Components of a Cognitive Theory of Problem Solving and Discovery with Law Encoding Diagrams

Peter C-H. Cheng.

ESRC Centre for Research in Development, Instruction and Training
Department of Psychology, University of Nottingham
Nottingham, NG7 2RD, U.K.
peter.cheng@nottingham.ac.uk

Abstract

The representational components for a cognitive theory of Law Encoding Diagrams, LEDs, are considered in this paper. Examples of problem solving and scientific discovery with LEDs are discussed and used to introduce the four schemas that appear necessary to account for the many different forms of reasoning that are possible with this class of representational systems. Possible forms of processing with LEDs are illustrated by analysing the diagrammatic nature of Galileo's kinematics discoveries.

Introduction

Law Encoding Diagrams are an interesting class of diagrammatic representations, the study of which can further our general understanding of reasoning with diagrams. They are different to the classes of diagrammatic representations typically studied in this field, because of the special way in which they encode information. But, they are common in many scientific and technical domains.

For a given domain, a Law Encoding Diagram, LED, is a representational system, that captures the laws governing a particular class of phenomena using the internal geometrical, topological or spatial structure of its diagrams, such that each instantiation of a diagram represents one instance of the phenomenon or one case of the laws. LEDs have some interesting properties that appear to make them suitable for a wide range of cognitive activities in various domains. To date the investigations have identified classes of LEDs for different domains, characterised the kinds of cognitive activity that are supported by them, and computer based systems have been developed to support selected activities. In the history of science, some important discoveries were made using LEDs (Cheng 1996b; Cheng & Simon 1995). Problem solving and learning with LEDs have been empirically studied (Cheng 1996f). LEDs are being exploited for computer based instruction in science and engineering (Cheng 1996a) and systems using LEDs have demonstrated benefits over "conventional" programs (Cheng 1996d). We have also begun examining the utility of LEDs, and other diagrams, for the acquisition of knowledge in the development of expert systems, for complex engineering processes (Cheng 1996c).

This work on LEDs has mainly concentrated on the formal structure of particular systems of LEDs, although various general explanations of the potential utility and benefits of LEDs have been made. For learning and instruction, LEDs may be considered as representations at an intermediate level of abstraction, providing a bridge to the conceptual gulf between the abstract general laws of a domain and concrete examples of phenomena (Cheng 1996a). For the inductive discovery of laws in science, the nature of the operators, regularity spotters and heuristics are transformed by LEDs, in contrast to conventional representations, with a consequent reduction of the size of the problem search space (Cheng and Simon 1995). In more general terms, some of the benefits of LEDs may come from the multiple *function roles* they simultaneously support (Cheng 1996e).

However, there is no cognitive theory, or theories, of reasoning with LEDs. Such a theory is needed for various reasoning: to underpin computational models of problem solving and discovery with LEDs; to give an understanding of how learning occurs with LEDs that in turn will provide the basis for establishing effective instructional principles for LEDs; to provide a basis for the design of effective tools and techniques for diagrammatic knowledge acquisition with LEDs.

From the perspective of human problem solving as the symbol processing in the form of heuristic search (Newell and Simon 1972), the first steps towards a cognitive theory of LEDs are taken in this paper. The proposal is that various schemas are required for cognitive activities with LEDs and four schemas are hypothesized for different aspects of reasoning with LEDs. Examples of LEDs for the domain of particle collisions will first be given and then used to introduce the schemas for LEDs. Some of the cognitive processes necessary for problem solving using the schemas are then discussed in the context of

In Anderson, M. (ed) AAAI Technical Report on the 1997 Fall Symposium:
Reasoning with Diagrammatic Representations II. (Technical Report FS-97-02) pp.85-93, 1997.

discoveries in Galilean kinematics. Some discussion of related work rounds off the paper.

LEDs for Particle Collisions

Many systems of LEDs have been discovered and some new ones invented in the course of the studies of LEDs. In this section one class of LEDs for elastic particle collisions is described to demonstrate some of the properties LEDs in general and to show some of the forms reasoning that they support.

