, when a precise regularity is found and lines remain, HUYGENS will apply the normalization operators, using the regularity found as the basis for standardization.

For example, when the CONSTANT regularity spotter is true and there are remaining lines, HUYGENS will use the distance found (labeled C in Table 3) as the basis against which to compare the remaining lines when applying the normalization operators. The two relative spotters are considered if no precise regularities are found. When a relative regularity is found, the SPOTTER-OPERATOR-MATCH heuristic suggests that particular operators may be appropriate. For example, suppose the BETWEEN regularity holds for three variables, A, B and C, so that B is always between A and C. Now, if other variables need to be considered, NORMALIZE-ADD-2 is a good choice, with the difference between A and C as the basis for standardization (i.e., labeled X in Table 2). The new term formed by this normalization operation will always be between the ends of A and C, so it is likely to be directly related to B.

The possibility exists that no regularity will be found in a group of diagrams. In such cases HUYGENS will apply the ADD and SUBTRACT operators to the selected pairs as a default. This is the DEFAULT-OPERATORS heuristic.

Heursitic	Action
SELECTION-BY-ATTRIBUTE	Select the most reasonable number of pairs and triplets of lines by considering those that match under different combinations attributes.
DEPENDENT-PROPERTY-FOCUS	Choose pairs and triplets of lines that have the same property as the independent line(s).
PREFER-PRECISE-SPOTTERS	Give regularities found by the precise spotters a higher priority than those found by the relative spotters.
SPOTTER-OPERATOR-MATCH	Specifies particular operator(s) when a particular regularity has been found.
DEFAULT-OPERATORS	When no regularities have been found try the ADD and SUBTRACT operators on the selected pairs.

Table 5 HUYGENS's Heuristics

Table 5 provides a summary of HUYGENS's heuristics. We now consider how the diagrammatic operators, regularity spotters and heuristics work together to make discoveries.

4.4 Simulation of a Discovery

In this subsection HUYGENS simulation of the discovery of the conservation of momentum is considered.

Symbol	Property	Type	Case 1	Case 2	Case 3
m_1	mass	independent	2	1	3
m_2	mass	independent	1	2	1
U_1	velocity	independent	3	3	3
U_2	velocity	independent	0	-3	1
V_1	velocity	dependent	1	-5	2
V_2	velocity	dependent	4	1	4

