Qualifying Causes as Pertinent

Giovanni Sileno\(^1\) (giovanni.sileno@telecom-paristech.fr)
Jean-Louis Dessalles\(^1\) (dessalles@telecom-paristech.fr)

\(^1\)LTCI, Télécom ParisTech, Université Paris-Saclay, 46 rue Barrault, 75013 Paris, France

Abstract

Several computational methods have been proposed to evaluate the relevance of an instantiated cause to an observed consequence. The paper reports on an experiment to investigate the adequacy of some of these methods as descriptors of human judgments about causal relevance.

Keywords: Actual Causation, Relevant Cause, Counterfactuals, Bayesian Inference, Relevance Theory, Simplicity Theory.

Introduction

Causes play a central role in the way we conceptualize the world. Seeking explanations of events, or, stated differently, attributing responsibility for their occurrence, is common in practically all human activities. Despite such widespread use, however, there is yet no established model about how people qualify a cause as pertinent (literally, holding together) to a specific event. As observed by (Glymour et al., 2010), most psychological literature focuses on judgments of general causation (about causal regularities), or, when investigating actual causation (about situational interpretation), it focuses on specific applications like the perception of causation and animacy in minimal “mechanical” settings—see e.g. the overview in (Rips, 2011)—or on higher-level tasks, like the attribution of moral responsibility. The logic-philosophical literature, for its part, focuses on the most on finding and incrementally resolving paradoxes on existing models, with little concerns about practical settings. A similar situation holds on the computational side, where several approaches compete—with mixed results—for producing human-like inferences about causation. In this context, this paper aims to assess the gap existing between theoretical models and empirical observations: it considers a short selection of methods and animacy in minimal “mechanical” settings—see e.g. (Yuan, Lim, & Lu, 2011) for a reasoned overview.

Bayesian inference Intuitively, part of the problem of the lack of sensitivity of counterfactuals may be due to determinism. Turning to probabilistic methods, and in particular Bayesian probability, a full research track investigates how to explain why certain variables are observed in certain states; see e.g. (Yuan, Lim, & Lu, 2011) for a reasoned overview.

Let us suppose to we are able to encode the domain model in a Bayesian network (BN). Amongst the proposed choices for computing the relevance of the occurrence of the event \(C\) to the occurrence of event \(E\), the most commonly used are the likelihood \(p(E|C)\), as in maximum likelihood (ML) estimation, or the product \(p(E|C) \cdot p(C)\) as in maximum a posteriori (MAP) estimation. Alternatively, observing that the occurrence of a cause increases the possibility of occurrence of the effect, we could capture the raise of probability by computing differences as \(p(E|C) - p(E)\) or \(p(E|C) - p(E|-C)\), or ratios of these terms. By dealing with co-occurrences, these approaches can be seen as conflating causes to evidential supports. Interestingly, an experimental study on measures of evidential supports by (Tentori, Crupi, Bonini, & Osherson, 2007) finds empirical alignment of human responses with respect to two measures:\(^1\)

\[
\log \frac{p(E|C)}{p(E|-C)} \quad (1) \quad \frac{p(E|C) - p(E|-C)}{p(E|C) + p(E|-C)} \quad (2)
\]

Returning to the causal domain, to avoid undesired effects—like consequences that “cause” causes—additional machinery is required, e.g. introducing time-related constraints in the inferences—see e.g. (Williamson, 2009).

Alternatively, (Pearl, 2000) proves that there is strict connection between structural equations and Bayesian networks,

\(^1\)From a theoretical point of view, however, the authors argue that (2) is the only one to satisfy desirable mathematical properties.
and proposes an unifying notation—causal Bayesian networks (CBNs)—introducing an explicit do notation to distinguish interventions from standard probabilistic events. An acknowledged problem with this method is that variables relevant to the computation may be unobservable (Zhang, 2008).

At a more fundamental level, however, the problem of model adequacy, i.e. of which variables to include in the model, concerns all these computational methods. Intuitively, a cognitive basis might offer a more robust solution.

