Explanatory reasoning for inductive confidence

David H. Landy
dhlandy@gmail.com
John E. Hummel
jehummel@illinois.edu
Department of Psychology, 603 E. Daniel St.
Champaign, IL 61820 USA

ABSTRACT

We present a novel strategy for combining a probabilistic logic with analogical inference. We apply the resulting inference engine to the task of categorical induction--deciding whether a category bears a property, given that other, related categories do or do not have that property. Following suggestions by Murphy, Sloman, and others, we model categorical induction as a process of causal reasoning, by which knowledge gleaned from possible explanations of the premises is used to form conclusions about the category.

INTRODUCTION

Explanation plays a central role in human cognition. People generate explanations for events quickly, flexibly, habitually and more or less explicitly, and the resulting explanations provide a powerful basis for reasoning and persuasion. Explanations help us to understand events (both expected and unexpected; e.g., "he's walking on the street because the sidewalk is closed") to solve problems (e.g., "the car won't start because it's out of gas") and to predict the future (e.g., "the frogs are dying off because of pollutants in their environments. If we don't act soon, other animals will start dying off as well."). They also play a central —perhaps *the* central—role in scientific reasoning.

People are also habitual inductive inference generators (see, e.g., Holland et al., 1989), and as such we need some basis for estimating the likelihood of our inductive inferences. This kind of likelihood estimation is typically discussed and modeled in terms of Bayesian (e.g., Griffiths & Tennenbaum, 2005; Kemp & Tennenbaum, 2009) or causal (e.g., Cheng, 1997; Novick & Cheng, 2004) reasoning. We present a model, Explanatory Reasoning for Inductive Confidence (ERIC), based on an alternative account of the process whereby we estimate the likelihood of our inductive inferences.

The central hypothesis motivating ERIC is that explanations, especially explanations of the more implicit variety, play an essential role in our estimates of the likelihood of our inductive inferences (Lombrozo, 2006; Medin, et al, 2003; Sloman, 1994). For example, imagine that you hear that many robins in a particular city are infected by some new disease. According to ERIC, upon hearing this news you would immediately attempt (more or less implicitly) to understand *why* they have the disease. That is, inductive generalizations are generated through an abductive explanatory process. We assume that this process is likely to be largely implicit because we assume that you generate many potential explanations (e.g., "perhaps this is something that affects all songbirds"; "perhaps it is something that afflicts all birds", "perhaps it is something about eating worms," etc.), implicitly assigning a likelihood estimate to each. If you are then asked to assess the likelihood of a related induction—for example, "what is the likelihood that sparrows will also be affected by this disease?"—we hypothesize that, as you did with the robins, you attempt to generate explanations for why sparrows might get the disease.

But this time, the knowledge you bring to bear on generating these explanations includes the explanations you generated for robins. The more successfully the knowledge in your long-term memory (LTM)—including your explanations for the disease-stricken robins—generates plausible and likely explanations for the inference, the more likely you will regard the inference that sparrows will get the disease.

Analogy plays a crucial role in ERIC: because knowledge about one category is never knowledge about another, explanatory knowledge must be adapted from one set of categories to any other. Knowledge, in ERIC, is never represented as rules over abstract variables; instead, knowledge is always represented as propositions about properties associated with specific categories, or about (generally causal) relations between such propositions. Analogical mapping is used to integrate surface similarity and structural commonalities between categories so that knowledge from old situation can be applied to current situations, to the degree that the situations correspond. That is, we treat analogical correspondence as a foundation of symbolic (rule-based) thought (see Gentner and Molita, 1998, who make just this suggestion).

While there is insufficient space in this paper to describe the model formally, the next section will attempt to provide an overview of its processes and assumptions.

MODEL

ERIC takes as input an *explanandum*—i.e., a thing to be explained: either a *premise*, which is assumed to be true (e.g., "robins get the disease"), or a *query*, whose probability is to be estimated (e.g., "sparrows will get the disease")—and generates explanations of why the explanandum might be true. Applied to property induction, the mechanism operates in two stages: First, ERIC explains the premise(s) and any knowledge gleaned from those explanations is added to the knowledge base. Next, it explains the query using that augmented knowledge. The result of these processes is an estimate of the likelihood that the query is true.

