Learning Qualitative Causal Models via Generalization & Quantity Analysis

Scott E. Friedman and Kenneth D. Forbus

Qualitative Reasoning Group, Northwestern University
2145 Sheridan Road, Evanston, IL 60208-0834 USA
{friedman, forbus}@northwestern.edu

Abstract

Learning causal models is a central problem of qualitative reasoning. We describe a simulation of learning causal models from exemplars that uses progressive alignment and qualitative process theory to derive plausible qualitative causal models from observations. We show how *protohistories* can be created via progressive alignment and used to infer causality. The result, a *causal corpus*, can make simple predictions and set the stage for more sophisticated qualitative models. The simulation has been successfully tested with learning causal mechanisms of three physical scenarios, with encouraging results.

Introduction

Forbus & Gentner (1986) proposed decomposing learning of physical domains from experience into four stages. (1) *Protohistories* are prototypical behaviors, generalized from multiple experiences. (2) The *causal corpus* consists of fragmentary causal models, created from protohistories. (3) These fragmentary models are organized into a *naïve physics*, which regularizes the fragmentary causal models by postulating broadly applicable mechanisms. (4) *Expert understanding* consists of deepening the naïve physics and tying it to mathematical and other formal models, typically culturally learned. Importantly, these stages are localized within the understanding of particular phenomena. For example, someone might have an expert understanding of electronics while having only a partial set of causal models for fluids.

This paper focuses on learning initial causal models of a domain from observations. We use qualitative process theory (Forbus, 1984) to formally represent causal models. Causal models are learned from symbolic representations of experiences via a combination of analogical processing (Gentner, 1983) and statistical methods. The simulation has been successfully tested on three scenarios; we use understanding floating versus sinking as a running example for illustration. We first review QP theory and the structure-mapping models we use. Then we discuss how protohistories are learned from experience via progressive alignment, proposing *generalization contexts* as a means of organizing experience around salient questions. Next we discuss *quantity analysis* strategies to develop fragmentary causal models by hypothesizing ordinal conditions, limit points, and new quantities. We summarize results from our simulation and close by discussing other related work and future plans.

Background

Our theoretical framework uses *qualitative process theory* as its account of modeling mechanisms of change. Changes are caused by continuous physical processes, which provide the notion of mechanism for causality (cf. Chi *et al* 1994; Ahn *et al* 1995). These changes propagate through the system via *qualitative proportionalities* which indicate causal relationships between quantities. Qualitative proportionalities provide only partial information about what will happen. This makes them particularly appropriate for representing local causal models, since models learned from one set of experiences can be more easily combined with others.

These causal laws are contextualized by belonging to either processes or *views*, and hold only when their *conditions* are true. Conditions are typically ordinal relations, involving parameters of the entities participating in the process or view. The values that a quantity is compared with in such relations are called *limit points*, since they help determine when processes start and stop, and when views hold or not. Postulating the existence of limit points is an important challenge in learning QP models, since they are crucial for prediction.

QP theory does not describe how these models are learned. We claim that statistical accounts of causality (cf. Pearl, 2000; Gopnik *et al* 2004) can be harnessed to produce QP models. We incorporate statistics via similarity, using structure-mapping operations to construct probabilities as a side-effect of assimilating experiences. The SEQL model of generalization (Kuehne *et al* 2000) constructs generalizations incrementally via analogical comparison. We simulate analogical matching via SME, the Structure-Mapping Engine (Falkenhainer *et al* 1989; Forbus *et al* 1994). Given two structured representations, the *base* and *target*, SME computes one or two *mappings* which describe how the base and target can be aligned.

Mappings include a set of *correspondences* that detail exactly which entities and statements in one description go with entities and statements in the other, a *structural evaluation score* which indicates the overall quality of the match, and a set of *candidate inferences* that are conjectures about the target, using the correspondences to project partially unmapped base structures. Candidate inferences allow predictions and explanations to be generated without rules, via analogy to prior experiences and explanations. This makes them particularly important for accounts of learning like ours that postulate localized, incrementally generated models.

SEQL operates by maintaining a list of generalizations and exemplars. Given a new exemplar, SEQL compares it with existing generalizations. If it is sufficiently similar to one of them, it is assimilated into that generalization. Otherwise, it is compared against the list of unassimilated exemplars. If a pair of exemplars is sufficiently similar, they are combined to form a new generalization.

