AI Magazine Volume 6 Number 3 (1985) (© AAAI)

"I Lied about the Trees" Or, Defaults and Definitions in Knowledge Representation

Ronald J. Brachman

AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974

Abstract

Over the past few years, the notion of a "prototype" (*e.g.*, TYPICAL-ELEPHANT) seems to have caught on securely in knowledge representation research. Along with a way to specify default properties for instances of a description, proto-representations allow the overriding, or "cancelling" of properties that don't apply in particular cases. This supposedly makes representing exceptions (three-legged elephants and the like) easy; but, alas, it makes one crucial type of representation impossible—that of composite descriptions whose meanings are functions of the structure and interrelation of their parts. This article explores this and other ramifications of the emphasis on default properties and "typical" objects.

Author's Note:

This short article has been circulating underground for quite some time. It was originally written several years ago, but somehow never managed to see the light of day. In the interim, much good work in knowledge representation has been done, and I had expected that most of what had been written in this note would be irrelevant by now. However, it seems that the simple main point here is just as cogent as ever: AI representation systems based on "frames" (and semantic networks, in many cases) can be very misleading—they are able to represent only a fraction of what it might appear they can. So it seemed reasonable, even after such a long time, to bring the article to light.

The article itself attempts to show how the uniform adoption of overridable default-style representations at the expense of supposedly less useful definitional representations has some quite serious flaws. As it turns out, while exceptional concepts like THREE-LEGGED-ELEPHANT are supposed to be the bread and butter of many frame systems, in those systems they really can't be represented at all. In a sense, frame systems are hardly what they advertise to be.

While I believe this to be an important point, this article was never meant to be the definitive work on logical distinctions in knowledge representation. Some of the notions mentioned here in passing (*e.g.*, analyticity) are perenially problematic. In addition, I have not really attempted to bring the body of the article up to date from its original form. The article is also generally nonconstructive. However, there is now ample evidence that this kind of analysis can lead to constructive suggestions for knowledge representation systems. In work pursued after the original version of this article was written, some suggestions of this sort were followed up quite successfully (see, for example, Brachman, 1983, and Brachman, Fikes,

[1]This article reports primarily on research done while the author was at Bolt Beranek and Newman Inc , Cambridge, Massachusetts. The work done at BBN was funded in part by the Defense Advanced Research Projects Agency, and was monitored by the Office of Naval Research under Contract No N00014-77-C-0378. The views and conclusions stated here are those of the author and should not be interpreted as necessarily representing the official policies, either express or implied, of the Defense Advanced Research Projects Agency or the U.S. Government I would also like to thank Fairchild Camera and Instrument Corp. and Schlumberger Palo Alto Research for providing productive working environments in which to continue the work begun at BBN. They probably don't want to have much to do with the views stated here, either.

Many of the issues raised in this article became clear in a series of discussions initiated in the Summer of 1979, in which several of us at BBN tried to figure out exactly what we had wrought with the KL-ONE primitives we had created Much credit should go to Ed Barton for raising the first doubts about the semantics of some of our links Hector Levesque, Jeff Gibbons, David Israel, and Rusty Bobrow were also in on the original discussions and, along with Brian Smith, have provided great intellectual stimulation over the course of the past several years I also want to thank Eugene Charniak, Scott Fahlman, and Stuart Shapiro for frank and stimulating discussions about their representation systems. I would also like to thank John Seely Brown, and several anonymous reviewers for their useful critiques I have tried to respond to as many of their comments as possible without changing the tone of the paper. David Israel and Hector Levesque were also invaluable in helping to make this article more readable and more accurate Finally, special thanks go to Pat Hayes for his advocacy of this paper over the years, and for his support in getting it into final publication form.

& Levesque, 1983; see also Brachman & Schmolze, 1985, and Levesque, 1981, for hints on how to integrate definitional representations with typicality statements). In addition, there have recently been many other interesting and relevant developments in representation. For example, with respect to the careful understanding of defaults and the like within semantic net and frame systems, I would direct the interested reader to at least Etherington and Reiter (1983) and Touretzky (1984), and most of the papers in the special issue of *Artificial Intelligence* on Non-Monotonic Logic (Bobrow, 1980). Elsewhere, Cherniak (1984) has made an interesting empirical argument for integrating both prototypical and deductive reasoning. With respect to the question of just what a "prototype" is, see at least Winograd (1978), Lehnert and Wilks (1979), and Bobrow and Winograd (1979). And for some related criticisms of representation systems and wishful thinking in AI, don't miss McDermott (1981).

Finally, in retrospect, I think there are two principal reasons for the kind of confusion addressed here—one that so often seems to accompany knowledge representation work. First, the different sources from which the field has sprung are generally incompatible (as detailed somewhat in Brachman, 1979; see also Winograd, 1978); those using representations as psychological models have very different concerns and criteria for adequacy from those who want special-purpose formal logics with interesting computational properties. Second, there seem to be rampant level confusions throughout the history of knowledge representation, in particular, between what Newell (1981) calls the *knowledge level* and the *symbol level.* For example, when discussing what a formal representation system is capable of representing, the notion of inheritance mechanisms really should not come up; this is usually an implementation, or symbol level concern (although it could be a concern of yet a third kind, about readability or maintainability of networks). Conversely, when discussing inheritance mechanisms, "active slot-values," and the like, what it is that is being represented is irrelevant. (This accounts for the confusion surrounding Fahlman, Touretzky, and van Roggen (1981), wherein the authors of an elaborate inheritance mechanism went searching for what it meant, and couldn't find a consistent, intuitive interpretation for their mechanism.) The distinctions raised and discussed in this little paper are principally addressed at the knowledge level. For more on this distinction and its impact on representation, please consult Brachman and Levesque (1984), Hayes (1979), Levesque (1984b), and Levesque (1985).

Clyde's Revenge

Remember the good old days of elephant jokes? Well, the knowledge representation community may have spent considerable time dissecting Clyde the Elephant and his friends (*e.g.*, see Fahlman, 1979), but it seems that the elephants may have the last laugh. Clyde's revenge comes to us in the form of an elephant joke, that goes something like this:

Q: What's big and gray, has a trunk, and lives in the trees?

A: An elephant—I lied about the trees.

Surprisingly enough, "lying" about properties, such as where elephants live, seems to be a basic feature of many current knowledge representation schemes—in particular, most of those semantic net and frame systems with "inheritance of properties." In this informal note, I will explain how a uniformly applicable facility to cancel properties that would normally be inherited is subtly but severely constraining. It limits us to interpreting nodes/frames and slots in these schemes as representing strictly default conditions, to the exclusion of definitional conditions, or even contingent universal ones. And without some definitional capability, frames cannot express even simple *composite descriptions*, like "elephant whose color is gray," "polygon with four sides," or "devil with a blue dress on."

This is not to say that an artificially intelligent system can get very far without being able to handle defaults and exceptions.[2] But the temptation has been to try to use default-oriented frame notations to do the complete job of knowledge representation—that is, for frames to represent arbitrary concepts and their interrelationships. This article addresses some of the consequences of succumbing to this temptation.

Despite what its title may lead you to expect, this is an article not so much about prevarication, but about distinctions—distinctions that ought to be captured by any knowledge representation language that attempts to cover reasonable conceptual ground. More generally, it is also about thinking through the consequences of intuitively appealing proposals in AI research. I will here point out, for example, how the most common approach to representing exceptions, while a reasonable one at first glance, is just not up to the task.

The exposition will go like this: First, I will explain briefly how some common forms of representation in AI give us the uniform ability to "lie about the trees." These forms then force us to interpret all properties as default properties—a regime, as it turns out, under which it is not possible to represent genuine universal truths. Then, after a brief digression or two, I address the main problem: Default representations with cancellation preclude any kind of a definition capability. And without some sort of definitional force, even the simplest composite descriptions cannot be formed—and every description in the language is doomed to be a primitive.[3] Put another way, despite

[2] Not to mention the fact that frames were originally conceived for default reasoning (Minsky, 1975) And see Minsky's comments about dead birds and birds with their feet set in concrete, in Kolata (1982). Reiter (1978) also makes a strong case for default reasoning.

[3] This is not a condemnation of representations based on primitives.

what we might be led to think, frame systems hardly provide anything framelike at all. Finally, the discussion concludes with a look at several other surprising consequences of exclusive reliance on default representations.