Table 6 Momentum Conservation Law Data

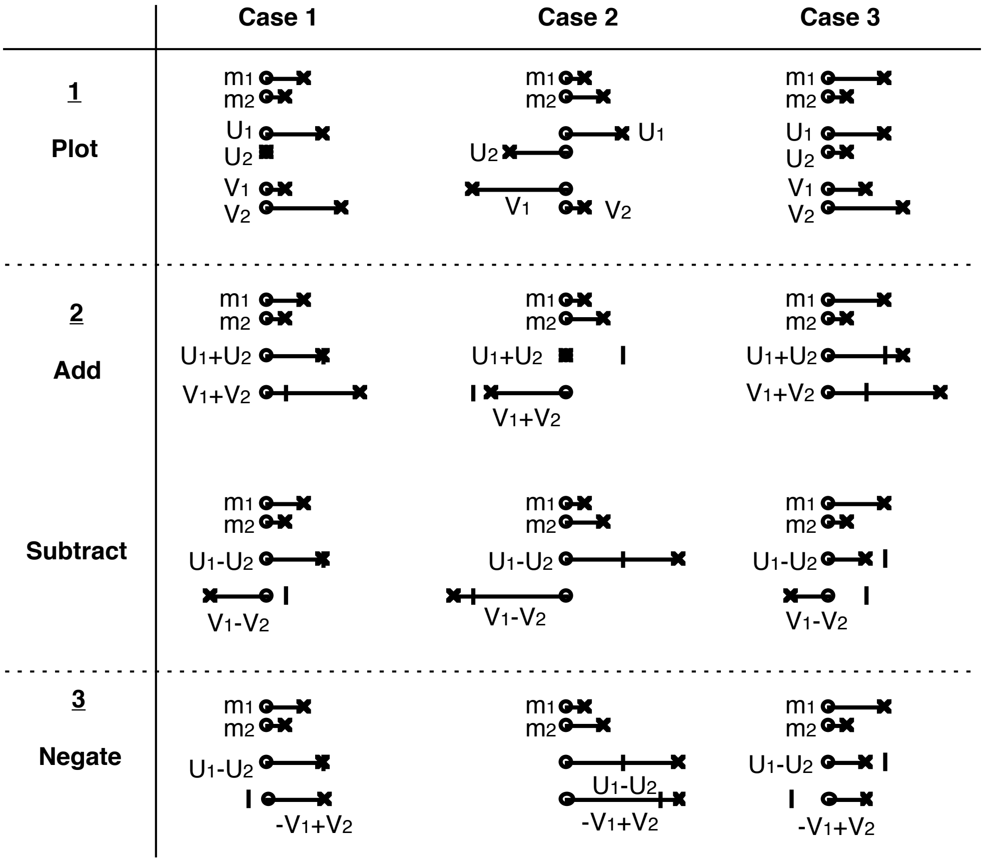


Figure 8 Beginning of the Momentum Law Discovery

The simulation of the discovery involved five cycles of regularity spotting and operator application. In the first cycle, the pairs of variables that were chosen by the SELECTION-BY-ATTRIBUTE heuristic were (m_1, m_2), (U_1, U_2) and (V_1, V_2), but these were narrowed down to the second and third pairs by the DEPENDENT-PROPERTY-FOCUS heuristic, because V_1 and V_2 are dependent and U_1 and U_2 share the same property. None of the regularity spotters found any common patterns for the pairs in the three diagrams, so HUYGENS resorted to the DEFAULT-OPERATORS heuristic. The lines generated by the ADD and SUBTRACT operators are shown in middle two rows in Figure 8, respectively. In the second cycle, the two new lines in each diagram were chosen by the SELECTION-BY-ATTRIBUTE heuristic. The NEGATIVE spotter was found to hold for the new lines generated by the SUBTRACT operator, the interest points of those lines in all three diagrams are on opposite sides of their shared origin. The SPOTTER-OPERATOR-MATCH then forced HUYGENS to generate a new set of diagrams by applying the NEGATE operator, as shown in row 3 of Figure 8. In the third cycle, the EQUAL regularity was found to hold for the new pair of velocity lines, so HUYGENS was able to find the relation among the velocities given by Equation 8.

