NETWORKS OF LAW ENCODING DIAGRAMS FOR UNDERSTANDING SCIENCE.

Peter C-H. Cheng

ESRC Centre for Research in Development, Instruction and Training.

Department of Psychology, University of Nottingham, Nottingham, NG7 2RD, U.K.

Email: peter.cheng@nottingham.ac.uk

(January 1998)

ABSTRACT

Understanding science involves the mastery of complex networks of concepts. To design effective computer based systems for learning science it is essential to adequately characterize the nature of those conceptual networks, so that clear and appropriate instructional goals can be defined and fed into the design process. This paper considers a novel class of representations for science instruction — Law Encoding Diagrams (LEDs) — and describes the nature of scientific understanding based on these representations. A framework of four classes of schemas has been proposed to characterizes problem solving and learning with LEDs. How the framework encompasses complex networks of concepts is discussed and the implications for the design of computer based learning environments based on LEDs are considered.

Cheng, P. C.-H. (1999). Networks of Law Encoding Diagrams for understanding science. *European Journal of Psychology of Education*, 14(2), 167-184.

INTRODUCTION

Visual representations may be effective for problem solving and learning. With new computer based technologies the variety of visual representation systems has dramatically increased with support for new forms of visualization and dynamic interactions (Kaput, 1992). The work described here is part of a research programming studying an interesting class of diagrammatic representations that can promote science learning — Law Encoding Diagrams (LEDs). For a given domain a LED 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 an instance of the phenomena or one case of the laws.

LEDs had a role in some major discoveries in the history of physics and chemistry (Cheng, 1996a). Their advantage for the making of discoveries compared to conventional propositional representations is in part due to their computationally efficiency (Cheng and Simon, 1995). A computer based discovery learning environment for particle collisions in physics has been built that exploits particular classes of LEDs. Figures 1 and 2 shows screen images of the system — ReMIS-CL (Representations: Multiple, Interactive, Structural for Conservation Laws). The system deals with head-on collisions between two elastic bodies moving in a straight line and may be used to introduce the important laws of energy conservation and momentum conservation. At the bottom of the screen there is an animated simulation of the collisions. In Figures 1 and 2 two LEDs are shown: the One-dimension property (1DP) diagram (left) and the Velocity-velocity (VV) graph (right). Both LEDs represent the masses of the bodies by the length of the lines **m1** and **m2**. The velocities before and after impact are represented as pairs of arrows (**U1 U2**) and (**V1 V2**). In Figure 1 the situation is one of equal bodies approaching from opposite directions with equal speed, and departing with the same speed but in opposite directions. Figure 2 shows a different configuration, with the two diagrams representing unequal bodies with unequal approach speeds colliding and then both going left at different speeds. The LEDs are linked in ReMIS-CL so that changes in the configuration of one diagram will automatically be reflected in the other.

Figure 1
Figure 2

Diagrammatic constraints govern the structure of LEDs. An example of one in the 1DP diagram is the *rectangle rule*, which specifies that the tails or heads of the pairs of velocity arrows lie at the corners of a rectangle and run parallel to the top or bottom of the rectangle (Figures 1 and 2, left). An example of a constraint of the VV graph is the intersection of the circle/ellipse and diagonal line (Figures 1 and 2, right). These constraints capture the laws governing the domain using the internal structure of the diagrams. In the experimental studies of learning with LEDs, subjects comprehended such constraints and were able to adopt a novel diagrammatic approaches when solving quantitative problems, which they previously answered poorly (Cheng, 1996b). An evaluation of ReMIS-CL has shown that subjects qualitative understanding of the domain improves, whilst a group using a matching learning environment with conventional representations did not (Cheng, 1996c).

A disadvantage of using LEDs in conjunction with conventional representations is the overhead associated with learning a new representation, but the potential advantages may out weigh this difficulty. There are various ways to characterize the possible benefits of LEDs. Like diagrammatic representations in general they can support more efficient inferences and searches for information (Larkin and Simon, 1987). They are specialized computational devices for reasoning about particular domains (Cheng, 1996b) and may be considered as effective tools for thinking (Kindfield, 1994). By combining diagrammatic rules that encode the laws of a domain with particular diagrams representing specific examples of phenomena, LEDs integrate levels of abstraction (Cheng, 1996b) and may help learners bridge the conceptual gulf between abstract general laws and concrete examples (White, 1989). LEDs may be considered as *limited abstraction representations systems* and possess the cognitive benefits conferred by the *specificity* of such representations (Stenning and Oberlander, 1995).

