UNDERSTANDING ADJECTIVES

Yawar Ali

Technical Report CSRI-167
January 1985

Computer Systems Research Institute
University of Toronto
Toronto, Canada
M5S 1A1

The Computer Systems Research Institute (CSRI) is an interdisciplinary group formed to conduct research and development relevant to computer systems and their application. It is jointly administered by the Department of Electrical Engineering and the Department of Computer Science of the University of Toronto, and is supported in part by the Natural Sciences and Engineering Research Council of Canada.

UNDERSTANDING ADJECTIVES

by

Yawar Ali

Department of Computer Science

A Thesis submitted
in conformity with the requirements
for the Degree of Master of Science
in the University of Toronto
January 1985.

Abstract

This thesis deals with the task of understanding certain kinds of descriptive phrases, viz. noun phrases containing sequences of prenominal adjectives. In order for a computational agent to understand such a noun phrase, two distinct yet related subproblems must be solved. The first problem is to determine exactly what each adjective modifies. In general, this can only be done by taking account of the semantic properties of the adjective in question, as well as those of other adjectives to its right, and of the noun itself. "Real world" knowledge and contextual factors also play a role in this process. The second problem is to construct a representation of the description embodied in such a noun phrase. Here, it is desirable that the structure of the representation correspond to the structure of modification within the phrase.

In this thesis, the first problem is addressed by developing a classification scheme for adjectives which allows us to substantially reduce the number of candidate interpretations, in some cases to a single one. A system is presented which takes account of the disparate semantic behaviour of different classes of adjectives, word order, punctuation in the noun phrase, and a frame-based store of "real world" knowledge, in order to determine the scope of adjectives within a noun phrase.

In approaching the second problem mentioned above, particular adjectives are taken to indicate restrictions on the values that objects may take on for associated properties. These properties may be featural, dimensional, or functional in nature. Frame-like structures are used to represent the generic concepts that are taken to be associated with noun phrases. Slots are used to represent stereotypical properties of generic objects, and interslot constraints enforce restrictions on the values taken on by inter-related properties. A single knowledge base is used to represent both the interpretations constructed for noun phrases and the "real-world" knowledge needed to resolve structural ambiguity in the noun phrase The contents of the knowledge base are organized along a classification axis and a generalization axis, as well as along an implicit aggregation axis. The use of these axes enables us to capture the relationships that various concepts bear to one another, and enables us to build up interpretations of noun phrases compositionally from the interpretations of their constituents.

Acknowledgements

I would like to take this opportunity to thank my supervisor Graeme Hirst for his invaluable contribution to the work described in this thesis. It was Graeme who originally suggested that I work on the problem that later came to be the topic of this thesis. During the course of working on this problem, I often had occasion to bounce ideas off Graeme; his comments greatly aided me in clarifying my thoughts, and in eliminating some of the more fanciful ones. Graeme is also to be thanked for his careful reading of various drafts of the thesis, and for many helpful suggestions towards making it more readable.

I would also like to thank John Mylopoulos for taking time from all his other responsibilities to read this thesis. In addition, I am grateful to John for providing me with the opportunity to get involved with the research activities of the AI group while still an undergraduate.

I thank Andrew Gullen for helping me to refine the ideas embodied in this thesis, and for much assistance in the implementation phase.

My initial supervisor Ray Perrault taught me much about Computational Linguistics, and for this I am deeply grateful to him. By always maintaining the highest standards of excellence in his work, Ray inspires his students to strive for the same in theirs. I also thank Ray for his continued interest in my work after his move to sunny California.

During my sojourn at this department, I came to know many weird and wonderful people, most of them fellow graduate students. I would like to express my appreciation to them, individually and collectively, for a wonderful two years.

Lastly, and most importantly, I must thank all the members of my extended family, and in particular my parents, for their tremendous support over the years, without which none of this would have been possible. This one's for them.

For my family

Table of Contents

CHAPTER 1

Introduction

1.1. Setting the Stage

"Some of us sometimes understand what is said. The problem is to understand how we manage to do this" (Paul Ziff).[1]

Broadly speaking, this thesis is concerned with the problem so succinctly stated by Ziff in the quotation above. This is not to say that we solve the problem, for as Ziff points out, "[it] is insoluble at present". He goes on to say: "But my attempt is not of the impossible. Neither have I been concerned to put forth a continuous complete analysis of the topic: a show of continuity could only be fraudulent, a claim to completeness at best presumptuous nonsense".[2] The reader stands warned: this thesis will not present a continuous complete analysis of understanding. Instead, it will attempt to develop a partial account of how a restricted class of utterances might come to be understood by a particular kind of language user. In the remainder of this section, I shall elaborate on each of these restrictions.

The utterances that we are interested in are those in which adjectives are used to construct descriptions. In particular, we will study descriptive noun phrases. Sidner (1983) points out that a given noun phrase may be interpreted in a number of different ways, depending upon the discourse context and the beliefs of the speaker and hearer. For example, consider (1) and (2) :

(1) I saw *the tall man* enter the building again today.

(2) Levis don't cater to *the tall man*.

While the definite noun phrase *the tall man* is present in both (1) and (2), in (1) it is being used to refer to a particular person (i.e., a SPECIFIC reading of the noun phrase), whereas in (2) it is to be interpreted as a description of a class of people (under a GENERIC reading of the noun phrase). Indefinite noun phrases also allow the specific/generic distinction, as demonstrated by (3) and (4):

(3) *An angry linguist* is banging on the door!

(4) *An angry linguist* is never at a loss for words, especially adjectives!

---

[1] From the opening of his essay "What Is Said"; in *Semantics of Natural Language*, Gilbert Harman and Donald Davidson (editors), D. Reidel Publishing Company: Dordrecht, Holland, 1972, pp. 709-721 (reprinted in Ziff 1972).

[2] Ziff (1972), Preface.

We shall be concerned with the problem of understanding noun phrases under the assumption that a generic reading is appropriate.

Both adjectives and nouns can appear as premodifiers of the HEAD (rightmost noun) of a noun phrase. When nouns act as modifiers of other nouns, they are said to exhibit NOUN-NOUN MODIFICATION, and sequences of nouns related through modification are termed NOUN COMPOUNDS. A considerable amount of work has been carried out by linguists and A.I. researchers on the task of determining the structure and interpretation of noun compounds. In what follows, we shall generally ignore the problem of noun-noun modification, and instead focus our attention on adjectival modification. However, it should be pointed out that our analysis has been strongly influenced by research on noun-noun modification by Finin (1980), D.B. McDonald (1982), and others, and we shall have occasion to refer to their work from time to time.

The kind of understanding that we are interested in here involves mapping from a generic noun phrase to some representation of the description it embodies. We shall refer to this process as SEMANTIC INTERPRETATION, although we shall be more concerned with capturing the structure of the concepts associated with noun phrases than in their abstract semantic characterization. Thus, we shall focus on studying the ways in which noun phrases may be used to construct descriptions, and how the interpretations of complex descriptive terms may be built up from the interpretations of their constituents. We also ignore in this work the difficult problem of REFERENCE: how concepts (and their computational representations) get mapped onto objects in the world.

We shall attempt to develop a computational account of the interpretation of generic noun phrases. The success of our approach will be judged by how well it captures our pre-theoretic intuitions about how descriptive terms should be interpreted. We shall make no claims for the psychological reality of the mechanisms that we develop. However, where psycholinguistic results seem to bear upon our work we shall discuss them and examine their import from our own perspective.

1.2. The Problem

In this work, we shall restrict ourselves to the problem of interpreting generic noun phrases consisting of a single noun preceded by a sequence of adjectives. Since our concern is with adjectival modification, and not with noun-noun modification, we exclude noun phrases with more than one noun, but place no restriction on the number of prenominal adjectives. Moreover, since both definite and indefinite noun phrases can have generic readings, we shall ignore the determiner sequences usually present in noun phrases.

The task that we have set ourselves may be roughly decomposed into two, related subproblems. The first subproblem is that of determining the scope of each adjective within a prenominal sequence. Each adjective in such a sequence may separately modify the head of the noun phrase, as in the examples below:

excited, red-headed student;
tall clever child;

young former Torontonian;
long-haired, slack-shouldered, whining beatnik;[3]
large, powerful engine[4]

Alternatively, the adjectives in a prenominal sequence may nest with one another to modify the noun in combination, as shown below:

former red-headed debutante;
tall young child;
good semantic theory;
typical spoilt brat;
fake automatic revolver

In some cases, both types of behaviour may be exhibited by adjectives within the same prenominal sequence, as shown below:

skinny excited former red-headed debutante;
big, black, fake automatic revolver;
polite, tall young blonde child

There are even examples in which some adjectives seem to "skip over" others in order to combine with distant adjectives, as in *tall red-headed young child*;[5] here, the interpretation seems to be "a child who is red-headed and tall for a young child".

The second subproblem to be solved is that of constructing adequate representations of the descriptions embodied in generic noun phrases. Since such phrases are taken to refer to classes of objects, their interpretations should be in terms of properties that are, in some sense, typical of objects of the relevant classes. The problem is further complicated by the fact that in many cases adjectives seem to refer to properties of objects other than those (objects) associated with their complements. For example, *short basketball player* is normally understood as "short for a basketball player", but *short engineer* is better interpreted as "engineer who is short for an adult". Similarly, we take a *clever woman* to be clever for a person, but a *clever idea* seems to be one which is due to a clever person.

1.3. Organization of this Thesis

The remainder of this thesis is divided into four chapters. Chapter 2 examines the syntactic and semantic behaviour of adjectives and the nature of adjectival modification. It introduces a classification scheme for adjectives, and describes how it may be used to help resolve ambiguity in the noun phrase. This chapter also examines the various kinds of modification relationship that can hold between an adjective and its complement.

---

[3] Example from Len Deighton's thriller *Only When I Larf* (p. 89), Sphere Books, London, 1967.

[4] Example adapted from Gil (1983).

[5] Example adapted from Siegel (1979).

Chapter 3 is concerned with the representation of knowledge. It introduces the PSN semantic network formalism, and describes how this language may be used to represent the concepts associated with noun phrases.

Chapter 4 describes the design of a semantic interpreter which employs techniques discussed in Chapter 3 to capture the knowledge discussed in Chapter 2. The chapter begins by examining the problem of interpreting noun phrases consisting of a single adjective modifying a noun. The latter part of the chapter shows how noun phrases with arbitrarily many prenominal adjectives may be interpreted by appropriate combinations of the methods used to interpret adjective-noun pairs.

Chapter 5 summarizes the work, unifying the ideas presented separately in the previous chapters. The implementation of the interpreter is briefly described, followed by an evaluation of its performance, and a discussion of how it might be improved. The chapter closes with an examination of some unresolved problems, and some suggestions for extending the work by linking it up with that of others.

CHAPTER 2

Adjectives and Adjectival Modification

We begin our study by discussing the syntactic and semantic behaviour of adjectives and the kinds of modification relationships that they enter into. After examining the behaviour of single adjectives, we consider the problem of interpreting complex noun phrases which include multiple prenominal adjectives.

2.1. Single Adjectives

2.1.1. Syntactic Classification

Grammarians have usually described adjectives in terms of their syntactic behaviour. In *A Grammar of Contemporary English* (1972), Quirk, Greenbaum, Leech, and Svartvik characterize adjectives as being able to appear in ATTRIBUTIVE position before nouns, as in:

(1) the successful businessman

(2) a dead penguin

(3) the old woman

or in PREDICATIVE position after INTENSIVE verbs (those co-occurring with subject complements), as in:

(4) That businessman seems successful.

(5) The penguin dropped dead.

(6) The woman appeared old.

Adjectives such as *successful*, *dead*, and *old*, which can appear in both positions, are considered to be CENTRAL to the class, whereas those that are restricted to one position or the other are regarded as PERIPHERAL.[1] According to this, *asleep* and *utter* would be peripheral adjectives, as demonstrated below:

(7) *the asleep man

[1] Levi (1978) studied a class of adjectives, which she termed NON-PREDICATING adjectives, that were characterized, in part, by their inability to appear in predicative position. We shall briefly look at non-predicating adjectives in the next section.

(8) The man seemed asleep.

(9) the utter fool

(10) *The fool seems utter.

A further characteristic of adjectives is that they cannot function as direct object in the absence of an article, nor in the presence of an indefinite article. So, we have:

(11) *We like successful.

(12) *It approaches dead.

(13) *She takes old.

(14) *Joan wants asleep.

(15) *That fool needs an utter.

Thus, on a superficial analysis, the noun *beef* might be classified as a peripheral adjective due to its ability to appear in attributive position, but not in predicative position:

(16) the beef steak

(17) *The steak seemed beef.

However, a more thorough analysis would disqualify it on the ground that it can appear as direct object in the absence of an article, as in:

(18) Joan likes beef.

It should be noted that some words, such as *criminal*, *classic*, and *noble*, take on characteristics of both adjectives and nouns, and thus cannot be distinguished using the tests we discussed above (Quirk et al., p.240).

For our purposes, the traditional, syntactic account is chiefly useful for pointing out the distinction between attributive and predicative uses of adjectives. We shall be concerned, primarily, with the former, although some of our subsequent analysis will also be applicable to adjectives in predicative position.

2.1.2. Semantic Classification

Based on the discussion in the previous section, we could countenance a division of adjectives among two classes: those that can appear in attributive position, and those that cannot. However, a purely syntactic analysis would not help us in further distinguishing between various kinds of attributive adjectives. Instead, we shall attempt to characterize such adjectives on the basis of their semantic behaviour.

There appears to be no clear consensus among linguists as to the best framework for analyzing the semantic properties of adjectives. We shall adopt a classification proposed by Siegel (1979), who employed three semantic categories of adjective: INTENSIONAL, ABSOLUTE, and MEASURE. In addition, we shall define a new category, namely that of ROLE adjectives.

Before we proceed to our semantic characterization, we shall briefly examine, and thereafter disregard, a class of adjectives studied by Levi (1978), who termed them NON-PREDICATING adjectives. These were distinguished by three properties: they could not appear in predicative position, their meanings varied with the nouns that they modified, and they could appear in syntactic environments where common nouns could occur. Some examples of non-predicating adjectives are:

alluvial	domestic	dramatic	editorial
electrical	financial	fraternal	linguistic
literary	monthly	musical	national
oceanic	presidential	rural	senatorial
solar	Tarskian	tropical	vocal

Levi argued, convincingly, that non-predicating adjectives were actually derived from underlying common nouns, and that their linguistic behaviour was more characteristic of nouns than adjectives. Accordingly, we shall exclude non-predicating adjectives from our analysis.[2]

2.1.2.1. Intensional Adjectives

INTENSIONAL adjectives are those whose meanings depend on the meanings of the nouns that they modify. Some examples of such adjectives are:

alleged	competent	consummate	fake
former	good	implicit	inner
inveterate	mock	original	ostensible
potential	rife	rightful	sincere
superior	terrible	typical	unusual

These include both EVALUATIVE adjectives, such as *good*, as well as ones like *former*, which denote more "objective" properties.

The intensional nature of such adjectives prohibits the substitution of co-extensional expressions in their complements. For example, if it is assumed that, by coincidence, all and only philosophers happen to be politicians, (19) does not express a valid statement,

[2] To be precise, Levi's arguments were applicable only to a subset of the non-predicating adjectives, excluding composite adjectives (such as *multicylinder*), and those (such as *potential*) that appeared to be derived from underlying adverbs. Therefore, we shall only exclude from our study those, "true", non-predicating adjectives that are not members of the two latter categories.

(19) Joan is a good philosopher if and only if Joan is a good politician.

since being good as a philosopher is quite different from, and probably inimical to, being good as a politician. Similar arguments would apply to the following:

(20) Joan is a competent philosopher if and only if Joan is a competent politician.

(21) Joan is a former philosopher if and only if Joan is a former politician.

(22) Joan is a terrible philosopher if and only if Joan is a terrible politician.

(23) Joan is a typical philosopher if and only if Joan is a typical politician.

Montague (1970) employed POSSIBLE WORLDS SEMANTICS to analyze intensional adjectives,[3] as part of his general approach of applying techniques of metamathematics to the study of natural languages. We shall not present a detailed account of Montague's theory here; the interested reader is referred to Dowty, Wall, and Peters (1981).

Montague took each word of a language to denote some object in a formal mathematical model, relative to some INDEX (a pair consisting of a possible world and a point in time). In particular, common nouns were taken to denote sets of individuals. The INTENSION of an expression was defined to be a function that gave, for each index, the denotation of the expression at that index. The intension of a common noun, called a PROPERTY, would be a function that gave, for each index, the set denoted by the noun at that index. For example, for each index, the common noun *writer* would denote a set of individuals consisting of all and only those who happen to be writers in that world at that time. Then, the intension of *writer* would be a function W that gave, for each index, the set of writers at that index.

