Diagrams as Physical Models to Assist in Reasoning

BALAKRISHNAN CHANDRASEKARAN

ABSTRACT. Diagrams serve a variety of roles in applied reasoning. At times, they are best viewed as “sentences” in a 2-D language, with specialized rules of inference that provide new diagrams. Other times, however, they are best understood as providing a physical model, much like an architectural model of a building or a 3-D molecular model of a chemical compound, for a state of affairs. In this paper, I discuss the notion of a physical model for a logical sentence, and the role played by the causal structure of the physical medium in making the given sentence as well as a set of implied sentences true. When the physical model is prototypical, it supports the inference of certain other sentences for which it provides a model as well. I also informally discuss a proposal that diagrams and similar physical models help to explicate a certain sense of relevance in inference, an intuition that so-called Relevance Logics attempt to capture.

1 Introduction

In this paper I wish to consider some issues related to the use of diagrams as aids in reasoning. A particular use of diagrams corresponds to treating the diagram as a physical model for the premises. The problem solver sees that the representation is also a model for another assertion that is not explicitly part of the premises, and concludes that the assertion follows from the premises. The phenomenon of interest is not specific to diagrams – other examples of external representations as part of problem solving include 3-dimensional architectural models during design and physical 3-dimensional molecular models in problem solving in chemical engineering – but diagrams are ubiquitous, so we will largely focus on them in the paper. Regarding a main concern of logic – accounting for justifiable inferences – this style of reasoning based on a physical model needs to be part of any account of natural reasoning. It puzzles me that more has not been said in logic about the use of such physical models as aids to reasoning, given how prevalent diagrams are in everyday as well as professional reasoning, and the role played by architectural and molecular models in their respective disciplines.
So this paper’s goal is to raise the profile of physical models in logic. I raise a set of issues for deeper consideration by logicians.

Before I get to the main issues of concern, I take a brief detour on the multiple, and at times contradictory, senses of the term “model” in science, logic, and cognitive science. This seems appropriate in a conference devoted to model-based reasoning.

2 Various senses of “model”

The term “model-based reasoning” is in a sense redundant – all thinking and reasoning is model-based in the sense that these processes involve making use of knowledge (or beliefs) about the domain of discourse, i.e., a model of the domain. Craik [Craik, 1983] specifically characterized cognitive activity in general as model–based in this sense, though he left unspecified the form in which the model was represented.

The term has been used in multiple, sometimes opposite ways, leading to much confusion. In one sense, a domain is a model of a description; in another usage, a description is offered as a model of a domain. For example, in logic, a domain provides a model of a set of axioms, e.g., arithmetic is a model of Peano’s Axioms and plane geometry is a model of Euclidean Axioms. If the axioms are a description, the domain that fits the description is a model of the description. A different but related usage is a model for a sentence that is constructed by assigning truth values to the elements of the Herbrand Universe.

In philosophy and practice of science, the direction is from domain to description. A description – a set of equations, e.g. – is a model of a domain if the description can be used to predict phenomena in the domain. Thus, Maxwell’s Equations model electro-magnetic phenomena and physicists speak of the Newtonian model versus the Einsteinian model.

In cognitive science the term “mental model” [Johnson-Laird, 1983] is used to describe a postulated internal diagram-like representation during syllogistic reasoning, containing a model of the situation. This is close, but not identical, to the sense of model in logic. There are additional usages, such as the phrase “model-based reasoning” as it was used, in the 1970-80’s, in the literature on diagnostic reasoning in AI. One type of representation was a collection of heuristic relations between observations and diagnostic hypotheses. This was called rule-based, and was contrasted with a model-based representation that made use of knowledge about the functions and the structure, i.e., the components and their connectivity, of the device being diagnosed [Chandrasekaran and Mittal, 1983; Chandrasekaran et al., 1989]. Perhaps the term “structural model-based” might have been more descriptive, since even the set of diagnostic rules in
the “rule-based” approach still constitutes a model of the device.

As mentioned, in this paper, I am concerned with the use of diagrams to assist problem solving and reasoning, where the diagram is a model in the sense of logic, a “situation” whose particulars make the sentence true. However, I wish to look at the diagram as a physical model.