We call a set of generalizations and exemplars that are being processed together by SEQL a *generalization context*. Generalization contexts can be defined bottom-up, via similarity-based retrieval, or by labeling, e.g., a learner might use a generalization context to process all examples that have been given a verbal label, like "cat".

Learning Protohistories

Protohistories are generalizations of specific observed behaviors. Observed behaviors are typically rich with perceptual information, and in new domains, impoverished with regard to explanations. We postulate that analogical generalization, as modeled in SEQL, is used to construct prototypical behaviors. Below is an example observation given to our simulation. It describes an adult female human, swimming (gliding) in a still pond, and floating:

```
(isa bodyInLiquid0 AdultFemaleHuman)
(isa container0 Pond)
(isa liquid0 (LiquidFn Water))
(in-UnderspecifiedContainer liquid0 container0)
(massOfObject bodyInLiquid0 (Kilogram 60))
(volumeOfObject bodyInLiquid0 (CubicCentimeter 62039))
(isa gliding0 MovementEvent)
(primaryObjectMoving gliding0 bodyInLiquid0)
(isa stillLiquid0 StandingStill)
(doneBy stillLiquid0 liquid0)
(in-Floating bodyInLiquid0 liquid0).
```

The vocabulary of concepts and relations is drawn from the ResearchCyc knowledge base[1], an independently developed representation system for common-sense knowledge. The predicate calculus was produced using a natural-language understanding system (Kuehne & Forbus, 2004) from simplified English, to reduce tailorability.

The simplified English that generates the above predicate calculus observation is:

> The woman *bodyInLiquid0* floats in water *liquid0* in a pond *container0*. The mass of the woman *bodyInLiquid0* is 60 kilograms. The volume of the woman *bodyInLiquid0* is 62039 cubic centimeters. The woman *bodyInLiquid0* is moving but the water *liquid0* is standing still.

For SME processing, `isa` statements are automatically translated into attributes (i.e., `(AdultFemaleHuman bodyInLiquid0)`). SEQL generalizations abstract specific individuals (e.g., `bodyInLiquid0`) into anonymous individuals, not variables. Numerical parameters (e.g., `(Kilogram 60)`) are also abstracted into anonymous individuals, but their values are preserved in a distribution for each quantity in the generalization. These distributions are used to conjecture limit points below. We ignore memory retrieval in this simulation, and provide as input a stream of observations like the above.

How many generalization contexts should be used? Since SEQL automatically constructs multiple generalizations according to similarity, one possibility is to use a single context. The drawback with a single context is that it may not provide enough discrimination for learning. For example, to learn why things float, the learner must distinguish between floating and sinking examples. We have observed that SEQL may, because of attribute information, cluster cases from both types of situations into the same generalization. Consequently, we create separate generalization contexts for each possibility. Every generalization context incorporates a set of *entry patterns* that are tested against new exemplars. When a new exemplar satisfies the entry pattern for a generalization context, it is processed in that context. The same example can be processed in multiple contexts, since a learner might be learning multiple concepts at once.

Consider a learner trying to understand the distinction between floating and sinking, as well as sailboats sailing. Figure 1 illustrates the three example generalization contexts that would be used. If an exemplar arrives with `(SinkingEvent sinking0)` as a constituent fact, with no mention of floating, it will be incorporated into the rightmost context alone. If another exemplar arrives with `(isa boat0 SailBoat)` and `(floating-in boat0 (LiquidFn Water))` as constituent facts, it will be incorporated into both leftmost and middle contexts.

[1] http://research.cyc.com/

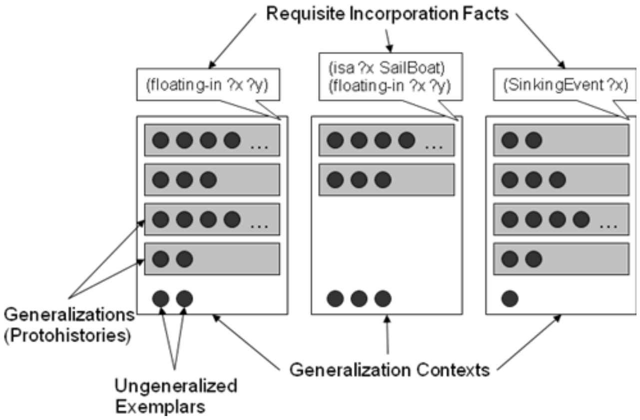

Figure 1: Example contextual protohistory organization

Learning a Causal Corpus

The causal corpus consists of a set of causal models grounded in, and connecting, protohistories. These causal models are local to particular protohistories or collections of protohistories. Restructuring these local models into general domain theories, of the kind typically used in qualitative reasoning, occurs only after a reasonable causal corpus has been constructed (Forbus & Gentner 1986). Even fragmentary causal models are quite powerful: Understanding what qualitative proportionalities hold in a protohistory yields a means of predicting the immediate consequences of parameter changes. Similarly, understanding quantity conditions that determine which protohistory represents the behavior that occurs in a situation enables predictions of state changes.