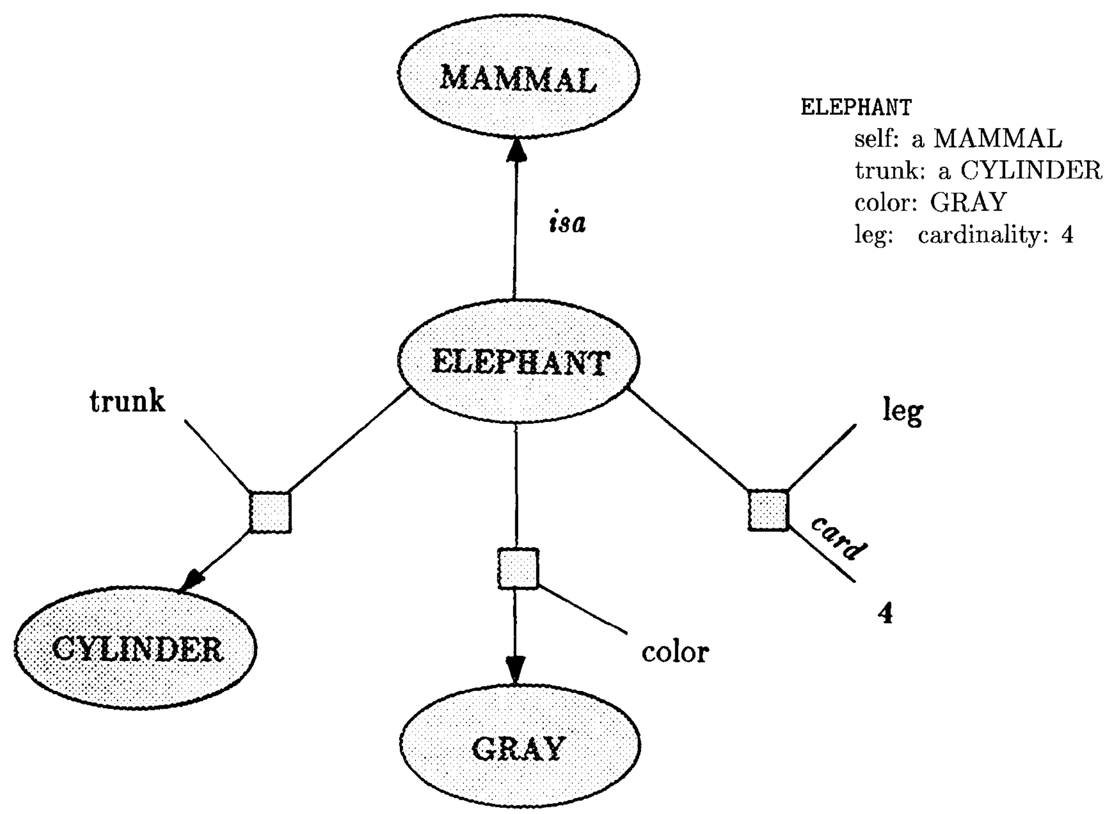
Background

Let's assume that we are talking about a typical representation scheme based on the notions of framelike conceptual units and slotlike role descriptions, or their equivalents, and a hierarchy expressing generalization relationships among the conceptual units.[4] In semantic nets, the conceptual units are "nodes;" in other systems, they are "concepts," "frames," "prototypes," "schemas," or "units." Here, for simplicity, I will use *frame* as a representative term for the supposedly structured units that all of these systems are based on (although I will occasionally lapse into node-and-link talk). In any case, the principal connective—"ISA" (or "*VC," "A-KIND-OF," or whatever)—makes one frame a subcategorization of another. The *inheritance of properties* (or the notion of "virtual copy" (Fahlman, 1979)) in such a scheme means that subframes inherit all the features of their parent frames.

Consider the simple frame, "ELEPHANT," which we might portray graphically and lexically as in Figure 1. The notation used in the figure is intended to be intuitive, and the details aren't important (the lexical notation is modeled after KRL (Bobrow & Winograd, 1977)). Among other things, the knowledge representation literature tells us that elephants are mammals, they are gray, they have trunks that are cylinders, and they have four legs. So we would see an ISA connection between the ELEPHANT frame (the ellipse labelled as such in the figure) and the frame, MAMMAL, with a slot on ELEPHANT (or four of them, perhaps) expressing that the number of legs of an elephant is four, a slot for the trunk, expressing its cylinder-hood, and a filled-in slot saying that the color

It is only to point out that representations with the uniform possibility of cancellation give you *nothing but* primitives Reasonable representations that are founded on primitives (*e.g.,* Schank, 1972, 1973; Wilks, 1973) in fact in the end depend (at least in part) on description-composition mechanisms of just the sort advocated in this paper, in order to allow the formation of something nonprimitive out of the primitives (and thereby to avoid an indefinite multiplication of primitives).

[4]It is possible to cast much of the recent work in the field into this mold without difficulty (including FRL, KRL, NETL, AIMDS, UNITS, FRAIL, etc.). KL-ONE (Brachman & Schmolze, 1985), on the other hand, is somewhat different, as is SNePS (S. Shapiro, personal communication, 1980; see also Maida & Shapiro, 1982).



A simple frame for "elephant".

Figure 1.

of an elephant is "gray" (the slots are pictured as small squares). One might be tempted to an informal reading of this as "an ELEPHANT is a MAMMAL that has four legs and a trunk that is a cylinder, and whose color is gray," or, alternatively, "*every* ELEPHANT is a MAMMAL that has four legs and a trunk that is a cylinder, and whose color is gray."

Property Inheritance and Necessary Conditions

As mentioned, the principal inference mechanism in these representation systems is property inheritance. It usually runs something like this: If Clyde is asserted to be an elephant by some kind of an ISA connection from CLYDE to ELEPHANT, then CLYDE would act as if it were a copy of the entire structure at ELEPHANT.[5] This being the case, we could conclude that Clyde is a mammal, he has a trunk that is a cylinder, the value of his color attribute is *gray,* and he has four legs.

Since these things accrue to Clyde in virtue of his being an elephant, slots and inherited ISA links seem to correspond to the consequents of universally quantified *if-then* statements. *If* you're put under the parent frame, *then* you get the properties (this goes for any instance, thus the implied universal quantification). While other kinds of inferences can be drawn from network representations of the kind we're discussing, it is pretty apparent that most people think first of property inheritance when they think of semantic nets. They are even often called "inheritance networks."

The fact that we think first of these as property inheritance systems makes it clear that they are meant to specialize in *necessary conditions.* The most common use of these representations takes advantage of the kind of "outwardness" of properties encouraged by necessary conditions: Once we know that Clyde is an elephant, then what elephant-properties he has can be found out by inheritance. Note that typically we don't use the properties present at ELEPHANT to see if Clyde is one.[6] Property inheritance as the fundamental method of inference enourages users to put frames and instances where they think they belong, and then to infer what follows. It is never up to the system to decide where to put CLYDE.