However, these general explanations do not fully capture the character LEDs and they only provide relatively weak constraints on the design of learning environments for LEDs. Similarly, the empirical evaluations of ReMIS-CL only demonstrate the benefit of computer based LEDs for a single limited domain, but knowing how best to apply the system to other domains, or to extend it beyond idealised elastic impacts between two bodies moving in one dimension, is problematic. Further, the studies have focused upon discovery learning in which subjects interactively manipulate the diagrams on screen to learn the rules governing the LEDs, but other forms of reasoning and learning are possible with LEDs. Thus, this research could benefit from a conceptual framework to help characterise the full spectrum of cognition with LEDs, so that the empirical evaluations can be placed on a more solid theoretical footing, and to provide a basis for the systematic exploration of the design space of effective instructional systems that exploit LEDs.

Fortunately, such a framework is has been formulated and is now being evaluated. Cheng (1997) used it to characterize Galileo's discoveries in kinematics and in another study (to be reported elsewhere) the framework was used to explain protocols of problem solving in electricity. Here, the nature of conceptual understanding in scientific domains will first be considered, to provide a map of the terrain that must be explored by learners. Such a map will be essential for defining instructional goals as targets for the design of effective learning systems. Second, the framework will be introduced and shown to adequately cover the landscape of scientific understanding, so that it can then be legitimately be applied to the design of systems. Finally, the general implications of the framework for computer based LEDs for science are explored by examining the benefits and limitations of ReMIS-CL and considering how it may be extended.

UNDERSTANDING SCIENTIFIC DOMAINS

The nature of scientific understanding is a question philosophers of science have wrestled with for centuries. For our purpose here, it will be sufficient to consider the more focused and concrete question of what constitutes a good conceptual understanding of a scientific domain with LEDs. The answer describes the terrain that learners must explore when learning science and which the framework must adequately characterize. Four different but related aspects of are considered in turn.

Diagrammatic Elements and LED Structure

A good understanding of a scientific domain with LEDs must be grounded on a good comprehension of the structure of the relevant LEDs. Basic aspects of this include what diagrammatic elements stand for which properties of the domain and how they represent the magnitudes of those properties. For example, the relative lengths of mass lines (**m1** and **m2**) in the 1DP diagram and the VV graph, Figures 1 and 2, stand for the relative masses of bodies. The geometrical, topological and spatial structure of the LEDs encode the laws of the domain, so knowing how the diagrammatic elements are interrelated is essential. For example, the rectangle rule governing structure of 1DP diagrams is an important constraint of that LED.

Knowing typical examples of diagrams that represent common situations will facilitate problem solving, because chunks of information will be made readily available through recognition of the situation (or diagram) followed by the simple recall of the diagram (or situation). For example, the chunk that incorporates 1DP diagram in Figure 3a is concerned with collisions between equal bodies, with the one initially moving body giving up all its motion (**U1**) to another that is initially stationary (**V2**). How this chunk can be used for problem solving will be seen below. Similarly, knowing examples of diagrams that have interesting patterns, such as symmetries, or that correspond to extreme cases of the domain, can aid comprehension of the underlying nature and scope of the laws, and also make more concrete the constraints on the structure of the diagram. For example, Figures 3b shows what happens when one body is very much bigger than the other, say a perfectly elastic planet (**m1**) colliding with a perfectly elastic pea (**m2**). This extreme case has the pea rebounding at no more than three times its initial speed (**U2**) in the opposite direction (**V2**), because the dashed diagonal line is not permitted to intersect the central mass line outside the rectangle. For most scientific domains there will be a many examples of phenomena with corresponding diagrams.

Figure 3

Encoding Laws and Global Conceptual Relations

Laws are the underlying relations that govern the nature of the phenomena in a domain, specifying permitted relations among variables in some formal representational system (e.g., algebra). Knowing what laws a LED encodes is obviously an important part of understanding a domain with LEDs.