In Montague's system, intensional adjectives were taken to denote functions from properties to properties. In other words, such an adjective would combine with the intension of a common noun to give the intension of a more complex common noun, representing the adjective-noun pair. The denotation of such a complex noun, at each index, would be the set of individuals described by the adjective-noun pair at that index. For example, the intensional adjective *good* would denote, at any fixed index i_1, a particular function, say G, from properties to properties. If this function G were applied to the intension W of the common noun *writer*, it would yield another function $F = G(W)$, such that, for each index i_2, F would give the set of good writers at i_2.[4]

Montague's analysis captures the relative nature of intensional adjectives, since they are applied to the intensions (or meanings) of the common nouns that they modify.

[3] Actually, Montague analyzed *all* adjectives in the same terms, as functions from properties to properties. Siegel (1979), Thomason (1974b), and others have argued against assigning a uniform semantic treatment to all adjectives.

[4] Note that the meaning of *good writer* will be as defined at index i_1, rather than at index i_2.

Unfortunately, it is not possible to directly implement Montague semantics computationally, since it is committed to a holistic view of the world that requires the availability of total information, not just about the actual world, but about all possible objects in all possible worlds (Barwise 1981, p.110).[5] As yet, there has been no adequate computational account of the semantics of such adjectives.

2.1.2.2. Absolute Adjectives

In contrast to intensional adjectives, ABSOLUTE adjectives are extensional, their meanings not being relative to the nouns that they modify. In addition, such adjectives yield readings that can be described as INTERSECTIVE, in the sense discussed below. Some examples of absolute adjectives are:

asleep	black	crested	dead
eternal	four-legged	grammatical	illegal
impossible	infallible	meaningless	nude
opaque	paradoxical	pastel	perfect
printed	spotted	true	unconscious

We can demonstrate the extensionality of such adjectives by assuming, again, that all and only philosophers are politicians, and then considering the validity of the statement expressed by (24).

(24) Joan is a nude philosopher if and only if Joan is a nude politician.

Clearly, the statement is valid. Thus, we conclude that the absolute adjective *nude* is extensional, since it permits substitution of co-extensional terms in its complement. Similar arguments apply to the following:

(25) Joan is a black philosopher if and only if Joan is a black politician.

(26) Joan is a crested philosopher if and only if Joan is a crested politician.

(27) Joan is a dead philosopher if and only if Joan is a dead politician.

(28) Joan is a four-legged philosopher if and only if Joan is a four-legged politician.

A further property of absolute adjectives is that in attributive position they yield readings that can be described as INTERSECTIVE (Montague 1970, Siegel 1979): absolute adjectives can be taken as denoting fixed sets of objects, as can common nouns; then, the denotation of a common noun modified by such an adjective is given by the intersection of

[5] Smith (1982, p.13) has argued that Montague semantics can be taken seriously computationally, on the grounds that a system making use of Montague semantics should only need to compute with (finite) designators of the infinite functions proposed by Montague, rather than with the actual functions themselves. However, it is not obvious how one might make use

the denotations of the adjective and the noun. For example, we may regard the adjective *pink* as denoting the set of all pink things, and the common noun *elephant* as denoting the set of all elephants. Then the denotation of *pink elephant* could be obtained by taking the intersection of the set of pink things with the set of elephants.

Such an operation would not be possible with intensional adjectives, since their denotations normally cannot be taken to be sets. For example, it does not make sense to speak of "the set of former things" or of "the set of good things" as being the denotations of the intensional adjectives *former* and *good*, respectively.

The semantics of absolute adjectives are straightforward, and a number of linguists and AI researchers have given essentially similar accounts of them in a variety of formalisms. For instance, Siegel (1976) presents a treatment of absolute adjectives as extensional one-place predicates within a Montague-style framework, and D.B. McDonald (1982), as an extension of his work on interpreting noun compounds, discusses his treatment of absolute adjectives[6] as type nodes (each representing the class of things described by such an adjective) of the semantic network formalism NETL (Fahlman 1979).

2.1.2.3. Measure Adjectives

MEASURE adjectives are those whose denotations must be determined with reference to a contextually determined comparison class, and, often, a measurement scale and a norm for the class with respect to the measured attribute (Siegel 1979). Examples of such adjectives are:

big	blunt	brief	broad
bulky	clean	cold	deep
dense	distant	fast	fat
hard	healthy	high	old
populous	remote	short	steep

Unlike absolute adjectives, measure adjectives do not yield intersective readings, and do not seem to denote fixed sets of objects. For example, one might try to conceive of the measure adjective *big* as denoting some fixed set of "big things", and hope to obtain the denotation of *big flea* by taking the intersection of the set of fleas with the set of "big things", which would presumably include the larger-sized members of all classes of objects. However, if one now tries to obtain the denotation of *big animal* by taking the intersection of the set of all animals with the same set of "big things", one will end up with a set consisting of the larger-sized members of every species of animal. This is clearly not the intended meaning of the noun phrase. For instance, one would not normally describe any group of fleas as "big animals", even though they may be big for fleas.

---

of this insight to actually design such a system.

[6] Actually, McDonald fails to distinguish between absolute adjectives and other, semantically more complex, adjectives. We have already pointed out that intensional adjectives cannot be interpreted in the same manner as absolute adjectives. In the next section, we will argue that neither can measure adjectives.

While the indeterminacy of measure adjectives has long been recognized in the literature, it has usually been ascribed to their being intensional noun-modifiers (for a recent example see Hellan 1981, p.46). In contrast, Siegel (1979) argues that measure adjectives should be regarded as a class of extensional predicates whose interpretation is dependent on contextual factors. Assuming, again, the co-extensionality of the class of philosophers and the class of politicians, does (29) express a valid statement?

(29) Joan is a fat philosopher if and only if Joan is a fat politician.

It might be argued, for instance, that philosophers, being inclined to sloth, could generally be expected to be fatter than politicians, it being a fluke, and a transitory occurrence, that the same set of people happen to be employed in both occupations. On this account, the statement expressed by (29) would not be valid. However, Siegel argues (and I agree) that the failure of substitution to preserve truth value in such a case is due to the varying influence of contextual factors on the semantic interpretation of different predicate-argument pairs, and should not be taken as evidence that these predicates are intensional (see Kamp 1975, p.126, for a similar argument). More precisely, different uses of a measure adjective may pick out different comparison classes, possibly resulting in different norms being selected for the application of the adjective, and yielding differing interpretations with different nouns. What is important is that the measurement scale for the attribute denoted by the adjective remains constant, even though the "reference point", or norm, may shift up or down the scale. Similar arguments would apply to the following:

(30) Joan is a distant philosopher if and only if Joan is a distant politician.

(31) Joan is a healthy philosopher if and only if Joan is a healthy politician.

(32) Joan is an old philosopher if and only if Joan is an old politician.

(33) Joan is a short philosopher if and only if Joan is a short politician.

In the preceding paragraphs our discussion has centred on those measure adjectives that denote single, measurable properties. However, there are many others which do not seem to denote single, objectively measurable properties. Examples of such adjectives are:

big; clean; healthy; large; strong;

Linguists differ over the treatment of such adjectives. For instance, McCawley (1981, p.187) has argued that the establishment of a fixed numerical scale for properties such as *healthiness* is not necessary, and that a semantic analysis of such adjectives should accommodate "subjective" scales and rankings for these properties. On the other hand, Klein (1980, 1982) attacks such assertions and suggests that they merely reveal the "bankruptcy" of the underlying semantic theories (1982, p.114). Siegel (1979) and Kamp (1975) have argued that the situation and linguistic context provide the information needed to interpret such adjectives.

I do not have a good solution for this problem. However, if it is agreed that lay judgements of such properties as *healthiness* are never very precise anyway, then one might be able to make a case for a representation that is based on some subjective, numerical scale permitting only very conservative estimates of precision. In the case of adjectives such as *big*, it often appears that context of use allows one to interpret them as referring to some specific, measurable attribute, such as *height*, or *volume*.[7]

In the past, AI researchers have tended to either ignore measure adjectives, or else to treat them just like absolute adjectives (see for example, D.B. McDonald 1982). Recently, Nishida (1983) has attempted to handle measure adjectives as part of his work on English-Japanese machine translation. Nishida treats measure adjectives as two-place predicates which require the explicit specification of a comparison class. For example, in his scheme *a big animal* would be represented as

$$\exists x\ [animal(x) \wedge big(x, animal)]$$ [8]

where the predicate $big(x,y)$ is true if and only if an individual denoted by x is "big *as* y", meaning that the bigness of x depends on y.

While Nishida recognizes the need to associate a comparison class with each use of a measure adjective, his analysis is unsatisfactory on two counts. Firstly, he seems to always take the comparison class as being the denotation of the complement of the adjective. We shall argue, in later sections, that this is justifiable only when the attribute denoted by the adjective is salient to the concept denoted by the complement. If this is not the case, then the appropriate comparison class must be inferred using real-world knowledge. For example, *tall basketball player* would probably be interpreted as "tall for a basketball player", but a *tall pastry chef* would more likely be tall for an adult.

Secondly, Nishida's analysis does not seem to adequately distinguish between intensional adjectives, such as *good*, and measure adjectives, such as *tall*. The meanings of the former depend upon the meanings of their complements, whereas the interpretation of the latter requires determination of the appropriate comparison classes, using the complements and real-world knowledge. However, the meanings of measure adjectives do not normally vary with different complements.

---

[7] Carey (1982) cites a number of studies which show that children learn adjective pairs such as *big/little* earlier than pairs like *tall/short*, *high/low*, and *thick/thin*. Some of these studies also seem to indicate that during the initial phases of acquisition of such adjectives, children seem to hypothesize adjectives such as *tall* and *high* to be synonymous with *big*, whereas adjectives like *short* and *thin* are treated as being identical to *little*. However, upon further work of her own, Carey concluded that individual children do not judge adjectives like *tall* to be synonymous with *big* indiscriminately, but rather only do so in certain kinds of tasks and not in others. These results are consistent with our treatment of adjectives like *big*; it would be interesting to test whether children systematically conflate only *tall* and *big* in certain tasks, conflate only *thick* with *big* in certain other, distinguished tasks, and so on.

[8] For expository purposes, I have slightly altered Nishida's original notation.

2.1.2.4. Role Adjectives

Many adjectives are derived from underlying verbs by means of affixation. One especially productive way to generate such adjectives is by adding a suffix to a base verb. There seem to be a fairly small number of such suffixes in English, with some of the most productive forms being these (Quirk et al., pp.994-1008):

-ive; -ed; -able; -ing

The following are some examples of adjectives that have been formed by adding one of these suffixes to a verb:

attractive	confused	admirable	enticing
connective	excited	desirable	frightening
impressive	pleased	lovable	insulting
persuasive	related	readable	pleasing
seductive	tired	thinkable	revealing

Such adjectives seem to refer to particular case roles of their underlying verbs, and so we will refer to them as ROLE ADJECTIVES.[9] Moreover, the case role which a particular adjective refers to seems to be determined by the suffix added to the verb. With regard to the particular suffixes mentioned above, we note the following characteristic relationships:

SUFFIX	CASE ROLE IT REFERS TO
-ive	AGENT
-ed	PATIENT
-able	PATIENT
-ing	INSTRUMENT

Thus, we would take *attractive* to refer to the AGENT role of an *attract* action, and *pleasing* to refer to the INSTRUMENT role of a *please* action.

We must be careful to distinguish role adjectives ending in *-ed* and *-ing* from participle forms of verbs, which often take the same lexical form. For example, an *escaped prisoner* is a prisoner who has escaped, not a prisoner who is the PATIENT of an *escape* action. Similarly, a *departing guest* is a guest who is in the process of departing, rather than the INSTRUMENT of a *depart* action (examples from Quirk et al., p.243).

2.1.3. Classifying Modification Relationships

2.1.3.1. Introduction

Having discussed the semantic properties of different classes of adjective, we now examine the various kinds of modification relationship that can hold between a single adjective and its complement, which is usually a noun. We shall see that in order to correctly interpret real-world data, it is necessary to carry out a detailed analysis of the

[9] We thus follow Finin (1980), who describes a similar class of nominalized verbs, and terms them ROLE NOMINALS.

kinds of modification relationship that adjectives can enter into.

2.1.3.2. Direct Modification

The simplest modification relationship involves the attribute denoted by an adjective being directly associated with the class of objects denoted by its complement. We call this DIRECT MODIFICATION.

Intensional adjectives, whose meanings depend on the meanings of their complements, seem to always be associated with direct modification. Thus, *former actress* would normally describe a person who was, but is no longer, an actress. On the other hand, there is no natural interpretation for a noun phrase such as *former black actress*, which seems to be describing a woman who has undergone a change in skin colour, rather than, say, a former actress who happens to be black.

In the case of absolute adjectives, we will say that direct modification holds only when the complement denotes the most general class of objects to which the adjective can apply. Thus, if we assume that the absolute adjective *dead* applies only to animals, then we would take the noun phrase *dead animal* to exhibit direct modification, but would not do so in the case of *dead elephant*;[10] the latter would be SUBCLASS MODIFICATION, to be discussed in the next section.

With measure adjectives, direct modification can occur when the attribute denoted by the adjective is salient to the class of objects denoted by its complement. For example, if we assume that basketball players tend to be taller than other people, then we would regard the phrase *tall basketball player* as exhibiting direct modification. On the other hand, assuming that height is not a distinguishing feature of the population of schoolteachers, *tall schoolteacher* would not exhibit direct modification.

Earlier we defined role adjectives to be those which refer to a case role of an underlying verb. The use of such an adjective usually involves filling the appropriate case of this underlying verb with the denotation of the complement of the adjective. We will say that direct modification is exhibited whenever the adjective's complement denotes the most general object capable of serving the indicated role in the action denoted by the verb. Thus, we would take *pleased person* to illustrate direct modification, assuming that a person, but no other entity, is capable of serving as the PATIENT of *please* actions. On the other hand, *pleased man* would not exhibit direct modification under our account.[11]

2.1.3.3. Subclass Modification

A more complex form of modification than the one discussed in the previous section occurs when the attribute denoted by an adjective is associated with a more general class of objects than those denoted by its complement. We call this SUBCLASS MODIFICATION.

---

[10] It might be argued that we are being excessively selective in not allowing phrases like *dead elephant* to be among those that can be described as exhibiting direct modification. We concede that there are probably no strong linguistic arguments favouring the position that we have adopted. However, from the standpoint of knowledge representation, we can readily justify our decision on the grounds that it permits us to economically represent the denotations of noun phrases (see Hayes 1977b for a similar argument).

[11] The arguments here are essentially the same as in the case of absolute adjectives, which we discussed earlier in this sec-

Most instances of modification by absolute adjectives involve subclass modification. This happens whenever the adjective is generally applicable to different kinds of objects, above and beyond those denoted by its complement. Thus, we could say that a *pink elephant* is an elephant and also a pink object (or animal, or mammal, or proboscidean, etc.). Similarly, a *drunk salesman* is a drunk person whose occupation is salesman.

With measure adjectives, subclass modification occurs when the attribute denoted by the adjective is not a distinguishing feature of the class of objects denoted by its complement, but is salient to a more general class of objects. For example, we would interpret *tall railway clerk* as describing a railway clerk who was tall for a person, rather than a person who was tall for a railway clerk (assuming, of course, that height is not a distinguishing feature of railway clerks). Similarly, *beautiful nurse* would be taken to denote the class of nurses who are beautiful for women, not beautiful for nurses.

In the case of role adjectives, we will say that subclass modification occurs when the complement does not denote the most general class of objects modifiable by the adjective. An example of this would be *annoyed customer*, which would be taken to denote someone who is a customer and an annoyed person. Other examples are *readable novel*, *attractive hostess*, *insulting gesture*.

2.1.3.4. Subpart Modification

Occasionally, it seems that the property denoted by an adjective, rather than merely failing to be salient to the class of objects that its complement denotes, simply cannot be attributed to them at all. Some of these cases can be resolved by noticing that the complement of the adjective denotes a physical or logical subpart of an object with which the attribute denoted by the adjective can indeed be associated. Then, the adjective is taken to apply to the "larger" object indirectly, through its subpart. We shall term this SUBPART MODIFICATION.[12] [13]

Subpart modification seems to be restricted to role adjectives, and, in a few instances, to measure adjectives. With measure adjectives, a few complex cases may be resolved by appealing to subpart modification. For example, a *clever idea* could be regarded as an idea that is due to a clever person or, perhaps more accurately, an idea which indicates that the person who conceived it was clever. Similar remarks could apply to *hungry mouth*, *angry look*, *intelligent eyes*.

With role adjectives, it might happen that the denotation of the adjective's complement cannot fill the indicated case role of the underlying verb, due to selectional restrictions. In some of these cases, the complement could denote a subpart of an object that would be able to fill the case role of the verb.