This impression of the perceived main use of frame/semantic network representations for necessary conditions is substantiated in various places in the literature. Charniak (1981), for example, proposes a combination of frames and predicate calculus notation in which he actually translates a frame into a typed universal statement, with explicit existentials for the slots. For instance, Charniak translates the frame

```
[elephant
      isa:  (mammal ?elephant)
      slots:  (head (elephant-head ?head))
      facts:  (has-part ?elephant ?head)...]
```

into

```
FORALL (?e (elephant ?e))
   EXISTS (?h (elephant-head ?h))
          [has-part ?e ?h] ... .
```

It appears that the principal use of the frame here would be to determine Clyde's properties once it was known that he was an elephant (the typed universal quantification could just as well have been a conditional). In a similar vein, Hayes (1979) interprets the slots of frames explicitly as the right-hand sides of conditionals. In his view, a frame represents a set of implicitly universally quantified conditionals. For example, "a frame representing the concept C, with slot-relationships $R_1, \ldots, R_n$" would become the following:

$$\forall x\, C(x) \supset R_1(x, f_1(x)) \;\wedge\; \forall x\, C(x) \supset R_2(x, f_2(x)) \;\wedge \ldots$$

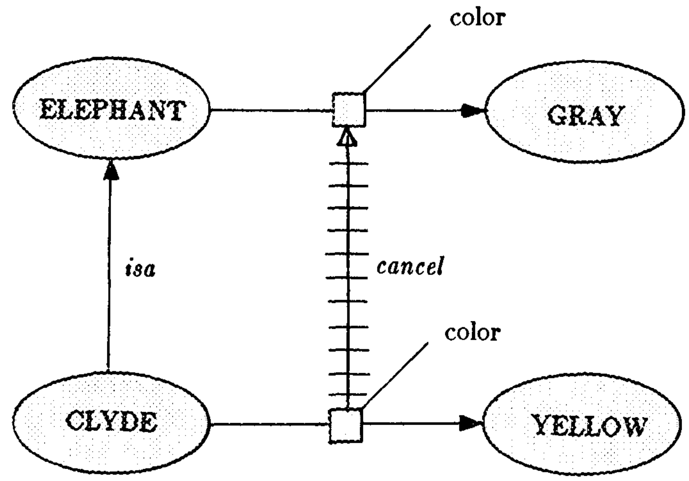
And much more recently, Fikes and Kehler (1985) have stated directly that most systems of this sort "specify only necessary conditions for class membership." In sum, then, between analyses like Hayes' and Fikes and Kehler's, and the way these representations are typically described in the literature, it is apparent that property inheritance dominates our thinking about them, and that most researchers would like to believe that frames have a universally quantified, *if-then* kind of import.

Is Nothing Sacred?

Now, despite the appeal of property inheritance as an inference mechanism, anyone who knows Clyde and his friends will tell you that not *all* instances of "elephant" need have the features specified in a frame like our ELEPHANT frame above. For example, it is quite conceivable that we might run across some poor elephant suffering from a bad case of hepatitis (picked up in his trek across the Alps, no doubt). To handle exceptions like yellow elephants, most frame notations allow the overriding, or *cancellation* of properties that a particular frame or instance would normally inherit. For example, in Figure 2, we see a typical way to say that Clyde is jaundiced, even though elephants "in general" are gray. The crossed line indicates

[5] The fact that with an inheritance mechanism one does not need to make an actual copy leads Fahlman to his "virtual copy" notion (Fahlman, 1977).

[6] Fikes and Kehler (1985) point out that this common interpretation of frame systems would allow *negative* inferences about class membership (Pat Hayes, personal communication, 1985, has also made this point). If Clyde failed to have one of the elephant-properties, then he couldn't be an elephant. However, as we shall soon see, because of what we can do to represent yellow or three-legged elephants, we couldn't in fact use the representation to determine Clyde's nonelephanthood definitively, even if we wanted to.



Cancelling a property.

Figure 2.

an explicit connection between the new slot and its overridden parent—the so-called "cancel link" provided by some systems (although most just leave it implicit).

The fact that we can override the normal color for elephants in favor of Clyde's special case means that the intuitive reading of ELEPHANT as including "every elephant is gray" is not quite right (nor, for that matter, are Hayes' and Fikes and Kehler's interpretations correct in general). Rather, it should be that elephants have such properties only as *defaults.* Thus, as Fahlman (1979) puts it, "I am using a weak sense of the word 'every' here: I mean that the property is true of every elephant for which it is not explicitly cancelled." Indeed, Fahlman has gone so far as to call his node "TYPICAL-ELEPHANT."

Now, this is not terribly bad if one were to believe that this kind of frame "is intended to represent a 'stereotypical situation' " (Hayes, 1979).[7] We just have to keep in mind that any notation that allows us to represent exceptions by overriding properties that would normally be inherited *must* be using its conceptual units in this "typically" sense. Otherwise, we are lying about the "every."[8]

Incidentally, notice that the implied sense of "typically" here is strictly along the lines of " 'In the absence of any information to the contrary, assume ....' " (Reiter, 1978), and has nothing to do with frequency of occurrence. In some dialects, "typically" is more closely synonymous with "usually." In the manner used in frame notations, however, a "typical" property could be violated in every single case! We do not treat the "usually" issue further, but merely note that it is yet another type of adverb of quantification that we might need to express.[9]

"It Ain't Necessarily So"

One consequence of the default interpretation of descriptions is that the slot notation for properties cannot be used, without alteration, to make an unequivocal universal statement, since we hedge on the "every." The minute I assume from the ELEPHANT frame that all elephants are gray, I could immediately find an exception just below (or anywhere below) that frame. And we certainly want to be able to make exceptionless universal statements, since there are many domains and situations where properties do hold for all instances of certain kinds. Indeed, this is the intent of necessary conditions, which it is now clear is *not* what frame-slots represent.

For example, as I look out of my office window at the parking lot, I can easily see that all of the vehicles parked here are cars (*i.e.,* there are no motorcycles or tractor-trailers). It would be most convenient to represent this contingent universal fact with a single statement. In fact, it might be *necessary* to represent it that way, since I might not have a representative for each individual instance in question (I don't even have a good idea of how many cars are in the lot—there must be between fifty and a hundred—so I need some noncommittal way to capture the universal[10]). One might suppose that we could use a

[7] It does seem palatable (at least at first) to take the frame ELEPHANT as representing something like a stereotypical elephant But when a frame stands for a nonnaturally occurring class, say, GÖDEL-NUMBER or HEPTAGON, a stereotypical interpretation doesn't seem so appealing (what does the typical Gödel number look like?). One might venture to say that this whole discussion is necessary because it is the "natural kinds" (Putnam, 1977) like *elephant* that have been emphasized by representation folks. In any case, keep this distinction in mind—it will arise again.

[8] I might also point out here that, to the extent that object-oriented programming languages like Smalltalk are used as representations (*e.g.,* see Borning, 1981), they fall into this class Their basic overridable inheritance makes them primitive representations indeed And even though some of these languages (*e.g.,* Flavors—see Carnese, 1985) provide great flexibility with "roll your own" inheritance, once it is used, the meaning of the entire representation framework comes into question. For more on this matter, see Israel and Brachman (1984).

[9] For some recent work inspired by the fuzzier aspects of prototypes, see Rich (1983), Zadeh (1983), Cohen and Murphy (1984), and Ginsberg (1984) See also Cheeseman (1985) for a direct contrast between frequency-oriented representations and default logics

[10] It is hard to ignore the fact that predicate calculus is really good

default frame notation like the one above, and always examine the set of instances of a frame to see if a description represented a true universal (provided that I create something like dummy instances for the vehicles I don't know much about as individuals). But this procedure would only be useful in knowledge bases where the information was expected to be complete (see Reiter, 1978, for more on the *closed world assumption,* which presumes "perfect knowledge about the domain being modeled"). And even then, checking all instances could be very costly, and it would have to be done every time.

At least one "proto-net" author has suggested a way to explicitly represent exceptionless properties: "...some fact [can be declared] sacred (uncancellable) and therefore true of all elephants without exception" (Fahlman, 1979). Still, the bottom line is, if the notation allows cancellation, but provides no mechanism for noting certain facts as uncancellable, then it simply cannot express universal truths.[11]

One is tempted to raise here a further distinction, between *contingent* universal truths (like the one about what's in the parking lot) and *necessary* ones (such as the truths of arithmetic). I would hardly lift an eybrow if someone told me that there was now a big rig outside of my office. On the other hand, I would think that something was fundamentally wrong if someone tried to convince me that two and two made five. While this is also an important distinction, and does bear on the import of "sacred" markings (*e.g.,* contrast "there *happen to be* no exceptions" with "it is *inconceivable* that there could be any exceptions"), it would take us off the track to delve into this point here. There is a philosophical gold mine (or endless maze, depending on your perspective) of subtleties here, in which the interested reader can find some wonderful epistemological and metaphysical nuggets in Kripke (1972).

Kinds of Cancellation

Before we get to the heart of the matter, let's briefly reconsider the "cancel" link described above. When using cancellation on slots, we need to address this question: What is being cancelled, the *value* (*e.g.,* having a cylindrical trunk), or the *attribute* (*e.g.,* having a trunk at all)? For example, we can say that Clyde has all of the properties of the typical elephant, except that his trunk is hexagonal. Here we're cancelling just the value on the trunk slot. But we might also want to say that Clyde has no trunk, in which case, we'd want to cancel the *attribute.*

at just this kind of noncommittal representation, for just the right reason (see Brachman & Levesque, 1982; Moore, 1982)—and that frames are notoriously bad at this (see Brachman *et al.,* 1983).