Given an explanandum, ERIC uses a base of knowledge—including propositions describing aspects of the world and causal connections between those aspects—to decide which of many possible circumstances are the actual cause of the explanandum. In some cases, ERIC evaluates potential causes already stored in its knowledge; in others it uses analogical mechanisms to postulate new facts or new causal links. In the formalism reported here, the ACME algorithm was used (Holyoak & Thagard, 1989); however, the character of the model is not tied to the particulars of the mapping algorithm. Each coherent set of structures which describes a possible cause of the explanandum is considered a potential explanation. The degree to which the original facts and causal connections are believed, together with the quality of the licensing inferences, are used to assign to each explanation a likelihood.

Consider Figure 1, which presents a possible state of knowledge. If C is the explanandum, then one possible explanation of C is A, which is weakly believed to be the reason that C is true. However, to the degree that C' expresses a state of affairs similar to C, another possible cause of C is generated by analogy to B' (which is strongly believed to be the cause of C'). Note that in this example, each node represents a group of statements (possibly a group of size 1); analogical inference is used to determine how and how well the entities and relationships referred to in one group correspond to those in another. Thus, the structures that serve to constrain the analogical inference are not detailed in Figures 1 and 2.

Figure 2 illustrates the result of the analogical inference process: a novel fact B is (weakly) hypothesized; a connection between B and C is inferred with a level of belief sensitive both to the strength of the connection between B' and C', and to the strength of the analogy; and stable analogical connections between B and B', and between C and C', and also between their component aspects are established (the effect of stable analogical connections will be discussed later). In the illustrated case, B is a novel fact, and the corresponding explanation is correspondingly weak; sometimes it is the case that the inferred explanatory fact, B, is already known (but not linked to A). In these cases, the explanation will tend to be much stronger.

The cause of each explanation (A and B in the example) is recursively explained (to a fixed depth), and analogical support, where available, is applied to each causal connection.

Explanations combine to lend credence to the explanandum, after the manner of a probabilistic argumentation system (Haenni et al, 2000). If the explanandum is a newly given fact (i.e., a premise), then knowledge involved in the explanation is updated using a Bayesian strategy: each piece of knowledge involved in any explanation is updated to the degree to which confidence in it would tend to lead to confidence in the (known) explanandum. This updated knowledge is used to evaluate queries.

This approach captures the intuitions that inductive judgments are based on (abduced) explanations, and explanations involve the application of analogical reasoning between distinct but related situations as a source of both novel guesses about causes, and of confidence in those causes. Furthermore, as we will demonstrate, the resulting model conforms to many observed patterns of human induction.

Two additional connectives are also used. First is the *causes relation* (⇒), which denotes a causal relationship.

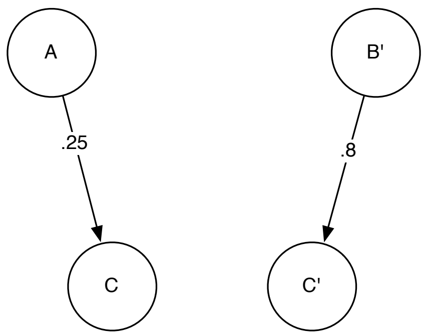

Figure 1: A possible state of the knowledge of ERIC. Facts A, B', and C' are known with certainty by the model; C represents an unknown explanandum. One possible explanation for C is that it results from A.

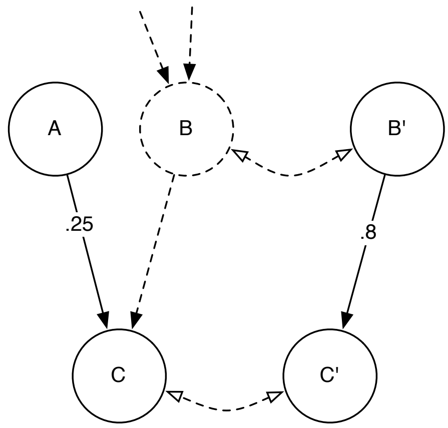

Figure 2: The result of ERIC's analogical inference mechanism acting on the knowledge in Figure 1. The dashed figures reflect the analogical mapping, and the inferences licensed by it. The arrows into B reflect ERIC's subsequent search for support for the hypothesized B.