Figure 1 shows 4 examples of an LED called the 1DP diagram. Each diagram shows a single collision between two bodies travelling in a straight line. The initial velocities before impact, U_1 and U_2, of the two bodies are represented by the **U1** and **U2** arrows in the diagrams. Similarly, the final velocities, V_1 and V_2, are shown by **V1** and **V2**. The masses of the bodies, m_1 and m_2, are depicted by the lines **m1** and **m2**. In elastic collisions both momentum and energy are conserved and the 1DP diagram encodes both conservation laws when various constraints on the structure of the diagram are met. For instance, the tips of the **U1** and **U2** arrows are adjacent, the overall length of the **U1–U2** and **V1–V2** lines are equal (velocity difference rule), and the middle diagonal/vertical line intersects the ends of the lines for the six variables (diagonal rule). (For a full list of the constraints on the 1DP diagram see Cheng 1996b)

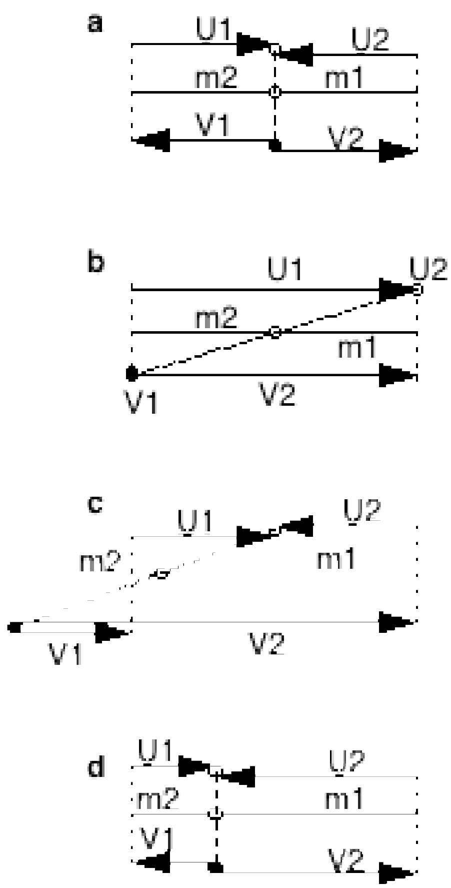

Figure 1. 1DP Diagrams

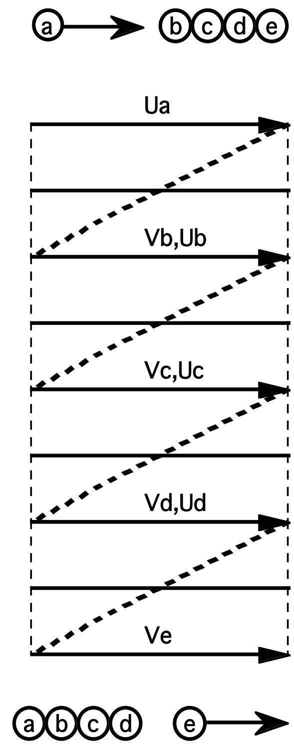

Figure 2. Composite LED

Figure 1a shows two bodies of equal mass approaching from opposite directions with equal initial speeds and rebounding with equal speeds. Figure 1b shows body-1 colliding with another stationary body of equal mass and transferring all its energy and momentum to the other. What is the result of a perfectly elastic planet colliding with a perfectly elastic pea, which approach each other from opposite directions with unit speed? This is a question that many undergraduate physicists get wrong. The answer is shown in Figure 1c, in which **m1** (planet) is very much greater than **m2** (pea) — the final velocity of the pea is three units in the opposite direction. The length of **V2**, and hence final speed of the pea, cannot be greater, without violating the constraints of this LED.