[11]More precisely, it can only represent universal truths extensionally (by explicitly indicating all cases), and then only under the closed world assumption.

Further, we probably need to extend the notion of "sacredness" proposed by Fahlman. It's one thing for Clyde to have lost his trunk. It's another for him to have lost all shape. Since he is a physical object, he may be able to lose parts, but he cannot lose properties essential to being a physical object (like the properties of having shape or mass). So a "sacred" marking will have to distinguish between the value's being uncancellable and the attribute's being uncancellable.

Interestingly enough, in this context the presence of an explicit cancel link may add something. Clyde the elephant's having no trunk (having lost it in the Punic Wars, say) is different than Tweety the canary's having no trunk.[12] So a cancelled attribute may indicate something about the history of a property or its basic applicability.

But there are further strange possibilities afoot. One such is illustrated by the proposed (admittedly bizarre) representation in Figure 3. Cancellation is just a syntactic mechanism. Thus it doesn't seem to rule out the possibility of cancelling the attribute name (the role being played) while leaving the value.

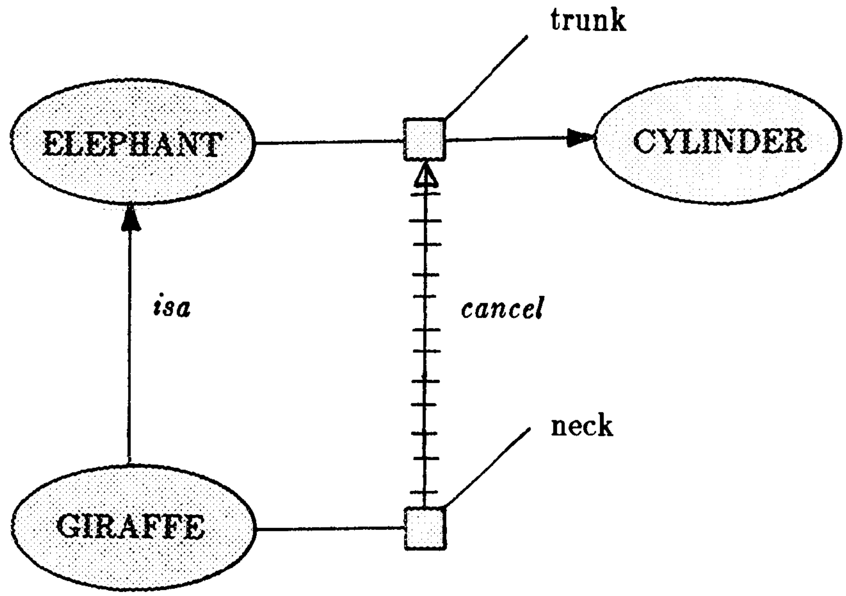
For that matter, it doesn't seem to rule out the cancellation of *every last attribute,* thus leaving us with possibilities like "A rock is an elephant, except that it is has no trunk, it isn't alive, it has no legs ...," and even more outrageous semantic anomalies that we dare not imagine here. If this is beginning to give you the feeling that the ELEPHANT frame doesn't really represent the concept of an elephant, then you're ready to go on to the next section. In any case, if we're going to admit cancellation at all, then someone had better think of a way to sort all these problems out.

Sufficiency and Compositionality: The Bad News Bared

Given the default interpretation of frames, and the realization that ISA links do not represent even the simple contingent universals that they seem to, one would suspect that even stronger statements like honest-to-goodness definitions are totally out of the question in standard frame systems.[13] One might also suspect that that's no big deal. Well, the first suspicion is right on the money; but the second is dead wrong. Let's investigate this in some detail.

[12]Fahlman inadvertently dives headfirst into this quicksand by introducing the possibility of Clyde's not having a mother by virtue of his being a clone (1979)–he indicates this by cancelling the MOTHER "role-node" for CLYDE But in that case, having a mother is not even relevant, rather than something that should be cancelled (whereas, if Clyde were a normal elephant who had lost his mother, cancellation would be a reasonable way to capture the fact)

[13]I am using "definition" here in its common intuitive sense as a (possibly complex) condition that is both necessary and sufficient (*e.g*, a universally quantified biconditional (*if and only if*)), and is also necessarily true. Thus, given our previous analysis, strictly default notations, which can't even express universals, are *a fortiori,* not able to express such conditions



A giraffe is an elephant whose trunk is its neck.

Figure 3.

Imagine trying to define two potentially similar frames, QUADRILATERAL and ELEPHANT. Something is a quadrilateral if and only if it is a polygon and has four sides: Easy enough. Elephants are mammals and they have four legs. Not so easy—even if it were necessarily the case, we could not conclude that any four-legged mammal was necessarily an elephant. That is, regardless of whether or not the complex property, "four-legged mammal," is necessary, it is certainly not*sufficient* for being an elephant. In fact, it is strongly believed that *no* combination of properties is sufficient to capture what it means to be an elephant—in other words, "natural kind" concepts (Putnam, 1977) cannot be defined. In contrast, there is nothing more to the story of quadrilaterals than four-sidedness on top of "polygonicity."

Now this contrast may not appear to be of any consequence; in fact, the distinctions raised here may seem like logical nit-picking. You would not be the first if you here resorted to an argument borrowed from language—that no lexical items of a natural language have complete definitions, not even the classic "bachelor" (*e.g.,* see Fodor, Garrett, Walker, & Parkes, 1980, or Bobrow & Winograd, 1979). Or, you might argue like many others that the vast majority of terms of interest to AI systems are like "elephant," for which criterial definition is impossible.[14] In any case, arguments like these just add insult to injury. We saw above how frame representations could not possibly represent necessary conditions (*i.e.,* simple universals) if properties could be cancelled. Arguments about the lack of importance of definition simply encourage us not to bother with even stronger compositional combinations of necessary and sufficient conditions.

Thus, we have most AI representation languages strongly favoring the nonmathematical cases. And with good reason: Why worry about definition if, at best, only quadrilaterals and the like can be defined? Well, consider this: Once we have the concept of an elephant—natural kind, primitive, or whatever—from it we can construct an indefinite number of composite concepts, each of which is in a relation to ELEPHANT that is surely definitional. For example, the concept of an elephant with three legs—call our frame for it "ELEPHANT-WITH-THREE-LEGS"—is a simple composition of two attributes, each of which is necessary and the pair of which is sufficient. That is, it is impossible to have an elephant with three legs that wasn't an elephant, and it should be impossible for an object that both was an elephant and had (exactly) three legs to fail to fall under ELEPHANT-WITH-THREE-LEGS.[15] Further, it

[14]For example, ". .aside from mathematics and the physical sciences, most of what we know about the world has associated exceptions and caveats" (Reiter, 1978) Minsky (Kolata, 1982) and Dreyfus (1981) have also made similar points

[15]Perhaps rather than calling it "definitional," it would be more accurate to call the the relationship between the concept of an elephant and the concept of elephant with three legs one of *analytic*

is undeniable that frames like ELEPHANT-WITH-THREE-LEGS will be useful: Any attributive modification[16] we can make to any frame (*e.g.*, any further specification of the type of a slot filler, such as "whose voltage is high," "whose argument is a string," or "whose ex-husband is Bobby Ewing") results in an analytically related frame.

In fact, a frame like ELEPHANT-WITH-THREE-LEGS (or try ELEPHANT-WITH-BLUE-EYES, if you prefer a less grisly example) should be quite analogous to QUADRILATERAL (*i.e.*, practically identical to QUADRILATERAL's brother, TRIANGLE). In the latter case, it happens that there is a nice atomic term to equate with a composite description, and its parent happens to be a clean, mathematical concept. No one will deny that the natural kind parent of ELEPHANT-WITH-THREE-LEGS is substantially different than the non-natural kind of parent (POLYGON) of QUADRILATERAL ("elephant" does not mean "mammal with four legs," and may not have a criterial definition at all). But we can no less deny that an elephant with three legs is an elephant. ELEPHANT-WITH-THREE-LEGS and its compositional brethren are more like QUADRILATERAL than they are like ELEPHANT.

