Explorations in the Contributors to Plausibility

Cynthia L. Loiselle
Paul R. Cohen

Department of Computer and Information Science
University of Massachusetts

ABSTRACT

In previous work, we identified a method for automatically deriving possible rules of plausible inference from a set of relations, and determined that the transitivity of underlying characteristics of the relations was a significant factor in predicting the plausibility of inferences generated from these rules. Recent work by other researchers has also focused on identifying these kinds of characteristics and examining their role in the ability to predict plausibility. We examine these sets of characteristics and conclude that those factors that preserve transitivity provide most of the power of these systems. We then show how inferences can be used to determine the intended semantics, and thus the appropriate set of representational features, of a relation.

INTRODUCTION

One important aspect of research on semantic relations is understanding their behavior in inferences. Studying inferences forces us to examine how we reason with these relations. Of particular interest are common sense or *plausible* inferences, inferences whose rules suggest conclusions that are not guaranteed to be true but are true often enough to be useful. Unlike deductive inference where, given the truth values of the premises, the truth value of the conclusion is determined by the syntax of the inference rule alone, plausible inference requires that we also know something of the semantic content of the inference rule. We have shown that by identifying characteristics of the relations used in inference rules, we can predict the plausibility of their conclusions.

Several recent papers (Cohen & Loiselle, 1988; Huhns & Stephens, 1988; Winston, Chaffin & Herrmann, 1987) have focused on binary relations used in inferences of the form

$$\text{Given} \quad \text{A } R_i \text{ B and B } R_j \text{ C}$$
$$\text{conclude either} \quad \text{A } R_i \text{ C or A } R_j \text{ C}.$$

These efforts analyze relations in terms of more "primitive" elements, which are used to predict the plausibility of these kinds of inferences. This paper will review these results and discuss their contributions, noting especially those factors that seem to provide most of the power behind the ability to predict plausibility. We then examine the role of a relation's interpretation and show that knowing the precise meaning of a relation is crucial to predicting plausibility. We conclude by discussing our current research, which explores how the meaning of a relation, defined to be the assignment of these more primitive elements, can be determined from the behavior of the relation in inferences.

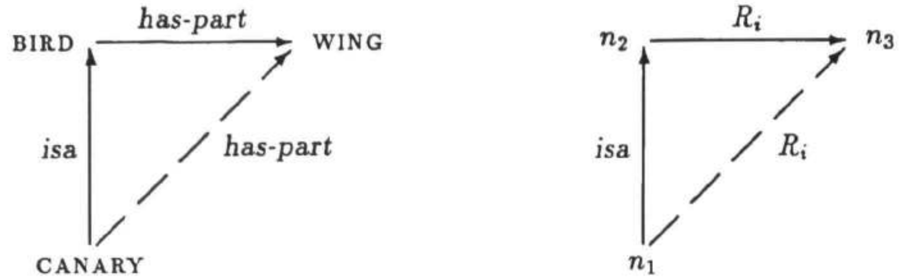

Figure 1: The triangular structure of property inheritance over *isa*

GENERATING RULES OF PLAUSIBLE INFERENCE

Cohen and Loiselle (1988) showed how the structure of property inheritance over *isa*, a common rule of plausible inference, could be generalized to generate other possible plausible inference rules. Figure 1 shows that property inheritance can be drawn as a triangle where the legs represent the known statements (premises) and the hypotenuse represents the conclusion. The left triangle illustrates a specific instantiation of property inheritance: the concept CANARY inherits the property "*has-part* WING" from its superclass BIRD. The right triangle shows the general form of property inheritance over *isa*: if n_1 is related to n_2 by *isa*, and n_2 is related to n_3 by any arbitrary relation R_i, we can infer that n_1 is also related to n_3 by R_i.