Our simulation uses three *causal learning strategies* – procedures that take protohistories and quantities as input, and generate causal hypotheses, expressible using the vocabulary of QP theory. We also describe a method for deriving complex quantities from constituent input quantities. We do not view this set of strategies as complete, but we believe they are a good starting point.

Analyzing quantity values enables us to hypothesize limit points, quantity conditions, and qualitative proportionalities. The *quantity condition strategy* identifies relevant ordinal relationships. The *limit point strategy* hypothesizes new causally-relevant values. The *quantity derivation strategy* hypothesizes compound quantities. We discuss each in turn.

Quantity Condition Strategy. Conditions for processes and views typically include ordinal relations between quantities. For instance, for a body to be floating in a liquid, its density must be less than the liquid's density. Quantity conditions are conjectured as follows:

1. Protohistories that summarize experience related to the target phenomenon are divided into two groups: those that express it (P^+) and those that do not (P^-).

2. For each protohistory p_i within ($P^+ \cup P^-$), the ordinal relationships $R_i = \{r_1, r_2, \ldots, r_n\}$ are identified that hold for every exemplar within P_i. The ordinal relationships tested are =, >, <, ≥, and ≤, over the set of exemplars that were used in forming P_i.

3. Conditions are identified that pertain to the entirety of P^+ and P^-, such that $R^+ = \{R^+_1 \cap \ldots \cap R^+_n\}$ and $R^- = \{R^-_1 \cap \ldots \cap R^-_n\}$.

4. Conditions that coincide with the phenomenon are the set $R_{cause} = R^+ - R^-$. Relationships that coincide with the absence of the phenomenon are the set $R_{prevent} = R^- - R^+$.

We use exemplars in step 2 because our encoding process does not automatically generate ordinal relationships from numerical values in observations. (The quantity value distribution information stored with generalizations cannot be used to compute this, because links to particular exemplars is not included.) This is a simplification: We believe that psychologically, encoding choices are driven in part by learning goals, which would propose encoding particular ordinal relationships in order to test conjectures via this strategy. Such goals might be generated based on trying various ordinals on a small number of exemplars, but that is left for future work.

Limit Point Strategy. Some physical phenomena occur when a quantity's value is above or below a specific limit point. Like the quantity condition strategy, the limit point strategy assumes that two sets of protohistories have been identified, such as water being heated and boiling and water being heated and not boiling. Recall that protohistories preserve the set of exemplar values $\{v_1, v_2, \ldots, v_n\}$ for each quantity. This information can be summarized via an interval V, where $V = [\min(v_1, v_2, \ldots, v_n), \max(v_1, v_2, \ldots, v_n)]$.

After calculating quantity intervals for individual protohistories, we first compute possible limit points by grouping protohistories into two sets: those that express the given phenomena $P^+ = \{p^+_1, p^+_2, \ldots, p^+_n\}$ and those that do not $P^- = \{p^-_1, p^-_2, \ldots, p^-_n\}$. For each quantity-type q, we merge the protohistory intervals so that

$$P^+_q = [\min(p^+_{1q}, p^+_{2q}, \ldots, p^+_{nq}), \max(p^+_{1q}, p^+_{2q}, \ldots, p^+_{nq})]$$
$$P^-_q = [\min(p^-_{1q}, p^-_{2q}, \ldots, p^-_{nq}), \max(p^-_{1q}, p^-_{2q}, \ldots, p^-_{nq})].$$

If the intervals P^+_q and P^-_q do not overlap for a quantity, it could be the case that a limit point exists within the interval $[\max(\min(P^+_q, P^-_q)), \min(\max(P^+_q, P^-_q))]$, or between the maximum point of the lower interval and the minimum point of the higher interval. This interval is then added to the causal corpus, as a limit point approximation.