The lesson here is that in order for a knowledge representation system to be able to handle any reasonable range of descriptions—*even the simplest composites constructed from natural kind-like concepts*—some type of definitional (*i.e.*, compositional—not of the "typical" kind) structuring capability is necessary. To form descriptions of a very common sort, necessity and sufficiency are demanded. Put another way, any knowledge representation inference mechanism worth its salt must be able to deduce—without fail—that a three-legged elephant has exactly three legs. The internal structure of a nonatomic concept like ELEPHANT-WITH-THREE-LEGS must be transparent to the system's inference mechanism, or else it can't tell if the network creator is lying about the frame's apparent composition. Note that this does not imply that any *lexical items* need have definitions (see Israel & Brachman, 1984, for a more detailed argument on this matter), only that one should not be able to cast any doubt on a three-legged elephant's being an elephant and its having three legs (and *vice versa*). Many of the frames we need to reason with are akin to noun *phrases* in natural language—it is silly to assume that we will be stuck just with the equivalents of simple nouns.

inclusion—the first is included in the second, which is constructed from it. Analyticity is somewhat problematic in the philosophy of language, but its thrust is appropriate and trusty enough for our purposes here. According to Quine, "Kant conceived of an analytic statement as one that attributes to its subject no more than is already conceptually contained in the subject Kant's intent can be related thus: A statement is analytic when it is true by virtue of meanings and independently of fact" (Quine, 1961)

[16]About attributive modification, Quine (1960) says "A composite general term thus formed is true of just the things of which the components are both true " This is, simply, the normal use of adjectives in natural language.

Oblivious to the Obvious?

It is a bit embarrassing to be sitting here writing something so patently obvious as "an elephant with three legs is an elephant." Unfortunately, most frame and semantic net representations are forced to miss such immediate inferences because of the way they are constructed. As mentioned, a default interpretation of properties is almost always adopted at the expense of any serious definition mechanism. Slots, since they can be overridden, can be taken to represent only properties that *typically* follow from being an instance of a frame (as pointed out earlier.) So we can't use the properties in both directions in the cases we should be able to—in all cases something has to be explicitly asserted to be under a frame before it can be determined what properties follow from that attribution. And even then, we cannot categorically draw the inference—I can't confidently conclude that Clyde is a mammal even if I've stated that he is an elephant, because such properties are cancellable, and I'm kidding myself if I assume that the ISA link means that truly all elephants are mammals.

As a result, the system cannot use the structure of frames to determine whether one is more general than another, even when it should be able to. That is, the typical frame system cannot tell if one frame is a specialization of another even if that fact should be transparent from its content. Say I want to create the frame corresponding to the composite concept of a rhombus—a polygon with four equal-length sides. Despite the fact that it should be self-evident that it bears an analytic relation to the frame for quadrilateral, unless I—*the user*—so specify, the system is blind to the blinding truth. All I can do is attach the RHOMBUS frame to the POLYGON frame (or to the QUADRILATERAL frame if *I* happen to see the obvious), and add that it has four equal-length sides, as if those were only incidental properties of the class of objects I was describing. It only makes matters worse that I could cancel one of those properties and my frame system would still think I had a rhombus in hand. The concept of a three-sided rhombus would look just as coherent to it as the concept of a three-legged elephant.

What this all goes to say is, in effect, that *every description in the network is primitive.* While the so-called "frames" look complex, they act more like atomic primitives, so doomed by the failure to express real universals and sufficient conditions.[17] The poor user must always explicitly tell the system *every* immediate superdescription

[17]It should be noted that Fahlman's default-based notation, NETL, has a construct called an "EVERY" node, which attempts to draw properties compositionally into the meaning of his nodes Apparently, however, the subsumption machinery that would make this a meaningful definition mechanism was not fully implemented (S. Fahlman, personal communication, 1980). Also, the KRL "perspective" (Bobrow & Winograd, 1977) looks like a composite description However, perspectives, unlike "units," do not exist in the general conceptual taxonomy for use by the KRL interpreter.

of a newly added description—even a composite one that, by all rights, ought to wear its meaning on its sleeve.[18] Some frames these are—they really can't hold much of anything. Frames in default representations are not structured objects at all; their wholes seem to be *less* than the sums of their parts.

Hyphens to the Rescue?

Representation schemes that suffer from definitional deficiency tend to try to hide the internal structure of descriptions within node/frame names. We might, for example, get the impression that POLYGON-WITH-FOUR-SIDES is a compositionally structured description—just what we would want for the concept of a quadrilateral. But in reality it would be just as primitive as ELEPHANT.[19] One interesting consequence is that the only sense in which these frame systems are "representations" is that their authors have somewhat arbitrarily assigned meaning to a set of atomic symbols. In a system without true compositional structuring, there is no notion of *representation by structured correspondence.*

Only if you aren't allowed to lie about properties can the system know automatically, for example, that a rhombus is a quadrilateral. Every composite description should really have a proper place in the network based only on its internal structure. One very important job of a representation system, then, ought to be to keep these things in their places (as in KL-ONE—see Brachman & Schmolze, 1985). In fact, the idea of a representation system that honors compositionality gives rise to the notion of a *classifier* (Schmolze & Lipkis, 1983)—a fundamental part of the inference mechanism that places frames where they belong with respect to all other previously defined frames, regardless of where the user chooses to start them out. Classification shifts an important burden from the user to the system (where it surely belongs), and also turns out to be extremely useful in knowledge acquisition and maintenance (Finin & Silverman, 1984; Neches, Swartout, & Moore, 1984)—not to mention that it allows the system to make inferences that it really ought to be making in the first place. In any case, no matter how hard we want to believe that the descriptions in our representations have intrinsic compositional meaning—and no matter how complex we make their names—unless the system can distinguish between defaults and definitions, "they all look the same" to it.

[18]If we had real universals (uncancellable), we could at least lay claim to a form of compositionality The system could then infer that an elephant with three legs was an elephant. But without sufficient conditions, it couldn't, for example, tell us that Clyde fell under the ELEPHANT-WITH-THREE-LEGS frame, even if we told it that he was both an elephant and had three legs. Thus the sad irony is that ELEPHANT-WITH-THREE-LEGS—one of the mainstays of the frames world—can't even be considered to legitimately represent the concept of an elephant with three legs (for if it did, then the system could tell that Clyde was one).

[19]This is a perfect example of the kind of wishful thinking in AI that McDermott (1982) has already raked over the coals

Truth or Consequences

If nodes in semantic nets and frames in frame systems do not represent definitions, or even simple universals, what in fact are they? Often we hear them referred to as "prototypes," because they are supposed to correspond somehow to typical objects (and, as mentioned, some formalisms even encourage the labelling of stuctures with terms like "TYPICAL-ELEPHANT"). Unfortunately, this whole area is rather murky, inhabited by strange beasts like "PROTO-stereotypes" and "general structures of typical events." The two articles that suggested the existence of these semantic creatures (Bobrow & Winograd, 1979, and Lehnert & Wilks, 1979) provide a more enlightening analysis of a part of this swamp (at least for KRL) than I can possibly offer here, so I will not attempt here to give the final word on these matters. Rather, given that the kind of "typicality" embodied in AI representations is a seemingly unending source of entertaining dilemmas, we here conclude our little diatribe against prototypes-at-the-expense-of-definitions with a few of the (perhaps) less serious consequences. I will leave you to draw your own conclusions about the ultimate value of not-quite-well-enough-thought-out knowledge representation frameworks.

When is a Typical Mammal Not a TYPICAL-MAMMAL?

A piece of a natural animal taxonomy might look like this:

INDIAN-ELEPHANT ISA ELEPHANT ISA MAMMAL

This kind of hierarchy is common in the literature, and looks eminently reasonable. But let us not be lulled into a false sense of security; if all of our descriptions are of typical objects, this one should really read,

TYPICAL-INDIAN-ELEPHANT ISA
TYPICAL-ELEPHANT ISA TYPICAL-MAMMAL

Now we're in for some trouble, since it is impossible to tell whether the second relation says

> The typical elephant [some mythical abstract individual] is a typical mammal.
>
> The elephant is the typical mammal. [In fact, the horse or the cow is probably the typical mammal.]
>
> A typical elephant [some real individual] is a typical mammal.
>
> Or even, The typical elephant is a mammal.
>
> Or An elephant is a mammal.