---

tion.

[12] Lakoff and Johnson (1983) point out that traditional rhetoricians used the term SYNECDOCHE to refer to the general linguistic device of substituting the part for the whole. We prefer to use our own term, subpart modification, since it lends itself to a natural interpretation in terms of PSN's PART-OF hierarchy.

[13] I have been unable to come up with any really compelling examples of what might be termed Whole-to-Part modification: in which the property denoted by an adjective is attributed to objects that are subparts of the objects that its complement denotes. The best examples that I have found are ones like *wooden screwdriver*, where "real-world" knowledge is used to

As an example, consider the phrase *satisfied voice*. It would be peculiar for a voice to be the PATIENT of *satisfy* actions. Indeed, the correct interpretation seems to be that the voice belongs to a person who is satisfied (i.e., a *satisfied person*). Moreover, the voice indicates that its owner is satisfied. Similarly, a *pleased look* could be interpreted as a look which is given by a person who is (characteristically) the PATIENT of *please* actions (i.e., a *pleased person*). Moreover, the look indicates that the person who gave it was pleased. Other examples are *satisfied smile* and *informed opinion*.

2.2. Multiple Prenominal Adjectives

In previous sections we have discussed the semantic classification of adjectives, and the varieties of modification relationship that could hold between a single adjective and its complement. We now turn to examine the problem of interpreting noun phrases with multiple prenominal adjectives.

As mentioned in Chapter 1, adjectives as well as nouns can serve as premodifiers of the heads of English noun phrases. There can be an arbitrary number of these modifiers in any given noun phrase, and they may combine with one another to form compounds which may themselves be subject to further modification. Noun compounds have been widely studied by both linguists and AI researchers. Notable among the linguists have been Lees (1960), Downing (1977), and Levi (1978). Among the AI researchers, the work of Brachman (1978), Finin (1980), and D.B. McDonald (1982) is prominent. However, very little attention has been paid to the study of multiple adjectives in prenominal position. The most comprehensive study that I am aware of was carried out by a linguist, Carl Bache, who examined a large corpus of data, consisting of some 4000 noun phrases having multiple prenominal adjectives, culled from sources as varied as novels, newspapers, popular magazines, and textbooks, in order to formulate rules governing the relative ordering of such adjectives. Although Bache's analysis was concerned mainly with syntax, and even though the goals of his work were very different from ours, we shall make use of many of the distinctions and classifications he proposed.

2.2.1. Syntactic Classification

In this section, we shall classify sequences of prenominal adjectives, on the basis of their syntactic form, into one of three different categories. Syntactic cues often serve to indicate the intended interpretation of noun phrases, and we shall make use of them whenever possible.

2.2.1.1. Broken Sequences

BROKEN SEQUENCES are prenominal sequences consisting of a number of adjectives which are all separated by commas and/or conjunctions. Some examples:[14]

cool, detached way (CSC 15)

---

infer that the object being described is a screwdriver whose handle is made out of wood.

[14] Data from Bache's study are labelled with a tag indicating the original source of the example, with reference to the tagged list of sources in Appendix A of this thesis, and the page number of the source.

tall, attractive gentleman (SCR 72)
red, smoking wreckage (KOFOCN 163)
ambitious and hard-working man (DE 4)
visual, auditory, gustatory or pain perception (MHFF 17)

2.2.1.2. Unbroken Sequences

UNBROKEN SEQUENCES consist of multiple prenominal adjectives which are not separated by commas or conjunctions. Some examples are:

deep emotional scars (Ms 16)
silly stupid negligible little lie (CSC 185)
long bony hand (CMMQ 113)
tall young blonde basketball player[15]
warm broad floorboards (PBASS 419)

2.2.1.3. Mixed Sequences

Sometimes one can have sequences in which some adjectives are separated by commas or conjunctions, but others are not. These are termed MIXED SEQUENCES. For example:

good-looking, fair young giant (CSC108)
glamorous, war-loving, dirty old men (VS5 17)
thin, pale young woman (Td 19)
ugly, little fat man (PBESS 130)
tall young, blonde woman

2.2.2. The Structure of Modification in Prenominal Sequences

Earlier, we discussed the various kinds of modification relationship that could hold between a single adjective and its complement. With multiple adjectives, a second, related problem presents itself: the problem of establishing the scope of each adjective in a sequence of adjectives. In the following sections we present various candidate structures for such modifier sequences and discuss the assignment of structure on the basis of syntactic cues in the noun phrase, the semantic classification of individual adjectives, and discourse pragmatics.

2.2.2.1. Flat Structures

The simplest pattern of modification arises when each adjective in a group separately modifies the head of the noun phrase. We shall say that such sequences have a FLAT

[15] Adapted from Levi (1978, p.248).

STRUCTURE.[16] In broken sequences, where the adjectives are separated by commas or conjunctions, only flat structures seem to be possible. Thus, a *tall, thin boy* is a boy who is tall for a boy and also thin for a boy, rather than being tall for a thin boy.

In many unbroken sequences, flat structures prevail also. For example, a *beautiful French wife*[17] is beautiful for a woman, not beautiful for a French woman, or for a French wife.

It seems that sequences of absolute adjectives always exhibit flat structures, since they denote absolute possession of an attribute. Thus, the only reasonable way to interpret *broken red dish*, for example, is in terms of separate modification of the noun by the two adjectives.

With measure adjectives, separate modification should prevail when the properties denoted by the adjectives are not dependent on one another. For example, *handsome young Italian doctor* (CSC 16) is not normally interpreted as referring to someone who is handsome for a young Italian doctor, nor to someone who is young for an Italian doctor. On the other hand, a *tall young boy* is likely to be tall for a young boy.

2.2.2.2. Nested Structures

In some cases, an adjective may modify a complex consisting of other adjectives and nouns to its right. We shall say that such phrases have a NESTED STRUCTURE. Siegel (1979) points out that this pattern seems to always hold when an intensional adjective precedes other adjectives in an unbroken sequence, as in *former young student*.

With measure adjectives, we may get a nested structure if the properties denoted by different adjectives interact with one another. For example, consider the noun phrase *small powerful engine*.[18] One reading of this would be: "an engine that is small for an engine and also powerful for an engine". However, if there is a positive correlation between the size of an engine and its power, then it would also be possible to interpret the phrase as "small for a powerful engine", since an engine that is small for a powerful engine might still be too large to be described as "small for an engine". In the absence of discourse information, both interpretations would seem to be equally plausible.

Now consider the phrase *large powerful engine*. Again, this could be assigned a flat structure and interpreted as "large for an engine and powerful for an engine". Alternatively, it could be assigned a nested structure and interpreted as "large for a powerful engine". In speech, the intended interpretation could be conveyed by prosodic cues. Thus, uttering

large POWerful engine

[16] Matthei (1982) studied children's interpretations of prenominal modifier sequences. He found that children tended to misinterpret complex noun phrases in a way which suggested that they were consistently assigning flat structures to the modifier sequences, even when such structures were inappropriate. On the basis of his study, he argued that children find it cognitively simpler to deal with flat structures than with nested ones.

[17] Bache (1978), p.58.

[18] Example from Gil (1983).

(where capital letters indicate stress) would indicate the first (flat) interpretation, whereas

LARge powerful engine

would indicate the second (nested) one. In writing, punctuation is often used to indicate the desired interpretation. Thus, the insertion of a comma or conjunction between the two adjectives would clearly indicate that the intended interpretation was a non-nested one.

Are we then to assume that the absence of commas and conjunctions is always indicative of modifier nesting? In the next section, we shall argue that the selective absence of punctuation within mixed sequences is, indeed, indicative of modifier sequence structure. However, in the case of totally unbroken sequences, it would be rash to draw such a conclusion. Examination of any reasonably-sized collection of non-technical writing will reveal that very often punctuation is entirely omitted within sequences of adjectives, even when the intended interpretation evidently corresponds to a non-nested modifier structure. In such cases, it appears that discourse and pragmatics often provide sufficient constraints to permit the intended interpretation to be inferred. For example, if it is known that powerful engines tend to be large for engines, then the most reasonable interpretation of *large powerful engine* is "large for a powerful engine", since the alternative interpretation, "powerful engine that is large for an engine", would carry redundant information, namely that the engine is large for an engine.

Our interpretation of these data is in accordance with Grice's (1975) maxims of Relation and Quantity, which stipulate that conversational contributions be relevant, and not more informative than is necessary. Further support is provided by the work of D.D. McDonald and Conklin (1982), who analyzed people's written descriptions of visual scenes, in an attempt to explain why some objects in a scene were consistently regarded as being more "salient" than others. One of their conclusions was that the "unexpectedness" of a feature is important in determining whether to include it in descriptions of objects or scenes. A similar point is made by Goodman (1983), who used transcripts of actual task-oriented dialogues to develop general criteria for relaxing the process of interpreting descriptions, in order to resolve failures in reference. He found that mention of unusual properties and features of objects in descriptions helped distinguish the referents from other objects in a scene.

2.2.2.3. Hybrid Structures

In certain cases, we get HYBRID STRUCTURES in which some adjectives exclusively modify the head of the noun phrase, and some modify a combination of other adjectives to their right and the head. Some examples:

ugly, little fat man (PBESS 130)
good-looking, fair young giant (CSC 108)
thin, pale young woman (Td 19)
glamorous, war-loving, dirty old men (VS5 17)

tall red-headed young child

Such patterns seem to occur with mixed sequences, where some of the adjectives are separated by commas or conjunctions, but others are not. Bache (1978) argues that the absence of commas and conjunctions in mixed sequences is symptomatic of structure, although this is not true in general for totally unbroken sequences, as discussed in the previous section. For example, we would interpret *tall red-headed young child* as "tall young child who is red-headed", with *tall* modifying *young child*, and *red-headed* modifying *child*.

CHAPTER 3

The Representation of Knowledge

3.1. Introduction

A semantic interpreter for noun phrases must have access to several different kinds of knowledge, including lexical information, knowledge about the structure of concepts, and knowledge of the relationships between different concepts. In addition, such a system must construct some sort of semantic representation for the concept associated with a noun phrase. In both cases, the system requires an appropriate formalism in which to represent knowledge.

According to one recent survey (Mylopoulos and Levesque, 1984), knowledge representation formalisms may be classified into three broad categories: LOGICAL schemes, PROCEDURAL schemes, and NETWORK schemes.[1] Logical representation schemes view knowledge bases as collections of logical formulae. On the other hand, procedural schemes see knowledge bases as being composed of active agents and processes. Yet another viewpoint is adopted by network schemes, which regard knowledge bases as collections of objects and relations defined over objects. Each type of scheme has its strengths and weaknesses, and Mylopoulos and Levesque (1984) provide a good summary of these.

Any one of the three types of representation schemes would be adequate, in principle, to represent the knowledge required by our semantic interpreter. However, Hirst (1983) has argued that network-like representations provide a firmer basis for semantic interpretation than the other schemes do. In particular, he suggests that network systems provide a deeper level of interpretation, allowing each syntactically well-formed component of an input sentence to be related to a semantic object, usually a piece of network structure. Another advantage of such a scheme, according to Hirst, is that it permits a more adequate account of reference, enabling synonymous words to be interpreted by the same semantic object. Finin (1980) and D.B. McDonald (1982), in their work on interpreting noun compounds, have provided similar arguments in favour of network representations. Accordingly, we shall adopt a network representation as the basis on which to design our semantic interpreter.

In the following sections, we shall examine one particular representational formalism, and discuss how it might be used to represent the concepts associated with noun phrases.

---

[1] Mylopoulos and Levesque (1984) also discuss HYBRID systems, that seek to combine features from a number of different kinds of knowledge representation schemes. The advent of such formalisms has been fairly recent, and many issues related to their design and use are still the subject of active research. The interested reader is referred to Brachman and Levesque (1982) and Levesque (1984) for a discussion of some of these issues.

3.2. The PSN Formalism

3.2.1. Basics

PSN (short for Procedural Semantic Networks) is a knowledge representation formalism developed at the University of Toronto (Levesque 1977, Levesque and Mylopoulos 1979). It is based on SEMANTIC NETWORKS (Quillian 1968), and provides facilities for creating and structuring large knowledge bases consisting of various kinds of OBJECTs. One of the basic types of object in a PSN knowledge base is a CLASS, which is used to represent a generic concept, or type. For example, we might define a class called US-PRESIDENT to represent the concept of a U.S. President. Particular concepts, or tokens, may be represented by PSN objects that are INSTANCES of classes. For example, we could define an object called, say, RICHARD-NIXON to represent one particular U.S. President.

Generic relationships are represented by means of RELATIONS, whose instances may be regarded as mappings from one object to another. For instance, we may define a relation called PREDECESSOR, having the class US-PRESIDENT as both its domain and range. Then we could assert that the PREDECESSOR relation holds between the objects RICHARD-NIXON and LYNDON-JOHNSON.

Most semantic network formalisms make use of a global interpreter to perform various operations on the network, which is regarded as a static data structure. In contrast, PSN adopts an object-oriented approach, providing classes and relations with a set of procedures, corresponding to the operations defined on objects of that type. The semantics of a class or relation are given by its behaviour under the operations defined on it, using the associated programs. There are four basic operations defined on classes:

(1) Assert that an object is an instance of a given class.

(2) Assert that an object is no longer an instance of a given class.

(3) Fetch all instances of a given class.

(4) Test whether an object is an instance of a given class.

Analogously, there are four operations on relations. These are:

(1) Assert that a given relation holds between two objects.

(2) Assert that a given relation no longer holds between two objects.

(3) Fetch all objects that are related to another by a given relation.

(4) Test whether a given relation holds between two objects.

3.2.2. Organization

Most semantic network schemes provide a number of organizational axes, or HIERARCHIES, for structuring a knowledge base. PSN offers three organizational axes, namely CLASSIFICATION, AGGREGATION, and GENERALIZATION.

Classification affords a means of associating objects with their generic types. PSN goes further than most other network schemes by considering every object to be a member of some class. In particular, it provides for METACLASSES: classes whose instances are other classes. For example, we could define a (meta-) class called PERSON-CLASS whose instances are classes of people, such as the class US-PRESIDENT. The provision of metaclasses allows PSN objects to be organized along what is called an INSTANCE HIERARCHY, in which objects that are "lower down" are instances of the objects "above" them (see figure 3.1). In this and later figures, instance-of links will be indicated by dotted arrows, PART-OF links by dashed arrows, and IS-A links by unbroken arrows.

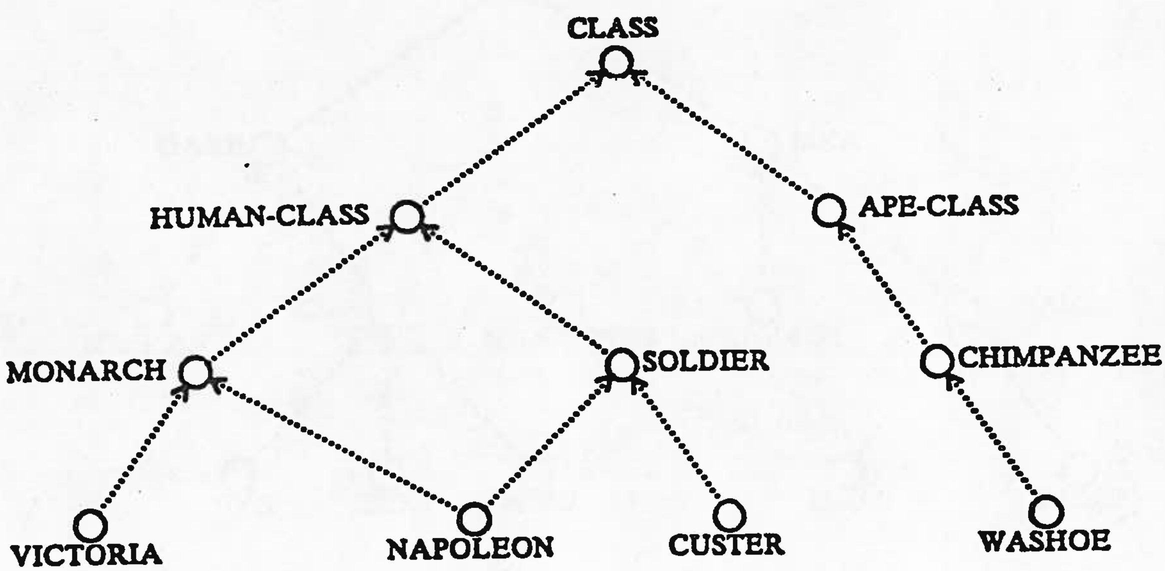

Figure 3.1 A small instance hierarchy

The aggregation axis can be viewed as relating objects to their (physical or logical) components, or PARTS. These, in turn, might be further decomposed into their own parts, giving rise to a PART-OF HIERARCHY. For example, a PERSON could have parts such as ARM, HEAD, etc. An ARM could have its own parts, such as HAND, FOREARM, ELBOW, etc. (see figure 3.2).