There are also global conceptual relations to be considered. Idealizations play a central role in scientific knowledge, but knowing how laws apply to the real world is also a big part of scientific understanding. The (elastic-)1DP diagram and (elastic-)VV graph are idealizations (Figures 1-3), because they represent perfect collisions in which no energy is lost. In real collisions energy is lost through the distortion of the bodies and as sound, but modified versions of the conservation laws still hold for these phenomena. LEDs for such *plastic* collisions and their underlying laws take into account the energy that is lost from the system. Figures 4 and 5 are examples of the *plastic-1DP diagram* and the *plastic-VV-graph*, for such collisions. The precise way these LEDs work is not important for the present discussions, but it should be noted that compared to their elastic counterparts: (i) they are more complex, requiring extra diagrammatic features and rules (e.g., the addition of a second inner circle in plastic-VV graph, Figure 5); (ii) their structure becomes the same as that of the non-plastic (elastic) LEDs when no energy is lost (e.g., the inner circle expands to overlap the outer circle). Knowing the relation between the different versions of the LEDs and laws is necessary for a complete understanding.

Figure 4

Figure 5

Simplifications are also important in learning about science, because they omit details that are not essential for an initial appreciation of the domain. The 1DP diagram is a good way to introduce the topic of particle collisions, because it simplifies the phenomena to one-dimensional motion. The *2DP diagram* can be introduced later to deal with motion in two-dimensions, Figure 6. The pairs of initial and final velocities are represented by **U1**, **U2**, **V1** and **V2**, respectively. This LED treats the velocities as components parallel and perpendicular to the plane of impact (**x** and **y** directions). The precise details of the structure of the LED does not need to be considered, except to note there is a (vertical) 1DP diagram embedded at the centre of the 2DP diagram and that the 2DP diagram reduces to the 1DP diagram when the motions are all in one direction, vertical in Figure 6.

Figure 6

The framework for LEDs will need to cope with such idealizations and simplifications, plus their interrelations; for example, the plastic-1DP diagrams and the 2DP diagrams can be combined to model plastic two dimensional collisions, *plastic-2DP diagrams*.

Modelling Complex Interactions of Components

Scientific phenomena are complex in different ways. Relations among the properties of the basic phenomena may be complex. For particle collisions both momentum and energy conservation laws, holding simultaneously, are needed to account for the phenomena, and their algebraic expressions involve multiple symbols and many mathematical operators.

Another type of complexity comes in the form of interactions among the basic units or components of a domain. These interactions may be governed by different laws that are largely independent of the laws governing the relations of variables for basic units. The basic intra-component relations of particle collisions deal with single impacts between two bodies and the inter-component interactions are collisions among multiple bodies. LEDs can model such interactions; for example, Figure 7 shows a sequences of multiple collisions between several elastic bodies moving in one dimension, as seen in

Newton's Cradle (the executive toy with suspended balls). Because the bodies are elastic the impulse travels through the bodies as a shock wave at the speed of sound, so complex collisions can be decomposed into to pair-wise collisions (interaction law). In terms of the 1DP diagram, this means that the individual diagrams like Figure 3a can be assembled to model complex situations, such as one moving ball hitting four stationary balls of the same mass, Figure 7. The same principle can be used to model interactions between bodies of unequal masses (see Cheng 1996b).

Figure 7

As with LEDs for the basic phenomenon, understanding the domain at this level involves knowledge of: the constraints for drawing composite diagrams; the laws governing the interactions; prototypical cases and their corresponding diagrammatic configurations; special examples for interesting diagrammatic patterns or extreme phenomena. The framework will need to cope with all these aspects of composite LEDs for complex interactions.

Multiple LEDs

There are usually different ways laws of a particular domain may be captured in diagrammatic form in an LED. The 1DP diagram and the VV graph are two examples for the same domain that use quite different geometric and spatial constraints to encode the same two conservation laws. To what extent multiple representations of the same domain may help learning and understanding in science is an open question. For example, Kaput (1992) argues that multiple representations may give a deeper understanding by allowing learners to identify the invariants of a domain independently of any particular representations.

Whatever the potential benefit, multiple LEDs for a single domain can exist, so the framework must be able to embody them. Further, some domains in science are governed by laws with the same underlying formal relations, even though their phenomena are quite different. For example, in dynamics the law 'force=mass*acceleration' has the same algebraic structure as the law 'voltage=current*resistance' in electricity. The LEDs for such laws will have the same diagrammatic structure but will apply to quite different phenomena. This is another aspect of understanding science the framework must cover.