The odd thing is that it is probably *none* of these that the prototype-based formalism is attempting to express. Rather, the intent is to say that elephants should be assumed to have all of the properties of mammals unless otherwise specified. (By the way, it's not clear that the first ISA relation—between INDIAN-ELEPHANT and ELEPHANT—has the same kind of intent.) Unfortunately, calling the nodes in semantic networks by names like "ELEPHANT"—or even "TYPICAL-ELEPHANT"—tempts us into thinking of the inter-node relation as an ISA relation. But in reality, it is not a classificatory relationship, but rather a simple line-of-inheritance specifier: It doesn't say whether elephants are mammals or not; it just says that elephants have mammal-properties unless told otherwise.[20]

Underlying the confusion here seems to be a failure to account for at least three different types of things, and the multiple kinds of ISA relations that might hold among things of those types:[21]

> The *concept* of a *kind* of a thing (*e.g.*, the concept of an elephant).
>
> A generic *description* that specifies the properties that *typically* apply to instances of a kind of thing.
>
> A *"prototypical" individual* that somehow typifies the kind.

Notice among other things a fundamental difference between the first two types of things—what we may consider bits of mental language (conceptual)—and the third, which is some actual or imagined member of the class. The last type of beast certainly seems to be the normal reading for phrases like "the typical elephant." Yet it is the middle type that seems most like the kind of thing denoted by nodes like TYPICAL-ELEPHANT in the type of networks we've discussed (see, *e.g.*, Lehnert & Wilks, 1979). Perhaps then such a node would more appropriately be called "ELEPHANT-WITH-ALL-DEFAULT-PROPERTIES."[22] While this seems to be a more true-to-life interpretation of type-nodes in proto-networks, it points out a fundamental problem: Using only such default node-types, there cannot be an independent concept of elephant that is without default properties (*i.e.*, no things of type 1 above). So we cannot distinguish between the normal senses of "Clyde is an elephant" and "Clyde is a typical elephant": "Everything that we have to say about 'the typical elephant' or about 'every elephant' is thus attached, in the form of a property or a statement-structure, to the TYPICAL-ELEPHANT node." (Fahlman, 1979).

Maybe it is not so disastrous to confuse "typical" and "every" for Clyde and other more or less standard mammals, but consider the poor platypus. In a non-prototype arrangement, a PLATYPUS surely ISA MAMMAL. Yet the platypus surely isn't a *typical* mammal! In fact, the platypus is probably the typical *atypical* mammal.

So how can we untangle this mess? Perhaps "typically" should be thought of as an operator on "normal" (compositional) concepts, in the style of Reiter's "A Logic for Default Reasoning" (Reiter, 1980). Then, rather than engendering confusion about the properties of "the typical elephant," we can express what we mean clearly: An elephant is a mammal, and *typically* is a creature with four legs.[23] In general it is probably a good idea to keep "typical" out of the names of our nodes.

How Many Kinds of Two-Toed Sloths Are There?

Along the same lines, careless attribution of meaning to descriptions because of seductive names can lead to another kind of confusion. For example, consider this: A two-toed sloth may not even be a two-toed animal. As Pat Hayes (personal communication, 1980) points out, there are several ways to interpret "two-toed" as an adjective:

> It can be characteristic of the type (as four-leggedness is with elephants and three-toedness is with three-toed sloths).
>
> It can be an inherent property of some particular anomalous individual born of a parent of a type with a different characteristic property (*e.g.*, a congenital defect—a three-toed sloth born with only two toes).
>
> It can be an incidental property accruing to some individual born normally of a type with a different characteristic property (*e.g.*, a three-toed sloth who lost one in the war).

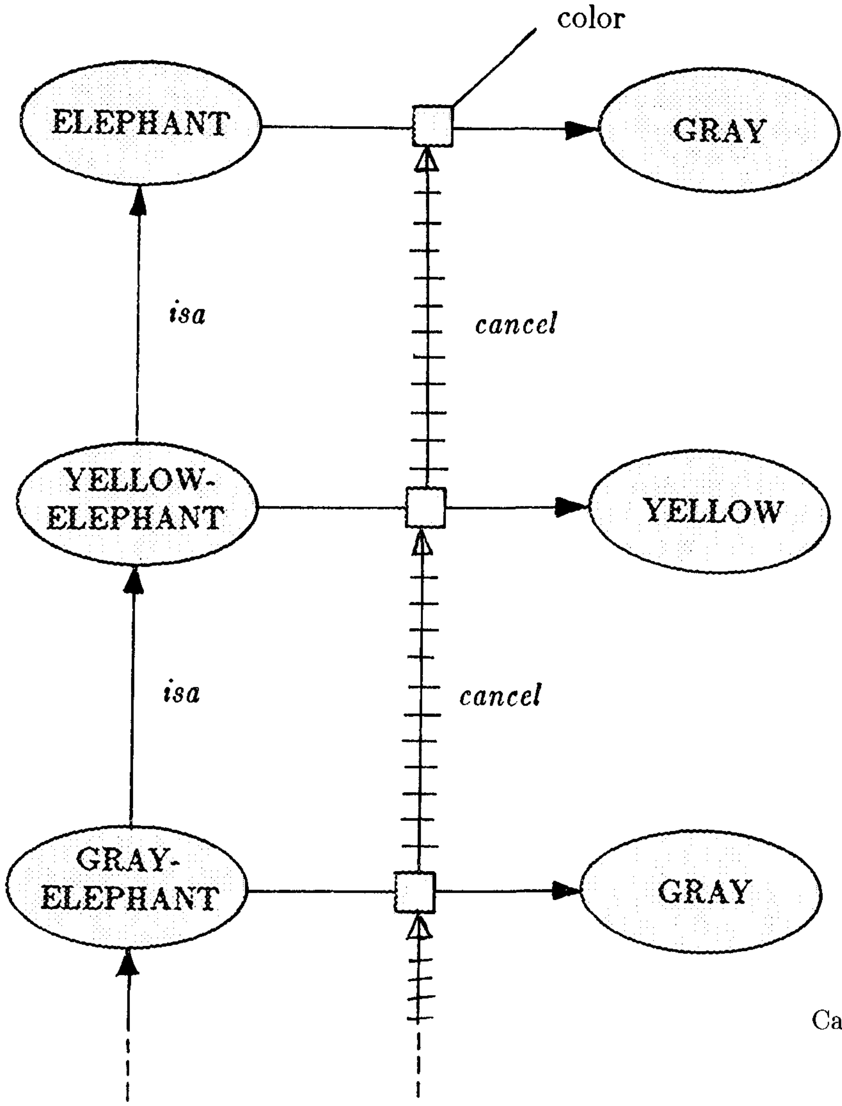
This being the case, we must be careful in even labelling a frame "TWO-TOED-ANIMAL" or "TWO-TOED-SLOTH" without attaching somehow a causal story of how

[20]Even this is being generous. What it really says is that every description attached in a certain way below (TYPICAL-)ELEPHANT inherits any property attached to (TYPICAL-)MAMMAL that isn't cancelled. Whether ELEPHANT represents elephants and MAMMAL mammals is another matter altogether

[21]I won't even mention here the distinction between the notion of an abstract individual like "the platypus" and the generic description of a platypus See Brachman (1983) for lots more on the ISA link Also, see Lehnert and Wilks (1979), Bobrow and Winograd (1979), and Winograd (1978) for some similar analyses of "prototypes"

[22]Given the point made in footnote 20, this node would really have to be something like "THING-WITH-DEFAULT-ELEPHANT-PROPERTIES" Whether or not a thing with all default elephant properties must be an elephant is up for grabs

[23]Don't forget that this implies that we need the machinery to construct the concept of a creature with four legs out of its constituent parts—with no prototypicality involved (a creature with four legs is *always* a creature and *always* has four legs). That is, even to state clearly the *default* for elephants, we really do need compositional structuring mechanisms



Cancel and Re-cancel.

Figure 4.

a so-described creature got that way. If we simply take some reasonable senses of the two English noun phrases then it is possible that a two-toed sloth (in the sense of type 3, describing a three-toed sloth that has had an accident) is *not* a two-toed animal (in the sense of type 1, the parent of the natural kind two-toed sloth concept).