Generalization recursively associates classes with their own generic types (metaclasses), resulting in an IS-A HIERARCHY. For example, the class US-PRESIDENT could be a subclass of the class PRESIDENT, which itself might be a subclass of the class PERSON, and so on. Similarly, we might define a class UNDERGRADUATE which is a subclass of the class STUDENT, which, in turn, could be a subclass of a class called ACADEMIC (see figure 3.3). By default, each instance of a subclass is also taken to be an instance of all the superclasses. Thus, the object RICHARD-NIXON, being an instance of the class US-PRESIDENT, would also be regarded as instantiating the classes PRESIDENT and PERSON, unless otherwise specified. One important benefit of a

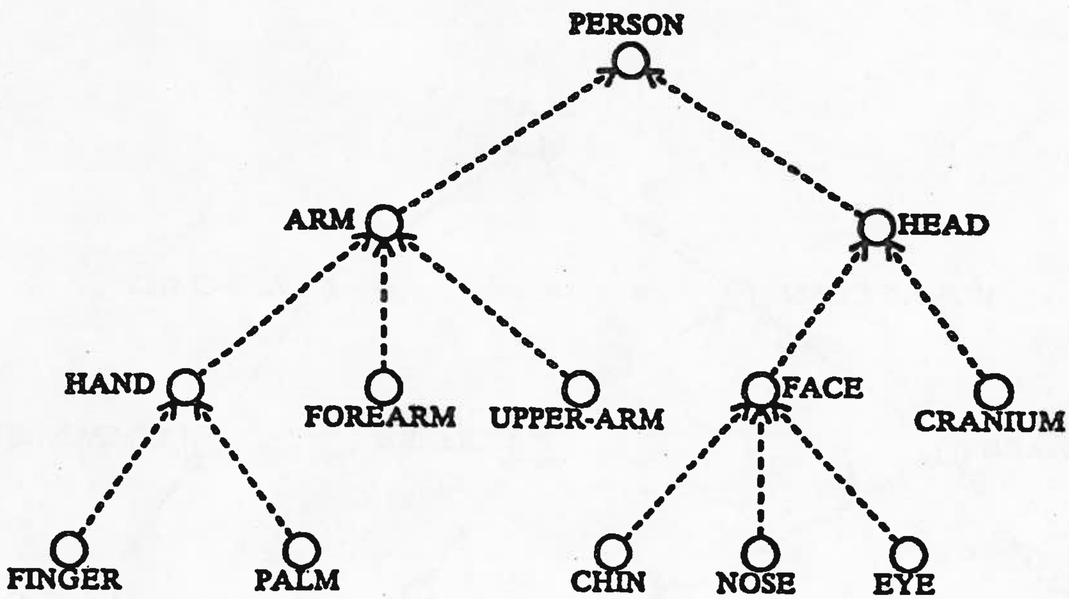

Figure 3.2 A small PART-OF hierarchy

generalization axis is that it obviates the need for having multiple representations of objects, by providing for classes to inherit, by default, properties that are associated with more general classes. Thus, the properties in question need only be explicitly represented once, yet they can be associated with many different classes. We will discuss the inheritance of properties in a later section.

3.2.3. Associating Properties with Classes

PSN provides a means for associating different kinds of properties with objects. Properties that are deemed, by the user, to be essential to the definition or meaning of an object are called its STRUCTURAL PROPERTIES. For example, sex and hair-colour might be regarded as structural properties of the class PERSON. The structural properties of a class are represented by objects called SLOTS, one slot for each property, and the collection of slots associated with a class is known as its STRUCTURE.

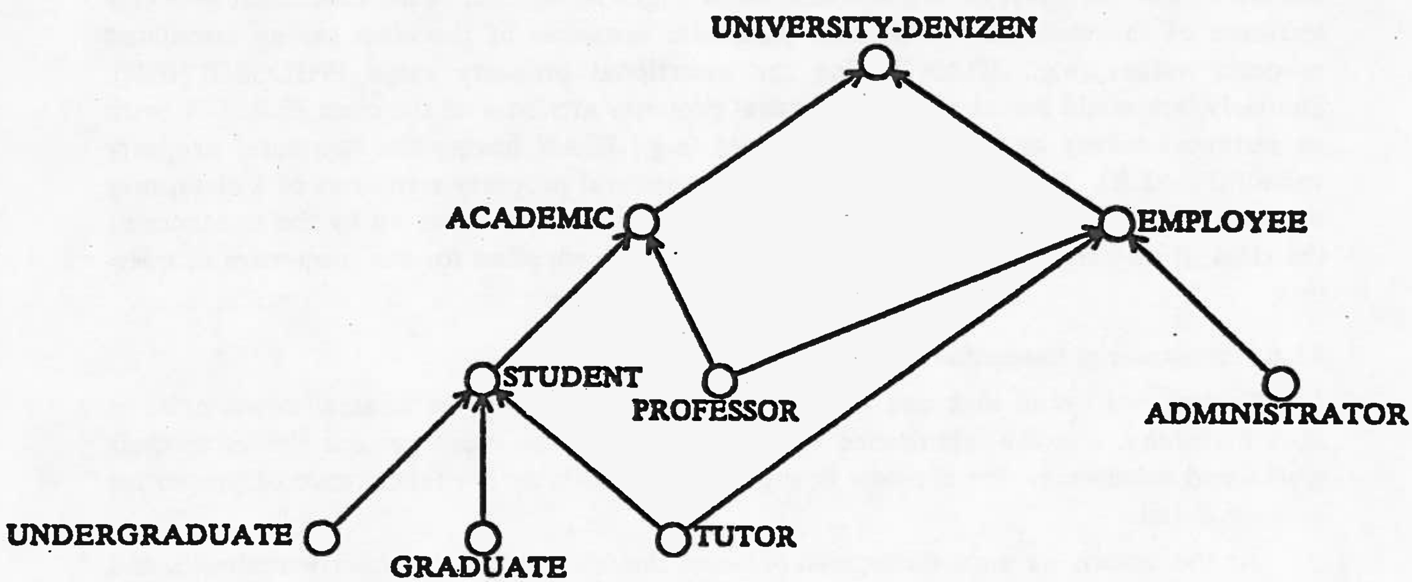

Figure 3.3 A small IS-A hierarchy

In addition to structural or defining properties, a PSN object may also be endowed with ASSERTIONAL PROPERTIES, which are considered to be incidental to the object. For example, having an occupation and a spouse might be viewed as assertional properties of the class PERSON, since individuals are, presumably, not required to have occupations or spouses in order to be considered persons. Assertional properties are represented by PSN relations, which we looked at earlier.

In the preceding discussion, we have described properties such as sex and occupation as being associated with classes, such as PERSON. However, the careful reader will have noticed that this is somewhat misleading, since properties such as sex are more accurately regarded as inhering in individuals, rather than in the class of all people. We shall now remove this imprecision, by taking account of another feature of PSN: the ability to distinguish between properties of classes and properties of their instances.

In PSN, particular properties of an object are called its PROPERTY VALUES, whereas properties of its instances are referred to as PROPERTY ATTRIBUTES of the class. The instances of a class may have property values corresponding to the property attributes of the class. For example, having an occupation might be defined as an assertional property attribute of the class PERSON, with particular instances of the class having associated property values (e.g., JOAN having the assertional property value PHILOSOPHER). Similarly, sex could be taken as a structural property attribute of the class PERSON, with its instances taking on corresponding values (e.g., JOAN having the structural property value FEMALE). Slots, which represent the structural property attributes of a class, may be given DEFAULT values. These default values may then be taken on by the instances of the class, if no other values are inherited or explicitly supplied for the properties in question.

3.2.4. Inheritance of Properties

Earlier, we noted that one of the important features of the generalization axis, or IS-A hierarchy, was the inheritance of properties from the more general classes to their specialized subclasses. We are now in a position to examine the inheritance of properties in more detail.

At the outset, we must distinguish between the inheritance of property attributes and the inheritance of property values. An object inherits attributes from all the classes that it is an instance of. Since instances of subclasses are implicitly considered to be instances of all the superclasses also, we can view property attributes as being inherited, between classes, down the IS-A hierarchy.

On the other hand, the inheritance of values cannot be cumulative, since a single value must be associated with each property of an object. The solution adopted in this case is for a subclass to inherit the property values of its immediate superclass[2], with the proviso that inherited values be pre-empted by any values that are explicitly supplied for the subclass. For example, the class RAVEN may be given a default COLOUR value of

[2] Notice that this does not address the case in which a class has more than one IS-A parent. We shall discuss this problem at some length in the next chapter.

BLACK, which would then be inherited by its subclass MALE-RAVEN. However, another subclass, say ALBINO-RAVEN, might be explicitly supplied with the value WHITE for its corresponding slot. In this case, the value BLACK inherited by the subclass would be pre-empted by the explicitly specified value WHITE.

It is also necessary to distinguish between structural and assertional properties when discussing the inheritance of property values. Since the structural properties of a class are definitional, its subclasses should inherit its structural property values. However, the same argument cannot be applied to the inheritance of assertional property values, since subclasses need not take on the incidental properties of their superclasses. Consequently, assertional property values are not inheritable.

Following Levesque and Mylopoulos (1979, p.111), we may summarize the rules governing inheritance as follows:

(1) Along the instance hierarchy,

(a) for each structural property attribute of a class, its instances must have a corresponding property value;

(b) for each assertional property attribute of a class, its instances can have zero or more corresponding property values;

(c) structural or assertional property values of a class are not inherited by its instances.

(2) Along the IS-A hierarchy,

(a) structural and assertional property attributes of a class are inherited by its IS-A children;

(b) each structural property value of a class is inherited, subject to pre-emption, by its IS-A children, provided that the IS-A children are instances of the metaclass having the corresponding property attribute;

(c) assertional property values of a class are not inherited by its subclasses.

3.3. PSN and Semantic Interpretation

As discussed in Chapter 1, we will assume that language users associate with each generic noun phrase some concept of what constitutes a typical entity of the kind referred to by the noun phrase. We shall now examine how we might represent such concepts using PSN.

A generic concept may be represented very simply in PSN, by defining a new class. For example, to represent the concept of a person, we may define a class called PERSON. We also need to associate with PERSON some information as to the characteristics of a typical person. For instance, a person has a height and an age, which are typically restricted to a certain range of values, and so on. At this point we are faced with a choice: we must decide to represent attributes such as height and age either as structural properties or as assertional properties of the class PERSON.

In order to resolve this question, we need to take account of exactly what we view the class PERSON to be representing. Since we want this class to represent the concept of a stereotypical person, we can take properties such as height and age to be structural rather than assertional. Thus, we would consider it part of the definition of a stereotypical person that her (or his) height lie within a certain range, and so on. It should be remarked that the adoption of such a stance with regard to characteristics of actual persons, as opposed to stereotypical ones, would be rather more problematic (see Putnam 1970, 1973 for arguments to this effect).

Having decided that properties such as height are to be regarded as structural, we now take up the question of how best to represent them. Let us assume that we have decided that the height of a typical person ranges over some fixed real-number interval. We would like a means to represent this interval and to associate its representation with the class PERSON. Conveniently, there happens to be a pre-defined PSN class, called INTERVAL, provided for representing such intervals, and having slots to stand for the upper and lower bounds of intervals. A particular interval, such as our stereotypical person's height range, may be represented by an object, call it PERSON-HEIGHT-RANGE, that is made an instance of the class INTERVAL, with appropriate values supplied for the slots representing the interval's bounds. Now we may associate PERSON-HEIGHT-RANGE with the class PERSON, as one of its structural property values.[3]

We have seen how we might associate individual properties, such as an age-range and a height-range, with generic concepts. However, it is often the case that such properties are inter-related. For instance, we might expect our stereotypical person's height to lie within a certain subrange, given that her age was similarly restricted. We can capture such dependencies in PSN by making use of INTERSLOT-CONSTRAINTS. These are expressions which can be attached to the slots defined by a class, and which serve to constrain the values that can fill each slot, in terms of the values taken on by the other slots to which the same interslot constraint has been attached. When the IS-A descendants of a class inherit property values from it, any associated interslot constraints are also passed down to them. As a simple example, consider the class INTERVAL, which represents real-number intervals, and defines slots corresponding to the lower and upper bounds of such intervals. For any instance of this class, we could define an interslot constraint, attaching it to both slots, which checks that the value supplied for the lower-bound of the interval is not greater than the value provided for the upper bound.

[3] We will assume that the class PERSON is an instance of a previously-defined class PERSON-CLASS, which has a slot HEIGHT-RANGE as the structural property attribute corresponding to the property value PERSON-HEIGHT-RANGE of PERSON.

All the properties we have examined so far have been DIMENSIONAL, in that they may take on values from a smoothly varying range of possible values. However, there are other properties that are better regarded as FEATURES, which an object either definitely possesses or else fails to possess. Such properties range over a small number of discrete values. For example, a person's hair may take on one of a small number of different colours. The ranges of such properties may be represented using the set facility of PSN. Each PSN set is a subclass of another class, called its TYPE-CLASS. The metaclass of a set is a subclass of the pre-defined class SET, as well as being an instance of another pre-defined class SET-CLASS. SET-CLASS defines a slot SET-TYPE, whose value is the type-class of the set. SET defines a slot CARDINALITY, whose value is a range of positive integers, giving the range-of-cardinality information to be associated with the set. For example, assume that there exists a class called COLOUR, whose instances are (the names of) colours. Then, we may define a class called COLOUR-SET-CLASS, a metaclass of sets, which will restrict its instances to have type-class COLOUR, and some given range-of-cardinality. COLOUR-SET-CLASS will be an instance of SET-CLASS, and a subclass of SET. Next, we may define a set HAIR-COLOUR, an instance of COLOUR-SET-CLASS, and a subclass of COLOUR. Various colours, such as WHITE, GREY, BLACK, GOLDEN, and BROWN, each an instance of COLOUR, may now be added to (made instances of) HAIR-COLOUR, subject to the cardinality restrictions placed on it.

CHAPTER 4

Designing a Semantic Interpreter for Adjectives

4.1. Introduction

In Chapter 2, we saw that a semantic interpreter for poly-adjectival noun phrases must take account of many different sources of information, including the semantic behaviour of adjectives, knowledge about the properties of entities referred to by common nouns, syntactic cues in the noun phrase, and contextual information. In Chapter 3, we examined a particular representation scheme and saw how it could be used to encode some of the knowledge that a semantic interpreter would require. In this chapter, we will show how to design a semantic interpreter that uses the techniques discussed in Chapter 3 to enable it to make use of the kinds of knowledge discussed in Chapter 2.

The basic idea is to supply our system with a PSN knowledge base that represents the concepts that we associate with common nouns. Then, upon being given a novel noun phrase consisting of a common noun preceded by adjectives, the system will make use of the knowledge base and linguistic knowledge to construct the representation of a new concept: one associated with the input noun phrase. The representation of this concept will then be added to the knowledge base, and appropriately related to the pre-existing contents. It will also be available for later use in constructing representations of other, yet more complex, concepts.

Before we proceed, a few remarks on terminology should be noted. We shall say that a common noun or a noun phrase REFERS TO or DENOTES (the concept of) a class of objects in the world, but is INTERPRETED BY a knowledge base object. This knowledge base object will also be said to REFER TO, or DENOTE, the same class of objects as does the noun. An adjective will be said to MODIFY a noun or part of a noun phrase, and will REFER TO, or BE ASSOCIATED WITH, some property or attribute of real-world objects. We will say that an adjective can APPLY TO real-world objects that possess the property associated with the adjective. For example, the common noun *woman* will be taken to refer to the class of women, and will be interpreted by the PSN object WOMAN, which will be taken to be coreferential with the noun. The adjective *tall*, which is associated with the property of height, may modify the noun *woman*, since it can apply to women (in virtue of the fact that they possess the property of having height).

We shall begin by considering the simple case of a single adjective modifying a noun. After this, we shall turn our attention to the problem of interpreting noun phrases containing multiple prenominal adjectives.

4.2. The Interpretation of a Single Adjective Modifying a Noun

We saw in Chapter 2 that there were three general forms of modification that could be exhibited by an adjective-noun pair. We shall take up each of these, in turn.

Before we go on, however, it should be remarked that we shall make no attempt to interpret intensional adjectives, since we have no adequate computational account of their semantics (see Section 2.2.2.1).[1] As a result of this restriction, we will be able to simplify the design of our interpreter considerably. In particular, by excluding intensional adjectives, we are assured that it will always be possible to interpret noun phrases COMPOSITIONALLY, in the sense that the concept associated with an adjective-noun pair will always be a more specialized, subconcept of that associated with the noun.[2]

4.2.1. Direct Modification

Recall that direct modification involves the attribute denoted by the adjective being somehow immediately relevant to the concept associated with the noun. In such a case, a new PSN object will be created, in order to represent the specialized concept associated with the adjective-noun pair. This new object will be added to the knowledge base as an IS-A child of the object associated with the noun; it will also be made an instance of the metaclass of its new IS-A parent. This will ensure that the new object can take on values for the properties defined by the metaclass, and that it inherits property attributes and structural property values from the object associated with the noun.