For example, $q \Rightarrow r$ should be read as "q (if true) would tend to explain (cause) r". This is similar to implication, but is constrained to causal implication, and has a rather different role in the inference algorithm (described later). Note that, in contrast to some prior models (e.g., Falkenhainer, et al 1989; Hummel & Holyoak, 1997, 2003), causal connections are treated as special types, and not as labelled two-place predicates (see also Hummel, et al., 2008). The second novel connective is the *mapping relation*, $q \leftrightarrows r$, which states that q maps (i.e., have been identified as at least homomorphic) to r. We will denote conjunctions in mapping and cause relations as lists, in order to emphasize the intended psychological structures and to improve readability. For example, we will say {isa(dog, mammal), warm-blooded(mammal)}⇒warm-blooded(dog): "together, the fact that a dog is a mammal and the fact that mammals are warm-blooded cause dogs to be warm-blooded".

Analogically Projectable literals as variables

In ERIC, unlike most traditional propositional systems, there are no variablized inference rules; the only general rules are those expressing the syntax of explanations and the probability estimates of combined terms. Rules governing inference are literal instances of prior knowledge. The applicability of known facts is governed first by their level of endorsement (i.e., the probability estimate of the statement) and the *projectability* of their categories to the corresponding categories in the candidate statement. Projectability is quite similar to similarity, but is intended to capture the internal generalizability of literals. Its formal definition will not be given here (a similar, similarity-based approach, has been used to implement rules in the CLARION architecture; Sun, 2006).

Casting rule application as a process of analogy over literals is consonant with several recent suggestions (Gentner and Molita, 1998; Hummel & Holyoak, 2003; Pothos, 2007). Replacing variables with projectable literals provides a natural way to create rules that are softly domain-limited. That is, rules are less and less applicable to more and more distant items and topics, but are not explicitly limited.

SIMULATIONS

For this paper, we chose to model reasoning in cases where knowledge is largely taxonomic. Inductions over taxonomic knowledge are well understood, and form a reasonable basis for evaluating models of induction (Heit, 2007). After presenting these results, we will provide initial simulations suggesting that the same model can also respond appropriately to questions about properties that are unlikely to be explained by taxonomic relations.

The simulations reported were run using arbitrarily chosen, rather than optimized values for the free parameters, such as the initial probability of an unsupported hypothesized fact (0.1) or causal relation (0.001). Depth 3 solutions were constructed for all testing and reporting purposes. We also explored a variety of solution depths and free parameter values; results were qualitatively similar for a variety of meaningful values of the parameters.

Two knowledge structures were used most of in these simulations. First, a taxonomic structure was constructed: several types of "animal" were described and several examplars of each were described (2 mammals, 6 birds, and 2 reptiles). Facts were included in knowledge stating that each animal and each type of animal was a variety of animal. So for example isakindof(robin, animal) was included in the knowledge base with a strength of 1.0. However, isakindof(animal, animal) was deliberately given a strength of 0.0, as was isakindof(animal, robin). Also, "animals" were described as a kind of "living thing". One causal relation given: a generic "property inclusion cause", corresponding to "If *x* is a kind of *y*, and property *z* holds for *y*'s, then *z* will tend to cause it to hold for *x*'s as well".

The second knowledge structure included both taxonomic and non-taxonomic knowledge. Approximately 200 facts were added to the knowledge base, including causal stories (e.g., "turtles are protected from attacks because of their hard shells"), typical properties of categories, and relations between properties. Results for each simulation are presented using each of these knowledge structures.

Heit (2000) reviews several benchmark properties of taxonomic inductions. We illustrate how ERIC produces several of these behaviors, and then consider some more subtle behaviors in non-taxonomic domains.

Similarity and inductive strength

In ERIC, property induction depends on the extent to which an assertion can be explained. Such explanations, however, apply not only to their target assertion, but to "nearby" assertions as well.