Admittedly, some portion of such slips can be attributed to the existence of a special group of animals that happen to have been labelled "two-toed sloths." But if all we have are frame names to go on, TWO-TOED-SLOTH really can be ambiguous in several ways. Note that this is only a problem in a strict prototype-representation, where a description does not wear its structure on its sleeve. In a carefully specified framework with description composition mechanisms, names are completely irrelevant; which kind of two-toed sloth we are talking about can be read directly from the description's representational structure (this is the case, for example, with the newer versions of KL-ONE (Brachman & Schmolze, 1985) and with KRYPTON (Brachman *et al.*, 1983)).

The Shape of Things to Come

Consider the following intuitively reasonable line of argument:

> An important point about the hierarchies we will want to use is that, while they may be very bushy, they are never very deep Probably the most elaborate hierarchy in the real world is the taxonomy of animals. This hierarchy contains levels for species, genus, family, order, class, and phylum. Here and there, there are things like sub-species, sub-class, and sub-phylum. Above the phylum level, we might have markers such as
>
> ANIMAL
> LIVING-THING
> NATURAL-OBJECT
> PHYSICAL-ENTITY

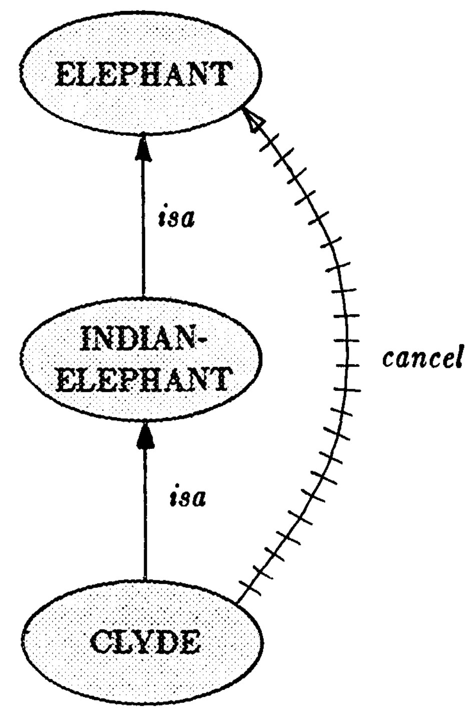
and finally just
ENTITY

> A classification hierarchy containing all these levels would still have a depth of only fourteen or fifteen. And surely common sense hierarchies are much shallower. I would think that COD IS-A FISH IS-A ANIMAL ... is typical. (Moore, 1975).

This appealing line seems to be predicated on an interesting assumption: As the cod goes, so goes everything else. The belief that (with perhaps a few technical exceptions) every concept of interest is natural kindlike has led to this assumption about the shape of the networks that will result when we "represent knowledge": "Knowledge bases consist mostly of short, bushy trees" (Fahlman, 1979). The feeling seems to be that the structure of the network will be dictated strictly by the relevant natural kind hierarchy and the depth of the hierarchy certainly won't exceed that of the "taxonomy of animals." In networks with no composite descriptions, this might well be true. However, from a formal *representation* standpoint—and not a psychological (or epistemological) one—there is absolutely nothing to determine *a priori* the depth of the network, except the grain of description. Description specialization is possible along any dimension of a concept whatsoever, *e.g.*,

ELEPHANT-WITH-THREE-LEGS
GRAY-ELEPHANT-WITH-THREE-LEGS
GRAY-ELEPHANT-WITH-THREE-LEGS-LIVING-IN-DALLAS
GRAY-ELEPHANT-WITH-THREE-LEGS-LIVING-IN-DALLAS-PLAYING-TIGHT-END-FOR-THE-COWBOYS,

etc., etc., etc.



Nonelephant elephants.

Figure 5.

More Fun With Cancel Links

Let's return to "cancel" links for a moment before wrapping up. Since we can override an attribute—say, the color of an elephant—why not turn around and override it again? Figure 4 shows a situation that notations that allow cancelling do not rule out. And why not continue to do so, alternating between GRAY and YELLOW to our heart's content? There's nothing to stop this kind of behavior, since the "implicit universals" aren't really universal![24]

Final Irony

By and large, we have been talking about cancelling *properties* (slots) of elephants and the like. We might also consider doing the same with ISA links (we wouldn't be the first—see Fahlman *et al.*, 1981). It at first seems plausible that an entity might have a "fundamental identity" (Fahlman, 1977), in which case it would be reasonable to forbid cancellation of ISA links. But notice that such a prohibition would make attributions like "elephants are mammals" uncancellable. That's fine for elephants, perhaps; but remember when whales were fishes?

On the other hand, if we were to allow cancellation of ISA links, then we would potentially be left in the untenable situation illustrated in Figure 5 (Fahlman, 1980): Something could be an INDIAN-ELEPHANT and yet not an elephant! Obviously, this can only happen in a situation in which the name attached to a description is not really telling the truth about the thing that it names.[25]

---

[24]Note that GRAY-ELEPHANT is really the same description as ELEPHANT. But with the usual network conventions pointed out in this article, the knowledge representation system wouldn't know it.

[25]I confess: After looking back at what I originally wrote, I find that this comment isn't quite fair. Determining whether whales are fishes

Conclusion

We have all from time to time enjoyed riddles like the one that initiated this little polemic. But imagine how unfunny it would be if the joke went like this:

Q: What's big and gray, has a trunk, and lives in the trees?

A: A giraffe—I lied about the color, the trunk, and the trees.

Or, worse yet:

Q: What's big and gray, has a trunk, and lives in the trees?

A: An idea—I lied about the color, the trunk, the trees, and about the "lives."

This begins to get boring, doesn't it? Imagine how tedious it would be if every single riddle you ever heard from now on was of this sort.

Well, pity the common frame or semantic net system, which has to live under these kinds of conditions. It seems that general arguments about the need to handle exceptions, and about the obvious way to represent them, have led to a breed of AI representation systems that don't allow the interpreter to predict when all bets will be off. While today's AI antilogicians might protest that definitions are impossible or irrelevant—after all, they say, "rules are made to be broken"—the pendulum seems to have swung too far in the direction of exception-handling. When *all* rules are made to be broken, then they aren't rules at all. The call to arms sounded by Fahlman (1979) and Minsky (Kolata, 1982) for the need to represent three-legged elephants and dead ducks (or those with their feet set in concrete), while crucially important if AI is ever to get a hold of "common sense," has led to naive mechanisms that both admit arbitrary, bizarre representations and force the ignorance of crucial, obvious facts (*e.g.,* the rule that elephants with three legs are elephants is *not* made to be broken). AI representation systems have thrown out the compositional baby with the definitional bathwater.

There are thus three simple ideas to take away from this paper: Frames are typically not very frame-like, definitions are more important than you might think, and cancellation is worse than it looks. If I have succeeded here, then perhaps the next time someone tries to sell you a representation system that allows cancellation, you'll think twice before buying. Or at least kick the tires.[26]

or mammals is a matter of discovery, the kind of thing usually left to biologists So our current theory about them could be wrong But, if we believe Kripke (1972), whales are *necessarily* mammals, if they are in fact mammals at all Thus, in the world it couldn't be otherwise So we have to be even more careful than we thought, and distinguish between representations reflecting our *having knowledge* about something (even about something that, if it is true, is necessarily so) and the metaphysically necessary truths themselves, which aren't matters of discovery or semantics at all We're in an odd situation in which a necessary truth (*e g.,* about whales) could be "cancelled," because our theory about it was wrong (such truths don't change but my knowledge about them is certainly non-monotonic) Despite this complication, the main point of this section is still valid—we can't do it all with one rule about cancellation—but the issue obviously needs much more thought

References

Bobrow, D. G (Ed.) (1980) Special Issue on Non-Monotonic Logic. *Artificial Intelligence* 13(1-2).

Bobrow, D G., & Winograd, T (1977) An Overview of KRL, a Knowledge Representation Language *Cognitive Science* 1(1):3–46.

Bobrow, D. G., & Winograd, T (1979) KRL: Another Perspective *Cognitive Science* 3(1):29–42.