For example, consider the noun phrase *blonde person* to exhibit direct modification, and assume that the initial knowledge base contains the object PERSON, an instance of the metaclass PERSON-CLASS. Moreover, let PERSON-CLASS have the structural property attributes (slots) AGE-RANGE and HAIR-COLOUR, indicating that classes of (stereotypical) persons may be distinguished on the basis of these properties, and let PERSON take on a particular real-number interval, PERSON-AGE-RANGE, as the property value corresponding to the AGE-RANGE attribute of its metaclass. Now, as described earlier, our interpreter will create a new class called BLONDE-PERSON, as the interpretation of the noun phrase *blonde person*, and will make it an IS-A child of the class PERSON, as well as an instance of the metaclass PERSON-CLASS. Since BLONDE-PERSON is an instance of PERSON-CLASS, it follows from the rules governing inheritance (see Section 3.2.4) that we may provide it with explicit values for the slots HAIR-COLOUR and AGE-RANGE. We would presumably want to restrict the value of HAIR-COLOUR for BLONDE-PERSON, but would not want to assert any value for AGE-RANGE, since a person's life-span is not normally correlated with her hair-colour. It would be desirable, in this case, to have the AGE-RANGE value of BLONDE-PERSON default to the value

---

[1] See Chapter 5 for discussion of how the coverage of our system might potentially be extended to (at least some) intensional adjectives also.

[2] Compositionality could not be maintained if we failed to exclude those intensional adjectives that did not satisfy the SUBPROPERTY ASSUMPTION (Montague 1970): that all objects in the extension of an adjective-noun pair also lie in the extension of the noun. For example, the adjective *red* satisfies the subproperty assumption, since every *red car* is a car, every *red house* a house, and so on, for every common noun that it could modify. However, things are different in the case of certain intensional adjectives, such as *alleged*, *possible*, *fake*, *reputed*, and *ostensible*, which do not satisfy the subproperty assumption. Thus, not every *alleged murderer* need actually be a murderer, every *possible winner* need not win, *fake guns* are not guns,

associated with PERSON. In fact, this is exactly what does happen, since the inheritance rules provide for structural property values to be inherited down the IS-A hierarchy, subject to pre-emption. Thus, by default, the structural property value PERSON-AGE-RANGE of PERSON will be inherited by its IS-A child BLONDE-PERSON.

We saw, above, that it may be necessary for a new class to assert values for some of the slots it inherits. This is an instance of a general problem that our interpreter faces: how to assign an internal structure to the new objects it creates. The form to be taken by this structure, in the case of interpreting an adjective-noun pair, will depend on the semantic class to which the adjective belongs. We shall now discuss the creation of this internal structure for each of the semantic classes of adjective recognized by our interpreter.

4.2.1.1. Direct Modification by an Absolute Adjective

In Chapter 2, we said that direct modification held between an absolute adjective and a noun only when the noun referred to the most general class of objects to which the adjective could apply. In order to enable our interpreter to recognize such cases, we will associate with each absolute adjective in its vocabulary the most general object in the knowledge base to whose referent the adjective could apply.[3] Then, a particular absolute adjective-noun pair will be taken to exhibit direct modification exactly when the noun is interpreted by the object thus associated with the adjective. For example, if the adjective *blonde* applies most generally to people, then we would associate it with the object PERSON in the knowledge base, and would be able to recognize *blonde person* as an instance of direct modification, since PERSON is also the interpretation of the noun *person*.

Once an instance of direct modification is recognized, an appropriate internal structure must be created for the PSN object representing the new concept. Now, absolute adjectives do not seem to apply to objects to various degrees; rather, such an adjective either applies to an object to an object, or else it fails to apply. For example, animals can't normally be described as "somewhat four-legged" or "slightly dead". In light of this, we shall assume that the properties associated with such adjectives range over discrete values, and that an object is described by such an adjective exactly when it takes on one or more of a fixed subset of values for the associated property. For example, we might assume that a person's hair-colour is restricted to one of a small number of colours, such as *auburn*, *black*, *brown*, *golden*, *grey*, *orange*, *red*, or *white*. Then, a *blonde person* might be one whose hair-colour was further restricted to be either *auburn* or *golden*.

In order to implement this scheme, we may make use of the set facility provided by PSN. We saw, in Chapter 3, how sets could be used to represent property values of classes. In particular, we will use sets as the values of property attributes that are associated with absolute adjectives. For example, the class PERSON might have a set, say PERSON-HAIR-COLOUR, as the property value corresponding to the HAIR-COLOUR attribute of its metaclass. The elements of this set will be instances of the class

and so on. It seems to be the case that all non-intensional adjectives satisfy the subproperty assumption.

[3] A PSN object A is said to be "more general" than another object B if A is located closer to the root of the IS-A (general-

COLOUR. Suppose PERSON-HAIR-COLOUR has the elements AUBURN, BLACK, BROWN, GOLDEN, GREY, ORANGE, RED, and WHITE. Then, in creating a new class BLONDE-PERSON, as the interpretation of the noun phrase *blonde person*, our interpreter will create a new set, BLONDE-PERSON-HAIR-COLOUR, a subset of PERSON-HAIR-COLOUR, having two elements: AUBURN and GOLDEN. This new set will be the value of the HAIR-COLOUR attribute for the class BLONDE-PERSON (see figure 4.1). In this and later figures, we will use a solid, double-headed, labelled arrow to indicate a slot-filler relationship; the arrow will be labelled with the name of the slot, and will be directed from the class that defines the slot towards the object that fills it. Recall from Chapter 3 that we use solid, single-headed, unlabelled arrows to represent IS-A links, dashed, single-headed arrows to indicate PART-OF links, and dotted, single-headed arrows to represent instance-of links. For example, in figure 4.1 the arrow labelled HAIR-COLOUR pointing from PERSON to PERSON-HAIR-COLOUR indicates that PERSON-HAIR-COLOUR fills the HAIR-COLOUR slot of PERSON.

4.2.1.2. Direct Modification by a Role Adjective

As with absolute adjectives, role adjectives exhibit direct modification when they are paired with a noun that is associated with the most general class of objects to which the

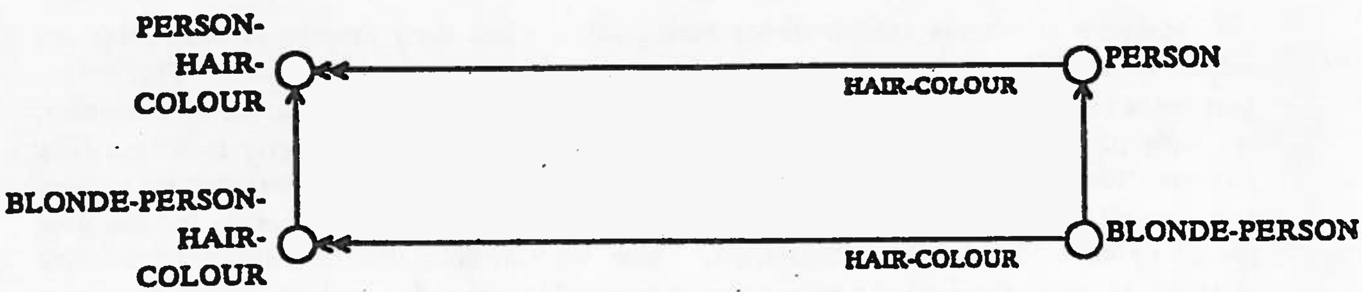

Figure 4.1 The result of interpreting *blonde person*.

ization) hierarchy than B is.

adjective would apply. Instances of direct modification in role adjective-noun pairs may be recognized using the technique discussed in the previous section.

Recall that a role adjective is associated with a particular case role of an underlying verb. When such an adjective modifies a noun, we take the noun to be filling the indicated case role of the verb underlying the adjective. In order to account for this, we shall define a PSN class called ACTION-CLASS, which has as instances various kinds of actions, which are intended to represent verb concepts. One particular instance of ACTION-CLASS will be the class ACTION, representing the class of which all particular actions are instances. A particular verb, such as *please*, will be represented by a class, say PLEASE, which is an instance of ACTION-CLASS, and a subclass of ACTION. In order to represent cases of verbs, we will have ACTION-CLASS define the slots AGENT, PATIENT, and INSTRUMENT, to be inherited by its instances, such as PLEASE, which may then assign values to them.[4]

Upon being given a role adjective-noun pair that exhibits direct modification, our interpreter will create a new class as its interpretation, making it a subclass of the class that serves as the interpretation of the noun. The new class will be related to the action associated with the adjective's underlying verb by filling the appropriate case role of that action. For example, assume that the noun phrase *pleased person* exhibits direct modification. Our interpreter will create a new instance of PERSON-CLASS, called PLEASED-PERSON, making it a subclass of the class PERSON. Among the linguistic information associated with the adjective *pleased* is the fact that it refers to the patient role of the verb *please*. Accordingly, the interpreter will designate PLEASED-PERSON to be the property value for the PATIENT attribute of the class PLEASE (see figure 4.2).

4.2.1.3. Direct Modification by a Measure Adjective

Measure adjectives exhibit direct modification when their associated properties are salient to the concepts associated with the nouns that these adjectives modify. This means that in order for an interpreter to be able to recognize instances of direct modification, we must provide it with some way of recognizing when a given property is salient to a concept. Now, in our representational system generic concepts are represented by classes, and properties of these concepts by the property values that a class takes on for the property attributes defined by its metaclass. Then, we may formalize our notion of salience in terms of restrictions on the values that it is possible for a class to take on for particular property attributes, as compared to other, related classes. Specifically, we will take a property to be salient to a concept exactly when the PSN class, call it C, representing that concept restricts the range of possible values of the relevant property attribute to a greater degree than do any of C's IS-A parents.[5]

---

[4] Particular actions, such as PLEASE, may use (intra-slot) constraints to enforce selectional restrictions on case roles. For example, PLEASE may constrain its PATIENT slot to be filled by instances of PERSON-CLASS.

[5] It may seem that, in talking about properties being considered salient based on the contents of our knowledge base, we have got things exactly backwards. Surely, the attribution of salience to properties is logically prior to the construction of representational structure, the latter being built so as to accord with our intuitions about the former. Although this point is well-taken, we should bear in mind the fact that our interpreter has direct access only to the structures in its knowledge base, and not to our mental concepts. Therefore, it is up to us, as designers of the system, to ensure that we provide it

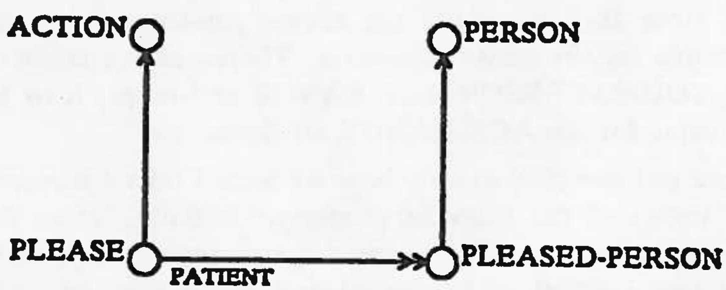

Figure 4.2 The result of interpreting *pleased person*

Since measure adjectives seem to apply to objects to various degrees, we shall represent their associated properties by slots that range over real-number intervals. For example, we could define the slot AGE-RANGE of the class PERSON-CLASS to range over real-number intervals. Then, instances of PERSON-CLASS, such as the class PERSON, could associate particular real-number intervals with the AGE-RANGE slot defined by their metaclass. For instance, let PERSON take on a specific interval PERSON-AGE-RANGE, with lower bound 0 and upper bound 80, where we take the numbers to represent age, in years.

Now consider the interpretation of the noun phrase *young person*, assuming that the measure adjective *young* directly modifies the noun *person*.[6] We would want to create a new class YOUNG-PERSON to serve as the interpretation of the noun phrase, and make it an IS-A child of the class PERSON, which interprets the noun, and an instance of the metaclass PERSON-CLASS. If we did nothing further, YOUNG-PERSON would be able

with an initial knowledge base that accurately captures our intuitions about the world, and that continues to do so under the kinds of transformations that we allow it to undergo. Assuming that we can maintain this correspondence, we will relax our phraseology somewhat, and speak as though the representations provide access to the intuitions underlying them.

Before we leave this point, it should be mentioned that we are skirting a number of deep philosophical questions concerning the epistemological status of representations, and the connection between language, mind, and the world. We make no attempt to address these issues in this work; the interested reader is referred to Dretske (1981, 1983), and Barwise and Perry (1983, 1984) for some promising new approaches to these problems.

[6] This would certainly be the case if the class PERSON, which interprets the noun, had no IS-A parents at all.

to take on values for the property attributes, including AGE-RANGE, that are defined by PERSON-CLASS. In addition, YOUNG-PERSON would inherit structural property values, such as PERSON-AGE-RANGE, from its superclass PERSON. However, it is clear that we would want YOUNG-PERSON to be more restrictive in the range of values it allows for the AGE-RANGE attribute than is PERSON. Fortunately, PSN provides us with a way to ensure this, since the rules governing inheritance state that property values that are inherited from IS-A ancestors are always pre-empted by any values that one might explicitly provide for the class in question. Therefore, we could define a new, more restrictive interval YOUNG-PERSON-AGE-RANGE and supply it to YOUNG-PERSON, as the appropriate value for the AGE-RANGE attribute.

So far, we have not specified exactly how we would take a measure adjective to restrict the range of values of the relevant property.[7] Initially, let us distinguish between those measure adjectives, such as *old*, that restrict a property to a subrange which is at the "upper" end of the base interval, and those, such as *young*, that indicate a subrange at the "lower" end. Should one let the subranges defined by such pairs of adjectives to overlap, or should they be strictly disjoint? For instance, could a person be, simultaneously, both young and old, both short and tall, and so on? It seems that the answer to this question is "no".[8] If this is the case, then it follows that an antonymous pair of measure adjectives serves to divide an interval into at least two disjoint subintervals, each corresponding to the range of property values staked out by one or the other of the two adjectives. There seems to be no reason, in general, to distinguish between the sizes of these subintervals, so we shall assume that they are of equal extent. Moreover, we will take the midpoint of the base interval as a natural dividing point, serving to separate the two subintervals.[9] In our example, then, YOUNG-PERSON-AGE-RANGE would simply be defined as an interval with the same lower bound as PERSON-AGE-RANGE, but whose upper bound is the midpoint of PERSON-AGE-RANGE.

In Chapter 3, we pointed out that properties such as age, height, and weight seem to be interrelated to some extent, and showed how interslot constraints could be used to capture these dependencies. Now, if a newly-created class places restrictions on the range of values that could be taken on by one of these properties, it is reasonable to expect that this might affect the values that it could take on for other, related properties. For example, if we know roughly how people's height varies with their age, then restricting the age to be, say, under 12 years would lead us to expect that the height would also tend to lie in a certain subrange, whose bounds we could estimate. On the other hand, if we decided to restrict the age of some subclass of adults, say U.S. presidents, to be more than 35 years, it would be reasonable to conclude that we would not need any further restrictions on

---

[7] In all cases, however, we would require that the restriction be to a proper subrange of the property values permitted to the superclass. This is necessary, on representational grounds, since PSN subclasses must always be specializations of their superclasses. On linguistic grounds, this is sufficient, since we are only dealing with adjectives that satisfy the subproperty assumption.

[8] Support for this view is provided by Kiefer (1978, p.142ff).

[9] We could have chosen any number of other ways of dividing up the base interval, subject only to the constraint that the two distinguished subintervals be of equal size. For our purposes, however, the particular subdivision chosen is not of great importance, as long as "reasonable" ranges of property values are associated with the adjectives in question.

the height-range than those already enforced on adults. The knowledge required to support such common-sense reasoning can be captured in interslot constraints. For example, a constraint on the slot HEIGHT-RANGE, for PERSON, might specify that the values taken on for the slot by subclasses of PERSON be restricted to certain subranges, depending upon the values they take on for the AGE-RANGE slot (see figure 4.3).

4.2.2. Subclass Modification

Subclass modification takes place in an adjective-noun pair when the property associated with the adjective is not regarded as being salient to the concept associated with the noun. We shall now discuss how a semantic interpreter might deal with subclass modification for each of our three classes of adjective in turn.