3 Roles of diagrams in reasoning

Diagrams perform many different types of assistance during problem solving. We identify four roles here: helping extend short term memory, helping organize problem solving by spatial organization of related variables, providing a model of the premises so that plausible subtasks in theorem proving may be hypothesized, and providing a model of the premises from which consequents can be inferred and asserted.

First, they extend short term memory, by providing a spatially organized external location in which to note down information. Second, they help organize problem solving. Larkin and Simon [Larkin and Simon, 1987] use the example of analyzing a pulley system – they show how the diagram of the pulley system helps the problem solver organize the sequence of equations to solve, or variables to assign values to. The problem solver can use his visual perception to locate the pulley that a strip of rope goes over, and thus to choose which tension variable to consider next.

A diagram might provide a model of the set of premises, i.e., it depicts a situation that satisfies the description. In the third role for diagrams, and the first role for diagrams as models, the model suggests hypotheses. This is exemplified by diagrams as used in proving theorems in Euclid. In this kind of use, a diagram is a model of the premises.

The problem solver knows that it is not a general model, i.e., while what is asserted in the premises is true in the diagram, all that is true in the diagram doesn’t necessarily follow from the premises. In any case, nothing can be asserted on the basis that it is true in the diagram. Nevertheless, it can provide information that can be used to set up and select subtasks. For example, the fact that two angles are adjacent, and the theorem involves one of the two adjacent angles might suggest to the theorem prover that perhaps stored theorems involving adjacent angles may be useful in advancing the proof. Lindsay [1988] provides a review of the issues in the use of diagrams in geometry theorem proving. It has been estimated that the use of the diagram in this way reduces the search space by an order of 300 or more. In the traditional use of Venn or Euler Diagrams in proving theorems in Set Theory, the diagrams play a similar role. It is important to emphasize that the information from the model is not asserted as conclusion, but only used to find other strategies for arriving at the general conclusion. It
may be pointed out in passing that the heuristic use of Venn and Euler diagrams in proving propositions in Set Theory stimulated a direction of research by Barwise and associates [Allwein and Barwise, 1996] on new proof procedures in which certain sequences of diagrams can be considered proofs in themselves. As I understand it, the diagrams are viewed as “diagrammatic sentences”, i.e., as two-dimensional syntax-controlled compositions of diagrammatic symbols\(^1\).

The fourth role for diagrams is that they directly support inference, when they are models of the premises. The information obtained from the model is asserted as conclusion in the general case, though exactly when and how much to generalize are issues for which answers differ from one diagrammatic application to another. This role of diagrams is my focus in this paper.

Let us consider two very simple examples. Given a simple addition problem in arithmetic, say to show \(1 + 3 = 2 + 2\), suppose we draw four points (or arrange four stones on the ground) as below:

\[
\begin{array}{cc}
\ast & \ast \\
\ast & \ast \\
\end{array}
\]

Figure 1.

Under the appropriate mappings, the situation is a model of \(1 + 3\). But it is also a model of \(2 + 2\). We can demonstrate to a child that \(1 + 3 = 2 + 2\) by using the above diagram. Here the generalization issue is trivial: the child could, but typically wouldn’t, say, “Maybe this is true when we add 1 star to 3 stars, but is it true when we add 1 slice of pizza to 3 slices of pizza?” . Our intuitions about numbers seem sufficiently robust that this issue doesn’t arise in a child or an adult. “Individuals that keep their distinct identity” seems to be the background intuition that is operational here, and using that we generalize from star marks on paper or stones on the ground to numbers in general.

