Scientific Theory Formation Through Analogical Inference

(Extended Abstract)

Brian Falkenhainer

Qualitative Reasoning Group
Department of Computer Science
University of Illinois at Urbana-Champaign
1304 W. Springfield Avenue, Urbana, Illinois, 61801
Arpanet: falken@a.cs.uiuc.edu

1 Introduction

In the course of trying to further understand the world around him, man repeatedly attempts to find explanations for observed physical phenomena. This scenario applies to both scientists working in the laboratory and non-scientists assimilating everyday experiences. People don't carry around a full theory of the world in their heads; they make conjectures as a result of everyday experiences. Theories are tentatively proposed, they are checked to see if they adequately account for observed behavior, and sometimes experiments are performed to confirm predictions sanctioned by the new theory. One of the goals of Artificial Intelligence is to construct intelligent, autonomous systems: They too must possess the flexibility to form and refine physical theories in the course of interacting with the world.

This paper presents an investigation into the process of scientific model formation; specifically, the discovery and refinement of qualitative models of the physical world. First, general principles underlying all scientific theory formation are discussed. A theory of analogical learning, called *Verification-Based Analogical Learning*, is then presented which adheres to these basic principles. This theory shows how analogy may be used to discover and refine scientific models of the physical world through simple observation and interaction with physical phenomena. It describes how an initial model of a domain may be constructed to explain a new, inexplicable situation and how a verification is constructed to demonstrate that the new model adequately explains the observed behavior. It goes on to show how a simple planner with a knowledge of naive physics may be used to verify certain types of predictions indicated by the new theory. Examples are taken from an implemented system, which uses the theory to discover and verify qualitative models of processes such as water flow and heat flow.

2 Model Formation and Verification

In general, we know that when two bodies, one hot and one cold, are placed in contact with each other, after a period of time they will reach the same temperature. What happens between the time the two objects are placed in contact and the time the two temperatures equalize? If the notion of water flow suggests itself, we may construct a model for the situation in which heat is seen "flowing" from a higher temperature to a lower temperature. Using the new model shows that it accurately explains the phenomenon. This is called *verifying the consistency*

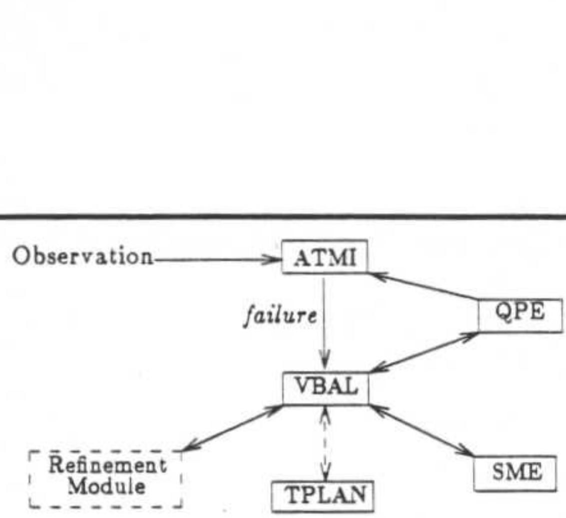
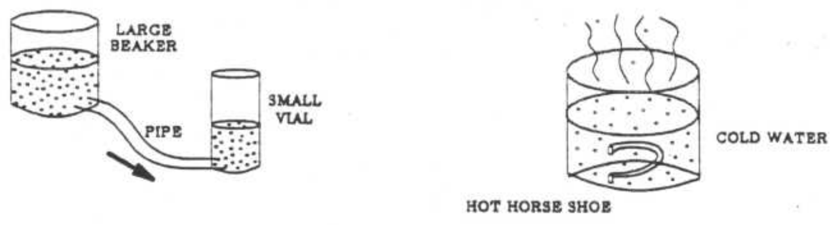


Figure 1: Block diagram of the Phineas system modules.

of the model. The new theory now predicts that certain other events must also be able to happen, such as the bidirectionality of heat flow. We attempt to recollect a prior experience (history) demonstrating this predicted behavior or we conduct simple experiments to explore the space of hypothesized behaviors. This is called *verifying the predictions* of the model. If we were to extend the analogy further by hypothesizing that heat was itself a type of liquid (i.e., the *caloric* theory of heat), a number of additional predictions may be made based upon the intrinsic properties of liquids and physical objects. For example, *conservation of matter* would lead to predictions based on *conservation of heat*. Exploring the consequences of these additional predictions is called *verifying the extension* of the analogy. This entire process of hypothesis formation, confirmation, refutation, and subsequent refinement is the essence of verification-based analogical learning (VBAL).

