Understanding Behavior Using Consolidation

Tom Bylander and B. Chandrasekaran

Laboratory for Artificial Intelligence Research
Department of Computer and Information Science
The Ohio State University
Columbus, Ohio 43210 USA

Abstract

In this paper, we wish to make three contributions to Naive Physics in the context of reasoning about devices. (1) We discuss some limitations of current qualitative simulation approaches with regard to a number of issues in understanding device behavior and point to the need for additional processes. (2) We introduce a new approach to deriving the behavior of devices called consolidation. In this approach, the behavior of a device is derived from the behavior of its components by inferring the behavior of selected substructures of the device. (3) We present an ontology of behavior and structure which is well-suited to the consolidation process. This ontology makes it possible to state rules of behavior composition, i.e., simple patterns of behavior and structure are used to infer additional behaviors.

1. Introduction

Naive Physics is the commonsense knowledge that people have about the world. This knowledge includes the ability to qualitatively understand the behavior of physical systems. Our investigation is presently concerned with a subset of physical systems, focusing on designed artifacts or devices. Ultimately, we are interested in developing a representation which is applicable to a wide variety of understanding problems. However, the research described here is conducted specifically on the problem of deriving the behavior of a device given its structural description and the behavior of its components.* We hope to integrate the results of this research with other work concerning the functional representation and diagnosis of complex devices [6].

One recent approach to this problem is *qualitative simulation* [2, 3, 5]. The description of the device determines the relevant quantities and constraints of the simulation, a simulation is performed, and the results are transformed into interpretations of the device's overall behavior.

This differs from quantitative simulation in several ways. Instead of assigning specific values to a quantity, only its ordinal relationship to important constants or other quantities is stated. Constraints are also qualitatively stated, e.g., proportionality may be asserted, but not a specific function. In addition to constraint satisfaction (the analogue of simulation by numerical methods), the techniques of qualitative simulation include constraint propagation, and matching descriptions of potential processes. The process of interpretation extracts state transition information, summarizing the possible behaviors and inferring causal relationships between device states.

*Other problems include design, diagnosis, planning (using devices to accomplish a goal), etc.

We propose an alternative approach that is a type of qualitative analysis. The behavior of the device is discovered by inferring the behavior of selected substructures from the behavior and structure of their components. Successful application of this process on increasingly larger substructures results in inferring the behavior of the device. This approach, called *consolidation*, has a number of desirable properties, including localized reasoning steps, causal analysis of behavior, and consistency of representation. Consolidation is not intended to be a complete solution to the inference-of-behavior problem; but where it can applied, we believe that it is a better alternative for analyzing and explaining behavior. We wish to emphasize at this point that an implementation of this approach is currently in progress.

The notion of reducing complexity by reasoning about a group of subcomponents as a single abstract component is shared by the work of Sussman and Steele [7] and is embodied in their notion of "slices." However, the aims and methods of their proposal make the details very different.

First, we argue that qualitative simulation has several undesirable characteristics as a Naive Physics theory. Next, we introduce consolidation, dividing the discussion into the description of components and the inference of behavior. A difficult example is then analyzed. Finally, unresolved issues are discussed.

2. Critique of Qualitative Simulation

2.1. Complexity

One desirable property of a Naive Physics theory is *simplicity of computation.* While current theories of qualitative simulation (QS) may be useful for providing upper bounds on the competence of qualitatively reasoning agents, they are unsatisfactory to account for human reasoning behavior due to the following two reasons.

First, QS is a *global reasoning process.* To perform the simulation for a particular point in time and to check if it has been done consistently, all the quantities and constraints must be taken into account. To go from one time point to another, the derivative of every quantity must be examined to update the quantities' values. This is true no matter the number of quantities and constraints the device has. A hierarchical breakdown is difficult because QS relies on nearly-closed systems (boundary conditions must be known or enumerable) and on constraint propagation. Forbus's notion of p-components [3] provides a method for subdividing a situation into independent parts. However, when the parts are more mutually dependent as in a device, additional techniques are called for.