Property inheritance requires that the first premise be *isa* and that inheritance occurs only over this *isa* link. By relaxing these requirements we can generate many other possible inference rules with the same triangular structure (Figure 2). Again, the left triangle gives a specific instantiation of one such inference rule while the right triangle shows the corresponding general structure. Since we no longer restrict which link can be "inherited over" we are free to infer either R_i or R_j in the conclusion. So this structure can be used to form two inference rules:

$$n_1\ R_i\ n_2,\ n_2\ R_j\ n_3 \rightarrow n_1\ R_i\ n_3$$
$$n_1\ R_i\ n_2,\ n_2\ R_j\ n_3 \rightarrow n_1\ R_j\ n_3$$

Clearly, although we can use this structure to combine any two relations to yield two possible plausible inference rules, not all the resulting rules will produce plausible conclusions. But if we are able to identify characteristics of these rules that will allow us to predict which rules will produce predominantly plausible conclusions, then this triangular structure is potentially a

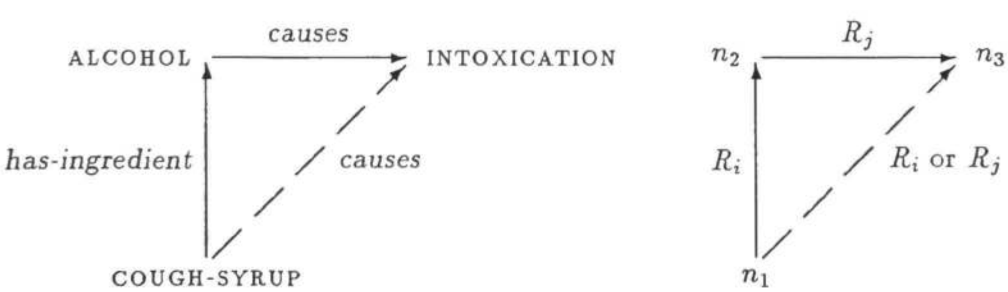

Figure 2: Extending the structure of property inheritance

powerful source of inference rules. The research discussed in the next three sections describes our attempts and those of other researchers to find the characteristics of inference rules that are highly correlated with plausibility.

TRANSITIVITY

Our initial experiments with these kinds of plausible inference rules identified two relation characteristics (Cohen & Loiselle, 1988). We studied a set of nine relations and determined that all had either an underlying sense of hierarchical inclusion, temporal ordering, or both. For example, the relation *component-of* conveys a sense of hierarchical inclusion since a whole includes its parts. Similarly, *caused-by* imposes a temporal order on the concepts it connects. When a relation has more than one interpretation both underlying senses may apply. For example, a mechanism may be either an instrument required prior to pursuing some activity, such as needing a key to unlock a door, or a subprocess subsumed by a superior process, as in respiration being a mechanism of maintaining life; therefore *mechanism-of* admits both a sense of hierarchical inclusion and temporal ordering. These underlying interpretations were used to determine the "deep structure" of the inference rule (Figure 3) where $n_3 \xrightarrow{h} n_2$ indicates that n_3 hierarchically includes n_2 and $n_1 \xrightarrow{t} n_2$ indicates that n_1 precedes n_2.