4.2.2.1. Subclass Modification by an Absolute Adjective

An absolute adjective-noun pair exhibits subclass modification when the adjective is applicable to a more general class of objects than the one that the noun refers to. Given

```
IF 0 < AGE-RANGE ≤ 1 YR. THEN
                    1.0 < HEIGHT-RANGE ≤ 2.0 FT.
ELSE
  IF 1 < AGE-RANGE ≤ 5 YR. THEN
                    1.5 < HEIGHT-RANGE ≤ 3.0 FT.
  ELSE
    IF 5 < AGE-RANGE ≤ 10 YR. THEN
                    2.5 < HEIGHT-RANGE ≤ 4.0 FT.
    ELSE
      IF 10 < AGE-RANGE ≤ 15 YR. THEN
                    3.0 < HEIGHT-RANGE ≤ 6.0 FT.
      ELSE
        IF 15 < AGE-RANGE ≤ 20 YR. THEN
                    3.5 < HEIGHT-RANGE ≤ 7.0 FT.
        ELSE
          IF 20 < AGE-RANGE ≤ 25 YR. THEN
                    4.0 < HEIGHT-RANGE ≤ 7.0 FT.
          ELSE
            IF 25 < AGE-RANGE ≤ 80 YR. THEN
                    4.0 < HEIGHT-RANGE ≤ 7.0 FT.
```

Figure 4.3 A simple interslot constraint for HEIGHT-RANGE such an adjective-noun pair, our interpreter would map the noun onto the knowledge base object, call it N, that interprets the noun. Similarly, the adjective would be associated with a knowledge base object, call it G, whose referent is the most general class of object to which the adjective applies. Then, subclass modification is recognized to be taking place exactly when N is a proper subclass of G. This is determined by performing a breadth-first search of the IS-A hierarchy above N, since in general a class may have many IS-A parents, and there may be more than one path between a pair of nodes in the hierarchy. Using breadth-first search ensures that a shortest path will be found between N and G, provided that a path does exist between them. The reason for searching up from N, rather than down from G, is that the hierarchy tends to have more of a "tree-like" structure than a "graph-like" structure; thus, while nodes may have multiple parents as well as multiple children, they typically have more children than they do parents. In light of this, it should require less search on average to scan up the hierarchy, looking for a particular node, than it would to scan downwards.[10]

Once an instance of subclass modification is recognized, the interpreter must construct an appropriate interpretation for the noun phrase. The manner in which this is done may be best illustrated by considering an example. Suppose that the absolute adjective *blonde* applies, most generally, to persons, and that we are faced with the task of interpreting the noun phrase *blonde woman*. If we examine the simple knowledge base shown in figure 4.4, we see that the class WOMAN, which is the object associated with the noun, is indeed a subclass of PERSON, the object associated most generally with the adjective. In constructing an interpretation for the noun phrase, we want to capture the

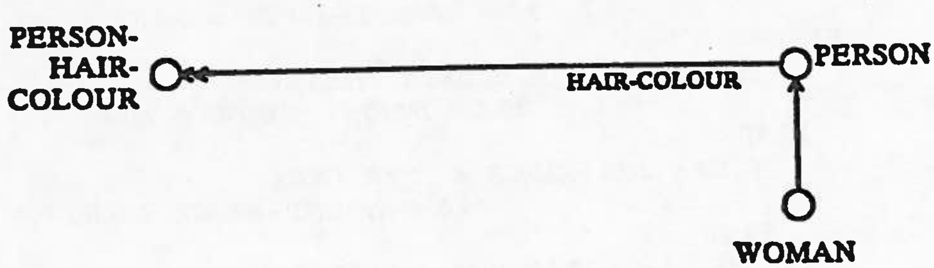

Figure 4.4 Initial state of knowledge base

[10] If it is the case neither that G and N are the same object, nor that G is an IS-A ancestor of N, then it will be concluded that there is no literal interpretation of the noun phrase.

intuition that a blonde woman is someone who is both a woman and a blonde person. We may do this by recursively invoking the interpreter to first construct an interpretation for the noun phrase *blonde person*, leaving the knowledge base in the state shown in figure 4.5.[11] Then the interpreter will create a new instance of PERSON-CLASS called BLONDE-WOMAN, as the interpretation of the original noun phrase *blonde woman*, making it an IS-A subclass of WOMAN as well as of the newly-created class BLONDE-PERSON (see figure 4.6).

It will not be necessary to create any internal structure for BLONDE-WOMAN, since it will automatically inherit the appropriate structure from its metaclass PERSON-CLASS. However, there may be a need for the interpreter to mediate in the inheritance of property values from the IS-A parents, since different superclasses may take on different values for certain property attributes. For example, BLONDE-PERSON and WOMAN will presumably take on different values for attributes like HAIR-COLOUR, AGE-RANGE, and HEIGHT-RANGE. In such a case, the interpreter must find a maximal consistent set of property values for the attributes in question, and provide these to the new subclass. Thus, for HAIR-COLOUR, it would find that BLONDE-PERSON had the value BLONDE-PERSON-HAIR-COLOUR, whereas WOMAN took on the value PERSON-HAIR-COLOUR. Now, since BLONDE-PERSON-HAIR-COLOUR is a subclass of PERSON-HAIR-COLOUR, it should be the one inherited by BLONDE-WOMAN.

So far, we have dealt with the case in which an absolute adjective modifies a noun that is interpreted by an immediate subclass N of the most general object G associated

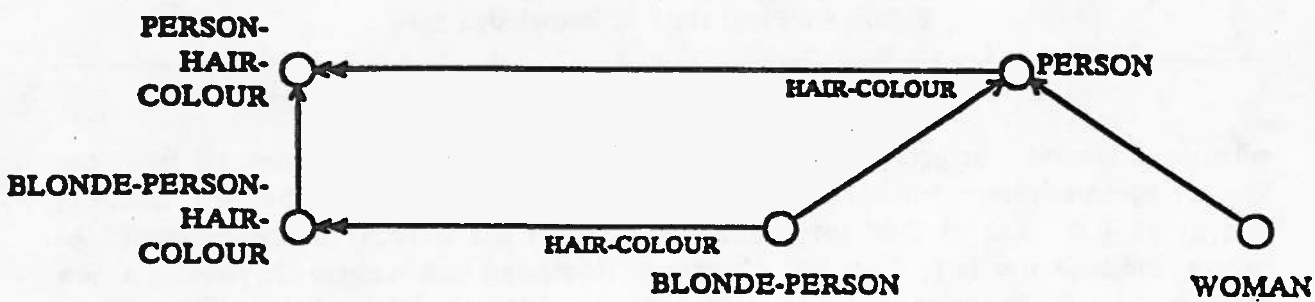

Figure 4.5 Intermediate state of knowledge base

[11] The noun phrase *blonde person* exhibits direct modification, and may be interpreted using the techniques discussed in the previous section.

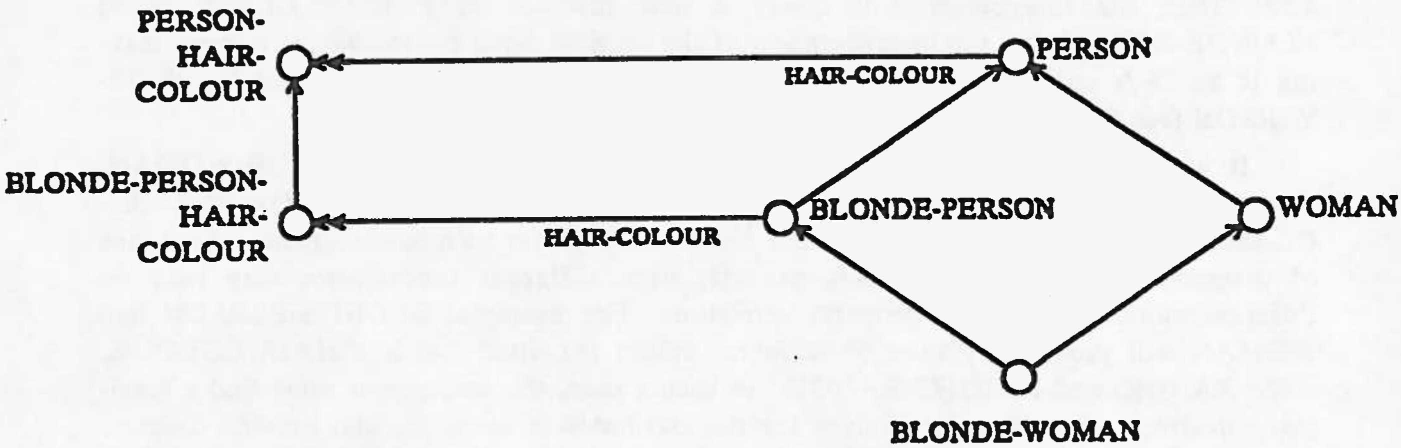

Figure 4.6 Final state of knowledge base

with the adjective. In general, however, these two objects may be separated from one another by an arbitrary number of other classes, $C_1, C_2, \ldots, C_n$, along the IS-A hierarchy (see figure 4.7). The method for dealing with such cases is fairly straightforward. As before, the adjective is first applied directly to (the noun that names) G, yielding a new subclass G' of G. The next step is to make a new class C_n', that is a subclass of G' and C_n, and which serves as the interpretation of the noun phrase obtained by applying the adjective to (the name of) C_n. During the third step, this procedure is repeated, where at each iteration some new class C_k' is created, with IS-A parents C_k and C_{k+1}', for k ranging from *n-1* down to 1 (see figure 4.8). Finally, a new class N', interpreting the given noun phrase, is created, with IS-A parents N and C_1', as before (see figure 4.9).

The main reason for creating the intermediate classes $C_1', C_2', \ldots, C_n'$ is to maintain the structure of the knowledge base, by preventing "semantically distant" concepts from being directly linked to one another. For example, suppose that at some point the knowledge base contains a representation of the concept *albino animal*, and we wish to create a representation for the concept *albino raven*. Suppose further that we directly link

Figure 4.7 Initial state of knowledge base

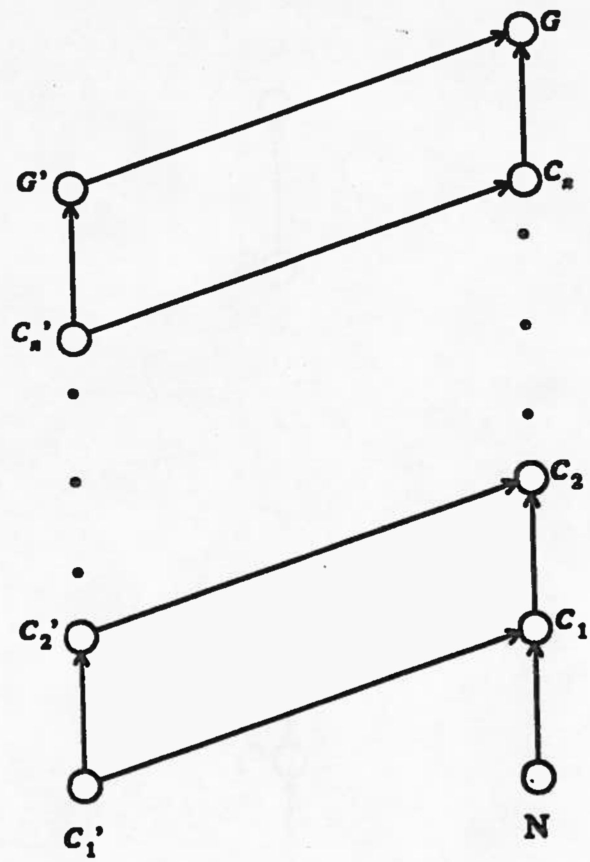

Figure 4.8 Knowledge base after third step

up the representation of the new, rather specialized concept with that of the general concept of albinohood. Now, if we continue to create representations of other "albino" concepts, and directly link each one up with the most general concept, we will end up with a knowledge base which would contain the information that the concepts *albino raven*, *albino crow*, *albino bird*, *albino jackal*, *albino mammal*, and *albino snake* are all indeed related to the general concept *albino animal*, but would not tell us much about how the more

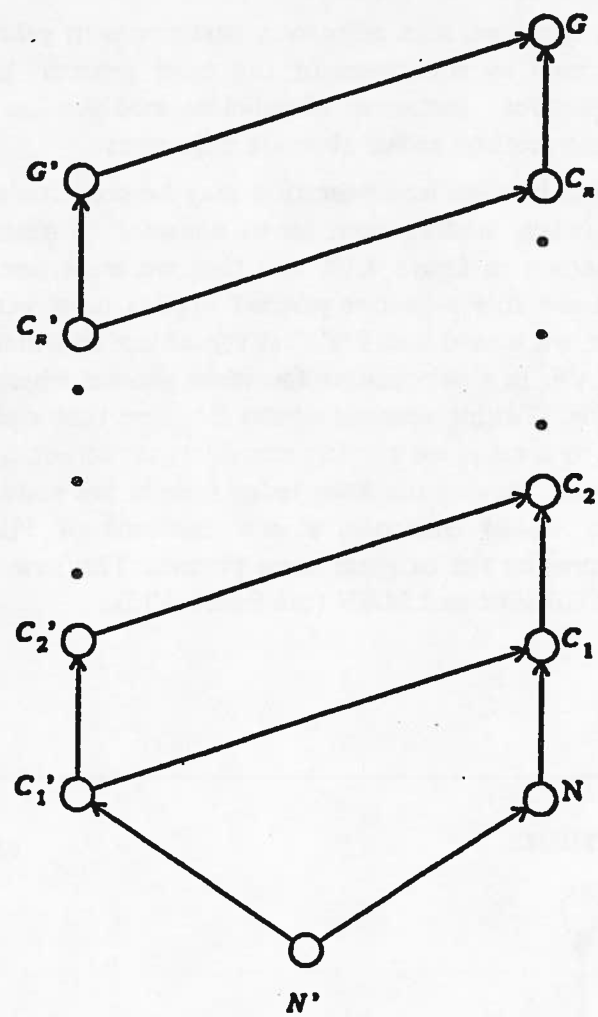

Figure 4.9 Final state of knowledge base

specialized concepts relate to one another. In such a case, it might require a considerable amount of inference to reach even such simple conclusions as "albino ravens and albino crows are both a kind of albino bird". On the other hand, our approach of creating intermediate classes, such as the ones representing *albino bird* or *albino mammal*, as part of a chain of links between specialized concepts and general ones allows such subsumption

relationships to be captured in the structure of the knowledge base.

4.2.2.2. Subclass Modification by a Role Adjective

As with absolute adjectives, role adjectives participate in subclass modification when their nouns are interpreted by subclasses of the most general knowledge base objects associated with the adjectives. Instances of subclass modification by role adjectives are recognized using the same method as for absolute adjectives.

In order to illustrate how an interpretation may be constructed for a role adjective-noun pair exhibiting subclass modification, let us consider an example. Assume that the knowledge base is as shown in figure 4.10, and that we must interpret the noun phrase *pleased man*, given that the role adjective *pleased* applies most generally to persons. As with absolute adjectives, we would first find that the object to which the adjective is being applied, in this case MAN, is a subclass of the most general object to which it could be applied, namely PERSON. Taking account of the intuition that a pleased man is a pleased person who happens to be a man, we would recursively construct an interpretation for the noun phrase *pleased person*, leaving the knowledge base in the state shown in figure 4.11.[12] Finally, the interpreter would construct a new instance of PERSON-CLASS called PLEASED-MAN, interpreting the original noun phrase. The new class would be made a subclass of PLEASED-PERSON and MAN (see figure 4.12).

Figure 4.10 Initial knowledge base

[12] The noun phrase *pleased person* exhibits direct modification, and may be interpreted using techniques discussed in section 4.2.1.2.

