Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Explorations in Understanding How Physical Systems Work

Permalink

https://escholarship.org/uc/item/2zw9r843

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 10(0)

Authors

White, Barbara Y.
Frederiksen, John R.

Publication Date

1988

Explorations in Understanding How Physical Systems Work

Barbara Y. White & John R. Frederiksen
BBN Laboratories

ABSTRACT

This paper presents a theory of how to enable people to understand how physical systems work[1]. Two key hypotheses have emerged from our research. The first is that in order to understand a physical system, students need to acquire causal mental models for how the system works. Further, it is not enough to have just a single mental model. Students need alternative mental models that represent the systems behavior from different, but coordinated, perspectives, such as at the macroscopic and microscopic levels. The second hypothesis is that in order to make causal understanding feasible in the initial stages of learning, students have to be introduced to simplified models. These models then get gradually refined into more sophisticated mental models. We will present a theory outlining (1) the properties of an easily learnable, coherent set of initial models, and (2) the types of evolutions needed for students to acquire a more powerful set of models with broad utility.

INTRODUCTION

We have been creating an intelligent, computer-based, learning environment that helps students to learn about the behavior of electrical circuits (Frederiksen & White, 1987; White & Frederiksen, 1986, 1987). The issues that have confronted us are relevant to the understanding of any physical system, including both designed and naturally occurring systems. By presenting the evolution in our attempts to foster understanding, we can raise some important issues concerning what it means to understand, and how to go about facilitating understanding.

"Causality", "mechanism", and "purpose" are perspectives that people often spontaneously impose on physical systems in order to understand how they work. These perspectives contrast with the type of constraint-based, quantitative reasoning that physicists often use in modelling physical systems, and that is presented to students in physics class and textbooks. This lack of connection between students' intuitive notions of causal mechanisms and quantitative circuit theory is, we believe, one reason why students typically have so much difficulty understanding the behavior of electrical circuits.

Our initial hypothesis was, therefore, that students would learn circuit theory more readily if it were introduced in the form of qualitative, causal models of circuit behavior. Causal models can simulate domain phenomena, and can provide a means of linking mathematical formalisms to these phenomena. In addition, they can introduce students to the basic concepts, laws, and causality of the domain. When internalized in the form of mental models, they can enable students to mentally simulate and to explain domain behavior. Such mental models can be used in solving a wide range of interesting problems, such as predicting circuit behavior, and designing and troubleshooting circuits. The computer system we built therefore incorporated causal models that could visually simulate and verbally explain the causality of circuit behavior. The system was designed to provide an external representation of the mental models that we wanted students to acquire.

---

[1] This research was funded by the Army Research Institute, under contract MDA-903-87-C-0545.

Causal models can embody different perspectives on system operation. For instance, such models can reason about circuit behavior at what might be called the macroscopic, or phenomenological level. That is, they can allow reasoning about circuit behavior through the application of a set of laws that govern the distribution of voltages and currents within the circuit. One can imagine, for example, closing a switch and hearing an explanation of how that causes voltages to be applied to devices within the circuit, and how that in turn causes a current through a device such as a light bulb, which causes it to be on. Alternatively, causal models can represent the behavior of circuits at the more microscopic level. For instance, one can imagine seeing how electrical forces within a circuit can cause mobile, charged particles to be redistributed when, for example, a switch is closed. Finally, causal models can take the form of functional models, which reason about the purpose of the circuit, and how subsystems within the circuit interact to achieve that purpose. Here, instead of reasoning about devices as potential conductors or sources of voltage, one reasons about devices as receivers, transformers and transmitters of information, and the causal propagation is in terms of information flow, instead of in terms of electrical charge (White & Frederiksen, 1987b). The question for instruction then is: which of these types of causal models should one start with and why? Further, why having picked one of them, are there any compelling reasons for introducing the others?