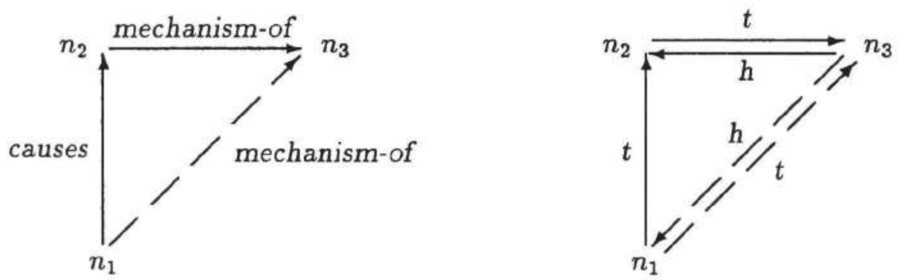

Figure 3: A plausible inference rule and its deep structure

We noted that some of our inference rules' deep structures preserved *transitivity*, that is, the same ordering, either temporal or hierarchical, was maintained between n_1 and n_3 in both the premises and the conclusion. The deep structure in Figure 3 is transitive because the temporal (t) links in both the premises and the conclusion indicate that n_1 comes before n_3. (The premises in intransitive rules do not imply any particular order between n_1 and n_3 nor is any required by the conclusion.) We also identified another characteristic of deep structures called *consistency*. Note that in Figure 3 some of the legs of the triangles are labeled with both temporal and hierarchical links, but only one of these forms a consistent interpretation, that is, we can choose between these two interpretations in such a way that allows us to label all three sides of the triangle with t-links (the consistent interpretation) but not with h-links. When a deep structure has multiple interpretations we use the consistent interpretation to determine transitivity. When no such consistent labeling is possible we call the structure (and its corresponding rule) inconsistent.

Our experiments with human subjects, who collectively viewed over 3000 inferences, showed that transitivity could be used to predict the plausibility of conclusions suggested by these inference rules with a fair degree of accuracy. Transitive rules yielded conclusions that were judged to be plausible in 77.4% of the inferences. For intransitive rules this figure was 38.8% and for rules having no consistent interpretation the results were near chance at 57.3%.[1] Thus with very little information about the specific inferences we are able to make modestly accurate predictions about the plausibility of their conclusions simply by knowing whether or not the rule is transitive.

It may be possible to improve the accuracy of our predictions by including additional information in our analysis. For example, knowing just the deep structure of a rule allows us to determine the rule's transitivity; seeing the transitive deep structure in Figure 3 lets us predict that approximately 77% of the inferences produced by this rule will be judged plausible. Knowledge about the specific relations used in this rule can improve this estimate, however. In this case, our data showed that only 73.6% of the inferences produced by the rule n_1 *causes* n_2, n_2 *mechanism-of* $n_3 \rightarrow n_1$ *mechanism-of* n_3 are judged plausible. If, instead, we knew that our transitive deep structure was derived from the rule n_1 *causes* n_2, n_2 *has-product* $n_3 \rightarrow n_1$ *has-product* n_3 we could predict a higher number of plausible conclusions because 87.9% of the resulting inferences were judged plausible in our experiment. Similarly, knowing the specific concepts that instantiate an inference rule also allows us to make more accurate predictions. The rule n_1 *has-ingredient* n_2, n_2 *causes* $n_3 \rightarrow n_1$ *causes* n_3 shown in Figure 2 seems generally plausible as does the instantiation shown there, but if we substituted AIR for COUGH-SYRUP the conclusion would certainly be judged unacceptable since the concentration of alcohol in air is too low to make us intoxicated.

The above discussion identifies a trade-off. With additional information about the relations in the rules, or the particular nodes used to instantiate the inferences, we could improve the accuracy of our predictions of plausibility. But acquiring and representing this additional information necessarily incurs additional costs. Therefore it is important to identify the amount and kinds of information required to achieve an acceptable level of predictability.

RELATION ELEMENT THEORY

Relation element theory (Chaffin & Herrmann, 1987) provides some of this additional information. By focusing on characteristics of the relations rather than on specific inference rules or instantiations, Chaffin and Herrmann are able to maintain a high degree of generality and incur little additional cost. Relation element theory holds that semantic relations should not be viewed as unitary semantic entities but rather as compositions of a set of simpler relation elements. Originally used to gauge the similarity of two semantic relations, relation element theory can also be used to predict the plausibility of an inference rule's conclusions.

Winston, Chaffin and Herrmann (1987) explore inferences based on the part-whole relation. They first note that although we ordinarily expect this relation to establish a strict partial ordering and thus be transitive many such inferences fail to produce plausible conclusions. For example, given the premises "Simpson's arm is part of Simpson," and "Simpson is part of the Philosophy department," it is not appropriate to conclude that Simpson's arm is part of the Philosophy Department (Winston, Chaffin & Herrmann, 1987). This apparent intransitivity is due to the use of two distinct senses of the relation *part-of* in the premises of the inference. The first statement expresses the relation between a component and the object to which it belongs whereas the

[1] The three classes of rules identified here do not account for all the data. See Cohen and Loiselle (1988) for a complete analysis.