---

\(^1\)The Stanford Encyclopedia of Philosophy entry on Model Theory (http://plato.stanford.edu/entries/model-theory/) says, “[...] the overwhelming tendency of this work is to see pictures and diagrams as a form of language rather than as a form of structure. For example Eric Hammer and Norman Danner (in the book edited by Allwein and Barwise) describe a ‘model theory of Venn diagrams’; the Venn diagrams themselves are the syntax, and the model theory is a set-theoretical explanation of their meaning”. It is worth noting, however, that the Hyperproof work by Barwise and associates is not in the “diagrams as sentences” framework. It is closer to the diagrams as models view of the current paper.
Let us consider another example, one that we will use often in the rest of the paper. Given “If $A$ is to the left of $B$, $B$ is to the left of $C$, is $A$ to the left of $C$?”, people often draw a diagram as in Figure 2:

![Figure 2](image)

There is a natural sense in which the physical diagram is a model of the problem situation\(^2\). The problem solver notices that indeed $A$ is to the left of $C$, and declares that the inference is true. Of course, the diagram only represents one specific way in which the points can be located to provide a model. Yet the problem solver makes bold to assert that the conclusion is true for all the specific ways in which the points could be located. Let us first consider how to relate the notion of physical models to the sense of model in predicate logic.

4 Physical fragments providing models for logical sentences

Let us start by restating the standard definition of a model for a sentence in Logic. An interpretation for a sentence $S$ consists of:

- A non-empty, possibly infinite, domain $D$ of individuals
- Assignment of specific individuals in $D$ to constant symbols in $S$
- Assignment to each $n$-ary function symbol in $S$ of an $n$-ary function that maps from $D^n$ to $D$.
- Assignment to each $n$-ary predicate symbol in $S$ of an $n$-ary function that maps from $D^n$ to \{True, False\}.

An interpretation for $S$ is a model for it if $S$ evaluates to True under the interpretation.

\(^2\)More precisely, it a model of the conjunct of the given premise with axioms that capture the structure of space in terms of which the predicate Left is defined. For someone for whom the semantics of Left is that of spatially left in ordinary language, the axioms are implicit, and the Figure provides a model for the premise.
4.1 Modeling physical things

Our goal here is to show how a physical entity may be used to provide an interpretation of a sentence.

**Domain of Individuals.** Let Π be a fragment of physical world. Let Δ_Π : \{π_1, π_2, \ldots\} be a (possibly infinite) set of entities, each π_i a part — a subfragment — of Π. The entities need not be physically disjoint — one entity may be a physical part of another entity; nor it is necessary for Δ_Π to exhaust Π, i.e., the totality of physical fragments represented by the elements of Δ_Π to be equal to the matter represented by Π.

We will use two concrete examples to illustrate the ideas: Π_1, the set of points constituting a finite physical horizontal straight line, say drawn on a piece of paper; and Π_2, a physical object intended to be an architectural “model” of a house.

Example Π_1: The entire finite straight line is Π. Each point in it is a π, thus Π has an infinite number of parts in this model. Another model for the same physical object might subdivide the line into various segments, each providing a π.

Example Π_2: The physical entity (the architectural “model”) as a whole is Π, and the physical matter corresponding to various rooms, walls, doors, etc., are the π’s.

**Functions and Predicates.** Let \{φ_i\}_{0 \leq i \leq k} be a finite set of functions of various arities, such that if φ_i is n-ary, it is a function from Δ^n_Π to Δ_Π. Similarly, let \{ρ_i\}_{0 \leq i \leq l} be a set of functions of various arities, such that if ρ_i is n-ary, it is a function from Δ^n_Π to \{T, F\}. The ρ’s are predicates defined on the physical variables, and thus the values that they take for their various arguments is determined by the causal structure of Π.

Example Π_1: The function, right_1(π_i), defined as “the point that is exactly 2 inches to the right; if there is no such point, the right end point” is a unary function. Example of a binary predicate ρ is: Left(π_i, π_j), with the obvious interpretation.

Example Π_2: Unary function Entrance-to(room_i), which takes values from the subset of parts of type “door”. Thus, e.g., Entrance-to(room_6) = door_6.

Example of ρ: Bigger-than(room_i, room_j) is a binary function which evaluates to True if the area of room_i is larger than that of room_j, and False otherwise.