Qualitative and quantitative problem solving with single collisions can be done simply with the 1DP diagram (see Cheng 1996a for examples), but as is the case with LEDs in general, complex interactions can also be analysed. Figure 2 shows the collision of one moving body and four stationary bodies in a line, all with the same mass, as in

Slot	1DP diagram
Diagram-features	**U1**, **U2**, **V1**, **V2** arrows; **m1**, **m2** lines.
Diagram-constraints	E.g., **U1** and **U2** heads adjacent, **V1** and **V2** tails adjacent, **m1** and **m2** end to end, velocity difference rule, diagonal rule.
Domain-properties	U and V - initial and final velocities; m - mass; subscripts for each body.
Encoded-laws	$m_1 u_1 + m_2 u_2 = m_1 v_1 + m_2 v_2$; $\frac{1}{2} m_1 u_1^2 + \frac{1}{2} m_2 u_2^2 = \frac{1}{2} m_1 v_1^2 + \frac{1}{2} m_2 v_2^2$.
Property mappings	**U** and **V** arrow lengths and orientation give *U* and *V* velocities; lengths of **m** lines give relative mass, *m*.
Interpretation-rules	**m**s are relative — consider ratio m_1/m_2; **U-V** topology same as domain structure.
Construction/manipulation-heuristics	E.g., draw **U**s and **V**s first because **m**s are relative.
Special-cases	Default, Reflected collision, Planet and pea.

Table 1 LED-schema for the 1DP diagram

Newton's Cradle (the common executive toy). The interaction can be broken down into four separate collisions, between pairs of balls, each represented by one 1DP diagram like Figure 1b. This analysis shows that energy and momentum are conserved through the interaction, as V_e equals U_a. Further, the decomposition of four 1DP diagrams reflects the underlying process, which permits such a pair-wise analysis. The impulse travels through the bodies as a shock wave moving at the speed of sound, much greater than the translation of the bodies, in effect isolating each pair of bodies for the others they happen to be touching. See Cheng (1996a) for other examples of composite 1DP diagrams, including interactions between unequal masses.

Modified versions of the 1DP diagram have been devised for inelastic collisions and collisions in two dimensions. Other systems of LEDs exist for this domain and a computer based discovery learning environment that exploits these LEDs has been built and successfully evaluated (Cheng 1996d).

Schemas for LEDs

Four schemas for different aspects of reasoning with LEDs are hypothesized. This section introduces the schemas using the cases from the previous section as examples. The schemas were developed by considering LEDs across different domains with respect to (i) the logical differences between the types of knowledge encoded and (ii) the similarities among the types of reasoning that are conducted with LEDs.

It is hypothesized that two pairs of schemas (frames) are needed for the characterization of LEDs: *LED schemas* and *LED-instance schemas*; *composite-LED schemas* and *composite-LED-instance schemas*. The first pair are for single basic LEDs that capture the laws of a domain for a particular set of circumstances. The second pair are for composite LEDs comprising multiple LEDs in one diagram, which capture the laws of interactions in complex situations of a domain. Within each pair, the first schema specifies the general nature of an LED or composite-LED and the second schema stores information about particular instantiations of the LED or composite-LED. An instantiation of a LED (or composite-LED) is not simply a specialization of the basic LED (or composite-LED) schema. A LED schema contains the information needed to generate a particular LED and to interpret it in terms of the domain. However, it does not itself contain a diagram nor does it directly apply to a specific domain condition; that is the job of the LED-instance schemas (composite-LED-instance schemas). The details of the four schemas are considered in turn.

A basic LED schema contains information, facts, about certain phenomena or cases in a domain governed by particular laws or relations and it specifies how those laws are diagrammatically encoded and to be interpreted. The details of the LED schemas for the 1DP diagram are given in Table 1. There are eight slots in LED schemas:

Diagram-features: the geometric, spatial or topological features of relevance in the diagram.

Diagram-constraints: rules defining the structure of the diagrams, with particular reference to relations among the diagram features.

Domain-properties: the domain variables of interest.

Encoded-laws: the laws and relations that govern the domain.

Slots	Default	Planet and pea	Reflected Collision
Diagrams	Figure 1a	Figure 1c	Figure 1a, 1d
Diagram-features	vertical & horizontal symmetry.	diagonal (nearly) intersects left vertical.	“diagonal” is vertical.
Domain-conditions	u_2=-u_1, m_1=m_2, v_2=-v_1, v_1=-u_1, v_2=-u_2 .	u_2=-u_1, m_1>>m_2, v_2≈-3u_2, v_1≈u_1 .	v1=-u1, v2=-u2, m1/m2=|u2/u1| .
Interpretation	Simplest case.	Limiting case on masses.	Overall momentum is zero.

Table 2. Some LED-instance schemas

Property-mappings: mappings between domain variables and diagram features.