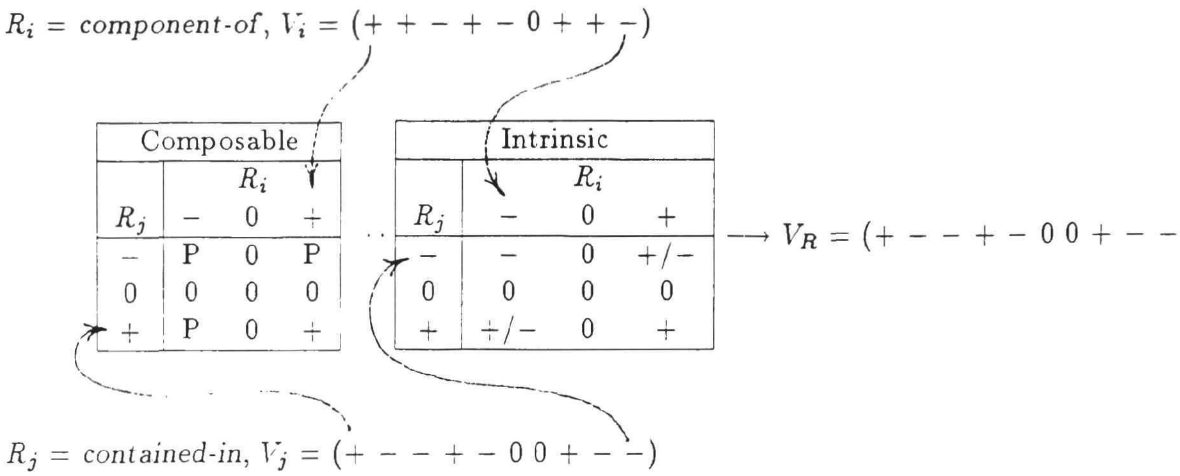

Figure 4: The algebra of extended composition.

second expresses the relation between a collection and one of its members.

The essence of this distinction is captured by relation element theory, which identifies three characteristic properties of the part-whole relation: whether the relation of part to the whole is functional, whether the parts are homeomerous, and whether the part can be, in principle, separated from its whole. According to the theory, all part-whole relations share the common element of *connection* between part and whole, this connection being modified by the values for the elements *functional*, *homeomerous*, and *separable*. Winston, Chaffin and Herrmann identify six kinds of part-whole relations and conclude that an inference is valid only if the same kind of *part-of* occurs in both premises as in the conclusion. This ensures that both the premises and the conclusion will have the identical set of relation elements. It also ensures transitivity.

EXTENDED COMPOSITION

Huhns and Stephens (1988) continue this line of research, identifying ten relation primitives, including several identified in Cohen and Loiselle (1988) and Winston, Chaffin, and Herrmann (1987). Examples of these primitives are composable, which indicates that the "fundamental characteristics" of a relation permits its use in these kinds of inferences; homeomerous, the domain of a relation is "the same kind of thing" as the range; and intrinsic, the relation specifies an intrinsic property of its domain or range. For each relation these primitives are assigned a value of +, meaning the characteristic is present, −, not present, or 0, if the primitive does not apply to the relation. Thus each relation can be represented by a vector of values for these ten primitives. Plausible inference rules are generated by the technique developed in Cohen and Loiselle (1988) and described above. (Huhns and Stephens call this technique "extended composition.") A corresponding algebra uses an operator table for each primitive to determine how the two vectors for R_i and R_j may be combined to yield a result vector for the conclusion (Figure 4). A match of the result vector to either or both of the premise relations' vectors is interpreted to mean the corresponding inference is plausible, provided the domain and range requirements of the relations are also met. For example, in Figure 4, V_R matches V_j so we predict that the inference rule n_1 *component-of* n_2, n_2 *contained-in* n_3 → n_1 *contained-in* n_3 will