If the intervals P^+_q and P^-_q overlap, there could still be an uninterrupted interval $[q_{min}, q_{max}]$ that represents a condition under which the phenomenon occurs. Instead of merging protohistory intervals into P^+_q and P^-_q, we test for exclusiveness, such that no protohistory intervals in P^+ overlap protohistory intervals in P^- for a quantity q.

Uninterrupted intervals in q are then added to the causal corpus as possible conditions for the target phenomenon.

Quantity Derivation Strategy. Understanding many physical phenomena requires introducing quantities beyond those observed. To understand why something floats versus sinks, for example, requires introducing the idea of density. If the quantity analysis fails to distinguish between two behaviors within the encoded quantities, the quantity derivation strategy proposes new quantities that are then searched for limit points and ordinal relationships. For all explicitly mentioned quantities a and b such that $a \neq b$, a set of new quantities C is derived:

$$C = \{a/b, b/a, a*b, a+b, a\text{-}b, b\text{-}a\}.$$

The units for the derived quantities may be identical to their constituent quantities ($kg + kg = kg$), or they may be combinations of their constituent units ($kg/cc = kg/cc$).

Simulation Results

We demonstrate how these methods combine to produce plausible causal corpus elements from a set of observations. We first go through a single learning task in detail, then summarize the results of others.

To investigate learning floating versus sinking, we encoded 30 unique exemplars – 16 floating and 14 sinking – in simplified English, which were fed into our natural language understanding system to automatically produce predicate calculus descriptions like our earlier example. Many factors used in the scenarios were based on Piaget's (1930) interviews with children: the motion of the water (still or wavy); the body in water (man, woman, log, cruise ship, or tree branch); the body of water (ocean, sea, lake, pond, bath-tub, or bowl); and autonomous motion of the body (moving/gliding or still). In all scenarios, a body floats when the body's density is less than 1 g/cc.

The simulation first generates protohistories from the exemplars. Two generalization contexts were used, with entry patterns `(in-Floating ?x ?y)` and `(isa ?x SinkingEvent)`, to model the focus on understanding when something floated or sank. The assimilation threshold for SEQL was set to 0.75. This yielded six protohistories, five for floating and one for sinking. All exemplars were assimilated into a generalization. Table 1 shows the protohistory abstractions with the generic entities in bold, and the protohistory size, |P|. Protohistories P_1 and P_4 preserved *tree branch* and *cruise ship* in their abstractions, respectively; the rest contain only generic entities.

The abstractions for P_2 and P_3 are identical, yet they are still distinct. This is due to uncertain facts within the generalization. Specifically, in P_2, P(body = man) = .66, and in P_3, P(body = woman) = .66. Thus, although the abstractions are identical, the underlying representations differ. Low-probability facts are considered for similarity processing, so they remained distinct.

Context	#	Protohistory Abstraction	|P|
Floating	1	Idle tree branch, wavy **water**	2
	2	Moving **body**	3
	3	Moving **body**	3
	4	Moving cruise ship	3
	5	Wavy **water**	5
Sinking	6	Idle **body**, still **water**	14

Table 1: Protohistories for floating and sinking

To generate causal corpus information for these protohistories, the strategies defined above were executed in the order given.

Given a set of protohistories, the simulation proceeds to analyze its quantities, searching for limit points and quantity conditions that help explain floating. The observable quantities yielded no causal hypotheses, so the simulation used the quantity derivation strategy to create new quantities and try again. One of the derived quantities does yield a limit point, as shown in Table 3. Since this limit point (which we know as density) was derived as the ratio of mass and volume, we also obtain the qualitative proportionalities shown in Table 3, imposing a causal direction on what was an algebraic relationship by assuming that observable parameters are more primitive than derived parameters. (This is a heuristic, of course, that could be incorrect – consider heat derived from temperature, for example.)

Causal Hypothesis Type	Formula
Derived Quantity	q = $\text{mass}_{\text{body}}/\text{volume}_{\text{body}}$
Limit Point	q < [0.001, 0.00102] kg/cc
Qualitative Proportionality	*floatability* α_{Q-} q
	q α_{Q+} $\text{mass}_{\text{body}}$
	q α_{Q-} $\text{volume}_{\text{body}}$
	floatability α_{Q-} $\text{mass}_{\text{body}}$
	floatability α_{Q+} $\text{volume}_{\text{body}}$

Table 3: Causal hypotheses generated about floating

In addition to floating/sinking, we tested the simulation on two other learning scenarios. This involved creating new stimuli descriptions and changing the entry patterns of the generalization contexts to suit the scenarios. The remainder of the learning process remained the same.