For instance, say that property f is asserted of category A. ERIC will induce several explanations of that fact. When presented with the query f(B), two factors will impact how "near" that situation is to the original f(A). First, the more A and B engage in similar statements, the more likely it is that explanations of f(A) will be to be adapted to strong arguments about f(B). Second, because ERIC uses projectability to map past examples to new situations, and projectability is quite similar to similarity, to the extent that B is similar to A, the analogical mappings will describe more similar situations, and therefore the explanations assigned a higher strength.

The fact that property inductions between categories are stronger when the categories are similar was one of the first properties of inductive arguments to be identified in the psychological literature (Rips, 1975).

Results from ERIC were obtained using a blank predicate (i.e., a predicate that appeared nowhere in knowledge, and was uniformly similar to other predicates). A series of inductive premise-query pairs was formed, of the form if Blank(A), how likely is it that Blank(B)?, for all distinct categories A and B.

Figure 3 (top panel) displays the relationship between the projectibility of A to B, and the resulting inductive strength.

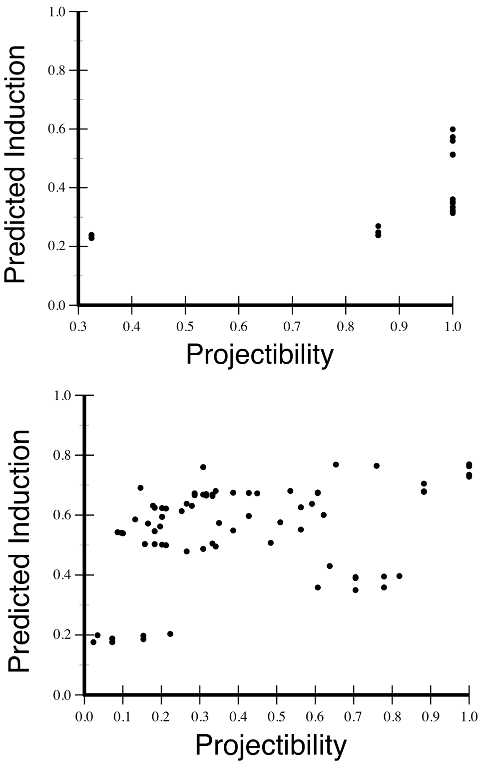

Figure 3: Relationship between projectibility and inductive strength for taxonomic knowledge only (left), and with both taxonomic and additional featural and explanatory knowledge (right)

As predicted, in the absence of other knowledge, projectibility and induction are strongly related. The bottom panel of Figure 3 displays the results of the same analysis, from simulations using the slightly richer domain knowledge.

The positive relationship between projectibility and induction is still apparent (r=.38, p<.01), but considerably more structure can be seen. This structure results from the fact that good explanations now do not so closely match featural overlap. Good explanations, in this knowledge base, largely result from taxonomic proximity (that is, most explanations are of the form "Probably robins have blank because birds generally have blank." Feature overlap tends to correspond to taxonomic category, but sometimes deviates. Thus, inductions between atypical members of categories may be stronger than their projectibility would suggest. Similarly, highly similar categories from different branches of the taxonomy may be highly projectible; however explanations in terms of taxonomic class will adapt only weakly between them.

Category Inclusion

ERIC does not represent categorical statements as absolute. That is, if it is true for ERIC that flies (bird), it is not necessarily the case that flies (robin). Therefore, it is possible for different supercategories of a particular category to differently generalize to that category. Indeed, in general, the farther up in the a type hierarchy a supercategory is from a category, the less properties of that supercategory will be generalized to the more particular category. This behavior can be seen in Figure 4, and is a well-known property of human categorization (Osherson, 1990; Sloman, 1998): close superodinates project their properties more strongly to a cagetory than far superordinates. In ERIC, this happens for two reasons. First, a category is likely to share more features and typical explanations of those features with a proximal superordinate category than with a distant supercategory, and these shared