Structural			
	R_i		
R_j	$-$	0	+
$-$	$-$	0	P
0	0	0	0
+	P	0	+

Temporal			
	R_i		
R_j	$-$	0	+
$-$	$-$	0	P
0	0	0	0
+	P	0	+

Intangible			
	R_i		
R_j	$-$	0	+
$-$	$-$	0	P
0	0	0	0
+	P	0	+

Figure 5: Operator tables for the transitivity-preserving primitives.

produce predominantly plausible conclusions. The results of this composition may be further pruned if the relations have incompatible domains and ranges. For example, although the operator tables may permit the composition of *subfield-of* and *subprocess-of*, the inference will be disallowed because it makes no sense to talk about a subfield of a process.

Huhns and Stephens apply their technique to a set of 21 relations (having a total of 861 possible compositions) to yield a composition matrix of 103 entries where the result vector matches the vector for either R_i or R_j and the corresponding inferred relation also satisfies the domain and range requirements established by the premise relations. That is, their algebra predicts that at least 103 out of 861 inference rules will produce predominantly plausible conclusions. (Since their algebra was designed for correctness instead of completeness, it is possible that some compositions not included the matrix might also produce plausible conclusions.) Huhns and Stephens claim validity for their results based on the plausibility of selected example inferences from the composition matrix.

TRANSITIVITY REVISITED

Three of the primitives in Huhns and Stephens' work indicate an ordering along a single dimension: structural indicates a hierarchical relationship in terms of physical structure, temporal indicates an ordering in time, and intangible indicates a hierarchical relationship in terms of ownership or mental inclusion. Since relations that indicate an ordering along a single dimension can be used transitively, these primitives capture the same kinds of underlying interpretations as our t-links and h-links (Cohen & Loiselle, 1988). Huhns and Stephens' temporal primitive corresponds to our t-link, whereas the structural and intangible primitives distinguish physical from mental inclusion, which are both represented by our h-link. This correspondence is also borne out by the operator tables for these primitives (Figure 5). For these three primitives, the values + and $-$ indicate the direction of this ordering rather than the presence or absence of the characteristic. A value of 0 indicates either that no such ordering exists or that the property does not apply. These tables preserve the ordering of the concepts when both premises have the same value (indicate the same ordering) and prohibit inferences when the premises have incompatible orderings (the value "P" means the inference is prohibited). Thus these operator tables ensure that only those inference rules that preserve transitivity will be generated by the algebra.

Since transitivity alone was shown to predict pretty well the plausibility of inferences in Cohen and Loiselle (1988), we were interested in how much the three transitivity-preserving primitives contributed to the power of Huhns and Stephens' method. To evaluate this, we implemented their algebra and used it to determine the number of matrix entries produced by every subset of three of the ten primitives. Any subset of the original primitives is guaranteed to produce at least the original 103 entries; fewer *additional* entries indicates that a particular subset of primitives comes closer to reproducing Huhns and Stephens' original composition matrix and thus contributes more power to the algebra.

The most powerful set of three primitives, structural, temporal and composable, produced 198 matrix entries. The set of three transitivity-preserving primitives ranked third with 213 entries, tied with the set temporal, intangible and composable. Huhns and Stephens (1988, p. 5) note that the composable primitive is also closely tied to transitivity. They state, "Assignment of values for this property can be guided by consideration of the transitivity of the relation, i.e., if a relation is not transitive (cannot be composed with itself), then it often cannot be composed with any other relation." The remaining set of three of these four primitives, structural, intangible and composable, produced 238 entries, ranking 22nd out of 120. For comparison, the least powerful set of primitives, near, connected and intrinsic, produced 351 entries, while considerations of domain and range incompatibilities alone yielded 444 entries. Based on these rankings we conclude that the transitivity component represented by the primitives structural, temporal, intangible and composable contributes the largest share of the power of Huhns and Stephens' representation and algebra, and that the cost of assigning values to the remaining primitives may often outweigh the slight increase in power they provide.