Relevance Theory  Relevance Theory (RT) identifies general principles that are supposed to govern successful communication. A statement is said to be relevant if the addressee is able to draw inferences from it. Since inferences may always be produced from any statement, RT prioritizes those which can be produced “effortlessly” (Sperber & Wilson, 1986). This principle is supposed to guide listeners in determining causal relations, as in the two following examples (Wilson & Sperber, 1998):

(i) John dropped the glass. It broke.
(ii) I got caught. My best friend betrayed me.

According to RT, causality is inferred because it enriches the context. In (i), the first statement is understood as a cause of the second one, because such a material relation is easy to access from experience. In (ii), the friend’s betrayal could be the consequence of the speaker’s having been caught. However, the converse causality is preferred because it is more “accessible” (and the speaker would have expressed things differently to mean otherwise). It has been observed that notions such as “effort” or “accessibility” are crucial for RT to make predictions and yet remain external to the theory (Levinson, 1989): there is a risk that RT’s principles be bent to justify any intuitively correct interpretation ex post facto.

Simplicity Theory Simplicity Theory (ST) has been introduced to account for interestingness, and offers an alternative definition of relevance (Dessalles, 2013). Relevant events must be unexpected, which means that they can be presented as more complex to generate than to describe. Cognitive complexities of generation and of description are measured as minimum description lengths (MDL) (Chater, 1999). In particular, generation or world complexity (denoted with $C_W$), generalizes the classical notion of (im)probability, as it computes odds without using set extensions. Considering two events $e_1$ and $e_2$, the complexity of their sequential composition (denoted by ‘$*$’) is (chain rule):

$$C_W(e_1 * e_2) = C_W(e_1) + C_W(e_2|e_1)$$  \hspace{1cm} (3)

where $C_W(e_2|e_1)$ is the conditional complexity of generating $e_2$ considering $e_1$ already realized.

Neglecting effects due to description complexity, a relevant (tentative) explanation, according to ST, is any piece of knowledge that diminishes generation complexity. The potential causal contribution of $c$ on $e$ is captured by $C_W(e) - C_W(e * c)$. When dealing with actual causation, $c$ is realized, so $C_W(c) = 0$. Thus, $c$ is a relevant actual cause for $e$ if the generation complexity $C_W$ of $e$ is smaller conditionally to $c$, i.e. $C_W(e|c) \ll C_W(e)$. A measure of relevance could then be:

$$C_W(e) - C_W(e|c)$$  \hspace{1cm} (4)

Consider again example (ii). “I got caught” is relevant as far as it can be perceived as unexpected: a complex set of circumstances was supposed necessary for this outcome to occur. The second statement “My best friend betrayed me” appears as a relevant cause as far as it makes the minimal causal path to being caught significantly shorter. Note that (4) is more constraining than RT’s principles. It states both that $C_W(e)$ is large and that $c$ provokes complexity drop once taken into account. When several causes are offered, the most relevant one should be the cause that provokes the largest drop.

Experimental test

Participants are asked to read five passages of a short story and to rank the pertinence of answers to simple why questions according to the scale NR: irrelevant, 1: low relevance, 5: high relevance (same ranking allowed). They are instructed to approach each answer as if it were the only answer produced by another locutor.

Passage 1 Johnny is 7 years old. In recent months his mother has been worried because he developed a craving for sweet things. She bought some pots of strawberry jam and put them into the larder (a small room near the kitchen). Then one afternoon she finds that Johnny has gone into the larder and has eaten half a pot of strawberry jam.

Q1. Why is "half a pot of jam gone"?
Q2. Why did "Johnny eat the jam"?
Q3. Why did "Johnny go into the larder"?

Each question had the same candidate answers:

- a. because of Johnny’s gluttony
- b. because Johnny ate it
- c. because mother has put the pot in the larder

save the second answer of Q3:

- b’, because johnny wanted to eat a pot of jam being there

Passage 2 The mother says: “That’s naughty. In the future you are never to enter the larder without my permission.” Several incidents then follow. First, Johnny gets a broom, hooks the pot of jam from above the shelf without entering into the larder and helps himself.