In summary, a good understanding of a domain based on LEDs must, at a minimum, encompass ten things:

(1) knowledge about elements and structure of LEDs;

(2) relate the laws of the domain to the structure of the diagrams;
(3) associate particular phenomenon with specific LED configurations;
(4) interrelate theoretical considerations with examples of phenomena;
(5) show how variants of laws (idealizations, specializations) are interrelated;
(6) classify phenomena and group them into related classes;
(7) distinguish relations within components from complex interactions among components;
(8) allow comprehension of the mutual constraints between inter-component interactions and intra-components relations;
(9) provide the means to interrelated different LEDs for the same domain;
(10) distinguish different domains that happen to have the same underlying relations and hence LEDs with the same diagrammatic structure.

A FRAMEWORK OF SCHEMAS

Various models or characterizations of the structure of mathematical and scientific knowledge have been formulated: Michener's (1978) ingredients and processes of understanding mathematics; Reif's (1987) types of inferences in physics; White and Frederiksen's (1990) progressions of qualitative causal models; Giere's (1994) model maps; the studies of expert/novice differences in physics problem solving by Larkin, McDermott, Simon & Simon (1980) and Chi, Feltovich, & Glaser (1981). The framework for LEDs shares many aspects of these analyses, but a significant difference is its focus on a diagrammatic representation rather than propositional or mathematical representations.

Cheng (1997) proposes four classes of schemas as (internal mental) memory structures for thinking with LEDs and uses them to explain how Galileo made kinematics discoveries with diagrams. Various criteria were used in the formulation of the schemas, including: general compatibility with schema theories (e.g., Ellis and Hunt, 1993); parsimony in the number and types of structures proposed; coverage of all the forms of knowledge and information processing central to LEDs, as presented in the previous section. The acceptability of framework has also been tested by using it to analyse problem solving protocols of experimental subjects learning about electricity with LEDs (to be reported elsewhere).

The four classes of schemas are: *LED schemas* (LS), *Meta-LED schemas* (MLS), *composite-LED schemas* (CLS) and *meta-composite-LED schemas* (MCLS). They can be distinguished on two dimensions, as shown in Table 1. The first dimension concerns the level of abstraction: LSs and CLSs hold information on examples of phenomena or classes of phenomena; MLSs and MCLSs hold general theoretical information on the laws governing the domain, the diagrammatic constraints of the representations, and how they interrelated. The second dimension distinguishes the loci of applicability of the schemas: LSs and MLSs cover individual components, the basic conceptual units of the domain; CLSs and MCLSs cover interactions among multiple components.

Justification of the framework, and its relation to schema theories in general, will be considered following the description of the structure of the schemas themselves.

LED schemas (LS) and Composite-LED schemas (CLS)

Table 2
Table 3

These classes of schemas are for particular phenomenon or classes of phenomena, with the LSs dealing with intra-component relations and the CLSs dealing with inter-component interactions. The slots of the two classes of schemas are comparable in that matching slots contain information of the same general type, but they are different in that the information pertains to either individual components or interactions among components.

Table 2 gives examples of LSs for the particle collisions domain. The Diagram slot stores an image(s) of the configuration of the LED for the schema's particular phenomenon (in whatever form that images are held in memory). The Diagram-configuration slot specifies the notable features that distinguish this diagrammatic configuration from the configurations for other phenomena. Domain-conditions give the particular circumstances that can be used to identify this phenomenon, which may be in the form of specific values of properties or as relations among them. Interpretation is a propositional description of what distinguished the particular phenomenon of the schema in domain relevant terms.

The slots of the CLSs are similar, but deal with interactions among components. Table 3 gives examples for multiple collisions in one dimension. For instance, the Composite-configurations slot defines the structure of the diagram as arrangements of LEDs for single components. Each slot of a CLS will have tests to check that the information applies to interactions among components, whereas a LS's slots will have tests to ensure that only information pertaining to an individual component can be held.

From the applications of the framework so far, it appears that the part LSs and CLSs have in reasoning is similar to the role that Koedinger and Anderson (1990) attribute to *Diagrammatic Configuration Schemas*, DCSs, in expert geometry problem solving. DCSs are perceptual chunks that store information for a particular geometry problem configurations. When the configuration of a diagram in a DCS matches a part of an external problem diagram, new information about the configuration can be asserted, if sufficient facts are available about the configuration. Like DCSs, when part of an external diagram matches the diagram in a LS or CLS, the rest of the information about the particular phenomenon represented by the diagram is made available for reasoning. However, unlike DCSs, information from a description of some phenomenon that matches the Domain-conditions or the Interpretation of a schema will make the Diagram (and the Diagram-constraints information) available, so that an external drawing can be made (or completed). This "two way" use of LSs and CLSs occurs because scientific problem solving is more diverse than the relatively unidirectional geometry proof problem solving studied by Koedinger and Anderson (1990).