PHENOMENOLOGICAL MODELS

We began by conducting studies of experts' reasoning about circuit behavior in the course of solving problems such as troubleshooting. We found that experts primarily reason about device states, and how a change in the state of one device can cause changes in the states of other devices. When the circuits dealt with were simple electrical circuits such as those found in an automotive electrical system, experts employed electrical concepts in predicting how a change in one device's state would propagate its effects to other devices within the circuit. This electrical reasoning incorporated the idea of causal interactions among devices, which are due to the effects of changes in conductivity of one device on voltages applied to other devices within a circuit. The experts reasoned solely about whether voltages would be present or absent across devices, not about incremental changes in those voltages or about their quantitative values.

To model this type of reasoning, the model we created reasons about gross aspects of circuit behavior, such as whether or not there is a voltage drop across a device. For example, in a circuit containing a battery, a switch, and a lightbulb, the expert reasons that (1) when you close the switch, that changes the conductivity of the switch causing a redistribution in the voltages within the circuit, (2) so that now there is a voltage drop across the lightbulb, whereas before there was not, (3) so the lightbulb goes from being off to being on. In creating this model, we initially adopted what might be called a "device centered" view of casuality. That is, reasoning about the effects of a change in circuit conductivity was done from the perspective of the individual devices in the circuit. For example, each device within a circuit would ask "What effect would closing that switch have on my behavior?". So, for instance, the lightbulb in this case would say "Before the switch was closed, I didn't have a good feed path to the battery, and now when the switch is closed, I do have a good feed path, and I also have a good return path, so I have a voltage drop across me, so I will be on". We found that such a model is easy to learn and is useful for introducing basic ideas, like the concept of a circuit.

One of the drawbacks of the device-centered model, however, is that it doesn't map readily to basic circuit laws (such as Kirchhoff's Laws) which are important in quantitative circuit analysis. These laws may be thought of as processes which govern the distribution of voltages and currents within an electrical circuit. We therefore went on to consider an alternative perspective (Forbus, 1985) in which the concept of "process" is regarded as central to understanding how systems work. For instance, you could reason that when you close a switch, there is a change in the switch's conductivity, which initiates a "voltage redistribution process". Instead of looking at what is going on from the perspective of each device, this redistribution process is run whenever a device changes state. Then, all devices within the effected part of the circuit simply ask whether the voltage drop across them has changed, and if it has, they alter their states appropriately. The voltage redistribution process takes the form of a set of qualitative rules such as, "If you have a circuit with an open in it, the only voltage drop in that circuit will be across the open. Whereas, if you have a circuit that is a complete conductive loop, there will be voltage drops across resistive devices in that circuit." These rules are, in effect, qualitative expressions of the circuit laws of quantitative circuit theory. Students can be easily motivated to develop this new, "process" perspective by presenting them with circuits (e.g., long series circuits) for which the device centered approach is extremely tedious and inefficient. For these circuits, the process-centered model represents a much-needed simplification.

Each of these alternative phenomenological models represents a different perspectives on circuit causality. Thus, in addition to being introduced to some fundamental concepts about circuits, such as voltage, conductivity, and the basic idea of circuit, students learn that there is more than one way to approach modelling circuit behavior that is consistent with the fundamental concepts. In each approach, we imposed a causality onto the behavior of circuits, such as a change in conductivity can cause a change in voltage, and a change in voltage can cause a change in device state. This causality maps nicely to ideas about electrical force and the behavior of charge carriers within a circuit, as well as to quantitative circuit analysis. However, there are serious limitations to only presenting student with such phenomenological circuit models. In particular, when exposed to these kinds of causal models, students would only have shallow answers to important questions such as, "What is a voltage drop, and how do voltages get redistributed within a circuit?" Beyond the set of rules for determining the distribution of voltages, students have no deeper sense of what is going on in a circuit. One of the reasons for this shallowness (which is also a characteristic of quantitative circuit models) is that in creating the voltage redistribution process, we had used qualitative versions of the steady state circuit laws, i.e., Ohm's Law and Kirchhoff's Voltage Law. Unfortunately, if you use steady state laws to determine what the next steady state will be, you have not reduced the process to fundamental principles like F=ma (predicting the effects of electrical forces on the movement of charged particles).