Q4. Why did "Johnny use the broom to hook the pot"?

- a. because Johnny wanted to take the pot
- b. because of Johnny’s gluttony
- c. because mother has put the pot in the larder
- d. because mother forbade Johnny to enter the larder
- e. because the broom was in the house
- f. because of gravity

---

2The story (the “legalistic child” case) is adapted from (Rissland & Skalak, 1989), in turn revisiting (Twining & Miers, 1982).
**Passage 3** Mother finds Johnny eating the jam, but he says to her: “I didn’t enter the larder”. Then another day, the cat enters the larder and attacks the salmon which mother has bought for a special occasion.

Q5. Why did "the cat enter the larder and eat the salmon"?
   a. because the cat was hungry
   b. because mother has put the salmon in the larder
   c. because the salmon was in the larder

**Passage 4** Mother, upstairs, hears Johnny laughing. She comes down to see Johnny standing outside the larder door watching the cat eating the fish. ‘I may not go into the larder’ he says.

Q6. Why did "the cat enter the larder and eat the salmon"?
   a. because the cat was hungry
   b. because mother has put the salmon in the larder
   c. because the salmon was in the larder
d. because mother forbade Johnny to enter the larder
e. because Johnny made fun of mother

**Passage 5** Finally, Johnny’s parents were out and Johnny was watched by his usually iron-willed babysitter, Maggie. Johnny’s parents forgot to tell Maggie anything about dinner. Supper was late and Johnny was hungry. Johnny asked permission from the babysitter to enter the larder. She said OK. Johnny feasted on jam. The questions were:

Q7. Why did "Johnny enter the larder"?
Q8. Why did "Johnny feast on jam"?

The candidate answers were for both:
 a. because he was hungry
 b. because of Johnny’s gluttony
 c. because Maggie granted him permission
d. because his parents forgot to tell Maggie about dinner
e. because supper was late

**Domain model**

Referring to the traditional terminology, each question specifies an *explanandum* (something which has to be explained), each answer proposes an *explanans* or explanation for the *explanandum*. The test implicitly builds upon three roles: X, the (virtual) person who asks the why question; Y, the (virtual) person who gives the answer; and Z, the (actual) respondent, who evaluates the response given by Y. In order to produce an answer or to evaluate its relevance, Z requires a model of the world in which the story has occurred. Such representation is not required to be isomorphic with the input in all details, but just to be an adequate synthesis.

**General action-scheme** All questions are about events—for the most, actions performed by an agent. Analyzing verbal reports of legal cases, (Pennington & Hastie, 1993) found that explanations of human behaviour in legal decision-making converge to the following *action-scheme*—here in the version proposed by (Bex & Verheij, 2011):

\[ \text{Motive} \Rightarrow \text{Intent} \Rightarrow \text{Action} \Rightarrow \text{Consequences} \]

To cover further cases, we considered additional elements (Sileno, 2016, Ch. 7); motive is interpreted as a situation perceived by the agent, triggering a preexisting motivation and producing an intent (here in the sense of specific desire); the intent develops into a certain action (or a course of actions) if coupled with perceiving the associated affordance and no inhibition is put in place by other motivational components; the action brings about a certain consequence depending on the actual environmental disposition. The following model components are then considered:

<table>
<thead>
<tr>
<th>M</th>
<th>motivation</th>
<th>T</th>
<th>motive</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>intent</td>
<td>F</td>
<td>affordance of I via A</td>
</tr>
<tr>
<td>A</td>
<td>action</td>
<td>H</td>
<td>inhibition of A</td>
</tr>
<tr>
<td>Q</td>
<td>consequences</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For instance, for a boy craving for sweets (M), the fact that there is a jam pot (T) “generates” a desire to eat that jam (I). If he is already able to eat it (F), he just does it (A). As a side effect, there will be less jam in the pot (Q). The relative dependencies of the model components are illustrated in the graph in Fig.1 (for simplicity perceptual with actual affordance are aligned). Inhibition, specified using an empty-circle arrow, is a *negative dependence*: the absence of the parent element enables the child element to occur.