Second, some theories of QS involves *substantial mathematical reasoning.* Quantities and their derivatives must be carefully handled so that constraints are not violated, and continuity is maintained. Since the constraints are stated in terms of

arithmetic and differential relationships, constraint propagation and checking for consistency require a considerable amount of mathematical reasoning. It must be stated, however, that Forbus's approach avoids much of the complexity of constraint propagation by restricting the paths over which propagation can occur.

2.2. Causality

Another desirable property is explaining the device's behavior in terms of the behavior and structure of its components; we want to know the *cause* of the device's behavior. Causality in some theories, especially the confluence theory of de Kleer & Brown [2], is identified mainly with the propagation of values through constraints.

The major problem with this position is that causality is viewed as a *"last straw" phenomenon*, i.e., the saying "the last straw broke the camel's back" would be translated, in this view, to "the last straw was the cause of the camel's broken back". De Kleer & Brown admit that their version of QS does not identify "the support which enables the causal action path to exist" [1]. However, it seems wrong to omit the support from a causal account, since the support may include the primary causal processes of the effect (e.g., most of the weight is already on the camel's back).

2.3. Representation

QS theories require descriptions of components to specify their outward structure, the quantities that are involved in interaction with other components, the constraints on those quantities, and the behavioral states.* This description may be thought of as the *behavioral laws* of the component. On the other hand, the representation of device behavior does not describe its behavioral laws, but is a network which shows the temporal (and causal) sequence of the components' states. If this process were to be repeated one more level (i.e., where the device at this level becomes a component at the next level), QS is not helpful, since it needs to have the behavioral laws of the device. This twin representation of behavior together with the global nature of QS limits the applicability of this approach.

Another problem with current representations is the *ontological impoverishment* of a theory primarily based on quantities and constraints. It is the burden of the model–builder to insure that the right types of quantities and constraints are represented and consistently defined. While there are guidelines for how to do this, these guidelines are *outside the representational system*. For example, Ohm's law is very significant for describing the behavior of electrical components. However, Ohm's law itself is not represented in QS, but is *compiled into* each component description that depends on it.**

3. Consolidation: Description of Components

Components interact with other components. The interaction is not just about components, but about the "stuff" or substances which potentially move between components and affect their behavior. What does a component have so that interactions can occur? We believe that a commonsense answer has two parts. One thing a component has is *structure*. On its exterior, it has places which are used to connect it to other components. On its interior, it has places which hold or contain substances. The other thing a component has is *behavior*, how it acts and is acted upon by substances. This section discusses how we represent the structure and behavior of components; the following section describes what inferences this representation supports.

3.1. Structural Primitives

Like de Kleer & Brown, we will use *connection* to signify that one component is attached to another component or is otherwise in meaningful spatial contact with it. An example of "meaningful spatial contact" is the relationship of the surface of a light bulb with the space around it, which in turn, might be in contact with something that the light affects. Note that we also include empty space as a type of component. This is essential for reasoning about movement though space, and about magnetism and gravity.

We also use *containment* as a structural relationship to represent the places inside components that substances can move from, move into, and be at rest. These places may or may not have significant capacity. The importance of this concept for Naive Physics theories was emphasized by Hayes [4].

For example, the light bulb in figure 1 has three connections: two electricity connections called "end1" and "end2", and a light connection called "surface". Inside of the light bulb, there are places where electricity passes through, and where light is produced. To model this, containers called "electrical" and "light source" are attributed to the light bulb, and are used in the behavioral description on the right.