LSs and CLSs can also be considered like schemas in standard schema theories (e.g., Ellis and Hunt, 1993). An existing schema may be specialized or generalised into new one when a when a new phenomenon is met. For example, the Zero-momentum LS (Table 2) is a generalization of the Default LS. The General-Newton's-Cradle for any number of pair-wise collisions can be specialized into the Simple-Newton's-Cradle for a specific number of collisions. Given a LED for a known phenomenon, say the VV graph in Figure 1, a new schema can be generated that inherits the contents of the Domain-conditions and Interpretation slots, say from a 1DP diagram LS, but that uses the VV graph as a new Diagram.

Meta-LED schemas (MLS) and meta-composite-LED schemas (MCLS)

Table 4
Table 5

Meta-LED schemas and meta-composite-LED schemas store information on: (i) the abstract general laws of a domain; (ii) the structure of LEDs for the domain; (iii) the relations between (i) and (ii). Like the LSs and CLSs, the slots of the schemas are similar, with the meta-LED schema covering relations within components and the meta-composite-LED schemas covering interactions among components.

The structure of meta-LED schemas is illustrated with examples in Table 4. There are seven slots: Diagram-features specify the diagrammatic elements of a particular class of LEDs for certain classes of phenomena. Diagram-constraints are rules defining the structure of the diagrams in geometric, spatial or topological terms. The *local* constraints are specific relations between elements and the *global* constraints determine the overall form an LED. For example, a global constraint is the rectangle rule we met above and another is the diagonal rule, Table 4, which specifies that ends of the lines not fixed at the corners of the rectangle must be co-linear diagonally or vertically (e.g., Figures 2 and 3). Domain-properties are the properties of interest in the domain. Encoded-laws are the laws and relations that govern the domain. Property-mappings indicate which domain variables correspond to particular diagram features. Interpretation-rules give a general description of the phenomena covered by the schema that distinguishes it from other schemas covering different classes of phenomena. Cases are pointers to LSs for particular phenomenon in the domain or important diagrammatic configurations that can be derived from the MLS.

Given some general abstract information about a problem, appropriate MLSs and MCLSs may selected to provide new information for solving the problem. For example, when in-elastic collisions are being considered, a match of this fact with the contents of Interpretation slot of the Plastic-1DP-diagram MLS (Table 4) makes available the laws governing this class of collisions and the rules for drawing LEDs.

Like schemas in general, MLSs and MCLSs can be generalized or specialized, with information inherited by the new schemas. For example, elastic collisions are a special case of collisions in general, so the 1DP diagram MLS can be considered as a special case of the Plastic-1DP diagram (Table 4), with fewer degrees of freedom in the equations and a more constrained LED structure.

Justification of Framework Structure

The framework builds upon typical schema theories (e.g., Ellis and Hunt, 1993) with the hierarchical interrelation of schemas within each of the four classes of schemas and *is-part-of* relations for LSs as sub-schemas of CLSs. However, why is it necessary to define the four classes of schemas, rather than attempt to accommodate scientific understanding with LEDs within a conventional account, is considered in this sub-section. The question can be broken down in to two separate issues underpinning the distinctions made in Table 1: (i) the phenomena level versus theoretical (meta) level; (ii) the separation of intra-component relations from inter-component interactions.

Two levels, LSs/CLSs and MLSs/MCLSs, are required because the knowledge of particular examples of phenomena is different in kind to the overreaching theoretical (abstract and general) knowledge about the relations governing all the classes of phenomena within a domain. This is apparent from the processes for generalizing descriptions of phenomena into general laws (LSs/CLSs into MLSs/MCLSs) and the processes for deriving descriptions of phenomena from the laws (MLSs/MCLSs into LSs/CLSs). The discovery of laws is not a simple process of induction where sets of variables replace sets of values or properties, or where constraints among variables are relaxed, as new examples are incorporated into memory. (That is the process of generalizing LSs/CLSs into new LSs/CLSs to cover a greater range of phenomena). Rather, the process is about finding the functional or structural relations among all the variables using some formal representational system, that typically requires an extended series of inferences. For example, to find the momentum and energy conservation laws from sets of data requires non-trivial problem solving techniques (Langley et al., 1987; Cheng and Simon, 1992). The same holds for the processes in the reverse direction, where applying laws to a particular case is more than assigning values to variables, but involves the computation of values within some formal system.