Actions usually have consequences relevant to other actions. For instance the agent might need an additional action A’ to bring about F (e.g. to go near the pot); in this case the new action should be added in in A’s action model as a new element •F, parent of F. Note however that the generation of the affordance might be independent from the agent; when the action is intentionally preparatory (i.e. part of a plan to perform A), •F depends on I as well. Small case letters will be used in case of ambiguities.

**Passage 1** (Fig. 2) The central event for passage 1 is “Johnny eating the jam”. Reading the propositional content provided in questions and answers through an action-scheme centered around this action, the following associations hold:

Q1 “half a pot of jam is gone” is a consequence (Q)
Q2, b “Johnny eats the jam” is the core action (A)
Q3 “Johnny goes into the larder” is a preparatory action to perform the core action (•F)
   a “Johnny is gluttonous” identifies the motivation (M)
c “mother has put the pot in the larder” is an event generating the motive starting the course of action (•T)
b’ “Johnny wanted to eat the jam” is the intent (I)
Applying informally the Counterfactuals question).

The models in the previous section serve as a common ground to take it directly (t), however, the prohibition of the mother to enter (H₁) is inhibiting this easier action. Hooking the pot (h) is an alternative plan. Then, using the broom (u) to hook the pot (to eat the jam) is a nested preparatory action. The presence of broom (b) is a necessary condition to perform it, while gravity (g) is a necessary condition to hook the pot. Q₄ ("use the broom to hook the pot") cannot be modeled simply as u, a possible option is the composite action u*h.

Passages 3 and 4 (Fig. 4) The core action is the cat eating the salmon. Passage 4 brings to the foreground the non-intervention of Johnny. The boy is normally (n) expected to stop the cat (s), overriding the prohibition issued by his mother (●H₂). Making fun of her (f) is a possible motivation behind the anomaly. Q₅ and Q₆ are modeled as ●F*A.

Passage 5 (Fig. 5) The action is centered again around Johnny eating jam. A new motivational state is added: hunger (h)—caused by the supper being late (l), in turn a consequence of the lack of instructions by the parents (f). Hunger does not enter directly in the action scheme, but only behind the scenes, as the reason why Maggie gives permission (p) to enter the larder (e), thus overriding the prohibition.

Computing relevance

The models in the previous section serve as a common ground for a direct operationalization of the methods presented in the introduction (except for RT, as it does not specify the notion of "effort"; its comparative evaluation is then left as an open question).

Counterfactuals Applying informally the but-for test on the model of passage 1, all answers qualify as causes. Following the formalization given by (Hitchcock, 2001), each element of our example is modeled as a binary variable associated to the occurrence or the non-occurrence of the event. The associated set of deterministic structural equations would be:

\[ I := M \land T \quad A := I \land F \quad Q := A \land F \quad F := \bullet F \quad T := \bullet T \quad \bullet F := I \]

Each equation can be seen as encoding counterfactual information related to a causal dependence. Also in this case, applying Hitchcock’s definition of active route for the determination of actual causes, we qualify positively all answers. Same results are obtained with the other passages.

Bayesian inference The application of Bayesian inference requires the domain model to be encoded in a Bayesian network. In principle, these graphs should be diagrammatically very similar to e.g. Fig. 2. A practical problem arises for deciding the parameters of the conditional probability tables. Even acknowledging subjective probability (i.e. capturing degree or strength of belief), it is not evident to provide solid backup for these numbers. Nevertheless, this is a required step to proceed with this method. With subjective estimations as parameters, we have extracted the relevance measures defined in the introduction, obtaining the results on Table 1.