3.2. Types of Component Behavior

Components act upon substances. We propose to describe these actions by a small set of relationships, using them as a foundation for representing additional knowledge about components and substances. They are:

- *Allow*. The component permits a specified kind of substance to move from one place to another.
- *Influence*. The component tries to move a specified kind of substance. There are two subtypes according to the spatial relationship of the **influence** with potential sinks and sources.
  - *Pump*. The component tries to move a substance through it, e.g., a battery has a **pump** electricity behavior from the negative to the positive terminal. The sink and source are external to a **pump** behavior.
  - *Expel*. The component tries to move a substance from (or to) an internal container, e.g., a balloon has a **expel** air behavior.
- *Move*. The component moves a specified kind of substance from one container to another along a specified path. **Move** behaviors are implicitly constrained by the amount and capacity of the containers.
- *Create*. The component creates a specified kind of substance in a container, e.g., a light bulb has a **create** light behavior.
- *Destroy*. The component destroys a specified kind of substance in a container, e.g., an acoustic insulator has a **destroy** sound behavior.

For example, the light bulb in figure 1 has **allow** electricity,

*Different "behavioral states" are associated with different sets of constraints. The total state of a component is its behavioral state and the values of its quantities.

**This criticism doesn't apply to Forbus's Qualitative Process theory [3], i.e., a single qualitative process description can be used to represent Ohm's law.

```
    gate|
        |        allow electricity between end1 and end2
    +--------+       thru electrical,
end2|        |       resistance 0, state closed
+---| switch |   allow signal from gate to sensor
|   |        |   change state to closed
|   +--------+       when [move signal from gate
|       |end1                     to sensor, rate on]
|       |
|       |
|  +---------+   pump electricity from negative terminal
|  |    +    |       to positive terminal thru electrical,
|  | battery |       amount positive
|  |    -    |   allow electricity between negative term.
|  +---------+       and positive terminal thru electrical,
|       |            resistance positive
|       |
|       |end2    allow electricity between end1 and end2
|   +-------+        thru electrical, resistance positive
|   |       |    allow light between light source
+---| light |        and surface
end1| bulb  |    create light in light source,
    |       |        rate [proportional (magnitude (rate
    +-------+               (move electricity between end1
        |                        and end2)))]
 surface|
```

Figure 1: Light Bulb Device

allow light, and **create** light behaviors (for the purposes of this discussion, other behaviors of the light bulb and other components have not been displayed). There are conditions on some of these behaviors, which are specified in the details of the description. For example, the **create** light behavior is dependent on movement of electricity. There is more discussion on this later.

Some components, such as the switch, have different behavioral states, where each state is associated with different behaviors. An additional type of behavior, *change state*, specifies a predicate on behavior and the next state of the component. For example, the switch in figure 1 has two states, open and closed, where the closed state has an **allow** electricity behavior, and the open state does not. The switch also has an **allow** signal behavior, and it will **change state** depending on the control signal that it receives.

3.3. Quantities

We use quantities to describe additional detail about behaviors and containers. Most of the behaviors have a natural measurement: **move** by rate of movement, **create** by rate of creation, **destroy** by rate of destruction, and **influence** by amount of influence. Also, some behaviors, especially **allow** behaviors, may have special quantities which are specific to the substance. Resistance, capacitance, and inductance are examples from electricity. The **allow** electricity behavior of the light bulb, for example, has a positive resistance.

Each container has quantities which describe its capacity and amount. The containers of the components in figure 1 can be modeled with infinitesimal capacity, so interesting issues concerning these quantities do not arise. In section 5, we will discuss an example in which these quantities have significant behavioral consequences.

Quantities can be used to express how some behaviors are dependent on other behaviors, i.e., how the component is acted upon by substances. For example, the rate quantity of the **create** light behavior of the light bulb is described as proportional to the rate of a **move** electricity behavior which goes through the light bulb.