To model learning how balance scales work (Siegler, 1983), we encoded nine scenarios using the methodology above, varying the kinds of objects on the balance and the posture of the object (e.g., sitting or kneeling or upright). Using two generalization contexts, one for *right-side sinking* and one for *left-side sinking*, the simulation generated two protohistories for each context. The quantity condition strategy creates the sensible quantity hypothesis

```
(> (massOfObject leftside0) (massOfObject rightside0))
```

to predict when the left side will sink.

In another learning experiment conjecturing when boiling would occur, six exemplars were encoded using the methodology above. The limit point strategy conjectures a limit point for temperature to predict when boiling occurs:

```
Hypothesis: phenomena occurs when
 (temperatureOfObject kettle0)
 is above some point in the range:
 [95.0-100.0] DegreeCelsius.
```

The lower bound of the range could be refined by more experience.

While the number of learning experiments conducted to date is small, the results obtained so far are very reasonable.

Related Work

The closest previous simulation is COBWEB (Fisher, 1987) which utilized conceptual clustering, but did not introduce causal models, nor was it tested on semi-automatically generated stimuli. Our quantity derivation strategy is inspired by Langley's (1981) BACON simulation.

Some of diSessa's (1983) p-prims (for "phenomenological primitives") can be viewed as causal corpus elements while others may be viewed as protohistories. No computational model for learning them was ever implemented.

Discussion and Future Work

Our simulation combines symbolic, relational representations with quantity analysis to learn causal models. We think this is a very promising approach to developing deep qualitative models of physical domains.

In cognitive psychology, many advocates of statistical accounts of causality do not include any notion of mechanism, and we obviously (along with Chi *et al* 1994; Ahn *et al* 1995) do not believe that is sufficient. As demonstrated in this paper, generalization and quantity analysis can be used to generate fragmentary qualitative models of these causal mechanisms.

This simulation is obviously only a beginning. In addition to testing the simulation on a broader range of learning problems, we also plan to incorporate retrieval, using MAC/FAC (Forbus *et al* 1995). Having the simulation generate its own distinctions to explore, perhaps via failed predictions made with protohistories, is also an important problem to investigate.

Acknowledgments

This work was funded by the Office of Naval Research under grant N00014-08-1-0040.

References

Ahn, W. k., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition, 54*, 299-352.

Chi, M., Slotta, J., & De Leeuw, N. (1994). From things to processes: A theory of conceptual change for learning science concepts. *Learning and Instruction*, **4**(1), 27-43

diSessa, A. (1983). Phenomenology and the evolution of intuition. In D. Gentner and A. Stevens (Eds), *Mental Models*. 15-33. Hillsdale, NJ: Lawrence Erlbaum.

Falkenhainer, B., Forbus, K. & Gentner, D. (1989). The Structure-Mapping Engine: Algorithm and examples. Artficial Intelligence. 41.

Fisher, D.H. (1987). Knowledge acquisition via incremental concept clustering. Machine Learning 2, pp. 139-172.

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

Forbus, K.D & Gentner, D. (1986). Learning physical domains: towards a theoretical framework. In R.S. Michalski, J.G. Carbonell and T.M. Mitchell" *Machine Learning: An Artificial Intelligence Approach, Vol. 2.*

Forbus, K., Gentner, D. & Law, K. (1995). MAC/FAC: A model of Similarity-based Retrieval. *Cognitive Science*, 19(2), April-June, pp 141-205.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2), 155-170

Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review, 111, 1*, 1-31.

Kuehne, S. & Forbus, K. (2004). Capturing QP-relevant information from natural language text. *Proceedings of the 18th International Qualitative Reasoning Workshop*, Evanston, Illinois, USA, August

Kuehne, S.E., Forbus, K.D., Gentner, D., & Quinn, B. (2000). SEQL: Category learning as progressive abstraction using structure mapping. *Proceedings of CogSci 2000*, August.

Langley, P. (1981). Data-driven discovery of physical laws. *Cognitive Science* 5, 31-54.

Pearl, J. (2000). *Causality:Models, Reasoning, and Inference*. Cambridge University Press.

Piaget, J. (1930). The Child's Conception of Physical Causality. Kegan Paul, London.

Siegler, R. (1983). Five generalizations about cognitive development. *American Psychologist* 38, 263-277.