Simplicity theory The chain rule formula (3) enables us to run through the models in search of the shortest path from cause to effect. Path lengths are measured by the sum of conditional complexities associated to the transitions and the complexity of nodes with no parents required to proceed in the path. Let us assume that all the dependencies belonging to the general action-scheme (Fig. 1) carry similar complexity of transition C₀. The graph implies:

\[ C_W(I) = C_W(M) + C_W(T) + C_0 \]
\[ C_W(A) = C_W(F) + C_W(I) + C_W(\neg H) + C_0 \]
\[ C_W(Q) = C_W(F) + C_W(A/F) + C_0 = C_W(A) + C_0 \]
Empirical results

The participants to our experiment were then considered ordering decreasingly the participants’ rank-deviation of rankings are reported on Table 3. These measured the ranking of answers from 0 (irrelevant, NR) up to 102 individuals (54% female, 71% age 31-50), mostly European researchers, recruited via social networks. The test was conducted online. Analyzing the responses, we initially quantified the ranking of answers from 0 (irrelevant, NR) up to 5 (highly relevant), as in the test. The average and standard deviation of rankings are reported on Table 3. These measures are however not necessarily the most illustrative for our study, as the resulting histograms have various shapes. We then considered ordering decreasingly the participants’ rankings and reading the (minimal) ranking value attained by 51% of the population. This method enables us to associate to the selected value a majority (simple or qualified) for which that answer has at least that ranking of pertinence. Table 4 illustrates that the relative ordering of such minimal rankings is consistent passing from 51% to 75% of the population.

We have also extracted the relative ordering of the rankings given by individual respondents for each question, in case there were cross-relations between answers that were lost by the previous analysis. Even if respondents were instructed to consider options as independent, one can indeed reasonably expect some repositioning effects due to the available choices. These relative orderings, reported on Table 5 (1 means ranked as most relevant by participants), are consistent with the previous results, although they lose information about the relative gap of pertinence between answers.

Comparative evaluation

As we can see on the tables, no measure of Bayesian inference is fully consistent with our experiments; likelihood and (1) are more aligned, followed by (2). In all cases, we observe a pathological response for Q2b (tautological answer). In many cases, even if the most pertinent cause is correctly identified, the relative order between the answers does not follow the empirical results. This may be due to a wrong choice of parameters or even to wrong dependencies in the model. Unfortunately, the framework is quite opaque to model correction tasks.

In contrast, the relevance measure computed via ST are quite aligned to the empirical results, with fewer (constraints on) parameters and at inferior computational cost. The irrelevance of Q2b is correctly captured. The relative ranking of Q1 and Q2, in the plausible hypothesis that $CW(M) \gg CW(T)$ (the child being glutonous vs putting jam in the larder), is correct. Q4 is also aligned: $a$ is necessarily the most pertinent cause; $d$ the second one, at the condition that $CW(H_e)$ is sufficiently high (e.g. by considering a plausible dependency of $H_e$ w.r.t. $H$); $b$ and $c$ are consistent with the previous ranking; $e$ and $f$ have low pertinence, because their complexity is very low. For Q4 and Q5, consistency holds if

Table 1: Relevance measures computed using Bayesian networks, with parameters: for passages 1, 2, 5, $p(*|T)$ = 0.5, $p(M | *T) = 0.2$, $p(*|PT) = 0.2$; for passage 2, $p(H_e) = 0.6$, $p(b) = 0.9$, $p(g) = 1$; for passages 3 and 4, $p(M | *T) = 0.2$, $p(*|T) = 0.5$; for passage 4, $p(f) = 0.1$, $p(H_e) = 0.8$; for passage 5, $p(f) = 0.3$.