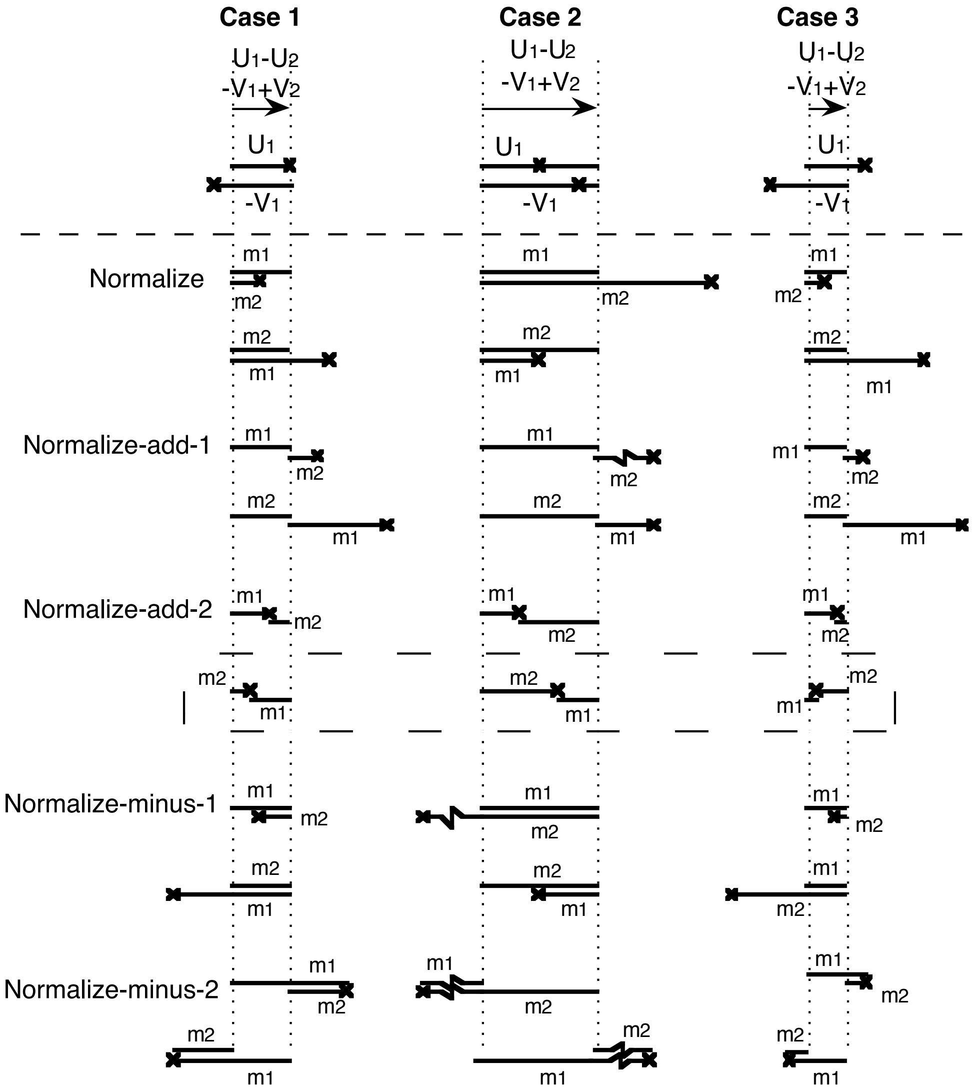


Figure 9 Momentum Law Discovery Continued

However, in the fourth cycle, it is realized that there were mass lines still to consider so HUYGENS had not finished. The common distance of U_1-U_2 and V_2-V_1 was used as the basis for the application of the normalization operators and the construction points were taken as the new interest points. In the fourth cycle the five normalization operators are applied, as shown in Figure 9 (the U_1-U_2 and the V_1-V_2 lines are shown at the top of each column, rather than repeat them next to each pair of mass lines). Each normalization operator was applied twice because either m_1 or m_2 may be associated with U_1. In the fifth and final cycle, HUYGENS found that the only regularity to hold was MEAN, for the groups produced by NORMALIZE-ADD-2, with m_2 associated with U_1. The distance from the interest point on the mass line to the ends of the lines for U_1 and -V_1 were equal. Thus, HUYGENS had found a relation between the velocities and masses, corresponding to the following equation:

$$\frac{m_2}{m_1+m_2}(U_1\text{-}U_2) = \frac{U_1 - V_1}{2} \quad . \qquad \ldots (9)$$

With a little algebraic manipulation and given Equation 8, previously found by HUYGENS, Equation 9 can be easily shown to be equivalent to the momentum conservation law, Equation 6.

HUYGENS has also been successfully run on data for a simplified version of Black's law on the temperatures of liquids.

4.5 Benefits and Limitations

Many systems already exist that successfully perform quantitative law induction (e.g., BACON, Langley, et al., 1987; ABACUS, Falkenhainer & Michalski, 1986). HUYGENS differs in that it searches the space of one-dimensional diagrams for regularities, rather than the space of algebraic terms. This alternative representation motivates HUYGENS's use of a greater range of operators and regularity spotters than the other systems. Some of the operators and spotters consider triplets of lines on each cycle, whereas previous systems usually consider relations between pairs of variables at any one time. There is a corresponding reduction in the size of HUYGENS's search space, because it can sometimes combine three lines into a single new line, when the other systems require two cycles to do the same. A manifestation of this can be seen in the need for the fifth version of BACON (Langley et al., 1987) to employ high level search control heuristics, which assume symmetry and conservation in the data, to improve the efficiency with which more complex laws are found.