Interpretation-rules: how the diagram’s structure should be interpreted in terms of the encoded laws or the nature of the domain.

Construction/manipulation-heuristics: recommendations of how best to construct or vary the configuration of the diagram.

Special-cases: LED-instance schemas for notable exemplars of the domain or important diagrammatic configurations.

In Table 1, the contents of the slots should be obvious from the foregoing descriptions of the 1DP diagram, with the exception of the interpretation-rules slot. The note about the topology of the diagram refers to the fact that there is a consistency from the left to right positioning of **U** and **V** in the diagram to the physical location of the bodies, before and after collision.

An LED-instance-schema contains information about a single instantiation of an LED schema for default or special cases of the domain. Its various slots are:

Diagrams: one or more typical images of the LED, in the “mind’s eye” or as symbolic labels/pointers to them in an external medium.

Diagram-configuration: notable arrangements of diagram features.

Domain-conditions: the particular circumstances under which the case holds, as specific values of variables or relations among them.

Interpretation: interpretation in terms of the encoded laws or the nature of the domain.

Some LED-instance schemas for the 1DP diagram are given in Table 2. The ‘planet and pea problem’ and the ‘reflected collision’ examples illustrate how striking diagrammatic configurations often correspond to important cases in the domain.

The second pair of schemas are for composite-LEDs. Whereas the basic LED schemas were concerned with the relations among diagrammatic features and laws for particular phenomena, the composite-LED schemas deal with relations among multiple LEDs and laws governing more complex interactions among phenomena. The composite LED-schemas operate on the next level of complexity to the basic LED schemas, but the two pairs of schemas are similar in many ways.

The composite-LED schema with 1DP diagrams for multiple body collisions is given in Table 3. The slots of the composite-LED schemas mirror those of the basic LED schemas. They include:

Component-LEDs: LEDs used in the composite construction.

Composition-constraints: rules for assembling the composite diagram specifying the permitted arrangements of the LEDs.

Slots	Multiple Body Collisions
Component-LEDs	1DP diagrams
Composition-constraints	In successive 1DP diagrams, **a** and **b**, for a given **m1**, **V1a** and **U1b** may share the same arrow.
Domain-description	Multiple successive collisions in one dimension.
Encoded-interaction-laws	Independent pair-wise collisions.
Mapping-rules	One 1DP diagram for each interaction.
Interpretation-rules	Free (not shared) **U** and **V** arrows are the overall initial and final velocities, respectively.
Construction-heuristics	E.g., begin with the first impact.
Special-composites	Newton’s Cradle single, (see Cheng 1996b for others).

Table 3. Composite-LED schema.

Slots	Newton's Cradle: One-onto-four
Diagrams	Figure 2
Composite-features	Column of 1DP diagrams
Domain-conditions	All but one body initially stationary
Interpretation	Simple/default case

Table 4. Composite-LED-instance schema

Domain-description: specification of the target interactions of interest.

Encoded-interaction-laws: the laws and relations governing the target interactions within the domain.

Mapping-rules: relations of component diagrams to interactions in the domain.

Interpretation-rules: interpretation of the diagram structure in terms of the domain.

Construction-heuristics: recommendations of how best to assemble composite diagrams.

Special-composites: notable exemplars of the domain or cases that show important diagram configurations.

The structure of composite-LED-instance schemas are similar to basic LED-instance schemas given the similarity of the composite-LED and basic LED schemas. A composite-LED-instance schema is presented in Table 4. The slots of the these schemas are:

Diagrams: one or more images of typical diagrams of the composite LED, in the mind's eye or a pointer to them in some external medium.

Composite-features: notable diagrammatic properties of the composite.

Domain-conditions: interaction conditions under which the case holds.

Interpretation: interpretation in terms of the nature of the domain interactions.

That completes the introduction of the schemas for LEDs. The next section considers how they can explain some of the forms of reasoning using LEDs.