The framework proposes different classes of schemas for intra-component relations as distinct from inter-component interactions, but the classes of schemas have the same overall structure with equivalent slots. Thus, there appears to be no psychological reason to double the proposed number of conceptual entities, because accounts of the processing of the schemas at the level of slot contents would be equivalent, whether there are two or four classes schemas. However, there is a logical basis for the distinction, because many scientific domains do have components whose internal relations are governed by different laws to those for the interactions between components. Such differences are reflected in the ways that the schemas are processed, with different patterns of reasoning being identifiable when thinking is concerned with intra-component relations versus inter-component interactions, or when the connections between the two levels are being made. A homogenous non-scientific domain that does not have such a natural division will not exhibit such patterns; but for domains that do, defining the different classes of schemas gives more explanatory power and greater clarity to the accounts of cognition. This claim is support by previous applications of the framework, which provided coherent characterizations of scientific discoveries and protocols of problem solving and learning in particular domains, with meaningful patterns of behaviour being characterized on precisely this basis. As will be seen below, this distinction also has implications for the design of instructional systems with LEDs.

From the use of the framework to date, it appears that the number and types of slots proposed for the four classes of schemas are both necessary and sufficient to account for problem solving and learning with LEDs. However, a further point should be made about definition of the slots. They have, quite deliberately, not been defined precisely using specifications of particular data formats, because the exact nature of the information may vary from domain to domain. For example, domains-conditions of LSs/CLSs may be specified as magnitudes of properties in one domain but as structures in another. It is sufficient for the purposes of the framework that the contents of the slots are separate expressions that enable the schemas to differentiate one phenomenon from another. The framework trades some precision in the definition of slots for a greater scope of applicability to diverse scientific domains.

The Framework and Aspects of LED Knowledge

From the previous sub-sections an implicit impression of how the framework characterizes scientific understanding will have been gained. This section makes the characterization explicit, with reference to the ten things, identified above, that are important for the understanding of scientific domains with LEDs. Figure 8 shows schematically relations within and between schemas of the four class, with each quadrant containing schemas of the same class. Boxes stand for schemas, with their labels naming examples in Tables 2 to 5 or referring to certain figures, and blank boxes representing other examples not discussed here. Plain lines connecting boxes are "conventional hierarchical" relations, as per schema theories in general. Other types of lines are explained below. Each of the ten aspects will be considered in turn, with examples in Figure 8 shown by appropriate numbers in circles.

(1,2) MLSs/MCLSs encompass knowledge about elements and structure of LEDs and relate the laws of the domain to the structure of the diagrams. (3) The association of particular phenomenon with specific configurations of LEDs is the job of the LS and CLS classes. (4) Knowledge that a law(s) applies to a particular phenomenon, or class of phenomena, is provided through the Cases slots in MLSs/MCLSs, but deriving a particular case will require the application of the other information from these theoretical schemas in a series of inferences. (In Figure 8, this is shown by solid arrows.) (5) Variants of laws in the form of idealizations and specializations are encoded by the interrelation of schemas within the hierarchies of MLSs and MCLSs. (6) Related phenomena and classes of similar phenomena are encoded by the interrelation of the schemas within the hierarchies of LSs and CLSs. (7) Relations within components are covered by MLSs and LSs, whilst complex interactions among components are handled by MCLSs and CLSs. The dashed lines, in Figure 8, between LS and CLS boxes indicates that it is possible to recognize perceptually a Diagram of a LS within a Diagram of a CLS, but this does not mean that LSs are sub-schemas of CLSs. (8) The interdependence of inter-component interactions and intra-components relations is embodied by MCLSs taking MLSs as sub-schemas in their Component-LEDs slot (dashed arrow). (9) LEDs from the same domain but with different diagrammatic structures will be related schema that inherit information from a common parent schema that holds general information about the domain (Domain-properties, Encoded-laws, Interpretation), but not about particular diagrams. (10) Diagrams with the same structure but applying to different domains are related through a common parent schema. That parent schema will hold information on Diagram-features and Diagram-constraints, and the abstract relations (Encoded-laws) common to both domains, but it will not carry information on how the diagrams map to domain properties (Property-mappings) nor any Interpretation in terms of the domain. Such a schema will not be applicable to any single domain, but would play role in providing insight into the similarities across different scientific domains.