4. Consolidation: Inference of Behavior

We propose to infer the behavior of a device by a form of composition. The behavior of selected substructures or *composite components* of the device is inferred from the behavior and structure of their subcomponents. Composites are used as *contexts* for forming intermediate points of understanding about the device. This composition is possible because the behaviors of components as represented above are themselves composable; certain behavioral and structural patterns give rise to additional behaviors. These *causal patterns* are also used to index into knowledge about the behavior of substances, i.e., knowledge about substances is organized around the possible generic situations in which behaviors are inferred.

4.1. Causal Patterns of Behavior and Structure

A causal pattern describes a situation in which a behavior may occur, asserting that if certain behaviors satisfy a specific structural relationship, then another behavior of a specified type may be caused.* For example, the **propagate pump** pattern specifies that a **pump** behavior in a serial relationship with an **allow** behavior will potentially cause another **pump** behavior, e.g., a **pump** electricity behavior between A and B, and an **allow** electricity behavior between B and C may cause a **pump** behavior between A and C. Whether this **pump** behavior actually occurs depends on the physics of the substance and the details of the sub–behaviors. The following are the causal patterns that we have discovered so far:

- *Serial/parallel allow.* An **allow** behavior caused by two serial or parallel** **allow** behaviors.
- *Parallel pump.* A **pump** behavior caused by two **pump** behaviors in parallel.
- *Propagate pump.* A **pump** behavior caused by a **pump** and an **allow** behavior in serial.
- *Propagate expel.* An **expel** behavior caused by an **expel** behavior and **allow** behavior in serial.
- *Serial/parallel move.* A **move** behavior caused by two serial or parallel **move** behaviors.
- *Pump move.* A **move** behavior caused by a **pump** behavior and an **allow** behavior, both on the same path from one container to another. In this pattern, the source and sink may be the same container in which case the movement is around a circuit.
- *Expel move.* A **move** behavior caused by an **allow** behavior which "connects" an **expel** behavior to another container.

We do not claim that this list is complete. Additional patterns may be required to reason about concepts like momentum, in which movement leads to additional influences. However, we believe that the number of additional patterns will be small.

---

*Currently, our theory does not handle situations in which the behaviors satisfying a pattern refer to different substances, e.g., oil and water.

**Roughly, two behaviors are "serial" if they share an endpoint; two behaviors are "parallel" if they have the same endpoints.

Suppose that a composite of the battery and the switch in figure 1 is chosen for processing. Behaviors are inferred based on the causal patterns as follows:

- Using the **serial allow** pattern, an **allow** behavior between the negative terminal of the battery and end2 of the switch is inferred. The resistance is determined to be positive from knowledge about electricity. Since the switch's **allow** behavior is active only during the closed state, the same is true of the inferred behavior. Since the states of the switch result in different behavior of the battery–switch, the battery–switch also has open and closed states.
- Using the **propagate pump** pattern, a **pump** behavior from the negative terminal of the battery to end2 of the switch is inferred. The amount is determined to be positive.

The causal patterns do not take into account that the battery–switch will also have the **allow** signal and **change state** behaviors of the switch. In general, those behaviors which affect the outward behavior of the composite, and which are not subsumed by an inferred behavior need to be copied to the composite. Also, note that none of the causal patterns refer to **create** and **destroy** behaviors. These kinds of behaviors are transferred to the composite if they are connected to the "outside" by **allow** behaviors.

The causal patterns are similar to Forbus's process descriptions [3]. Both describe the conditions necessary for some behavior to happen. One important difference is that the causal patterns are generic to all substances. While a process description can be stated at a high level of generality, there is no commitment by the theory to any particular level of generality. In practice, there are different process descriptions for different types of substances such as liquid, gas, heat, etc. Also, the process descriptions can be used only when changes occur, while the causal patterns can handle situations, such as two batteries connected serially, in which no physical change takes place.