The current implementation of VBAL, called Phineas (Figure 1), is designed to operate as a passive observer, relating observed physical phenomena to known theories of the world.[1] These theories are expressed as qualitative models of various physical processes, such as moving, bending, and liquid flowing, using Forbus' *Qualitative Process theory*. When a situation is witnessed which the program's current models cannot explain, the VBAL control module is invoked to generate a new or revised model that accounts for the new observation. The system uses Forbus' measurement interpretation program (ATMI) to monitor the world and relate observation with known or conjectured theories. Learning is triggered when ATMI fails to adequately interpret the events. The VBAL program interacts with an analogy module, the Structure-Mapping Engine (SME) and a knowledge refinement module to construct a new or revised model. Forbus' Qualitative Process Engine (QPE) takes the model and produces a new envisionment, which ATMI in turn applies to the current situation. This cycle of discovery and refinement will continue until an accurate model has been formed or until the system has exhausted all the possibilities. Once a consistent model is found, predictions implied by the

[1]Due to differences in data syntax and lisp dialects among the four implemented modules, the current implementation of Phineas is not yet fully autonomous. Some of the links must be assisted by hand, such as translating a QPE envisionment into one which ATMI can read. While effort has been made to insure that the hand translations modify only the syntactic properties of program data, I cannot be certain that nothing has been overlooked until these modules are properly integrated. This integration is expected to be completed shortly. The planner has worked properly on the example given here, but it is still under development and I refrain from strong statements about its implementation status. The knowledge refinement module is currently unimplemented.



(a) (b)

Figure 2: Two examples of water flow and heat flow.

model are explored using Hogge's time-based planner (TPLAN), which uses its knowledge of naive physics to make appropriate transformations to the current situation. A more detailed description of the system's operation will now be presented in conjunction with an example of how the system learns a new model of heat flow by drawing an analogy with a similar water flow experience.

2.1 Model Formation

The VBAL process begins when a situation is encountered for which current domain theory fails to account. First, a prior experience that appears to exhibit similar behavior is accessed from memory and SME is used to form a match between the changes observed in the prior experience and the changes taking place in the current situation. This analogy serves to explicitly indicate the object and quantity correspondences between the two domains. Once a satisfactory experience has been retrieved, the domain theory used to account for the prior situation is fetched. Analogy is applied again to map the potentially analogous domain theory to the new domain of interest. This second analogy is generally a pure mapping of structure from one domain to another, appropriately transformed according to the object correspondences provided by the prior analogy between the two histories.

For example, suppose that the program was presented with measurements of the heat flow situation in Figure 2 and described in Figure 3. If the program has no theories of heat flow, Phineas will be unable to interpret the new observation. Using SME, the program is able to establish an analogy with the previously encountered water flow experience shown in Figure 2 (see also Figure 3). This match serves to establish which things from the two situations are behaving in the same way. It is seen from Figure 3 that the roles of the beaker and the vial in the water flow history are found to correspond to the roles of the horse shoe and water in the heat flow history, respectively. Those correspondences which provide a mapping between entities or between their quantities (e.g., `Pressure` and `Temperature`) are stored for later reference.

When it is satisfied that the chosen water flow history is sufficiently analogous to the current situation, the Phineas program fetches the relevant domain theory which led to its prior understanding of the water flow experience. This model, expressed as a QP theory process definition, states that if we have an aligned fluid path between the beaker and the vial, and the pressure in the beaker is greater than the pressure in the vial, then a fluid flow process

Water Flow History

```
(Situation S0)
(Decreasing (Pressure (At beaker S0)))
(Increasing (Pressure (At vial S0)))
(Decreasing (Amount-of (At beaker S0)))
(Increasing (Amount-of (At vial S0)))
(Greater (Pressure (At beaker S0))
      (Pressure (At vial S0)))

(Situation S1)
(Meets S0 S1)
(Constant (Pressure (At beaker S1)))
(Constant (Pressure (At vial S1)))
(Constant (Amount-of (At beaker S1)))
(Constant (Amount-of (At vial S1)))
(Equal-To (Pressure (At beaker S1))
       (Pressure (At vial S1)))

(Function-Of (Pressure ?x)
          (Amount-of ?x))
```

Heat Flow History

```
(Situation S0)
(Decreasing (Temperature (At horse-shoe S0)))
(Increasing (Temperature (At water S0)))
(Greater (Temperature (At horse-shoe S0))
       (Temperature (At water S0)))

(Situation S1)
(Meets S0 S1)
(Constant (Temperature (At horse-shoe S1)))
(Constant (Temperature (At water S1)))
(Equal-To (Temperature (At horse-shoe S1))
        (Temperature (At water S1)))

(Function-Of (Temperature ?x)
          (Heat ?x))
```

Match

Pressure	↔	Temperature
Amount-of	↔	Heat
S0	↔	S0
S1	↔	S1
beaker	↔	horse-shoe
vial	↔	water

Figure 3: Analogical match between water flow history and heat flow history.

will be active. This process has a flow rate which is proportional to the difference between the two pressures. The flow rate has a positive influence on the amount of water in the vial and a negative influence on the amount of water in the beaker.

Using SME a second time, the liquid flow theory is matched to the current heat flow situation, producing the model of heat flow shown in Figure 4. The analogy at this stage is highly constrained, due to the set of entity and function correspondences established when the water flow and heat flow histories were matched.