REDUCTIONISTIC, PHYSICAL MODELS

We decided that we needed to unpack the voltage redistribution process via a dynamic physical model. This model would introduce a simple causal mechanism that would drive the process. The causal mechanism would be based on what diSessa (1983) refers to as phenomenological primitives, i.e., things that you can experience with your own body, such as pushes and pulls and a sense of equilibrium. Furthermore, it would be based on more fundamental physical laws such as F=ma and Coulomb's Law, rather than on the steady state circuit laws. Another important property of this model would be that by watching this physical mechanism run, the steady state circuit laws would emerge. In other words, it would be a model based on Coulomb's Law, and Ohm's and Kirchhoff's Laws would be emergent properties of the physical process embedded in the model.

It is interesting to note that when we surveyed physics text books, electrical engineering text books, and technical text books that are used in training electrical technicians, we were unable to find such a dynamic physical model. In two advanced physics text books, we did find a mathematical derivation of how you can start with F=ma and get Ohm's Law, but these derivations still focused on reasoning directly about the properties of circuits in a steady state, not about transient phenomena. The fact that we didn't find such a mechanism presented in any of the text books suggests that, from the perspective of physicists, either it is unnecessary to develop such a model of transient phenomena within circuits, or such dynamic models are so complex as to confuse rather than enlighten. However, while such transient models are not available for electricity, they are fairly common in the study of gas diffusion and heat flow. Moreover, while physicists may accept abstract, algebraic derivations of circuit laws without any necessary link to a transient, physical model, in the initial stages of learning most students find that mathematical abstractions and algebraic derivations make sense only after the domain is understood in causal terms. If a dynamic model is not supplied, they will attempt to build their own, or to interpret the circuit laws in terms of their prior conceptions of electricity.

We therefore focused on creating a model that would embody a simple mechanism which would help people to understand the origins of circuit behavior such the voltage redistribution process. The model is based on a simplification of Coulomb's law in which we quantize distance by dividing resistors up into little slices. We also ignore surface charges and regard the resistor as essentially one dimensional. Within each slice of the resistor, we represent the charge of that slice. For example, the middle slice (with the holes filled with minuses) shown in Figure 1 is neutral, the one on the right has a negative charge, and the one on the left has a positive charge. These circles and minuses are not meant to represent individual atoms and electrons. They are just a representation of the charge of that slice. If you prefer, you can think of them as being replaced with a number, so the middle slice has a charge of 0, the right is -5, and the left slice is +5. Now if you put two of these slices next to one another, and if there is a difference in their charge, then there will be an electrical force exerted on charged particles within the two adjacent slices. This can be thought of as due to the negative charges repelling one another and the positive charges attracting the negative charges. These forces will accelerate the mobile charges (the minuses), causing them to migrate (i.e., be redistributed) from the relatively more negatively charged slice to the more positively charged slice until both slices have the same charge.

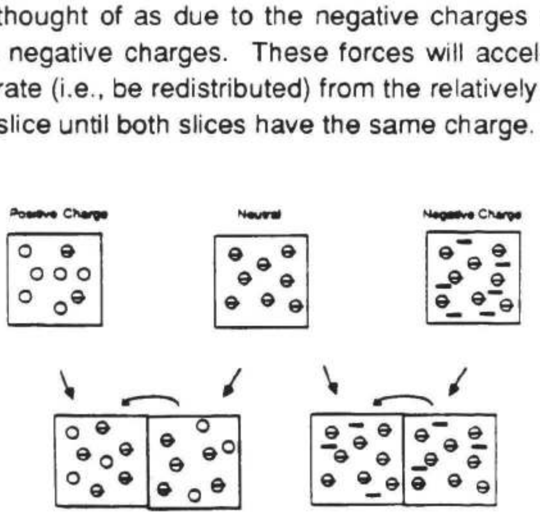

Figure 1. Slices of a resistor.