Processing Schemas for LEDs in Galilean Kinematics

Modelling processes of scientific discovery with diagrams is a somewhat neglected area in the diagrammatic reasoning and computational scientific discovery research communities. But such models are important, because of the novel and ingenious ways diagrams are used in discovery problem solving and the central role that diagrams play in science. LEDs have had a role in some important discoveries in the history of science (Cheng 1996b), including those of Galileo's *Two New Sciences*, TNS (Galileo 1638/1974) and Newton's *Principia*. Some models of discoveries with LEDs have been built, including the induction of the 1DP diagram (Cheng and Simon 1995) and the finding of selected propositions from the TNS, (Cheng 1992). However, thoroughgoing cognitive models of discoveries with LEDs has not seemed feasible, because no general theory existed. The proposed set of schemas for LEDs appears to fill this gap and holds out the prospect that coherent and general computation models can be developed. As a preliminary to this, this section considers the derivation of a representative set of propositions from the TNS. The aim is to further clarify the nature of the four schemas and to illustrate some of the forms of information processing that can be done with them.

We start with an example of an basic LED schema, which corresponds to the second corollary of Proposition II of Galileo's TNS. This theorem states that the ratio of distances covered by objects moving under uniform acceleration from rest is equal to the ratio of the squares of their respective times. For example, quadrupling the distance will take twice as long. The LED for Proposition II will be called the *Time-squared* LED, Figure 3. The *times* of descent from rest over two different *heights* are the domain-properties of this LED schema. The lines **T**, **X**, and **Y** are the diagram-features and the diagram-constraint fixes their relative lengths thus: **X/Y=T/X**. With the descents starting from the horizontal **S**, lines **T** and **Y** represent the two distances and their respective times are given by the lengths of **T** and **X**. Alternatively,

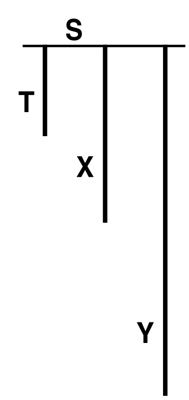

Figure 3. Time-squared LED

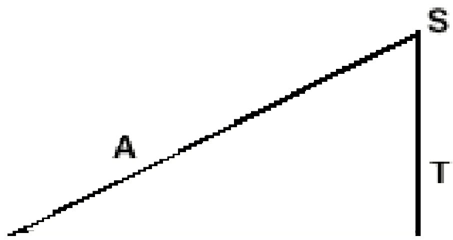

Figure 4. Incline-plane LED

the times to cover **T** and **Y** are given by the lengths of **X** and **Y**, respectively. Galileo used the same diagrammatic elements to represent two different domain variables (distance and time) and there are two alternative sets of mappings in the diagram for the time variables. These property-mappings are interesting for two reasons. First, they demonstrate the need for distinct diagram-feature, domain-property and property-mapping slots in the general structure of LED schemas. Second, it illustrates the ingenuity with which Galileo formulated this LED schema, which as we will see enabled later inferences based on this LED schema to be made simply and directly. (Similar examples are found in Newton's *Principia* [Cheng 1996b].) The encoded-law captured by the diagram is the times-square law, which can be illustrated by letting **T** (or **Y**) represent unit time, so from the given equation, distance, **Y** (**T**), is proportional to time squared, $\mathbf{X}^2$.

Although Galileo does not provide any special-cases/configurations, an obvious LED-instance schema is one in which the proportions of **T**:**X**:**Y** are 1:2:4, as happens to be the case in Figure 3. Typically, the proportions are not rational numbers.

The next example shows how a new LED schema can be generated by incremental refinement. The culmination of this line of reasoning is Proposition III, which concerns motions down an inclined plane. The proposition states that objects accelerating uniformly from rest down inclined planes of equal heights will have a ratio of times of descent equal to the ratio of the respective distances along the planes. Proposition III will be called the *Inclined-plane* LED schema, Figure 4 and Table 5. The important diagram-features are the diagonal, **A**, and the vertical, **T**, and the diagram-constraints specifies that the triangle must have a horizontal base and a right angle. Clearly the property-mappings are from the hypotenuse and vertical to the path on the inclined plane and a vertical drop.