This example demonstrates a number of points. First, the "analogy" here is composed almost entirely of analogical inferences, since the system had no prior model of heat flow. Hence, the model was *constructed* by analogy rather than augmented by analogy. This shows the power of SME's candidate inference mechanism. In addition, it shows the utility of the candidate inferences' skolemized entities. The results produced by SME (Figure 4) contain the entity `(*skolem* pipe)`. This indicates that, at the moment, the heat path is a conjectured entity. Further experimentation could be used to identify the actual heat path, a knowledge of paths in general could be used to indicate that physical contact is a likely path, or the path could be left as a conjectured entity. This last choice corresponds closely to the period in

```
Gmap #1: { (AMOUNT-OF-35  HEAT-WATER)  (AMOUNT-OF-33  HEAT-HSHOE)
          (PRESSURE-BEAKER  TEMP-HSHOE) (PRESSURE-VIAL  TEMP-WATER) }
    Emaps:  { (beaker  horse-shoe)  (vial  water) }
    Weight:  2.675
    Candidate Inferences: (IMPLIES
                              (AND (ALIGNED (*skolem* pipe))
                                   (GREATER-THAN (A TEMP-HSHOE) (A TEMP-WATER)))
                              (AND (Q= (FLOW-RATE pi) (- TEMP-HSHOE TEMP-WATER))
                                   (GREATER-THAN (A (FLOW-RATE pi)) zero)
                                   (I+ HEAT-WATER (A (FLOW-RATE pi)))
                                   (I- HEAT-HSHOE (A (FLOW-RATE pi)))))
```

Figure 4: An Analogically Inferred Model of Heat Flow Produced by SME.

science when a substance, called the *ether*, was believed to exist in order to provide a medium for the flow of light.

2.2 Verifying Consistency

The consistency of the new model is verified by using it to account for the original situation. When a qualitative process model has been constructed, an analysis of the model by the Qualitative Process Engine produces a description of all possible behaviors for the current physical configuration, called an *envisionment*. An envisionment describes physical states and the possible transitions between them. The behavior of the system through time may then be represented as a single path through the envisionment. The consistency of the new model is verified by using it to account for the original situation. It is able to provide an explanation for the observations if a path through the envisionment formed from the model can be found which corresponds to the measurements. In this example, ATMI finds that the new thoery accurately models the heat flow situation. The program has thus verified that the theory provides a consistent explanation for this and functionally similar instances of heat flow. If the verification step fails, model refinement may be used to account for slight imperfections in the analogy or the process may be repeated using a different analogous situation.

2.3 Verifying Predictions

Given that the consistency of the model has been confirmed by finding a path in the total envisionment that explains the current situation, what may be said about the other paths in this new envisionment? The new model states that the system should be able to exhibit the behavior described by all of the paths in the envisionment. Those paths which describe behavior not yet seen represent the most basic form of prediction sanctioned by the new theory. These predictions are easily explored through simple experiments, constructed by a planner which uses its knowledge of qualitative physics to manipulate the world and achieve its goals. This planner is able to establish any situation in the envisionment, allowing us to manipulate the system through every path and confirm or disconfirm the validity of the model's predictions.

2.4 Extending the Analogy

If an analogy proves useful in understanding a given phenomenon, it would be wise to extend the analogy further and explore the limits of the analogy's validity. For example, the water flow – heat flow analogy may be extended by hypothesizing that heat is itself a type of liquid and possesses the properties known to hold for liquids (the *caloric* theory of heat). By extending the analogy in this manner, we are forced to conjecture a law of conservation of heat which states that heat can never be lost nor created. In the early nineteenth century, the caloric theory was widely believed and evidence for or against conservation of heat was sought. It was the phenomenon of friction which led to the eventual downfall of the caloric theory of heat and gave rise to the energy interpretation. While the original flow model may remain intact, its theoretical underpinnings originating from extending the analogy to conjecture a heat liquid must be replaced by a notion of heat *energy* flowing.

3 Discussion

The formation of a model through analogy and the verification of its consistency has been fully implemented in the Phineas system. This system has been used to discover qualitative models of various types of water flow and heat flow phenomena. Work is currently underway to fully implement prediction verification. A time-based planner possessing the power to reason with models of qualitative physics is being used to construct experiments that investigate the validity of predicted behavior. In addition, work is in progress to integrate Rajamoney's directed experimentation system, giving Phineas the ability to conduct further and more creative experimentation and knowledge refinement. It is hoped that this will enable Phineas to explore the consequences of extending analogies to make further predictions.

This work shows that analogy may serve as a useful "inventive" mechanism, enabling a reasoning system to construct an initial theory of some domain which knowledge refinement methods may subsequently make adjustments to. Analogy offers a technique for making large leaps in current knowledge. However, the validity of analogical inferences is very tenuous and requires a cautious investigation into the learned concepts. One way to ensure that the inferences make sense is to compare the consequences of these inferences against observed physical behavior - hence, verification-based analogical learning. Theories produced by analogy are evaluated by their ability to predict observed physical phenomena. This methodology applies equally well to any form of theory formation, in which conjectures of uncertain validity are made.