We can increase the resistance of a resistor by putting more and more of these slices next to one another. We model a battery by simply saying that it is a device which maintains a constant positive charge on one of its terminals, and a constant negative charge on the other of its terminals. Then, for example, if you attach the negative terminal of a battery to one end of a resistor a resistor as in Figure 2 (so you don't have a complete circuit), you can watch what happens over time. Within the model, time is also quantized. Within each time interval, a certain proportion of mobile charges will be redistributed, the actual number depending on the difference in charge between the two adjacent slices. As this model runs, one can see that on each time increment, adjacent slices go part way towards reaching equilibrium. In this way, one can watch the system settle down into a final, steady state. In this case, there will be no voltage drop across the resistor and the source voltage will be developed across the open segment of the circuit. Furthermore, if you put together a complete circuit, like the one shown in Figure 3, and let the process run until it reaches a steady state, you see how you get Ohm's Law and Kirchhoff's Laws from a system that behaves in accordance with Coulomb's Law. This is a nice example of how you can start with some very simple ideas about attraction, repulsion, and equilibrium, and show how a system that operates according to a mechanism based on these ideas will settle down into a steady state that obeys Ohm's Law and Kirchhoff's Laws.

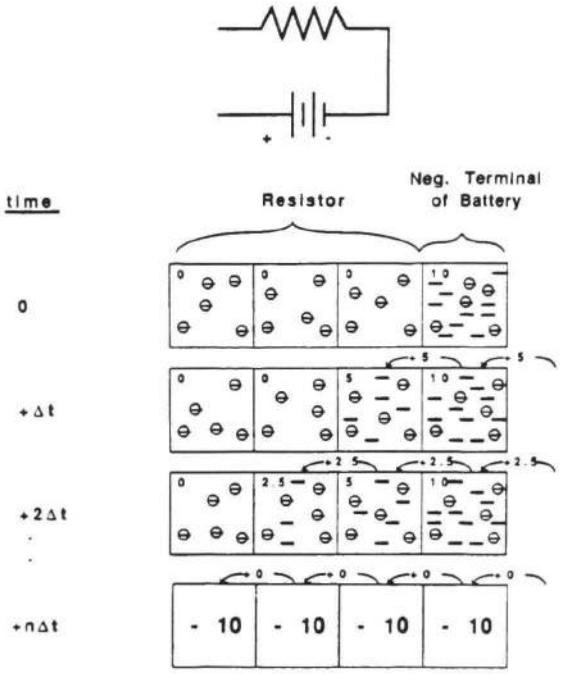

Figure 2. An open circuit.

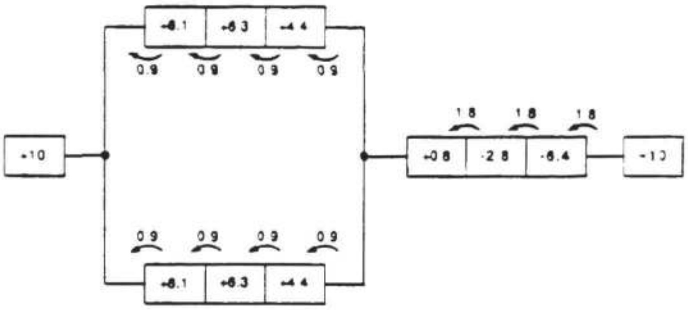

Figure 3. A series-parallel circuit.

This model makes the qualitative models described earlier "make sense". It provides a sense of mechanism, so that the macroscopic models, combined with this unpacking of the voltage redistribution process, provide a coherent, although simplified, view of how circuits work. This model is, however, far from "complete". Firstly, there is no model provided to explain how electrical force gets produced by a battery. Secondly, the holes and minuses are not meant to represent individual atoms and electrons. Instead they provide a representation of the charge of each slice of a resistor. And, unless you actually represent the individual electrons and animate them by showing them moving and bumping into atoms, you do not get a good sense of what resistance is. So, one could hypothesize that we need to unpack one more level, and show the behavior of atoms and electrons. However, we argue that this is unnecessary. People already have good intuitions about resistance. The hard and key ideas are voltage drop and the process by which voltages get redistributed. This model represents that concept and process succinctly.

By interacting with such models, students can also come to appreciate some important distinctions among physical theories: The laws that are used to describe the steady state are usually redundant. They are, if you will, multiple constraints on the configuration of charges in the steady state. Given their redundancy, if several statements are given, one can derive the remaining ones (this is an example of constraint-based reasoning). In this way, students can be introduced to constraint-based reasoning in a qualitative form as a precursor to its algebraic forms. In general, students can come to appreciate epistemological distinctions among forms of physical theories and ways of reasoning, and learn how they can exist side by side as alternative ways of viewing the physical world.