The developed of this schema begins with the only assumption (postulate) in the TNS, the *Equal-speed-assumption*. This states that the terminal speeds are equal for descents down inclined planes of equal height but of different lengths, including the vertical (e.g., **SA**, **ST**). Initially the Inclined-plane LED only encodes this fact, Table 5, and in the derivation of Proposition III the LED schema is elaborated. Given that the mean speed is one half the final speed and that the final speeds are equal, the mean speeds must also be equal. So mean speeds are constant for planes of a particular vertical height. As the product of mean speed and time gives the distance travelled, times of descent are in proportion only to the lengths of the planes, including the vertical. Thus, the encoded-laws slot of the incline-plane schema is

Slot	Equal speed assumption	Inclined-plane LED/Proposition III
Diagram-features	Side **T**, hypotenuse **A**.	Side **T**, hypotenuse **A**.
Diagram-constraints	Right angle triangle, base horizontal.	Right angle triangle, base horizontal.
Domain-properties	Terminal speed, height, (distance along plane).	Height; distance along plane, *d*; descent times, *t*.
Encoded-laws	Terminal speeds are equal.	$t_{\mathbf{A}}/t_{\mathbf{T}}=d_{\mathbf{A}}/d_{\mathbf{T}}$.
Property mappings	Triangle height represents common elevation of planes. Lengths and directions of **T** and **A** give distance and directions of the motions.	Triangle height represents common elevation of planes. Lengths and directions of **T** and **A** are size and directions $d_{\mathbf{T}}$ and $d_{\mathbf{A}}$. Lengths of **T** and **A** are magnitudes of $t_{\mathbf{T}}$ and $t_{\mathbf{A}}$.
Interpretation-rules	Topology of diagram and domain are similar.	Topology of diagram and domain are similar.
Construction/manipulation-heuristics	Make one of the paths vertical.	Make one of the paths vertical.
Special-cases	—	—

Table 5. Derivation of the Inclined-Plane LED Schema

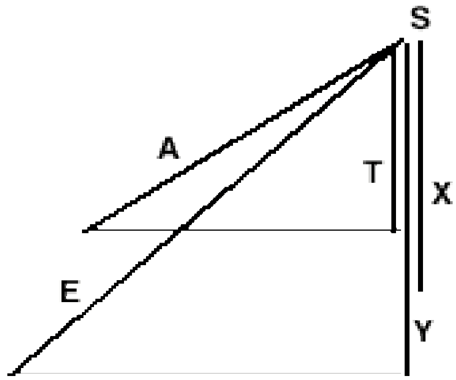

Figure 5 Two-inclined-planes LED

augmented with the relation $t_{\mathbf{A}}/t_{\mathbf{T}}=d_{\mathbf{A}}/d_{\mathbf{T}}$, where t is time and d is distance. Table 5 summaries how the various slots of the LED schema are augmented in the derivation of the Inclined-plane LED schema.

The inferences beginning with the Equal-speed-assumption and ending at Proposition III shows that discoveries can be made by the refinement and extension of an LED schema, under a goal of finding contents of slots that cover more of the variables that are relevant in the domain. The final version of the Inclined-plane LED encodes laws for distances and times in addition to final speeds.

The next step in the analysis is the construction of a composite-LED schema from basic LED-schemas. The schema is the *Two-inclined-planes* LED and corresponds to Proposition V of the TNS. The proposition states that the ratio of times of descent down planes with different inclines, lengths and heights equals the product of the ratio of their lengths and the inverse ratio of the square roots of their heights. For example, increasing the length of a plane and decreasing its height will greatly increase the time of descent. The Inclined-plane and Time-squared LED schemas are the component-LEDs and the domain-description states that descents down two inclined planes with different length and height are being considered, Figure 5. The composition-constraint is the sharing of a common apex, at **S**. Notice that Figure 5 is assembled using one set of components from Figure 3 and the components from Figure 4 twice over. The times of descent down the two inclines are being compared, the domain-interpretation-rule. The derivation of the encoded-law uses the laws in each of the component LEDs. Its starts with the elaboration of the ratio of times; i.e.,

$$t_A/t_E = t_A/t_T \cdot t_T/t_Y \cdot t_Y/t_E.$$

Substituting ratios of distances for the ratios of times using the encoded-laws of the Inclined-plane LED and the Time-squared LED gives,

$$\begin{aligned} t_A/t_E &= d_A/d_T \cdot d_T/d_X \cdot d_Y/d_E \\ &= d_A/d_E \cdot d_Y/d_X \\ &= d_A/d_E \cdot \sqrt{(d_Y/d_X)} \,.\, \sqrt{(d_Y/d_X)}. \end{aligned}$$