A present deficit of HUYGENS is its inability to cope simply with laws with power terms. To deal with squares and square-roots Galileo and other early physicists employed the two-dimensional geometry of conic sections and circles (see Cheng, 1992b). A one-dimensional technique to find power laws is to take successive differences between the values of dependent variables until a linear series is obtained. However, this technique only works when the independent variable increases as an arithmetic progression and the index of the power law is a positive integer. Further, substantial amounts of drawing and redrawing are required.

Other abilities that HUYGENS will need are the means to recognize and define terms for intrinsic properties and to cope with noisy data; abilities that BACON possesses. Noisy data can be handled naturally under the diagrammatic approach by using error bars, similar to those used when plotting experimental errors on graphs. Briefly, the spotters would be allowed to match interest points so long as they fall within an interval, centered on the target position, whose length is some given percentage of the line of concern. Schemes for coping with noisy data can be devised using this technique.

5 BEYOND ONE-DIMENSIONAL DIAGRAMMATIC LAW INDUCTION

Huygens's and Wren's diagrams encode the fairly complex momentum conservation laws in a deceptively simple manner, and in a way that makes it easy to reason qualitatively and quantitatively about different collisions (see Figures 6 and 7). The HUYGENS system has successfully modelled the diagrammatic discovery of the laws. Some of the consequences for the study of scientific discovery and for understanding the processes of creative cognition will be considered.

The HUYGENS system demonstrates how one-dimensional diagrammatic law induction can be performed in a manner consistent with the view that scientific discovery is problem solving characterized by heuristic search. In themselves, HUYGENS's operators, regularity spotters and heuristics are quite straightforward, but in combination they provide an effective model of law induction. This is a further step towards understanding the processes of scientific discovery, which although seemingly mysterious can be rationally understood. In particular, HUYGENS provides further computational evidence for the view that switching back and forth between representations is an effective way to enhance creativity. From given numerical data, HUYGENS switches to a space of diagrams in its search for regularities by looking for patterns in the diagrams. When patterns have been found, the regularities are simply transformed back into equations. The change to diagrammatic representations permits different operators, regularity spotters and heuristics to be employed that are more effective than those used in the direct search of a space of algebraic terms. The reasons for this are that diagrammatic representations often encode or index information in ways that help to reduced search (Larkin & Simon, 1987) and that they enable perceptual inferences to be made (Larkin, 1989) in problem solving .

The work with HUYGENS shows how discoveries that use diagrammatic representations can be modelled. The wider implication is that creative discovery with visual imagery, such as the episodes described by Shepard (1978, 1988), may be amenable to computational modelling in a similar fashion. One-dimensional diagrammatic law induction can be characterized as heuristic problem solving, but it is the subject of current research whether the same is true for diagrammatic discovery more generally, including one-dimensional deduction and two-dimensional induction and deduction. However, there are good reasons to think that the problem solving view will cover other forms of diagrammatic discovery.

It seems possible to model two dimensional diagrammatic discovery. Consider Koedinger and Anderson's (1990) *diagram configuration model* of expert problem solving. The central idea is that chunks of perceptual knowledge are stored as *diagrammatic configuration schemas*. Each schema holds various pieces of information, including: a *configuration*, in the form of a diagram of the given situation; a *whole-statement*, which expresses the main theorem or idea of the schema; *part-statements* or *part-properties*, which indicate important features of the configuration; and, *sufficient conditions* or *ways-to-prove*, which define sets of part-statements that are sufficient to prove the whole-statement. The way problem solving proceeds with the schemas involves mapping the configuration into a suitable part of the problem diagram and searching for part-statements that fulfill the sufficient conditions. If one set of the sufficient-conditions is complete, then the whole-statement is applicable to the problem. Successive applications of different schema may fill in all the steps required for a complete problem solution.