4.2. Simplification of Structure

If a composite simply inherited the structure of its subcomponents, the description of larger composites would become increasingly complex, making it harder to reason about them. This is allayed in two ways. First of all, only the external connections of the composite become part of its behavioral description. For example, the positive terminal of the battery and end1 of the switch would not be referenced in the battery–switch's description.*

Second, *composite containers* may be instantiated as combinations of several other containers. In the battery–switch, the electrical containers of the battery and switch are combined to form a single electrical container. The creation of composite containers is governed by the inference of behaviors, under a constraint that restricts behaviors to reference only a limited number of connections and containers. For example, the "thru" attribute of the inferred **allow** electricity behavior of the battery–switch may only reference one container, thus a composite container is instantiated.

4.3. Physics of Substances

The physics knowledge contains the procedures that are used to validate inferred behaviors and determine the values of their quantities. Each substance has procedures which are associated with the causal patterns, and with other known situations such as dependencies and the inference of composite containers. For example, when the **serial allow** pattern matches on behaviors involving electricity, the resistance of the caused behavior is determined by summing the resistances of the causing behaviors.

The reasoning is more complicated when dependencies are involved. Suppose that we chose a composite consisting of the light bulb and switch. This composite will also have a **create** light behavior, which should have a rate quantity specified as:

proportional (*magnitude* (*rate*
  (**move electricity** *between* **end1 of the switch**
    and **end2 of the light bulb**)))

The places mentioned by the dependency must be part of the simplified structure of the composite. To do this, there must be knowledge of what paths through the composite will also go through the light bulb, and the dependency must be modified accordingly.

4.4. Light Inference

The primary effect of the light bulb system is that light is produced when the switch is closed. Consider now a composite which consists of the battery–switch and the light bulb. This inference can proceed as follows:

- The **allow** electricity behaviors of the battery–switch and light bulb satisfy the **serial allow** pattern, resulting in an **allow** electricity behavior around the electrical circuit. The resistance is positive. The behavior is active only during the closed state.
- The **pump** electricity behavior of the battery–switch and the **allow** electricity behavior of the light bulb satisfy the **propagate pump** pattern, from which a **pump** electricity behavior around the circuit is inferred. The amount of the behavior is positive. The behavior is active only during the closed state.
- The two behaviors inferred above satisfy the **pump move** pattern, so a **move** behavior around the circuit is inferred. The rate of the move is positive. The direction depends on how electricity is modeled.
- This **move** behavior satisfies the dependency expressed in the **create** light behavior of the light bulb. The rate of creation is calculated as positive.

In the inference of the rate of creation, every behavior of the components and element of structure which plays some role in the creation of light has been used in the consolidation process. *The explanation of this inference provides a complete causal account of the creation of light in the light bulb system in terms of the components' behavior and the device's structure.*

Also note that all the electrical connections are internal to the device. Thus no electricity behavior becomes part of the final description of the device's behavior. *The device's behavioral description states only what the outward behavior of the device is, not how it is accomplished.*

5. Another example

To further illustrate how consolidation works and to explain additional features of this analysis, consider the situation in figure 2. The source and sink components have containers of water of differing temperatures. The source component has an **expel** water behavior. There is a connection between the components which permits the flow of water.

*Connections which connect two or more components are assumed to be internal to the device, unless declared otherwise.

```
+--------+
|        |  container a of water, temperature x
| source |  expel water from a
|        |  allow water between a and hole
+--------+
    |hole
    |
    |hole
 +------+
 |      |  container b of water, temperature y
 | sink |  allow water between b and hole
 |      |
 +------+
```

Figure 2: Water Containers

A difficulty in modeling this device is representing temperature. We will say that the water within the **a** and **b** containers *contain* heat, and that water has a **expel** heat behavior. The amount of the **expel** behavior corresponds to the water's temperature. This extended notion of containment can also be used to model concentrations of dissolved material.

When water movement is inferred, heat movement should also be inferred. The movement of heat will not only affect the amount of heat, but will affect the amount of the **expel** heat behavior within the containers. These inferences are general enough to be codified as causal patterns:

- A **move** behavior of a substance S1 which contains a substance S2 causes a **move** S2 behavior along the same path.
- A **move** behavior of a substance S1 which contains a substance S2 and which has an **expel** S2 behavior affects the **expel** S2 behavior at the sink of the move.