**Properties and Causal Structure of Π.** For the purpose at hand, the physical structure is modeled in terms of a set of variables, selected attributes of the
physical system. A specific physical instance will have specific values for these variables. Let \( \Theta_i : \{ \theta_{i1}, \theta_{i2}, \ldots, \theta_{ik}, \ldots \} \) be a set of variables in terms of which entity \( \pi_i \) is modeled, and let \( \Theta = \bigcup_{i=1}^{n} \Theta_i \). The causal structure of \( \Pi \), which constrains the values of the variables in \( \Theta \), determines the truth values of the various predicates for various values for their arguments, and thus the truth values of sentences composed out of these predicates.

Thus, part of modeling a physical fragment for the purpose of providing an interpretation for a sentence involves identifying a physical system with the right properties to provide an interpretation, and then setting its parameters to where the fragments provides a model for the sentence.

Example \( \Pi_1 \): Let part \( \pi_i \) be modeled in terms of a single variable \( x_i \), the \( x \)-coordinate of \( \pi_i \) from some origin. \( \text{Left}(\pi_i, \pi_j) \) is defined by the values of \( x_i \) and \( x_j \). Additionally, the constraints of the physical line result in constraints between predicates: if \( \text{Left}(\pi_i, \pi_j) \) and \( \text{Left}(\pi_j, \pi_k) \) are both True, then \( \text{Left}(\pi_i, \pi_k) \) is constrained to be True.

Example \( \Pi_2 \): The parts of the house may be modeled in terms of their length, width, height, area, etc. Color and material out of which a part is made may also be in the set of variables. Whether room5 is larger than room3 is fully determined by the physical dimensions of \( \Pi \); there is no additional freedom to assign \( T \) or \( F \). Additionally, if in a physical architectural model room1 is larger than room2 which in turn is larger than room3, the model will necessarily satisfy the predicate, larger-than(room1, room3).

In order to avoid confusion between different usages in science and engineering on one hand and in logic on the other, we use the term \( p \)-model to refer to a description of a physical entity as in the next definition.

**Definition.** A \( p \)-model of a physical fragment \( \Pi \) consists of the following specifications:

- \( \Delta_\Pi \), a set of individuals consisting of parts of \( \Pi \).
- A set \( \{ \phi \} \) of functions of various arities, such that an \( n \)-ary function is a mapping from \( \Delta_\Pi^n \) to \( \Delta_\Pi \).
- A set of functions \( \{ \rho_i \} \) of various arities, such that an \( n \)-ary function is a mapping from \( \Delta_\Pi^n \) to \( \{ T, F \} \).
- A set of variables \( \Theta \) in terms of which \( \Pi \) and elements of \( \Delta_\Pi \) are modeled; a causal structure \( \Pi_{ax} \), that defines the causal constraints between the variables in \( \Theta \).

**Remark.** There is an infinity of \( p \)-models for a given physical entity.
4.2 A physical entity supporting a logical model

Let a p-model $M_\Pi$ of a physical fragment provide an interpretation for a sentence $S$, i.e., each constant symbol in $S$ is mapped to a specified element in $\Delta_\Pi$, each $n$-ary function symbol in $S$ is mapped to a function in $\{\phi_k\}$, mapping from $\Delta^n_\Pi$ to $\Delta_\Pi$, and each $n$-ary predicate symbol in $S$ is mapped to a function in $\{\rho_l\}$, mapping from $\Delta^n_\Pi$ to $\{\text{True, False}\}$.

**Definition.** If a sentence $S$ evaluates to True under the interpretation provided by a p-model $M_\Pi$ of a physical fragment $\Pi$, we say that $\Pi$ provides a physical model for $S$.

**Remark.** What makes a predicate true or false in a physical model is that the variables take specific values in the physical fragment, and the causal structure $\Pi_{ax}$ constrains values between variables.

**Examples**

Consider the following sentence $S$:

$$\forall x \forall y \forall z (L(x, y) \& L(y, z) \rightarrow L(x, z))$$

Let $\Pi$ be a physical 1-D spatial line fragment, and let the following be a p-model $M_\Pi$ for $\Pi$.

- $\Delta_\Pi$: the (infinite) set of points in the line fragment, $\{x_i\}$.
- $\Theta$: a single attribute, the co-ordinate of a point $x_i$ with respect to some origin.
- $\{\phi\}$: null set.
- $\{\rho_l\}$: a single function, $\text{Less-than}(x_i, x_j) = \text{True}$, if the co-ordinate of $x_i$ is less than that of $x_j$; False, otherwise.