Now, it is possible to view the discoveries in Galileo's *Two New Sciences* as cases of problem solving under the Koedinger and Anderson model, treating each of the 38 propositions as a diagrammatic configuration schema. For example, how might Galileo have known that the sixth proposition, Figure 2, was the right one to use for the problem of least time of descent? Quite simply, he could have found the proposition 6 by searching through the space of all the diagrams encoding particular laws and theorems, which he had already discovered. Problem 6 can be reformulated as diagrammatic configuration schema. The configuration diagram is Figure 2. The whole-statement says that the times of descent down all the inclined planes are equal. The part-statements would include facts, such as, (i) each inclined plane runs from the circumference to the bottom of the circle, (ii) each inclined plane runs from the top of the circle to the circumference, and (iii) the circle is vertical. Sets of sufficient-conditions are (i) and (iii), or (ii) and (iii). Hence, the schema for proposition 6 can be applied to the quickest descent time problem, Figure 1, by adding a vertical circle to the problem diagram, Figure 3, from which it is immediately seen that the first set of sufficient-conditions is satisfied.

New configuration schemas can be defined when a new proposition has been discovered using existing schemas. For example, a new configuration schema might be defined for proposition 30. Its configuration diagram would be a triangle representing an inclined plane. The whole-statement would say that this inclined plane covering a fixed horizontal distance has the minimum descent time. The part-statement would indicate that the angle of the plane is 45°, or that the vertical height of the plane is equal to its horizontal length. A sufficient condition in this case would be either part-statement.

The propositions of the *Two New Sciences* can therefore be recast as diagrammatic configuration schemas, but it remains to be seen whether the discovery of all 38 of Galileo's propositions can be modelled under this approach. However, it seems that two-dimensional diagrammatic discovery is possible and that it does fall within the problem solving as heuristic search paradigm.

6 CONCLUSIONS

Diagrams have an important role in scientific creativity, because their representational properties make them effective for problem solving and discovery. This chapter has considered examples of the creative use of diagrams in the history of science and has described the

HUYGENS discovery system, which uses one-dimensional diagrams to inductively discover laws. The possibility of discovering laws with two-dimensional diagrams has also been considered. This research and other investigations have typically focussed on the properties of diagrammatic representations in isolation or in comparison with other representations. However, scientific discovery probably does not usually rely on a single representation, whether diagrammatic or otherwise, but often involves the use of multiple representations in an integrated manner. Diagrams predominated in the Galilean examples, but they were not the only representational formalism he used. Galileo's knowledge of arithmetic and algebra were important in his discoveries, especially in the interpretation of the empirical data in his early work (Drake, 1987), and they clearly interact in a complementary way with the diagrammatic representations in his discoveries. Future research must also consider discovery with multiple representations; how the best representations for a given problem can be chosen from those available, and the ways in which different representations complement each other in problem solving and discovery.

ACKNOWLEDGEMENTS

PCHC was supported in part by a U.K. Science and Engineering Research Council postdoctoral fellowship, which was held at Carnegie Mellon University, and by the U.K. Economic and Social Research Council, which is funding the Centre for Research on Development, Instruction and Training. HAS was supported by the Defense Advanced Research Projects Agency, Department of Defense, ARPA Order 3597, monitored by the Air Force Avionics Laboratory under contract F33615-81-K-1539. Reproduction in whole or in part is permitted for any purpose of the United States Government. Approved for public release; distribution unlimited.

REFERENCES

Barbour, J. B. (1989). Absolute or Relative Motion: The Discovery of Dynamics; Volume 1. Cambridge: Cambridge University Press.

Cheng, P. C.-H. (1991). Modelling experiments in scientific discovery. In Proceeding of the 12th International Joint Conference on Artificial Intelligence (pp. 739-744). Mountain View, CA: Morgan Kaufmann.

Cheng, P. C.-H. (1992a). Approaches, models and issues in computational scientific discovery. In M. T. Keane & K. Gilhooly (Eds.), Advances in the Psychology of Thinking (pp. 203-236). Hemel Hempstead, Hertfordshire: Harvester-Wheatsheaf.