Another interesting property of this process model emerged when describing it to physicists. We would say, "This is a diffusion model which represents a first approximation to what happens in electrical circuits. It is the kind of model that you might use, for example, to explain heat flow to someone." The physicist would say, " Oh yes, you could use it to explain what happens in gases as well". These discussions implied that there are a small number of key processes, like this diffusion process. These key processes are understandable because they can be grounded in phenomenological primitives (i.e., can introduce a simple mechanism that relates to things you can do with your own body). They also have the property that they can be used to help students understand a wide range of physical domains. Thus they can play a key role in enabling students to learn.

CONCLUSIONS

To summarize the viewpoint that has evolved in this research: We hypothesize that students need to start with a dynamic, physical model that provides a sense of mechanism and introduces domain causality. Such models are, however, not particularly useful for problem solving. Thus, in addition, one needs to acquire more abstract, behavioral models that are reasoning at a more macroscopic level. At the level, for example, of switches closing causing lights to go on and off. These kinds of models are useful not only for understanding, but also for problem solving, and need to be introduced in both qualitative and quantitative forms. However, causality and mechanism are not the only perspectives that people employ to understand how systems work. There is another perspective that people often impose on physical systems, and that is the perspective of "purpose": What is the system designed to do, and how do the subsystems within it work together to achieve that purpose? This represents an important way of coming to understand systems, particularly in the case of designed artifacts, and we have been doing further research concerning how to introduce functional models into our intelligent learning environment.

The intelligent learning environment incorporating a progression of qualitative models has been tried out with seven subjects. It was successful in all seven cases in enabling novices to mentally simulate and to troubleshoot circuits. When subjects were not provided with a version of the reductionistic, physical model, they spontaneously invented their own reductionistic explanations which were incorrect and usually inconsistent with the causality embedded in the macroscopic models (White & Frederiksen, 1987b).

In helping students to acquire this set of mental models, we have the same objectives that physicists have in evolving such models: We want students to possess a set of coherent models that will enable them to predict and understand what will happen across a wide range of domain phenomena. In the case of electrical circuits, the models we have described allow students to mentally simulate circuit behavior from different, but coordinated, perspectives. For instance, the causality of our microscopic models is consistent with that of our macroscopic models of circuit behavior. Students can employ this set of models to make predictions about circuit behavior, as well as to design and troubleshoot circuits. The models thus provide a coherent, linked, knowledge structure that can generalize across a range of situations to solve problems and to generate explanations.

REFERENCES

Brown, J.S. & deKleer, J. (1985). A qualitative physics based upon confluences. In Bobrow, D.G. (Ed.). Qualitative Reasoning about Physical Systems. Cambridge, MA: MIT Press.

diSessa, A. (1983). Phenomenology and the evolution of intuition. In D. Gentner & A. Stevens (Editors), Mental Models. Hillsdale, NJ: Lawrence Erlbaum Associates.

Forbus, K.D. (1985). Qualitative process theory. In Bobrow, D.G. (Ed.). Qualitative Reasoning about Physical Systems. Cambridge, MA: MIT Press.

Frederiksen, J., & White, B. (in press). "Intelligent tutors as intelligent testers." To appear in Diagnostic Monitoring of Skill and Knowledge Acquisition. N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Editors), Lawrence Erlbaum Associates, Hillsdale, New Jersey.

White, B., & Frederiksen, J. (1986). "Intelligent tutoring systems based upon qualitative model evolutions." In the Proceedings of the Fifth National Conference on Artificial Intelligence, Philadelphia, Pennsylvania.

White, B., & Frederiksen, J. (1987a). "Qualitative models and intelligent learning environments." In AI and Education. R. Lawler & M. Yazdani (Editors), Ablex Publishing Corporation.

White, B., & Frederiksen, J. (1987b). "Causal Model Progressions as a Foundation for Intelligent Learning Environments." Report No. 6686, BBN Labs, Cambridge, MA. To appear in Artificial Intelligence.