Now substituting for one of the square root distance terms using the alternative distance ratio given by the law form the Time-squared LED,

$$\begin{aligned} t_A/t_E &= d_A/d_E \cdot \sqrt{(d_Y/d_X)} \cdot \sqrt{(d_X/d_T)} \\ &= d_A/d_E \cdot \sqrt{(d_Y/d_T)}. \end{aligned}$$

Thus the encoded-law for the Two-inclined-planes composite-LED schema states the time of descent is in proportion the length of the plane and inversely proportional to the square-root of its height.

An example of a special-case of this Two-inclined-planes composite-LED schema is Proposition IV, a composite-LED-instance schema in which the lengths of the inclines are equal, but the heights are unequal (domain-condition). The order of the propositions in the TNS implies that the Two-inclined-planes LED is a generalization of Proposition IV, but it is as plausible that the discovery of Proposition IV followed the specification of the Two-inclined-planes LED.

The final LED to be considered is also derived from the Two-inclined-planes composite-LED schema, the *Equal-time-circle* LED, Figure 6. It considers motion down inclined planes, **P** and **Q**, running from the top or bottom of a vertical circle to points on the circumference (diagram-features and diagram-constraints). Descents on all such paths are completed in equal times (encoded-law). The derivation of the Equal-time-circle LEDs involves the application of the encoded-law from the Two-inclined-planes LED, specializing it for the particular configuration of paths in the circle. The details of the derivation are not important to consider here, rather it is the general nature of the derivation that is of interest. Although, on first sight, it appears that the new LED is a composite-LED-instance, that would be named in the special-cases slot of the Two-inclined-planes composite-LED schema, it is really a (basic) LED schema in its own right. The description of the configuration in terms of

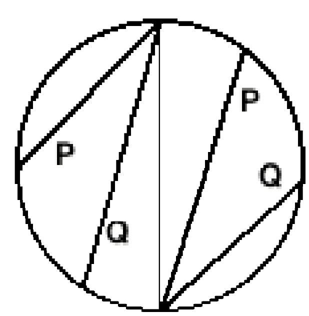

Figure 6 Equal-time-circle LED

individual diagrammatic elements rather than component-LEDs, and the simplification of the encoded-law to a statement of the equality of times, means that the Equal-time-circle LED must be a (basic) LED schema, not a composite-LED-instance schema.

This case is an interesting contrast to the earlier example, in which the new Inclined-planes LED schema was generated by refining the (basic) LED schema for the initial assumption/postulate of the TNS. That LED was derived by augmenting the contents of existing slots rather than completely replacing them, which is the case for the Equal-time-circle LED.

The analysis of this series derivations demonstrates how some forms of reasoning with LEDs can be plausible explained in terms of the processing of schemas, at least in this class of scientific discovery task. Despite this informal analysis there do appear to be at least two forms of problem space search. First, there are searches within a single schema with the goal of filling or extending slots, as in the development of the Inclined-plane LED. Second, there is the search of the space of schemas, in which new schemas are generated from existing ones, such as the derivation of the composite Two-inclined-planes LED from the Time-squared and Inclined-plane LEDs. Clearly, much work is required to develop this characterization but the modelling of LED based scientific discovery now appears feasible.

Discussion

The cognitive theory of LEDs is still under development, so it is premature to consider the general implication of these ideas in detail. However, the previous work most closely related to schemas for LEDs will be considered. Koedinger and Anderson's (1990) *diagram configuration model* of expert geometry problem solving proposes a single class of diagram *configuration schemas*, DCSs. These are clusters of geometry facts associated with a single prototypical geometrical diagram, a configuration. The facts include: geometry statements for the configuration as a whole; relations among the parts of the configuration; conditions that determine whether inferences can be made about the configuration. Geometry proofs are made by (i) parsing a given problem diagram to find familiar configurations, (ii) encoding the given and goals in canonical terms of the relations among parts of a configuration, and (iii) iterative searching through the space of diagram configurations to find a link between given and goal statements.