Under the interpretation $M_\Pi$, $S$ True in $\Pi$. A physical 1-dimensional line fragment is thus a physical model for $S$.

As more complex example, consider $S'$:

$$[\forall x \forall y \forall z (L(x, y) \& L(y, z) \rightarrow L(x, z)) \&
(\forall x \forall y (L(x, y) \& L(y, x) \rightarrow \text{Eq}(x, y))) \&
L(A, B) \& L(B, C)$$

Consider the physical diagram in Figure 2, with the following $M'_\Pi$. 
$M'_\Pi$: $M_\Pi$ as defined in (2) plus the following assignments: Constant $A$, $B$ and $C$ assigned to the points in the 1-dimensional line fragment corresponding to the coordinates as in the Figure. Eq($x, y$) assigned to function “Equal($x_i, x_j$) = True iff $x_i$ is the same as $x_j$, and False otherwise”.

(4)

Under the interpretation $M'_\Pi$, $S'$ evaluates to True, so the diagram in Figure 2 is a physical model for $S'$. Readers will recognize (3) as a simple axiomatization of left-ness plus the premises of the problem we stated at the beginning. $M'_\Pi$ is also a model for the following:

\[
[(\forall x \forall y \forall z (L(x, y) \& L(y, z) \rightarrow L(x, z)) \&
[(\forall x \forall y (L(x, y) \& L(y, x) \rightarrow Eq(x, y))) \&
L(A, C)
\]

(5)

Remark. In applied reasoning, the agent is reasoning in some domain of interest, $D$, and he is interested in making a model of a sentence, say $S$. Let $D_{ax}$ be the set of axioms that describe the relevant aspects of the domain of interest. Thus, the agent is looking for a physical model of $D_{ax} \& S$. When we say that Figure 2 is a model of Left($A, B$) & Left($B, C$), it is because we interpret Left in the spatial meaning of the terms. This interpretation assumes $D_{ax}$. If instead $S$ were Goo($A, B$) & Goo($B, C$), we wouldn’t see Figure 2 as its physical model. Successfully making a physical model of $S$ when the agent is reasoning in $D$ involves finding a physical medium such that its causal structure $\Pi_{ax}$ has the right kind of homomorphism relation with $D_{ax}$.

There is no requirement that an arbitrary $\Pi$ have a p-model that provides an interpretation for an arbitrary sentence $S$. In fact, it is a special situation where a physical model can be constructed so as to provide an interpretation for a sentence. In the next section, we discuss how such physical models are often used.

Warrant for generalization

Figure 2 is a model for $S'$, but it is just one model. There are infinitely many configurations of points for which the corresponding physical diagrams will provide a model for $S'$. Nevertheless, we generalize the inference to a class of situations. Figure 2 also provides a model for “$A$ is farther left of $B$ than $B$ is of $C$”, but we know that this inference cannot be generalized.

This is quite common in applied reasoning. A chemist, who is considering whether $S_1 \rightarrow S_2$, where $S_1$ and $S_2$ are sentences in his domain, might construct a chemical reaction which is a model of $S_1$ (really a model of his
domain axioms and $S_1$), see if it is also provides a model for $S_2$, and, though the specific chemicals in interaction model only instances of $S_1$, generalize to the larger class. Of course, a good chemist would know just want sort of model to construct that would bear the generalization.

This style of proof might be called physical-model-based proof. The Model-Based Rule of Inference may be stated as follows:

Given an inference problem, $S_1 \rightarrow S_2$, where $S_2$ is not a logical truth, in domain $D$ with domain theory $D_{ax}$, and given a physical fragment $\Pi$ such that it provides a p-model $M_\Pi$ that satisfies $D_{ax} \& S_1$, if $M_\Pi$ also satisfies $D_{ax} \& S_2$, and if $M_\Pi$ has warrant for generalization with respect to the inference $S_2$, conclude $S_1 \rightarrow S_2$ in the general case in $D$.

One might use the term prototypical to describe a model that provides such a warrant for generalization.