Figure 8

DESIGNING SYSTEMS FOR LEARNING WITH LEDS

This section considers the implications of the framework for the design of computer based systems for learning with LEDs by first identifying the features of ReMIS-CL that appear to help or hinder learning and explaining why in terms of the framework. Then the framework is used to make design suggestions for the future development of the system.

The present version of ReMIS-CL was built before the development of the framework. The design was guided by various ideas (Cheng, 1993). The first was to provide interactive LEDs for students to discover the constraints of the LEDs, with the system automatically maintaining the correct structure of the diagrams. ReMIS-CL allows students to examine many different configurations and there is some evidence that students who do more diverse exploration gain a better qualitative understanding of the domain (Cheng, 1996b, 1996c). For example, as the ratio of the masses changes, they are better able to predict what will happen to the final velocities. The second idea was to provide two LEDs for the same domain, the 1DP diagram and the VV graph, to see whether multiple representation could facilitate the discovery of the invariants (laws) of the domain. Two LEDs may help the learners to spot those aspects peculiar to each LED and those that are common, so identify the underlying relations by a process of triangulation. However, in the evaluations of ReMIS-CL, there was little evidence to show that learners were attempting to integrate the two LEDs in some way. For instance, the amount of switching between LEDs of the systems was unrelated to subjects' learning gains.

The framework can explain both findings. Diverse exploration of the configurations of a LED, rather than the amount of activity on the system, leads to better qualitative understanding. ReMIS-CL seems to support learning of particular LSs by providing a simulation of the collisions that closely matches the configuration of the LEDs. The size of the bodies are depicted as different size circles and the animation speed is in proportion to the magnitude of the velocities. All the pieces of information needed to fill the slots of a LS are simultaneously present. The system was designed so that users directly manipulate elements representing particular properties within the diagrams, so subjects could easily examine a series of similar configurations systematically by making small incremental changes to the diagram or explore very different configurations by making dramatic changes. The different goals of these approaches will influence the organization of the acquired LSs in a way that reflects the similarity of phenomena or the similarity of LED configurations, rather than embodying arbitrary relations among the features of the LEDs, such as the width of the diagram. Thus, subjects qualitative reasoning, in the experiments, may have been supported by an appropriate hierarchy of LSs, with an overall organization of the schemas reflecting the important relations in the domain.

The lack of effect of giving two LEDs to the subjects can be explained in terms of the conjunction of the relatively short time subjects spent on the system and the amount of knowledge required before commonalties between two different LEDs can begin to be considered. To gain a deeper understanding of the domain by combining both the 1DP diagram and the VV graph requires a process of generalization that yields the common parent schema for the two LEDs — MLS box '(9)' in Figure 8. However, it is unlikely given the relatively short time subjects had on the system that they would have acquired full MLSs for the two LEDs, so it would be difficult for them to comprehensively compare the slots of the MLSs to discover what contents were, or were not, shared in common. Further, the framework makes it clear that attempting to learn a MLS from sets of LSs is a difficult task in itself, but ReMIS-CL does little to support this process. For example, it prevents incorrect diagrams from being generated but does not deliberately focus attention onto the diagrammatic constraints. Although in hindsight it is clear that subjects had insufficient time or support to fully understand the LEDs, this problem has only made conspicuous by the application of the framework.

A limitation of ReMIS-CL, which is obvious from the descriptions of the domain presented here, is its narrow focus on simple two-body one-dimensional elastic collisions. This is not a criticism that follows from the framework, but the framework does raises the question of whether covering more of the domain would also have an effect of enhancing learning about this particular part of the domain. By acquiring related schema, say for plastic or two-dimensional collisions, this could help learners refine the contents of the schema for elastic collisions by making more explicit constraints that were only implicit when schemas for a single class of phenomena were considered in isolation. For example, the fact that the lines in the 1DP diagram are always parallel stands in stark contrast to the lines in the 2DP diagram (Figures 3 and 6). For this reason, and also simply to broaden the scope of ReMIS-CL, covering other aspects of the particle collisions domain is desirable.