There are clearly some similarities between DCSs and LEDs. The clusters of facts in DCSs are similar to some of the slots in the LED schema and LED-instance schema. However, the differences between them are more revealing. First, LEDs represent classes of phenomena, so they contain information about the mappings from the features and structure of the diagram to the variables and laws of a domain, but such mappings are not present in DCSs that represent abstract mathematical facts. Second, neither the LED schema nor the LED-instance schema are the same as DCSs, but slots equivalent to the contents of DCSs are distributed across the two schemas. This is due to the fundamental role of diagram construction and manipulation in reasoning with LEDs, that requires explicit knowledge of how generate and interpret LEDs, i.e. LED schemas, which must be distinguished from particular instances of LEDs that apply to specific phenomenon in a domain, i.e. LED-instance schemas. Third, two pairs schemas for LEDs are proposed, for basic LEDs and composite LEDs. This is in contrast to the single DC schema and is a reflection of the narrower focus of Koedinger and Anderson on geometry proof problem solving. The aim in the current approach is to develop a theory not only for reasoning in relatively simple domains but for domains that involve complex interactions.

Four schemas for LEDs have been introduced and used to analyse LEDs from two different domains. The coherence of the analysis provides some support for the validity of the claim that LEDs can best be characterized as schemas and that the particular schema structures posited are correct. Analysis elsewhere of LEDs for quite different domains, including electrical circuits and the atomic structure of molecules, lends further support for the hypothesized set of schemas for LEDs. The next stage in this work is to more fully investigate the nature of the information processing that occurs with these schemas by developing models for a range of different tasks, including discovery, problem solving and learning. Further, the schemas for LEDs are being used for the cognitive analysis in the development of principles of instruction and techniques for knowledge acquisition that exploit LEDs.

Acknowledgements

This research was supported by the UK Economic and Social Research Council and by a grant from the UK Engineering and Physical Sciences Research Council. I am grateful to the members of the ESRC Centre for Research in Development Instruction and Training for their help and support in my pursuit of this research, with special thanks going to David Wood.

References

Cheng, P. C.-H. (1992). Diagrammatic reasoning in scientific discovery: Modelling Galileo's Kinematic Diagrams. In H. Narayanan (Eds.), *AAAI Technical Report on Reasoning with Diagrammatic Representations (Report No. SS-92-02)* (pp. 33-38). Menlo Park: CA: AAAI.

Cheng, P. C.-H. (1996a). Law encoding diagrams for instructional systems. *Journal of Artificial Intelligence in Education*, **7**(1), 33-74.

Cheng, P. C.-H. (1996b). Scientific discovery with law encoding diagrams. *Creativity Research Journal*, **9**(2&3), 145-162.

Cheng, P. C.-H. (1996c). Diagrammatic Knowledge Acquisition: Elicitation, Analysis and Issues. In N. R. Shadbolt, H. O'Hara, & G. Schreiber (Eds.), *Advances in Knowledge Acquisition: 9th European Knowledge Acquisition Workshop, EKAW'96*. (pp. 179-194). Berlin: Springer-Verlag.

Cheng, P. C.-H. (1996d). Learning Qualitative Relations in Physics with Law Encoding Diagrams. In G. W. Cottrell (Ed.), *Proceeding of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 512-517). Hillsdale, NJ: Lawrence Erlbaum.

Cheng, P. C.-H. (1996e). Functional Roles for the Cognitive Analysis of Diagrams in Problem Solving. In G. W. Cottrell (Ed.), *Proceeding of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 207-212). Hillsdale, NJ: Lawrence Erlbaum.

Cheng, P. C.-H. (1996f). Problem solving and learning with diagrammatic representations. In D. Peterson (Eds.), *Forms of Representation* (pp. 47-66). Exeter, U.K.: Intellect Books.

Cheng, P. C.-H., & Simon, H. A. (1995). Scientific Discovery and Creative Reasoning with Diagrams. In S. Smith, T. Ward, & R. Finke (Eds.), *The Creative Cognition Approach* (pp. 205-228). Cambridge, MA: MIT Press.

Galileo Galilei (1638/1974). *Two new sciences* (S. Drake, Trans.). Madison, Wisconsin: University of Wisconsin Press.

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

Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.