Remark. The rule of inference blocks asserting a logical truth based on the physical model. This is because we wish the physical model to play a role in the assertion. This is related to our remarks on Relevance Logics in the next section.

In many cases, the applied reasoner has limited or no access to $D_{ax}$ is an explicit form. However, he has a body of intuitions and practices that help him construct prototypical models for classes of $S$'s that provide warrant for generalization and help him scope them.

Let $S$ be a sentence in a domain $D$ characterized by axioms $D_{ax}$. Let $\text{Closure}(D_{ax} \& S)$ be the set of all inferences that are deducible from $D_{ax} \& S$. If $\Pi$ provides a model for $S$, it will also provide a model for all elements of $\text{Closure}(D_{ax} \& S)$. However, it will also provide a model for many other inferences that are not in $\text{Closure}(D_{ax} \& S)$. The reasoning agent needs to know how not to make the inferences that are not in $\text{Closure}(D_{ax} \& S)$, even though the model supports it, e.g., not to infer “$A$ is farther left of $B$ than $B$ is of $C$” from Figure 2.

5 Prototypical models

What makes a prototypical model? The specifics depend on the domain and the predicates of interest, but some general intuitions may be useful. The following ideas might help in the development of a more formal account.

The first idea is minimality. Let $S$ be $\text{Left}(A, B) \& \text{Left}(B, C)$ (we are implicitly in the domain of 1-d space with a directional axis). Just as Figure 2, Figure 3 also provides a model for $S$. However, it provides a model for unrelated things such as $\text{Inside}(D, E)$. Clearly, any inference based on this model, such as $\text{Left}(A, B) \& \text{Left}(B, C) \rightarrow \text{Inside}(D, E)$, would be
a mistake. Figure 2 is in some sense minimal compared to Figure 3 for \( \text{Left}(A, B) \& \text{Left}(B, C) \).

![Diagram](image)

Figure 3.

The next idea is that of multiple prototype models. Let \( S' \) be \( \text{Left}(A, B) \& \text{Left}(A, C) \). While Figure 2 provides a model for \( S' \), it doesn’t seem prototypical for another reason: it only accounts for a subset of instances. Figure 4 provides another model.

![Diagram](image)

Figure 4.

Figure 2 provides a model for \( \text{Left}(B, C) \) and Figure 4 provides a model for \( \text{Left}(C, B) \), neither of which follows from \( S' \), thus neither of these inferences have a warrant for generalization. Applied reasoning in this case requires that two models be set up, each of which allows certain inferences, say \( \text{Right}(B, A) \), but not \( \text{Left}(B, C) \) or \( \text{Left}(C, B) \).

The third idea is a revisit of what we mentioned earlier, that Figure 2 doesn’t support the generalization of “\( A \) is farther left of \( B \) than \( B \) is of \( C \)”, though the figure provides a model for it. Suppose a new predicate \( \text{boogoo}(x, y, z) \) is defined as “\( x \) is farther left of \( y \) than \( y \) is of \( z \)”. Consider \( S'' \): \( \text{Left}(A, B) \& \text{Left}(B, C) \& \text{Left}(C, D) \& \text{Left}(D, E) \), modeled by Figure 5. Figure 5 also provides a model for \( \text{boogoo}(A, C, D) \), which has a warrant for generalization, unlike \( \text{boogoo}(A, B, C) \) in Figure 2.

![Diagram](image)

Figure 5.

It can be seen, from a metatheory of \( \text{Left}(x, y) \), that it provides an or-
der between $x$ and $y$. Thus any conjunctions of $\text{Left}(x, y)$’s would specify an order or alternate possible orders. Only information that follows from the order information has warrant for generalization. In every domain, in principle such metatheories may be constructed, but in practice, an agent performing applied reasoning in some domain usually has no access to such metatheories. Even when he has access to the axioms for his domain, such as in the sciences, they are provisional and potentially revisable. So, reasoning in the practical world is aided by models constructed and interpreted with the aid of intuitions based on experience, training and partial theories. Such models play a large role in commonsense reasoning as well, as evidenced by the research of [Johnson-Laird, 1983] on how people solve syllogistic problems by constructing mental models that have many points of contact with the models that we describe here. Because people lack access to fully worked-out metatheories, some of the reasoning errors that occur in practical reasoning are due to mistakes in the application of generalization and in the construction of prototypical models.