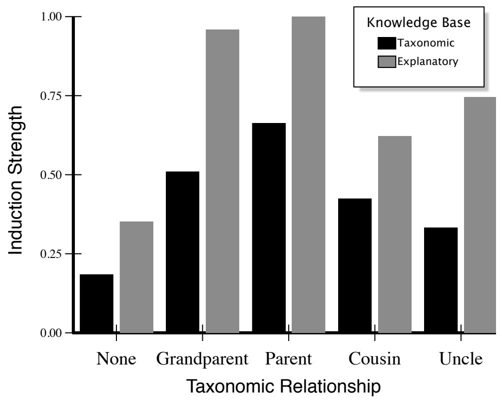

Figure 4. The average strength of an induction to a species category from various other categories.

features will generalize well. Second, because the close superordinate can be explained by a more distant taxonomic ancestor, asserting a property of the close superordinate tends to lead ERIC to assume that the property holds of the distant one as well. Thus, there are more good,
supported explanations of the query category when the close superordinate is the premise.

Typicality of premise categories

We'll consider two different ways of characterizing the extent to which a subcategory is typical of a basic category. First, we'll use *feature typicality* to refer to the idea that a category is typical of its superordinate to the extent that they simply share features, without regard to reasons why both have those features. This is the kind of typicality employed in Osherson et al's similarity-coverage model (1990) and the connectionist feature overlap model (Sloman, 1993).

The other notion of typicality we'll evaluate explicitly considers explanatory structure as a determiner of causality. In this account, robins are not typical birds because they have bird features, they are typical birds explicitly *because they got their bird features in the right way* (see Sloman, 1994). On this kind of typicality, which we'll call *causal typicality*, a subordinate is typical of its superodinate to the extent that it shares features with the superordinate, and that it shares those features either in virtue of being a member of its category, or for the same reasons that apply to the superordinate category. We report results using the former: typical features of category members are those explained in terms of the superordinate category. Kemp and Tenenbaum (2009) come close to this approach when they assume that a category features result from a generating mechanism common to both super- and sub-ordinates.

ERIC is capable of representing both these kinds of typicality, and attributes to them different kinds of effects on inductions from a category to its superordinate. To examine these differential effects, we constructed four "bird" subtypes: "eagle", "sparrow", "blue heron", and "penguin". The bird category was given four features (i.e., generic unary predicates were added to knowledge of the form X_i(bird)), and also made a subtype of animal (and the taxonomic animal knowledge was included on all typicality tests). The feature overlap of the subcategories was as follows: sparrow had all of the bird features, and no additional features. Blue heron had two of the four bird properties, and no additional properties. Eagle, like sparrow, had all four bird properties, but also had two additional idiosyncratic properties. Finally, penguin had only one typical bird feature, and had three additional idiosyncratic features. Thus, both penguin and sparrow have four features, and vary only in the typicality of those features. Blue heron, sparrow, and eagle show the relationship between number of features and typicality.

To examine *causal typicality*, we used the same feature knowledge set, but added to knowledge simple explanations of the source or justification of each feature. The explanations in this set were all taxonomic, e.g., *isa*(eagle, bird), *sings*(bird) ⇒ *sings*(eagle). For the idiosyncratic features of eagle and penguin, alternative super-ordinates were created (e.g., *isa*(penguin, polar_animal), *lives_on_ice*(polar_animal)⇒*lives_on_ice*(penguin)). All typical features were explained with reference to the bird category.

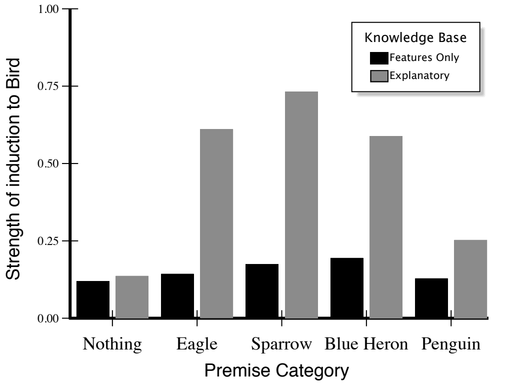

Figure 5: Inductive strength of a blank property to bird, from premise categories of varying typicality