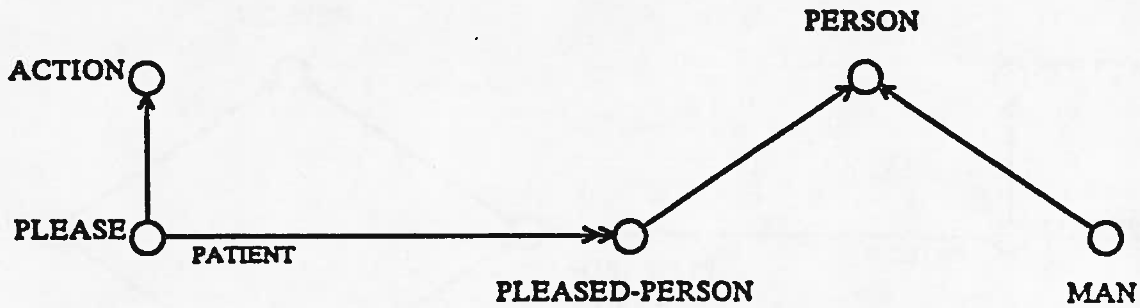

Figure 4.11 Intermediate state of knowledge base

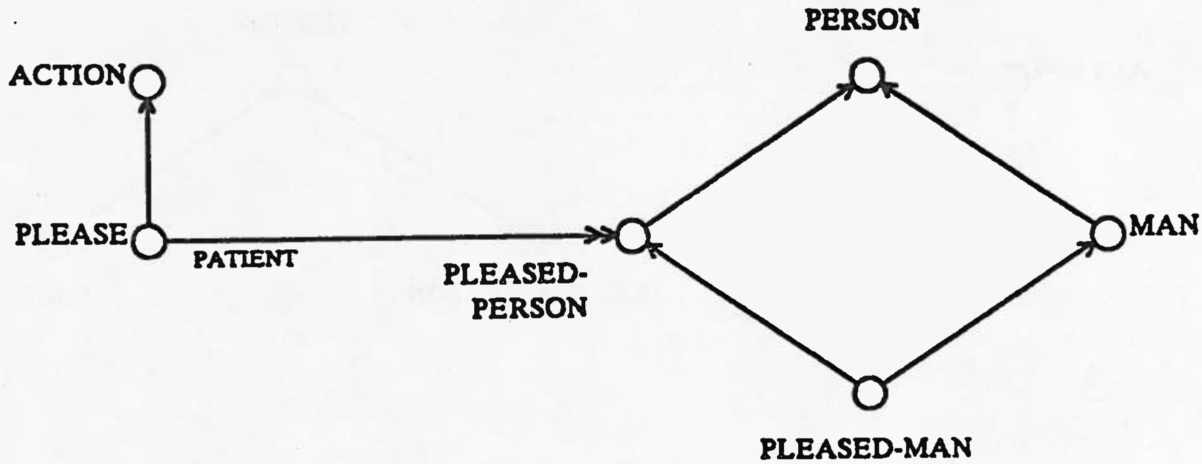

Figure 4.12 Final state of knowledge base

4.2.2.3. Subclass Modification by a Measure Adjective

A measure adjective-noun pair exhibits subclass modification when the property associated with the adjective is not salient to the concept associated with the noun. In our representational system this would correspond to the case in which the knowledge base object that interprets the noun does not take on a more restrictive range of property values for the relevant attribute than the corresponding values taken on by its IS-A parents. In other words, at least one of its IS-A parents takes on the same range of values as it does, for the attribute in question.

In constructing an interpretation for a measure adjective-noun pair that exhibits subclass modification, the interpreter proceeds in much the same fashion as it does for absolute adjectives. Let us consider an example. Suppose that the knowledge base is initially as shown in figure 4.13, and that we are faced with the task of interpreting the noun phrase *tall waiter*. This would be taken as an instance of subclass modification, since, presumably, the class WAITER would take on the same value for HEIGHT-RANGE as does its IS-A parent, MAN. The interpreter would perform a breadth-first search up the IS-A hierarchy from WAITER, looking for a class C that takes on a more restrictive

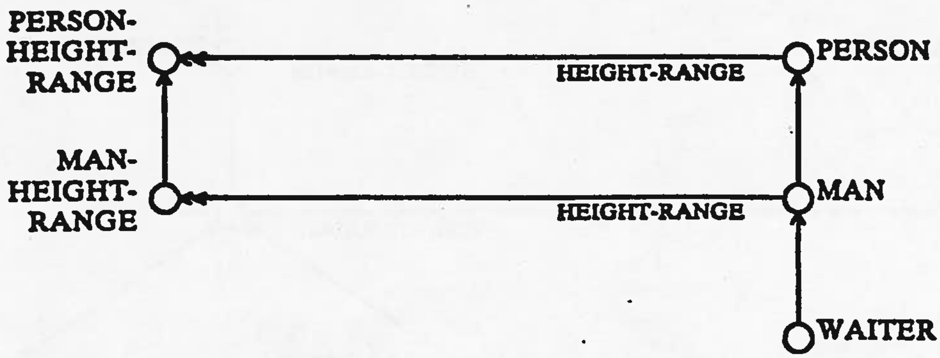

Figure 4.13 Initial state of knowledge base

value for HEIGHT-RANGE than do any of C's IS-A parents. Let us assume that the class MAN restricts HEIGHT-RANGE to a greater degree than does its IS-A parent PERSON.[13] In this case, we want to capture the intuition that a tall waiter is a waiter who is tall for a man; in other words, a waiter and a tall man. We proceed, as we did with absolute adjectives, by recursively invoking the interpreter to construct an interpretation for the noun phrase *tall man*, leaving the knowledge base in the state shown in figure 4.14.[14] Next, the interpreter will create the class TALL-WAITER, a new instance of PERSON-CLASS, and make it a subclass of WAITER and TALL-MAN (see figure 4.15).

It might be argued that *tall waiter* would better be interpreted by an object that was an instance of both WAITER and TALL-MAN, rather than one that was a subclass. Although we agree that the concepts associated with nouns like *waiter* seem to be different in nature from those associated with "natural kind" terms, such as *person*, it does not follow that the organization of knowledge must mirror the distinction. Carey (1982) cites psychological studies which demonstrate that people organize their concepts

[13] Of course, the class C may be arbitrarily distant along the IS-A hierarchy from the object named by the noun. The section dealing with subclass modification by absolute adjectives contains a discussion of the techniques used to handle such complex cases.

[14] The noun phrase *tall man* may be interpreted using the techniques discussed in the section dealing with direct modification.

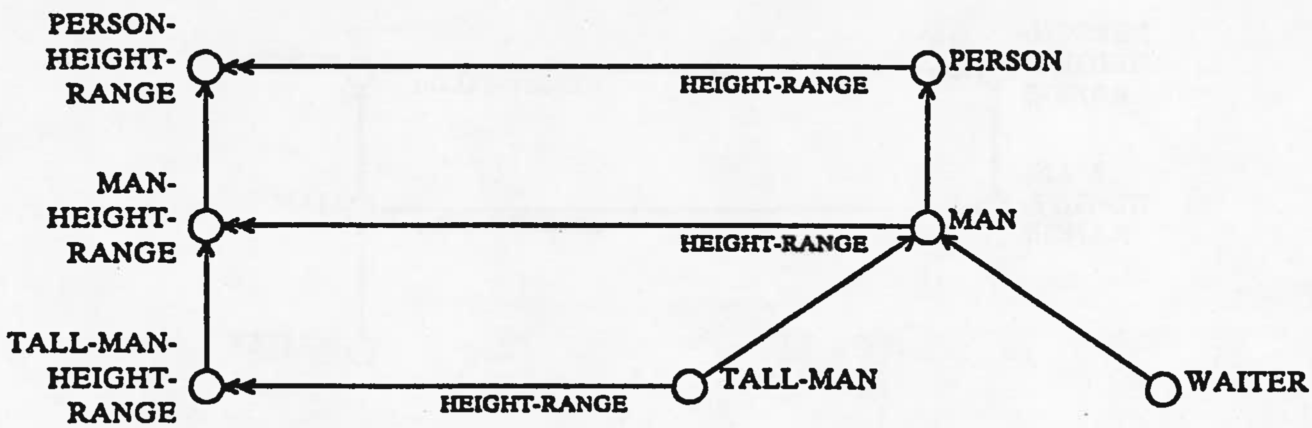

Figure 4.14 Intermediate state of knowledge base

around prototypes, without regard to whether the concepts have definitions or not.[15] Thus, even concepts like the one associated with the term *odd number* seem to have their internal structure organized around prototypes, as shown by the fact that subjects "agreed on *seven* as the best exemplar of odd numberhood, verification times reflected subjects' rankings, and so on" (p.353). In the light of such evidence, and given that we are more interested in capturing the structure of concepts than their abstract semantic characterization, we may readily justify our own reliance on prototypes to organize knowledge, since these provide us with the ability to construct psychologically plausible interpretations.

As with subclass modification by absolute adjectives, the interpreter may need to mediate in the inheritance of property values by the new class. For instance, the values taken on by TALL-MAN and WAITER for the HEIGHT-RANGE attribute will differ. The interpreter will have to find a common subrange of values to be inherited by TALL-WAITER. In this case, TALL-MAN-HEIGHT-RANGE will itself be a subrange of the HEIGHT-RANGE value inherited by WAITER. Therefore, TALL-WAITER will inherit TALL-MAN-HEIGHT-RANGE as its property value. In a more complicated case, the property values of the superclasses might be only partially overlapping. In that instance, the interpreter would have to find the common values and supply them to the new

[15] Carey points out that "if a concept has a best exemplar, in that subjects agree on a prototypicality ranking and this ranking predicts behavior with respect to linguistic hedges, ease of learning, reaction times in verification tasks, and so forth, then that concept has an internal structure organized around a prototype" (1982, p.353).

subclass. This may also involve the construction of a new interslot constraint for the subclass, by intersecting the constraints of the IS-A parents.

4.2.3. Subpart Modification

Subpart modification takes place when an adjective is applied to a noun that refers to a physical or logical subpart of the kind of object with which the property denoted by the adjective is normally associated. Absolute adjectives do not seem to be amenable to subpart modification, while measure adjectives only seem to be used in this way occasionally. However, role adjectives seem to be involved in subpart modification fairly frequently. Accordingly, we shall discuss subpart modification only in the context of role adjectives.

A role adjective-noun pair exhibits subpart modification when the referent of the noun cannot fill the indicated case role of the verb underlying the adjective. Moreover, the noun must refer to a subpart of an object that would be able to fill the case of the verb. When the interpreter finds that the object G directly associated with the adjective is

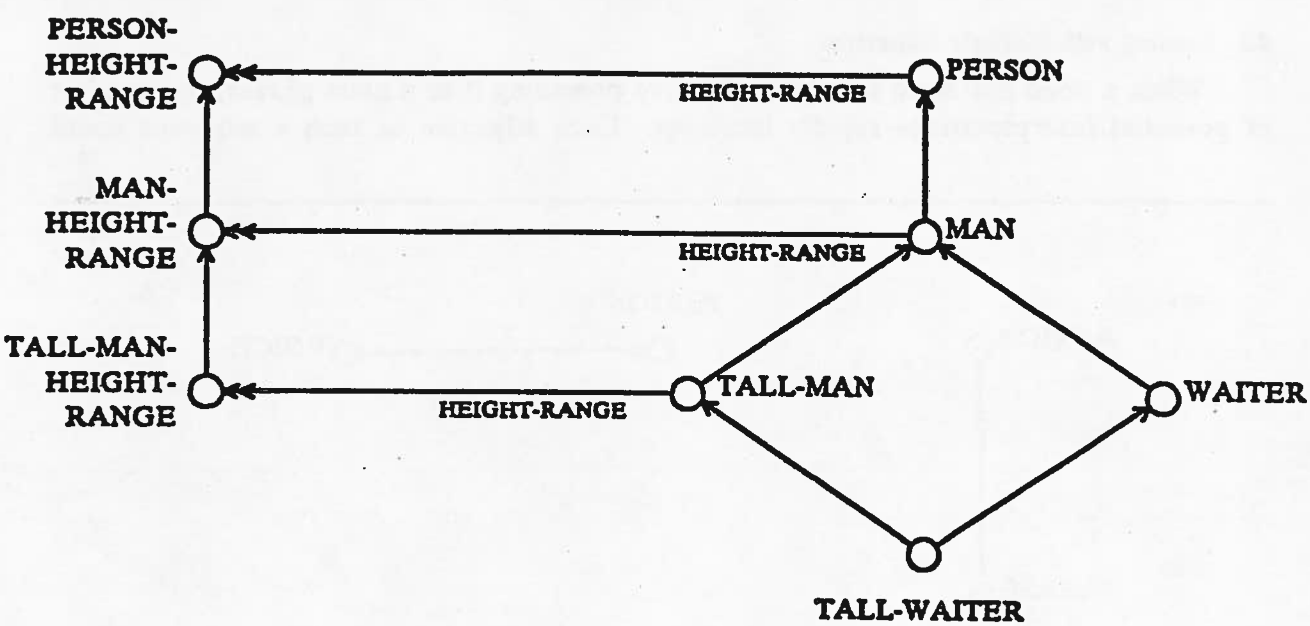

Figure 4.15 Final state of knowledge base

not the same as, or a superclass of, the object N associated with the noun, it checks to see whether the PART-OF parent P of N is identical to G, or is one of its subclasses. If this turns out to be the case, subpart modification would be indicated. Otherwise, it would be concluded that the adjective was probably being used in a metaphorical fashion.

Let us consider an example. Suppose that the knowledge base is as shown in figure 4.16, and that we are trying to interpret the noun phrase *pleased voice*. Since we take the role adjective *pleased* to apply, most generally, to persons, and since VOICE is not identical to, or a subclass of, PERSON, we rule out direct modification and subclass modification. However, since VOICE is a PART-OF PERSON, subpart modification is indicated. Here, we would like to capture the intuition that a pleased voice is a voice that belongs to a pleased person.[16] We may do this by recursively invoking the interpreter to construct an interpretation of the noun phrase *pleased person*, as discussed in the section dealing with direct modification. Next, we create a new class PLEASED-VOICE, as the interpretation of our original noun phrase. This new class will be made an IS-A child of the class VOICE, and a PART-OF child of the class PLEASED-PERSON (see figure 4.17).[17]

4.3. Dealing with Multiple Adjectives

When a noun has more than one adjective preceding it in a noun phrase, the number of potential interpretations rapidly increases. Each adjective in such a sequence could

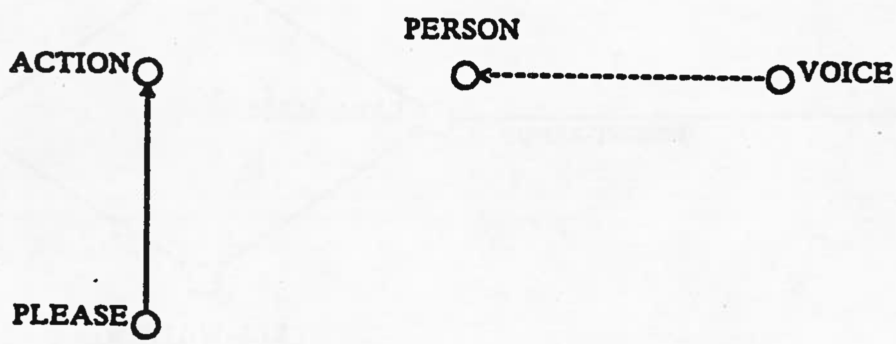

Figure 4.16 Initial state of knowledge base

[16] Of course, we use the term *pleased voice* to indicate that the voice somehow carries the information that its owner is pleased. However, we shall only try to capture the weaker fact here, namely that the voice merely belongs to a pleased person.

[17] The present implementation of PSN does not support explicit PART-OF links. Therefore, the interpreter deals with subpart modification using a modified version of the scheme discussed here.

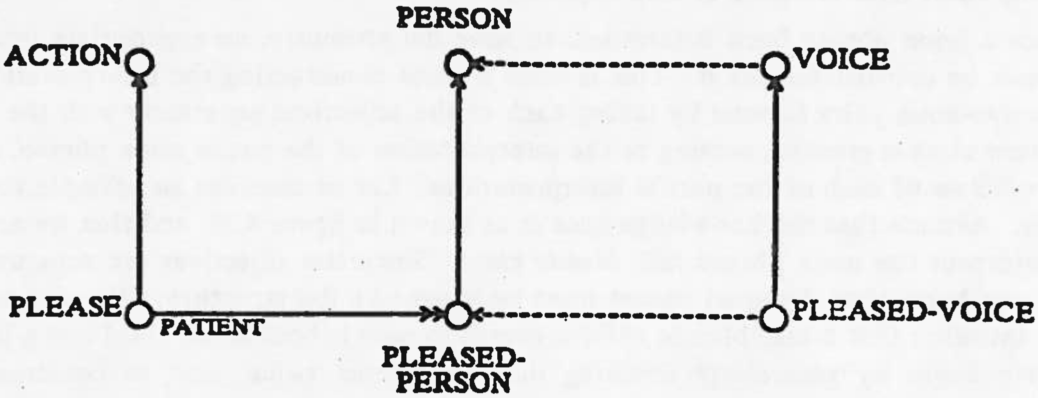

Figure 4.17 Final state of knowledge base

either be modifying the noun by itself, or in combination with some other adjective, or adjectives. We shall make use of all the information at our disposal, including punctuation in the noun phrase, semantic properties of the adjectives, and discourse conventions, in order to come up with a unique interpretation. So far, we have only talked about constructing an interpretation for an adjective modifying a common noun. However, since we use PSN classes to interpret noun phrases, as well as common nouns, we can use the same techniques to interpret adjectives modifying noun phrases as we used to interpret adjectives modifying nouns.

As discussed in Chapter 2, sequences of prenominal adjectives may be assigned one of three different kinds of structure. We shall now examine the methods we use to interpret poly-adjectival noun phrases, with respect to the different kinds of modification structures they display.