ONTOLOGY MAINTENANCE: USING INFERENCES TO DETERMINE RELATION SEMANTICS

The work by Winston, Chaffin and Herrmann on the part-whole relation discussed above makes it clear that often what we consider to be a single semantic relation may be used in several different ways with corresponding differences in meaning. Furthermore, it shows that the plausibility of inferences using such a relation cannot be reliably determined unless the intended meaning is known. We cannot say whether the rule n_1 *part-of* n_2, n_2 *part-of* $n_3 \rightarrow n_1$ *part-of* n_3 will produce plausible conclusions unless we know whether both premises use the same type of part-whole relation. While relation element theory, and its extension in Huhns and Stephens' set of relation primitives, gives us a representation for specifying these intended meanings, it doesn't tell us how to determine the correct definition (assignment of primitive values) of a relation. Ontology maintenance offers a solution for this problem.

Ontology maintenance is concerned with assuring that the definitions of relations are correct. "Correct" means that we are able to accurately predict the plausibility of inferences using these relations. Thus, when we add a definition of a new relation to a knowledge base, or modify an existing one, we can check whether the definition is correct by generating inferences we expect to be plausible.

We are currently developing an ontology of semantic relations based on their behavior in inferences. This ontology includes a hierarchy of relations determined by their primitive assignments (Figure 6). Relations inherit primitive values from their parents, therefore their placement in the hierarchy determines the kinds of inferences predicted to be plausible for each relation. Evaluating these inferences thus evaluates the (possibly partial) definition of a relation suggested by its placement in the hierarchy. The following example illustrates how the relation

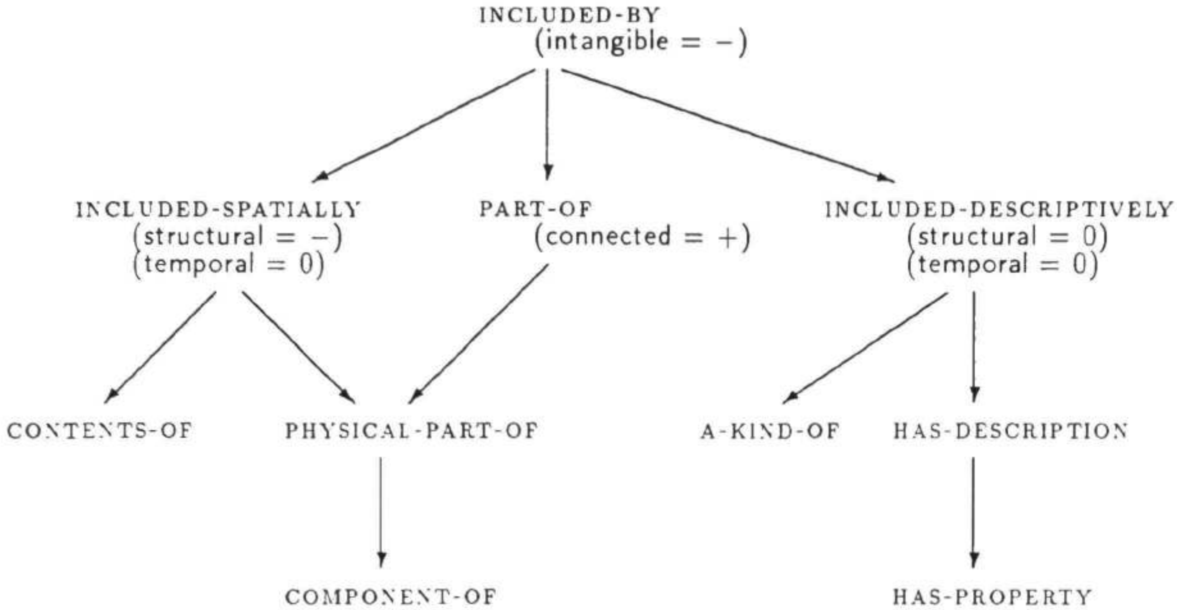

Figure 6: A partial hierarchy of relations with primitive values.

material-of, which indicates the main substance of which an object is made, is placed in the hierarchy, and thus how we determine the correct primitive values for this relation.