Results are displayed in Figure 5. Consider first the explanatory knowledge base: this base matches the empirical finding that more typical members more strongly support inductions to the superordinate (Osherson et al, 1990; Rips, 1975): penguin only very poorly licenses inferences about birds in general. Sparrow—the most typical category member—appropriately licenses induction to bird best. We can see importance of the linking explanations, however, when we turn to the feature-based knowledge set. Here the blue heron category generalizes best, and penguin is only slightly behind eagle. This happens for the reason we highlighted earlier: extra knowledge about a category decreases its projectability, all else being equal. The knowledge pattern on which taxonomic reasoning is based in the features-only knowledge set is restricted to a very simple, general pattern: isa(x,y), f(y) ⇒ f(x). Since Blue Heron has fewer features than Sparrow, this rule projects better onto it, leading to the reversal in the first two groups. Thus, without linking explanations, extra features block the application of general knowledge, even when the premise and conclusion category tend to share features. It would seem that ERIC requires linking explanations to match empirical demonstrations of typicality.

Number and Diversity of premise categories

One of the more controversial properties of property inductions concerns how the number and diversity of premise categories affects the strength of an induction. In general, it appears that the more premise categories are given to have a property, the more strongly that property will be extended to categories unrelated to the premises (Osherson, 1990; Nisbet, 1983). Furthermore, in such cases, the more diverse the premises, the more powerful the argument, at least for certain reasoners in appropriate domains (Osherson, 1990; López, 1997). Osherson et al. observed what they called *nonmontonicity effects*. For example, an argument from flies having a property to bees having it was judged as stronger than an argument from flies and orangutans to bees (see also Sloman, 1993). The absence of diversity-based reasoning has frequently been observed when reasoners have rich causal knowledge about the content domain (López et al, 1997; Medin et al, 2002).

The role of rich causal knowledge in explanation will be discussed shortly. First, we consider more traditional cases of diversity and its exceptions. As shown in Figure 6, in general a larger number of premises increases ERIC's estimate of the power of an argument, and more diverse premises are beneficial over similar premises. However, this pattern interacts strongly with the relationship between the premises and the conclusion. When one premise is similar to the conclusion, arguments are little benefited by additional, less similar premises. Although the model shows a decreased benefit in these cases, under its current formulation it did not produce the reversal of diversity found in human judgments.

A final feature of property inductions is that the property being extended (more precisely, the reasoner's knowledge about the property) substantially affects inductions made. To simulate this fact, we created knowledge corresponding to the taxonomic and predatory structures explored by Shafto et al (2008).

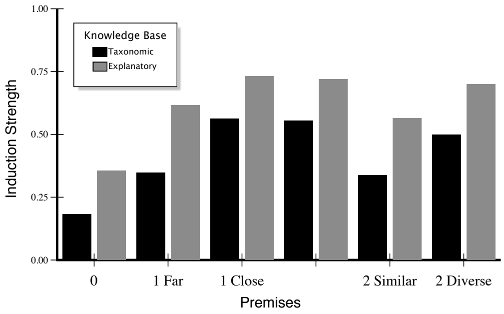

Figure 6: Inductive strength of a blank property based on the number, similarity, and diversity of premise categories.

Property induction in knowledge-rich contexts

Beyond the predatory and taxonomic knowledge, we provided basic background knowledge about two properties. One property was identified as "having a particular bone", by having five features common to a generic "body parts" category. Knowledge about property distribution was restricted to a single general statement that body parts of super-ordinates are often shared with subordinate types. The second property was identified as a "disease", by having five features common to a generic. Explanatory knowledge about diseases consisted of a general transmission-based case. Both diseases and body parts were identified as kinds of properties, and as single general fact about properties stated that they are often shared from super-ordinates to sub-ordinates. Then, for each pair of animals, each property was generalized from each member of the pair to the other. Results are presented in Figure 7.

Induction on having a particular bone was always primarily taxonomic. Explanations involving eating occurred (remember that *all* conceivable explanations are generated) but were accorded very low strengths. On the other hand, induction over disease susceptibility was more complex: when predation relations went from premise to query, inductions were strong and the strongest explanations tended to be in terms of predation.