Bolinger, D (1965) The Atomization of Meaning *Language* 41(4):555–573.

Borning, A. (1981) The Programming Language Aspects of ThingLab, a Constraint-Oriented Simulation Laboratory *ACM Transactions on Programming Languages and Systems,* 3(4):353–387

Brachman, R J. (1979) On the Epistemological Status of Semantic Networks. In N. V. Findler (Ed.), *Associative Networks: Representation and Use of Knowledge by Computers.* New York: Academic Press.

Brachman, R. J. (1983) What IS-A Is and Isn't: An Analysis of Taxonomic Links in Semantic Networks. *IEEE Computer* 16(10):30–36

Brachman, R J , & Levesque, H J. (1982) *Competence in knowledge representation* AAAI-82, 189–192

Brachman, R. J., & Levesque, H. J. (1984) *What Makes a Knowledge Base Knowledgeable? A View of Databases from the Knowledge Level.* Proceedings of the First International Workshop on Expert Database Systems 30–39.

Brachman, R. J., & Schmolze, J. G. (1985) An Overview of the KL-ONE Knowledge Representation System *Cognitive Science* 9(2):171–216.

Brachman, R. J , Fikes, R E. & Levesque, H. J. (1983) KRYPTON: A Functional Approach to Knowledge Representation *IEEE Computer* 16(10):67–73.

Carnese, D. J (1985) Multiple Inheritance in Contemporary Programming Languages MIT/LCS/TR-328 Laboratory for Computer Science, Massachusetts Institute of Technology.

Charniak, E (1981) A Common Representation for Problem-Solving and Language-Comprehension Information. *Artificial Intelligence* 16(3):225–255

Cheeseman, P (1985) *In defense of probability* IJCAI-85.

Cherniak, C (1984) Prototypicality and Deductive Reasoning. *Journal of Verbal Learning and Verbal Behavior* 23(5), 625–642.

Cohen, B., & G. L. Murphy (1984) Models of Concepts. *Cognitive Science* 8(1): 27–58

Dreyfus, H. L (1981) From Micro-Worlds to Knowledge Representation: AI at an Impasse. In J. Haugeland (Ed), *Mind Design* Cambridge, MA: The MIT Press

Etherington, D., & Reiter, R (1983) *On Inheritance Hierarchies with Exceptions* AAAI-83, 104–108.

Fahlman, S. E. (1977) A System for Representing and Using Real-World Knowledge. Ph D Dissertation, Massachusetts Institute of Technology

Fahlman, S. E. (1979) *NETL: A System for Representing and Using Real-World Knowledge.* Cambridge, MA: The MIT Press

Fahlman, S. E (1980) *Exploring the Limits of Marker-Passing Parallelism* Workshop on Knowledge Representation and Language Comprehension, Brown University.

Fahlman, S. E., Touretzky, D S & van Roggen, W (1981) *Cancellation in a Parallel Semantic Network.* IJCAI-81, 257–263.

Fikes, R , & Kehler, T (1985, in press) Control of Reasoning in Frame-Based Representation Systems. *Communications of the ACM, Special Issue on Architectures for Knowledge-Based Systems.*

Finin, T., & Silverman, D. (1984) *Interactive classification: a technique for building and maintaining knowledge bases* Proceedings of the IEEE Workshop on Principles of Knowledge-Based Systems, 107–114

[26]Sorry, representation systems don't *really* have tires .

Ginsberg, M. L. (1984) *Non-monotonic reasoning using dempster's rule.* AAAI-84, 126–129

Hayes, P J. (1979) The Logic of Frames In D. Metzing (Ed), *Frame Conceptions and Text Understanding.* Berlin: Walter de Gruyter and Co.

Israel, D. J , & Brachman, R. J (1984) Some Remarks on the Semantics of Representation Languages. In M L. Brodie, J Mylopoulos, & J. W Schmidt (Eds.), *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages* New York: Springer-Verlag.

Kolata, G. (1982) How Can Computers Get Common Sense? *Science* 217:1237-1238.

Kripke, S. A. (1972) Naming and Necessity. In G Harman & D. Davidson (Eds.), *Semantics of Natural Language.* Dordrecht, Holland: D. Reidel, 253–355

Lehnert, W., & Wilks, Y. (1979) A Critical Perspective on KRL *Cognitive Science* 3(1):1–28

Levesque, H J. (1981) A Formal Treatment of Incomplete Knowledge Bases. Ph D. dissertation, University of Toronto Also published as FLAIR Report No 3, Fairchild Laboratory for Artificial Intelligence Research, Palo Alto, CA.

Levesque, H. J. (1984a) *A fundamental tradeoff in knowledge representation and reasoning.* Proceedings of CSCSI/SCEIO-84, 141–152

Levesque, H J. (1984b) Foundations of a Functional Approach to Knowledge Representation. *Artificial Intelligence* 23(2): 155–212.

Levesque, H J. (1985) *Overview of Knowledge Representation.* Islamorada Workshop on Large Scale Knowledge Base and Reasoning Systems, Islamorada, FL.

Maida, A S., & Shapiro, S C. (1982) Intensional Concepts in Propositional Semantic Networks. *Cognitive Science* 6(4): 291–330.

McDermott, D. (1981) Artificial Intelligence Meets Natural Stupidity In J. Haugeland (Ed.), *Mind Design.* Cambridge, MA: The MIT Press.

Minsky, M. (1975) A Framework for Representing Knowledge In P. H. Winston (Ed), *The Psychology of Computer Vision.* New York: McGraw-Hill

Moore, R C. (1975) A Serial Scheme for the Inheritance of Properties. *SIGART Newsletter,* 53: 8–9

Moore, R C. (1982) *The role of logic in knowledge representation and commonsense reasoning.* AAAI-82, 428–433.

Neches, R., Swartout, W. R., & Moore, J (1984) Enhanced Maintenance and Explanation of Expert Systems Through Explicit Models of Their Development. *Proc IEEE Workshop on Principles of Knowledge-Based Systems,* Denver, IEEE Computer Society Press, 173–183.

Newell, A. (1981) The Knowledge Level *AI Magazine* 2(2):1–20.

Parkes, C H. (1980) Against Definitions. *Cognition* 8(3):263–367.

Putnam, H (1977) Is Semantics Possible? In S P Schwartz (Ed.), *Naming, Necessity, and Natural Kinds.* Ithaca, NY: Cornell University Press.

Quine, W. V. (1960) *Word and Object* Cambridge, MA: The MIT Press.

Quine, W. V. O. (1961) Two Dogmas of Empiricism. In W. V. O Quine, *From a Logical Point of View.* New York: Harper

Quine, W. V. O (1976) Truth by Convention. In W V O. Quine, *The Ways of Paradox and Other Essays* Revised and enlarged edition. Cambridge, MA: Harvard University Press

Reiter, R. (1978) *On Reasoning by Default.* Proceedings of TINLAP-2, Urbana, IL, 210–218

Reiter, R. (1980) A Logic for Default Reasoning. *Artificial Intelligence* 13(1,2):81-132.

Rich, E. (1983) *Default reasoning as likelihood reasoning.* AAAI-83, 348–351

Schank, R. C. (1972) Conceptual Dependency: A Theory of Natural Language Understanding. *Cognitive Psychology* 3(4):552–631.

Schank, R. C (1973) Identification of Conceptualizations Underlying Natural Language Understanding In R C. Schank & K. M Colby (Eds.), *Computer Models of Thought and Language* San Francisco: W H Freeman and Company.

Schmolze, J G , & Lipkis, T A. (1983) *Classification in the KL-ONE Knowledge Representation System.* IJCAI-83, 330–332.

Touretzky, D S. (1984) *Implicit Ordering of Defaults in Inheritance Systems* AAAI-84, 322–325

Wilks, Y. (1973) An Artificial Intelligence Approach to Machine Translation. In R. C Schank & K. M. Colby (Eds.), *Computer Models of Thought and Language* San Francisco: W. H Freeman and Company

Winograd, T. (1978) On Primitives, Prototypes, and Other Semantic Anomalies. *Proc. TINLAP-2,* Urbana, Ill., 25–32.

Zadeh, L A (1983) Commonsense Knowledge Representation Based on Fuzzy Logic *IEEE Computer* 16(10):61–65.