At first, we might believe *material-of* to be a kind of *part-of* relation; indeed, Winston, Chaffin and Herrmann (1987) claim that the "stuff-object" relation is a type of part-whole relation. Therefore we begin by placing *material-of* under *physical-part-of* in the relation hierarchy. This results in *material-of* inheriting the primitive assignments structural = −, temporal = 0, intangible = −, and connected = +. We then generate inferences predicted to be plausible. For our experiments these were derived from a knowledge base we are developing to represent common sense information about a house. One such inference is

Given:	WOOD	*material-of*	AXE-HANDLE,	and
	AXE-HANDLE	*component-of*	AXE	
Infer:	WOOD	*material-of*	AXE.	

Immediately we see that to evaluate the inference we must know more precisely the intended meaning of *material-of*. Will we allow it to indicate a substance in any area of an object or do we require it to refer to the entire object? If we had intended the former then this would seem a reasonable inference, but since we intended the latter the inference is unacceptable, and therefore, the value for at least one of these primitives must be incorrect.

Examining the hierarchy, again we decide that perhaps a material is more like a property of an object than it is part of an object. This suggests placing *has-material* (the inverse of *material-of*) under *has-description* in the relation hierarchy. Now *has-material* inherits the primitive values structural = 0, temporal = 0, and intangible = − and we generate inferences like

Given:	BOARD	*has-material*	WOOD,	and
	WOOD	*has-property*	FLAMMABLE	
Infer:	BOARD	*has-property*	FLAMMABLE.	

This time the inference is acceptable, indicating that these primitive values are correct, and we keep *has-material* under *has-description* in the relation hierarchy.

CONCLUSION

Certainly the more information we have about an inference, the better we will be able to judge the plausibility of its conclusion. But for tasks that do not require a high degree of accuracy in such judgments we may realize a savings by placing ourselves relatively low on the information/accuracy trade-off. The cost of assigning values to many different primitives for a large number of relations may cause us to want to limit the set of primitives used. Therefore, it is important to examine the sources of power in our representations. The results presented here suggest that primitives that represent different kinds of transitivity contribute most of the power in predicting plausibility.

Our ability to predict the plausibility of inferences is determined by our ability to define relations correctly. Our research in ontology maintenance explores how we can verify a relation's definition by examining inferences we expect to be plausible.

ACKNOWLEDGMENTS

This research is funded by the Office of Naval Research, under a University Research Initiative Grant, Contract #N00014-86-K-0764. We are indebted to Dave Hart for his comments on drafts of this paper.

REFERENCES

Chaffin, R. and Herrmann, D. (1987) Relation element theory: A new account of the representation and processing of semantic relations. In *Memory and Learning: The Ebbinghaus Centennial Conference*, ed. D. Gorfein and R. Hoffman, 221-245. Hillsdale, NJ:Lawrence Erlbaum Associates.

Cohen, P. R. and Loiselle, C. (1988) Beyond ISA: Structures for plausible inference in semantic networks. In *Proceedings of the Seventh National Conference on Artificial Intelligence*. St. Paul, MN.

Huhns, M. and Stephens, L. (1988) Extended Composition of Relations. Technical Report ACA-AI-376-88, Microelectronics and Computer Technology Corporation, Austin, TX.

Winston, M., Chaffin, R., and Herrmann, D. (1987) A taxonomy of part-whole relations. *Cognitive Science* 11(4):417–442.