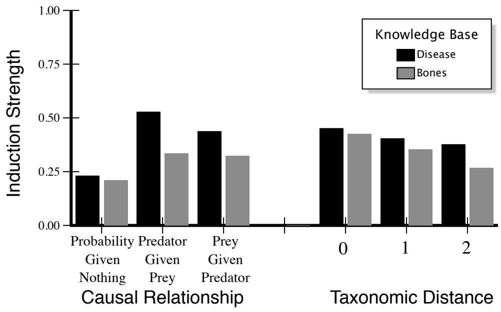

Figure 7: Dependency of inductive strength on property and category relationships.

When predation relations went from query to premise, inductions were generally weaker and the strongest explanations were frequently taxonomic. Similarly, when predation relations did not connect the premise to the query, generalizations were primarily taxonomic. These patterns are quite similar to those found to obtain in human judgments (Shafto et al, 2008), and demonstrate the power of ERIC to adjust its application of "rules" over different domains of its knowledge.

Because ERIC "knows" only two reasons why a creature would get a disease (taxonomic similarity to another creature that is susceptible, and transmission through eating), when simply asked to explain why a predator gets a disease, ERIC forms the hypothesis that perhaps its prey gets the disease. Also, notice that both properties degrade taxonomically. This is because both kinds of knowledge are in the system, and so both affect, to some degree, the same judgments. The model predicts that people will also blend different theories and domains of knowledge when making inductions.

DISCUSSION

ERIC uses a probabilistic logic to encode the intuitions that knowledge-driven explanations are generated in support of property inductions, and that novel explanations are generated via analogical adaptation from nearby cases. We demonstrated that, using these intuitions, ERIC is capable of capturing many known phenomena in property inductions, including the specific role of knowledge about property distributions (e.g., predator-prey relations versus taxonomic distance relations).

Compared to other analogical inference systems, ERIC is distinctive in eschewing complex analogies between large schematic structures. Instead, ERIC builds explanations through many, smaller analogies between analogs often consisting of just two or three propositions. While ERIC's complete lack of large-scale structure surely reflects a simplification, postulating analogical extension as an implementation of symbolic inference suggests that many, if not most, analogies are of just this sort.

ERIC occupies a space quite close to a recent model by Kemp and Tenenbaum (2009). Their model makes property inductions in two steps: first information about a domain is used to construct a functional form for the distribution of properties across the domain; second that function is used to generate prior probabilities for the property distribution. These priors (after the usual application of probability theory using given premises as data) produce induction strengths. We are also using domain knowledge to predict estimates of induction strength, and assuming that different knowledge about property distributions is responsible for differential inductions.

Kemp and Tenenbaum's model postulates that individual features or characteristics are used to produce domain-scale property distribution functions, and these functions are responsible for producing induction estimates. The problem of relevance—of deciding which properties or features should be used to construct the domain theory—is sidestepped, and the authors focus on evaluating whether the use of premises to bias estimates based on well-constructed domain theories can produce human-like responses (and they demonstrate quite effectively that it can).

Our approach, in contrast, assumes that the extension of properties is based on the confluence of local, relevant properties, and that individual properties are extended on the basis of families of arguments. This set of assumptions allows the application of analogical reasoning, and the construction of individual explanations of particular extensions. Implicitly, the knowledge in ERIC constrains estimates of the probability of different property distributions, and it does so in domain-dependent ways. However, rather than constructing those distribution probabilities explicitly, as in Kemp and Tenenbaum's model, ERIC uses knowledge directly to generate estimates. Our hope is to account for some of the ways people decide what knowledge to use in the service of inductions.

Limitations

ERIC leaves much undone. First, at this point it will not make arguments *against* hypotheses. Generally, this means that probabilities rise and do not fall. Worse, because there is only one probability value associated with any statement, the model has no way to assert that a fact is certain to be false, even if it might be explained somehow.

Another fundamental feature of real explanations beyond the scope of the current modeling effort is the role the purposes and target audience play in constraining the form and content of explanations (Van Fraessen, 1980; Chin-Parker & Bradner, 2008).