Our everyday reasoning as well as reasoning in professional disciplines is full of implicit and explicit guidelines about how to construct diagrams and physical models that give the desired information perceptually and how to generalize. The ubiquity of such diagrams in our reasoning should not keep us from appreciating the hard-won nature of discoveries of appropriate diagrams for classes of problems. Such discoveries are prized – transmitted culturally for everyday reasoning, and made part of training in professional disciplines, with discoverers often honored with awards.

6 Relevance logics and physical models

Relevance Logics are the subject of a substantial body of work, a summary of which is available in Stanford Encyclopedia of Philosophy [Encyclopedia, 2005]. My knowledge of the area is limited and thus my goal will be correspondingly modest – to consider, informally, possible connections between the idea behind Relevance Logics and physical models so as to invite further attention to this connection.

Relevance logics are a response to what some people take to be paradoxes of traditional implication. The paradox arises in that in some of the inferences authorized by the semantics of traditional implication the antecedent doesn’t seem relevant to the consequent. Examples of paradoxes mentioned in the Stanford Encyclopedia page are:

Material implication paradoxes:

- $p \implies (q \implies p)$
- $\neg p \implies (p \implies q)$
• \((p \rightarrow q) \lor (q \rightarrow r)\)

Strict implication:
• \((p \& \sim p) \rightarrow q\)
• \(p \rightarrow (q \rightarrow q)\)
• \(p \rightarrow (q \lor \sim q)\)

Now suppose a reasoning agent performing domain-specific reasoning uses prototypical physical models to assert consequences, as described in earlier sections of the paper.

The consequences he asserts based on physical models, at least for the examples we will consider, will not fail the test of relevance. That is, given \(p\), he constructs a prototypical model for it in his domain \(W\). If the model is also a model for \(S\), and if \(S\) has a warrant for generalization, \(p \rightarrow S\) will be a relevant inference in \(W\). Technically, any model is also a model for logical truths, so we add a rule forbidding asserting a logical truth as a consequence of any sentence.

Consider as an example:
\[p \rightarrow (q \rightarrow p)\] (6)

In this case \(p\) doesn’t seem to play a relevant role in making \((q \rightarrow p)\) true. However, suppose there is a domain \(D\) in which \(q\) is a possible cause for \(p\), and that \(q\) would definitely cause \(p\). In that domain, it wouldn’t be a surprise to assert that if \(p\) is true, then if \(q\) is known to be present, \(q\) caused \(p\), and consequently that the truth of \(q\) would imply the truth of \(p\). If \(D_{ax}\) denotes the axioms characterizing \(D\), the following would not fail the test of relevance:
\[(D_{ax} \& p) \rightarrow (q \rightarrow p)\] (7)

Suppose \(D\) describes a physical domain in which we construct a physical fragment \(\Pi\) that provides a \(p\)-model for \(q\). \(\Pi\), under the same \(p\)-model mappings, would also provide a model for \(p\). That is, there is no way to construct a model for \(q\) in \(D\) without it being a model for \(p\) as well. In this case, there is no failure of relevance to assert (6). Similarly, consider an example of strict implication,
\[(p \& \sim p) \rightarrow q\] (8)

There would be no way to make a physical fragment in any domain that provides a model for \(p \& \sim p\). No one will be able to assert (8) based on a physical model.
A physical model constructed for \( \sim p \) will by definition not provide a model for \( p \rightarrow q \), so based on the Model-Based Rule of Inference, we won’t be able to assert \( \sim p \rightarrow (p \rightarrow q) \).

Suppose in some domain \( D \) it is the case that \( (p \rightarrow q) \lor (q \rightarrow r) \), i.e., \( D_{ax} \rightarrow ((p \rightarrow q) \lor (q \rightarrow r)) \). In this domain, any model that supports \( p \) will be a model for \( q \), and any model that supports \( q \) will be a model for \( r \). The inference calls for two prototypical models, one created to support \( p \) and the other for \( q \). The agent will correctly assert \( (p \rightarrow q) \lor (q \rightarrow r) \), and the implications are relevant.