For heat, the effect is that the **expel** heat amount of the sink will tend towards the **expel** heat amount of the source, i.e., the temperature of the sink will move towards the temperature of the source.

With these additional patterns, the inference of the **move** water behavior will lead to inferring a **move** heat behavior, and to consideration of the effect upon the **expel** heat behavior of the sink. In this case, the capacities of and amounts within the two water containers will affect what will actually happen. When initial values for these are chosen, the behavioral description can then be *run* to determine what sequence of events will occur.

6. Summary

In Section 2, we discussed several problems with qualitative simulation in understanding device behavior. Here, we discuss how consolidation overcomes some of those difficulties.

In consolidation, the reasoning occurs in well defined, locally contained steps. Mathematical ability determines the sophistication of analysis, but is not required to perform it. To claim simplicity, an implementation is necessary to demonstrate that the analysis can be done efficiently in real situations. The primary complication is deciding what composites should be analyzed. We are developing general heuristics for making this decision. Also, domain knowledge could provide additional heuristics, such as a library of precompiled composites that are instantiated when they are recognized.

Causality is directly linked to the idea that the components of the device cause the behavior of the device. Reasoning about behaviors using the causal patterns leads to an inference structure showing the behavioral relationships between the components and the device. A complete account includes the dependencies between behaviors, and how they are satisfied.

The representation distinguishes between different types of quantities and constraints. Different quantities are associated with the behaviors and substances that they describe. Constraints are embedded within dependencies and the physics knowledge of substances. The device's behavior is represented in the same manner as a component's behavior.

A possible problem is that states of components almost always become state of composites. A combinatorial problem might occur when several components have multiple states. One alternative is to use simulation in this kind of situation. A more interesting alternative is to simplify the description of composites using various means such as inferring that certain states are impossible, combining related states into a single state, etc.

Consolidation is but one of the multiplicity of processes and representations that are a part of Naive Physics reasoning. Further research is called for in describing the relationship between consolidation and qualitative simulation, in expanding the richness of the structural primitives, and in representating and integrating, for example, discreteness of motion, temporality, and mixing of substances.

Acknowledgments

We would like to acknowledge John Josephson, Jon Sticklen, and Ken Forbus for their comments on earlier drafts of this paper. Of course, this doesn't imply endorsement of the final product. This research is funded by the Air Force Office of Scientific Research through AFOSR grant 82–0255.

References

1. de Kleer, J., and Brown, J. S. Assumptions and Ambiguities in Mechanistic Mental Models. In *Mental Models*, Lawrence Erlbaum, Hillsdale, New Jersey, 1983, pp. 155–190.

2. de Kleer, J., and Brown, J. S. A Qualitative Physics Based on Confluences. *Artificial Intelligence 24* (1984), 7–83.

3. Forbus, K. D. Qualitative Process Theory. *Artificial Intelligence 24* (1984), 85–168.

4. Hayes, P. J. Naive Physics I: Ontology for Liquids. 35, Institut pour les Etudes Semantiques et Cognitives, Unversite de Geneve, Geneva, 1978.

5. Kuipers, B. Commonsense Reasoning about Causality: Deriving Behavior from Structure. *Artificial Intelligence 24* (1984), 169–203.

6. Sembugamoorthy, V., and Chandrasekaran, B. Functional Representation of Devices and Compilation of Diagnostic Problem Solving Systems. Technical Paper, AI Group, Dept. of Computer and Information Science, The Ohio State University, 1984. To appear in *Cognitive Science*.

7. Sussman, G. J., and Steele, Jr., G. L. CONSTRAINTS -- A Language for Expressing Almost–Hierarchical Descriptions. *Artificial Intelligence 14* (1980), 1–39.