Despite these limitations, ERIC does successfully apply fundamental intuitions from theory-theory and analogical reasoning literatures. From the former comes the idea that explanations of known or given facts are responsible for property induction evaluations. From the latter comes the intuition that reasoning often derives from analogies to similar situations. ERIC comprises one attempt to instantiate these theoretical ideas in a formal framework, and demonstrates that such a formalization can capture much of what we know about property inductions.

ACKNOWLEDGEMENTS

We thank Derek Devnich, Eric Taylor, and Brian Ross for substantial suggestions. This research was funded by AFOSR Grant # FA9550-07-1-0147.

REFERENCES

Cheng, P. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367–405.

Chin-Parker, S. Bradner, A. (2008). The pragmatics of explanation. In V. Sloutsky, B. Love, K. MacRae (Eds.), *The 30th Annual Conference of the Cognitive Science Society*. Washington, D.C.

Falkenhainer, B., Forbus, K. D., and Gentner, D. (1990). The structure mapping engine: Algorithm and examples, *Artificial Intelligence, 41*:1-63.

Griffiths, T. L., Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51*, 334-384.

Haenni, R., Kohlas, J. Lehmann, N. (2000). Probabilistic argumentation systems, in: J. Kohlas, S. Moral (Eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Vol. 5: Algorithms for Uncertainty and Defeasible Reasoning*. Kluwer Academic, Dordrecht.

Heit, E. (2000). Properties of inductive reasoning, *Psychonomic Bulletin and Review*, *7*(4), 569-592.

Heit, E. (2007). Models of inductive reasoning. In Ron Sun (ed), *Handbook of Computational Psychology*. New York: Cambridge University Press, 322-338.

Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. (1986). *Induction: Processes of inference, learning, and discovery*. Cambridge, MA: MIT Press.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295-355.

Hummel, J., Holyoak, K. J. (1997). "Distributed Representation of Structure: A Theory of Analogical Access and Mapping." *Psychological Review*, *104*(3), 427-466.

Hummel, J. & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*(2), 220-264.

Hummel, J.E., Landy, D., Devnich, D. (2008). Toward a process model of explanation with implications for the type-token problem. *The Association for the Advancement of Artificial Intelligence Fall Symposium*, Arlington, VA.

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological Review*, *116*(1), 20-58.

Lombrozo, T (2006). The structure and function of explanations. *TRENDS in Cognitive Sciences*, *10*(10), 464-470.

López, A. (1995). The diversity principle in the testing of arguments. *Memory & Cognition*, *23*, 374-382.

Medin, D. L., Coley, J. D., Storms, G, Hayes, B. K., (2003). A relevance theory of induction. *Psychonomic Bulletin and Review*, *10*, 517-532.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review*, *90*, 339-363.

Novick, L. R., & Cheng, P. W. (2004). Assessing interactive causal inference. *Psychological Review*, *111*, 455–485.

Osherson, D. N., Smith, E. E., Wilkie, O., López, A., & Shafir, E. (1990). Category-based induction. *Psych Review*, *97*,185- 200.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning & Verbal Behavior*, *14*, 665-681.

Shanahan, M. (1989). Prediction is deduction but explanation is abduction. *Proceedings IJCAI*, 1055-1060.

Sloman, S. A. (1993). Feature-based induction. *Cognitive Psychology, 25*, 231-280.

Sloman, S.A. (1994). When explanations compete: the role of explanatory coherence on judgments of likelihood. *Cognition 52*, 1-21.

Sloman, S.A.(1998). Categorical inference is not a tree: The myth of inheritance hierarchies. *Cognitive Psychology*, *35*, 1-33.

Sun, R. and Zhang, X., (2006). Accounting for a variety of reasoning data within a cognitive architecture. *Journal of Experimental and Theoretical Artificial Intelligence*, *18*(2), 169-191.

Van Fraassen, B. C. (1988). The pragmatic theory of explanation. In J. C. Pitt (Ed.), *Theories of explanation.* New York, NY: Oxford.