Cheng, P. C.-H. (1992b). Diagrammatic reasoning in scientific discovery: Modelling Galileo's Kinematic Diagrams. In H. Narayanan (Eds.), Working Notes of the AAAI Spring Symposium on Reasoning with Diagrammatic Representations (pp. 33-38). Stanford University, California.

Cheng, P. C.-H., & Simon, H. A. (1992). The right representation for discovery: Finding the conservation of momentum. In D. Sleeman & P. Edwards (Eds.), Machine Learning: Proceedings of the Ninth International Conference (ML92) (pp. 62-71). San Mateo, CA: Morgan Kaufmann.

Drake, S. (1978). Galileo at Work. Chicago: University of Chicago Press.

Falkenhainer, B. C., & Michalski, R. S. (1986). Integrating quantitative and qualitative discovery: the ABACUS system. Machine Learning, 1(4), 367-401.

Galileo (1974). Two new sciences (S. Drake, Trans.). Madison, Wisconsin: University of Wisconsin Press.

Gorman, M. E. (1992). Using experiments to determine the heuristic value of falsification. In M. T. Keane & K. Gilhooly (Eds.), Advances in the Psychology of Thinking (pp. 147-176). Hemel Hempstead, Hertfordshire, UK: Harvester-Wheatsheaf.

Hutton, C., Shaw, G., & Pearson, R. (1804). The Philosophical Transactions of the Royal Society of London. London: C&R Baldwin.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. Cognitive Science, 12, 1-48.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. Cognitive Science, 14, 511-550.

Langley, P., Simon, H. A., Bradshaw, G. L., & Zytkow, J. M. (1987). Scientific Discovery: Computational Explorations of the Creative Process. Cambridge, MA: MIT Press.

Larkin, J. H. (1989). Display-based Problem Solving. In D. Klahr & K. Kotovsky (Eds.), Complex Information Processing: The Impact of Herbert A. Simon (pp. 319-341). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65-99.

Narayanan, H. (Ed.). (1992). AAAI Spring Symposium on Reasoning with Diagrammatic Representations: Working Notes. Stanford University, California; March 25-27 1992.

Novak, G. S. (1977). Representations of knowledge in a program for solving physics problems. In Proceedings of the 5th International Joint Conference on Artificial Intelligence. Mountain View, CA: Morgan Kaufmann.

Qin, Y. (1992). From language to mental images to equations. Doctoral Thesis, Department of Psychology, Carnegie Mellon University.

Qin, Y., & Simon, H. A. (1990). Laboratory replication of scientific discovery processes. Cognitive Science, 14, 281-312.

Schunn, C. D., & Klahr, D. (1992). Complexity management in a discovery task. In Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, (pp. 177-182). Hillsdale, New Jersey: Lawrence Erlbaum.

Shepard, R. N. (1978). Externalization of mental images and the act of creation. In B. S. Randhawas & W. E. Coffman (Eds.), Visual Learning, Thinking, and Communication. New York: Academic Press.

Shepard, R. N. (1988). The imagination of the scientist. In K. Egan & D. Nadaner (Eds.), Imagination and Education. New York: Teachers College Press.

Shrager, J. (1990). Commonsense perception and the psychology of theory formation. In J. Shrager & P. Langley (Eds.), Computational Models of Scientific Discovery and Theory Formation (pp. 437-470). San Mateo, CA: Morgan Kaufmann.

Shrager, J., & Langley, P. (Eds.). (1990). Computational Models of Scientific Discovery and Theory Formation. San Mateo, CA: Morgan Kaufmann.

Tabachneck, H. (1992). Computational difference in mental representations: the effects of mode of data presentation on reasoning and understanding. Doctoral Thesis, Department of Psychology, Carnegie Mellon University.

Zytkow, J. (Ed.). (1992). Proceedings of the ML92 Workshop on Machine Discovery. Aberdeen, Scotland.