Consider the strict implication:

\[
p \rightarrow (q \rightarrow q)
\]  

(9)

A prototypical model constructed for \( p \) would not be a model for \( q \), but it would trivially be a model for \( q \rightarrow q \). However, because of the rule that logical truths cannot be asserted as consequences of anything from a physical model, the agent will not be able assert (9). If \( p \) is \( \text{Left}(A, B) \) and \( q \), \( \text{Inside}(C, D) \), with the obvious interpretations for the predicates in the spatial domain, the prototypical model for \( p \) will simply be two objects with labels \( A \) and \( B \), with \( A \) to the left of \( B \). This diagram will not provide a model for \( \text{Inside}(C, D) \) and thus the applied reasoner will not assert (9) based on the physical models. The same rule will also block the agent from asserting: \( p \rightarrow (q \lor \sim q) \).

Our discussion above leads to the conclusion that judgments of relevance with respect to implications are with respect to specific domains, whose structures (causal structures if domains are physical) then can be judged to play or not play a role in the antecedent making the consequent true. That is, the issue of relevance in the various implications that we considered arises in reasoning in specific domains. Further, making such judgments is facilitated in the specific domains by constructing physical models when possible.

7 Concluding remarks

My interest in the issue of models in reasoning arose in the context of my interest in the use of diagrams. Diagrams are ubiquitous reasoning aids in many situations. I have been interested in the idea that diagrams are just the most prominent example of larger class of reasoning aids that provide physical models to premises in some domain. My goal in this paper has been to explore the logic of such physical models. Applied reasoning, where reasoning agents are concerned with inferences in specific domains rather than abstract notions of validity, has not drawn as much attention from logicians as it should. I think, however, that the use of such physical models
in applied reasoning raises important issues in logic. I have attempted to formalize the notion of some piece of physical matter providing a model for a sentence. I have identified a proof technique called physical model based inference in which the prototypical models in specific domains are constructed that support useful generalizations. Developing reasoning skills in various domains includes learning or developing intuitions about how to construct prototypical models for specific reasoning situations. I also related such physical model-based reasoning to issues in Relevance Logics, where the goal is to identify when and how antecedents can be said to have a role in the consequent being true. Physical models, by incorporating the underlying causality of the domain, make it possible, under many conditions, to see whether or not the antecedents play a role in an implication being valid. My goal in this paper is to invite the attention of logicians more expert than I to look into what I think are important problems.

Model-based reasoning of the kind I have discussed is not merely an issue in logic, but in artificial intelligence. AI has focused almost exclusively on what might be called linguistic representations, mirroring the logical form of natural language sentences. However, real reasoning in humans is multi-modal, with perceptual and kinesthetic modalities often contributing to problem solving. Diagrams provide an important window into such multi-modal representations. In [Chandrasekaran et al., 2004], we describe a diagrammatic representation and reasoning architecture that integrates traditional symbolic reasoning with diagrammatic reasoning.

Acknowledgments

This paper was prepared through participation in the Advanced Decision Architectures Collaborative Technology Alliance sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAD19-01-2-0009, and by federal flow-through by the Department of Defense under contract FA8652-03-3-0005 (as a subcontract from Wright State University and Wright Brothers Institute). I am indebted to Peter Schroeder-Heister for his comments on an earlier draft, specifically his suggestions on prototypical models and Relevance Logics. My comments here are elaborations on his basic points. I thank Gerard Allwein, Neil Tennant, Stewart Shapiro, and the anonymous reviewer that Lorenzo Magnani asked to read the paper, for helping in various ways to make this paper have fewer errors than would be the case otherwise. Finally, I thank Lorenzo for inviting me to the conference and for his organizational leadership in this area of research.
BIBLIOGRAPHY


Balakrishnan Chandrasekaran
Laboratory for AI Research
Department of Computer Science and Engineering
The Ohio State University
Columbus, OH, USA
Email: chandra@cse.ohio-state.edu