As already noted, the framework predicts that learning about an aspect of a domain that only requires the acquisition of a schema from existing schemas of the same class will generally be easier than attempting to learn that same schema by invoking schemas from an other class. Knowledge of the 1DP diagram MLS is easier to obtain by specializing the 2DP diagram MLS than by generalizing many examples of LSs for different configurations of 1DP diagrams. Similarly, generalizing the CLSs for Simple-Newton's-Cradle is easier than deriving the General-Newton's-cradle CLS (Table 3) from the 1DP-series-collisions MCLSs. The extension of ReMIS-CL to cover more of the domain will, thus, requires the provision of more instructional support for students than is given in the discovery learning approach used so far. The next version of the system will need to show, amongst other things, how the diagrammatic constraints of the LED defined in each MLS translates into the various diagrammatic configurations of known LSs. This might be done visually by temporarily highlighting the diagram's features covered by a particular constraint. For example, the rectangle rule of the 1DP diagram, described above, could be drawn to the attention of the learner by superimposing a red rectangle over each 1DP diagram. Selecting a different constraint, such as the diagonal rule, would mean highlighting a different part of each 1DP diagram.

The framework also provides some guidance on the order in which interactions among components and relations within components should be introduced. One can image that considering inter-component interactions first would provide a good context for the later introduction of intra-component relations. However, in the framework MLSs are found as sub-schemas within MCLSs, so trying to learn MCLSs first will mean there will be no contents for one its slots. Hence, learners must return to the MCLSs again to fill in the gaps after the MLSs have been learnt. However, by dealing with MLSs first, it is more likely that one pass though the MCLSs might be sufficient. Learning about interactions through compositional analysis may overall be easier, because it is more constrained with fewer unknown entities having to be entertained than under a decomposition approach.

CONCLUSIONS

To investigate design of effective computer based systems of learning science with LEDs, this paper initially considered what it is to have a good understanding of a scientific domain. Such an understanding involves a complex network of concepts with many different kinds of information at different levels of abstraction and generality. The framework of schemas has been shown to adequately characterize such complex networks of concepts, so it can be used as a basis for considering the design of instructional systems for learning science. By applying the framework to ReMIS-CL explanations were provided for why the system seems to be effective for promoting some aspects of scientific understanding but not others. Further, suggestions for the future development of the systems were derived from the framework to illustrate how it can be used to help constrain the design of LED based instructional systems.

ACKNOWLEDGEMENTS

This work was supported by the U.K. Economic and Social Research Council through the Centre for Research in Development, Instruction and Training. Thanks should to everyone in the Centre and in particular David Wood. My gratitude to Wolfgang Schnotz and two anonymous reviewers whose comments greatly helped improve this paper.

REFERENCES

Cheng, P. C.-H. (1993). Multiple interactive structural representation systems for education. In R. Cox, M. Petre, P. Brna, & J. Lee (Eds.), *Proceedings of the Workshop of Graphical Representations, Reasoning and Communication of the AI-ED 93 World Conference on Artificial Intelligence in Education*. (pp. 25-28). Edinburgh: University of Edinburgh.

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

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

Cheng, P. C.-H. (1996c). Learning qualitative relations in physics with law encoding diagrams. In G. W. Cottrell (Eds.), *Proceeding of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 512-517). Hillsdale, NJ: Lawrence Erlbaum.

Cheng, P. C.-H. (1997). Components of a cognitive theory of problem solving and discovery with law encoding diagrams. In M. Anderson (Ed.), *AAAI Fall Symposium: Reasoning with Diagrammatic Representations II.* Cambridge, MA: AAAI.

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.

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.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, **5**, 121-152.

Ellis, H. C., & Hunt, R. R. (1993). *Fundamentals of Cognitive Psychology* (5th ed.). Madison, WI: Brown and Benchmark.

Giere, R. N. (1994). The cognitive structures of scientific theories. *Philosophy of Science*, **61**, 276-296.

Kaput, J. J. (1992). Technology and Mathematics Education. In D. A. Grouws (Eds.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 515-556). New York, NY: MacMillan.

Kindfield, A. C. H. (1994). Biology diagrams: Tools to think with. *Journal of the Learning Sciences*, **3**(1), 1-36.

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: Computation Explorations of the Creative Processes*. Cambridge, MA: MIT Press.

Larkin, J. H., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Models of competence in solving physics problems. *Cognitive Science*, **4**, 317-345.

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

Michener, E. R. (1978). Understanding understanding mathematics. *Cognitive Science*, **2**, 361-383.

Reif (1987). Interpretation of scientific or mathematical concepts: Cognitive issues and instructional implications. *Cognitive Science*, **11**, 395-416.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science*, **19**(1), 97-140.

White, B. (1989). The role of intermediate abstractions in understanding science and mathematics. In *Proceedings of the Eleventh Annual Meeting of the Cognitive Science Society*. (pp. 972-979). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

White, B., & Frederiksen, J. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, **24**, 99-157.