Table 2: Relevance strengths computed as drops of generation complexity ($C_0$ is the transition complexity of action-scheme dependencies) and then constraints as: $CW(I) > CW(M)$ and $CW(I) > CW(T)$, etc. Turning upon the passage models\(^1\), by applying (4) we obtain the expressions reported on Table 2, from which we can extract similar constraints. Note how this analytical form does not require to decide the parameters upfront.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>0.15</td>
<td>0.53</td>
<td>0.11</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Q2</td>
<td>0.16</td>
<td>1.01</td>
<td>0.11</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>Q3</td>
<td>0.15</td>
<td>0.09</td>
<td>0.12</td>
<td>0.04</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>Q4</td>
<td>0.67</td>
<td>0.13</td>
<td>0.11</td>
<td>0.12</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td>Q5</td>
<td>0.33</td>
<td>0.18</td>
<td>0.19</td>
<td>0.07</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>Q6</td>
<td>0.85</td>
<td>0.47</td>
<td>0.16</td>
<td>0.66</td>
<td>0.51</td>
<td>0.14</td>
</tr>
<tr>
<td>Q7</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
<td>0.12</td>
<td>0.13</td>
<td>0.04</td>
</tr>
<tr>
<td>Q8</td>
<td>0.13</td>
<td>0.16</td>
<td>0.12</td>
<td>0.11</td>
<td>0.12</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 3: Ranking attributed to answers: mean ± st.dev.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>4.3</td>
<td>5.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q2</td>
<td>5.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q3</td>
<td>3.2</td>
<td>5.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Q4</td>
<td>5.4</td>
<td>3.2</td>
<td>2.1</td>
<td>5.4</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q5</td>
<td>5.4</td>
<td>2.1</td>
<td>4.2</td>
<td>5.4</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q6</td>
<td>5.3</td>
<td>3.1</td>
<td>4.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q7</td>
<td>5.3</td>
<td>3.1</td>
<td>5.4</td>
<td>2.1</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>Q8</td>
<td>5.3</td>
<td>5.4</td>
<td>1.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

Table 4: Minimal rankings for simple (qualified) majorities.

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>4.3</td>
<td>5.5</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q2</td>
<td>5.4</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q3</td>
<td>3.2</td>
<td>5.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
<td>4.3</td>
</tr>
<tr>
<td>Q4</td>
<td>5.4</td>
<td>3.2</td>
<td>2.1</td>
<td>5.4</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q5</td>
<td>5.4</td>
<td>2.1</td>
<td>4.2</td>
<td>5.4</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q6</td>
<td>5.3</td>
<td>3.1</td>
<td>4.3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Q7</td>
<td>5.3</td>
<td>3.1</td>
<td>5.4</td>
<td>2.1</td>
<td>3.1</td>
<td></td>
</tr>
<tr>
<td>Q8</td>
<td>5.3</td>
<td>5.4</td>
<td>1.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
</tr>
</tbody>
</table>

\(^1\)In a context in which $A$ would occur if not inhibited by $H$: $CW(A | \neg H) = C_0$, the double inhibition $B \rightarrow H \rightarrow A$ translates into $CW(\neg H | B) = C_0 + CW(B)$ and so $CW(A) = 2C_0 + CW(B)$.
the fundamental problem of contextualizing interpretation by
deciding the set of variables under study upfront. This study
was necessarily limited to models made by hand to test the
adequacy of methods; however, the positive confirmation of
good judgment prediction of ST theory is a strong motivation
towards automatizing the model construction process, by
inverting the problem: an action-scheme is nothing more than
tentative answers to a sequence of why questions.

References

Bex, F., & Verheij, B. (2011). Solving a Murder Case by Asking
*Algorithmic probability and friends*, 7070 LNAI, 119–130.
Glymour, C., Danks, D., Glymour, B., Eberhardt, F., Ramsey,
Hitchcock, C. (2001). The Intransitivity of Causation Revealed in
Linguistics*, 25(2), 455–472.
Pennington, N., & Hastie, R. (1993). Reasoning in
explanation-based decision making. *Cognition*, 49, 123–
163.
based and rule-based reasoning: A heuristic approach. In
Proceedings of the 11th IJCAI (pp. 524–530).
Sileno, G. (2016). *Aligning Law and Action*. Doctoral dis-
sertation, University of Amsterdam.
Comparison of confirmation measures. *Cognition*, 103(1),
107–119.
*Relevance theory: Applications and implications* (pp. 1–
Yuan, C., Lim, H., & Lu, T. C. (2011). Most relevant explana-
tion in bayesian networks. *Journal of Artificial Intelligence