4.3.1. Noun Phrases with Flat Structure

Flat structures are those in which each adjective in a sequence separately modifies the head of the noun phrase. Such structures are always displayed by broken sequences, in which all adjectives are separated by commas or conjunctions. Sequences that consist only of role adjectives and/or absolute adjectives also have a flat structure, since these adjectives do not seem to interact with one another. Things get more complicated when a sequence contains measure adjectives. For each such adjective, we must check whether its associated property is related to the properties associated with other measure adjectives to its right. If these properties are unrelated, then a flat structure is indicated. On the other hand, if the properties are related, then the structure would be taken to be flat only if

punctuation separated the adjectives in question.[18] Of course, if there was only one measure adjective in the noun phrase a flat structure would be indicated, since measure adjectives do not nest with absolute or role adjectives.

Once a noun phrase been determined to have flat structure, an appropriate interpretation must be constructed for it. This is done by first constructing the interpretations of the adjective-noun pairs formed by taking each of the adjectives separately with the noun. Then a new class is created, serving as the interpretation of the entire noun phrase, and is made a subclass of each of the partial interpretations. Let us consider an example to illustrate this. Assume that the knowledge base is as shown in figure 4.18, and that we are trying to interpret the noun phrase *tall, blonde child*. Since the adjectives are separated by commas, we know that the noun phrase must be assigned a flat structure. We wish to capture the intuition that a tall, blonde child is someone who is both a tall child and a blonde child. We begin by recursively invoking our interpreter twice, first to construct the interpretation of the noun phrase *blonde child*, then to interpret *tall child*, leaving the knowledge base in the state shown in figure 4.19. Next, we create a new class, TALL-COMMA-BLONDE-CHILD as the interpretation of the entire noun phrase, and make it a subclass of both TALL-CHILD and BLONDE-CHILD, as well as an instance of PERSON-CLASS (see figure 4.20). Incidentally, the same interpretation would have been constructed if the noun phrase had been free of punctuation, since it contains only one measure adjective, which must therefore modify the head noun. In the next two sections

Figure 4.18 Initial state of knowledge base

18 As it stands, this condition is too strong, since it would force nested interpretations even for noun phrases like *small powerful engine*, which seem to be genuinely ambiguous, as pointed out in Section 2.2.2.2. Instead, nesting should be inferred only when the alternative, flat interpretation would carry redundant information. The current implementation always selects a nested interpretation over a flat one. See Chapter 5 for some discussion of how the interpreter could be modified to achieve more flexible behaviour.

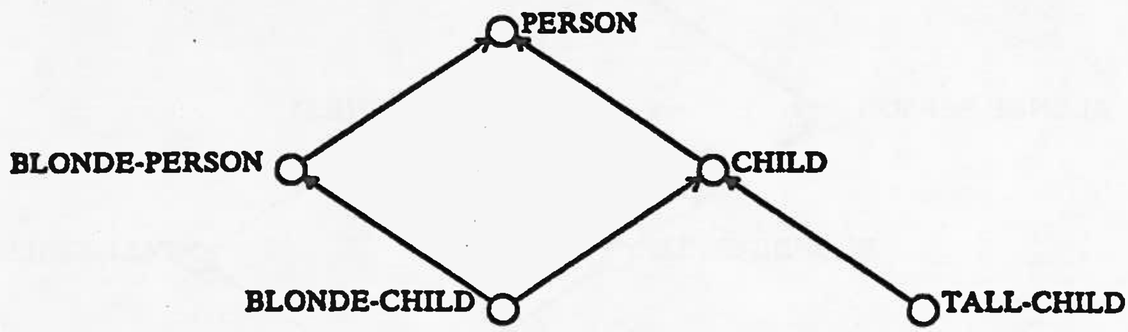

Figure 4.19 Intermediate state of knowledge base

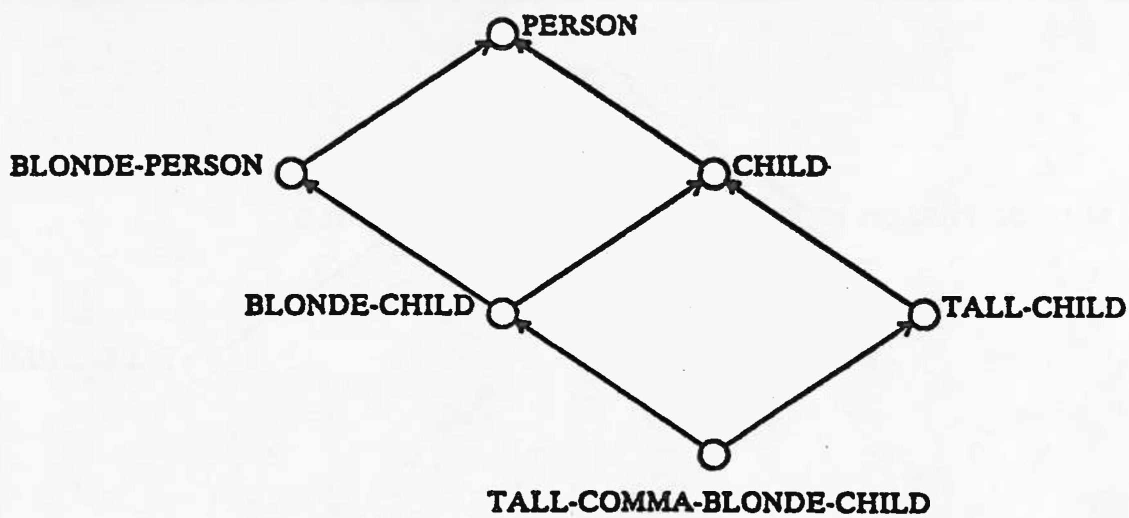

Figure 4.20 Final state of knowledge base

we discuss how noun phrases with multiple measure adjectives are interpreted.

4.3.2. Noun Phrases with Nested Structure

Nested structures are those in which each adjective in a sequence modifies the noun only in combination with all the other adjectives to its right. Such structures are sometimes displayed by unpunctuated sequences that consist entirely of measure adjectives. If the properties associated with these adjectives are interrelated, the noun phrase has a potentially ambiguous interpretation, depending on whether the adjectives are taken to modify the noun independently, or in combination. For example, assuming that size is correlated with power, the noun phrase *large powerful engine* may be interpreted either as "large for an engine and powerful for an engine", or as "large for a powerful engine". In Section 2.2.2.2, we argued that the second, nested interpretation is the more appropriate one. Accordingly, we shall assign nested interpretations to noun phrases exactly when they have unpunctuated prenominal sequences that consist only of measure adjectives whose associated properties are interrelated.[19] In order to construct an interpretation for such a noun phrase, we recursively construct interpretations for the pair consisting of the noun and the rightmost adjective, then for the next adjective, moving leftward, modifying the pair whose interpretation was just constructed, and so on, for all remaining adjectives in the phrase. At each stage of this process, we construct the interpretation of the phrase obtained by prefixing another adjective to the phrase interpreted during the previous stage. For example, assume that we are to interpret the noun phrase *tall young child*. Since, as discussed in Section 4.2.1.3, height and age are related, we must construct a nested interpretation, corresponding to the reading "tall for a young child". We begin by constructing an interpretation for the phrase *young child*, consisting of the rightmost adjective and the noun. This leaves the knowledge base in the state shown in figure 4.21. Next, the adjective *tall* would be examined; since this is a measure adjective, the interpreter would first attempt to use it to extend any existing sequences of measure adjectives that might have been found to its right in the noun phrase.[20] In this case, there is

Figure 4.21 Knowledge base after interpreting *young child*

[19] As pointed out in the previous section, it may not always be desirable to force a nested interpretation in such cases. See Chapter 5 for discussion on how the interpreter might be modified to achieve more flexible behaviour.

[20] If the adjective could not be used to extend any sequence of measure adjectives, or if there were no such sequence to its right, it would be taken to modify the head of the noun phrase. On the other hand, if more than one sequence could potentially be extended by the new adjective, the rightmost one would be selected, since it would be the one closest to the adjective, and therefore the most plausible candidate. There do not seem to be any examples in which one would be forced to take an adjective to simultaneously be a member of more than one sequence. However,

exactly one such sequence, *young child*, that could potentially be extended by prepending the adjective *tall*. Since the adjective *tall* is known to refer to the HEIGHT-RANGE attribute, the interpreter would check whether the class YOUNG-CHILD, which interprets the subphrase *young child*, restricts the value of the HEIGHT-RANGE attribute to a greater degree than does its immediate IS-A ancestor CHILD. This would indeed be found to be the case, assuming that, as discussed in Section 4.2.1.3, we had used an interslot constraint to capture the interdependence of the HEIGHT-RANGE and AGE-RANGE attributes of the class CHILD. Therefore, the interpreter would apply *tall* to the previously interpreted subphrase *young child*, resulting in the situation illustrated in figure 4.22.

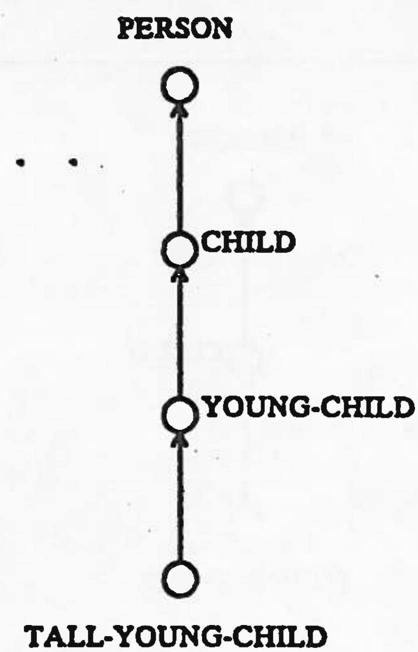

Figure 4.22 Final state of knowledge base

---

in cases where there is more than one potential complement for an adjective, it would probably be better to treat the noun phrase as genuinely structurally ambiguous; in Chapter 5 we discuss how the interpreter might be modified to deal with such cases.

4.3.3. Noun Phrases with Hybrid Structure

Hybrid structures are those in which some adjectives in a sequence exclusively modify the noun, while others modify a combination of other adjectives and the noun. In other words, hybrid structures are obtained by combining subsequences that possess flat structure with others that have a nested structure. Unfortunately, it is not always easy to separate these subsequences, since they sometimes overlap. For example, the noun phrase *impressive large black powerful engine* could be analyzed to contain three subsequences: two single-adjective flat sequences, consisting of the adjectives *impressive* and *black*, respectively; the third, with nested structure, consisting of the adjectives *large* and *powerful*.[21]

In order to construct an interpretation for such a noun phrase, we would first construct interpretations for all its subsequences, using the techniques discussed in the previous two sections. Then, each of these partial interpretations would be made an IS-A parent of a new class, which would serve as the interpretation of the whole noun phrase. For example, consider the noun phrase *tall blonde young child*. We take this to have a hybrid structure, with the adjective *blonde* modifying the noun independently, and the adjectives *tall* and *young* modifying the noun in combination. The interpreter would process the phrase from right to left, starting with the noun. It would begin by constructing an interpretation for the subphrase *young child*. Next, it would encounter the adjective *blonde*, recognize it to be an absolute adjective, and therefore interpret it to be exclusively modifying the head of the noun phrase, forming the subphrase *blonde child*. Next, the adjective *tall* would be examined, found to be a measure adjective, and used to extend the previous measure adjective sequence *young child*, as described in the previous section,[22] resulting in the state shown in figure 4.23. At this point, after processing all the adjectives in the noun phrase, it would create a new class to interpret the entire noun phrase, making it a subclass of the partial interpretations previously constructed (see figure 4.24). Appendix B of this thesis contains some more examples illustrating the operation of the interpreter.

---

[21] There is an alternative analysis in which the adjectives *impressive* and *black* form a single, flat-structured sequence. Since the adjectives in such a sequence would be taken to modify the noun independently anyway, we shall simplify matters by always treating them as though they formed separate flat-structured sequences.

[22] Since *blonde* is an absolute adjective, and thus not capable of forming a nested sequence, the interpreter would not need to check the "tall for a blonde child" reading.

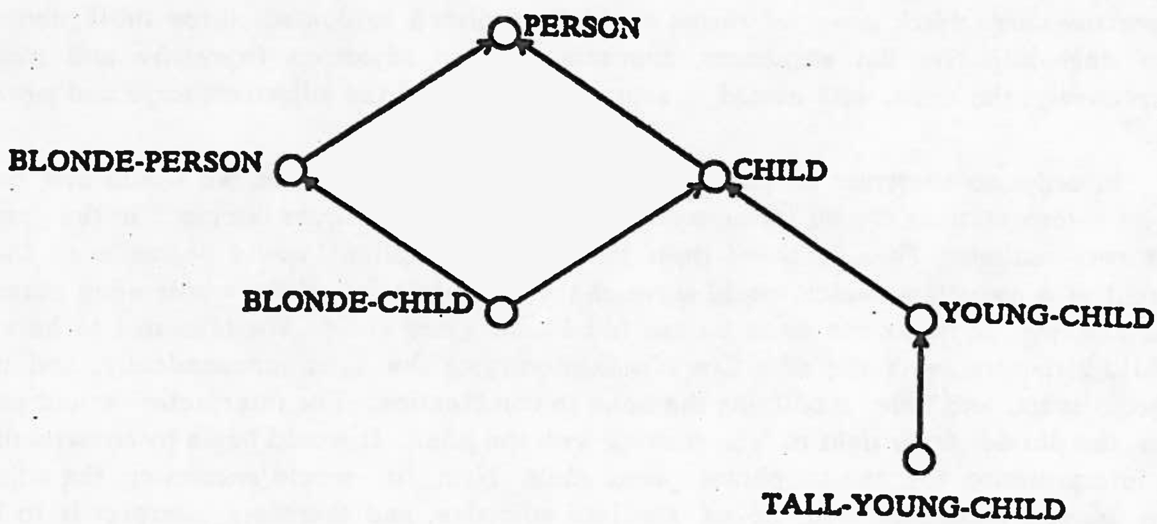

Figure 4.23 Intermediate state of knowledge base

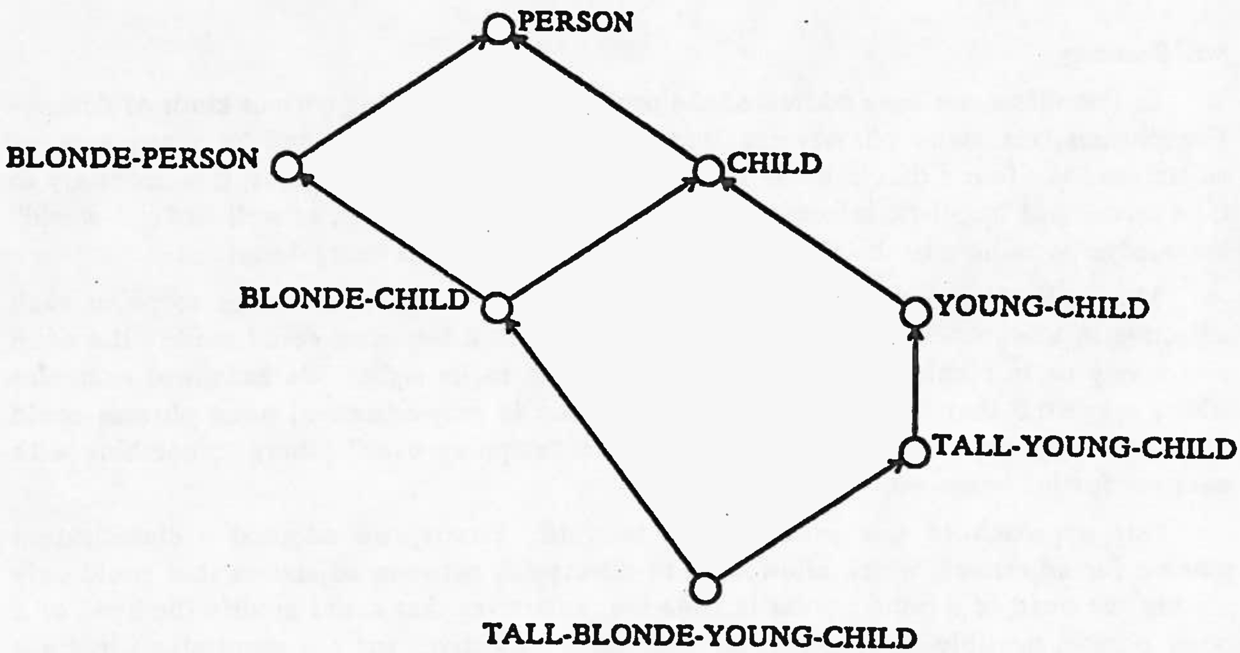

Figure 4.24 Final state of knowledge base

CHAPTER 5

Conclusion

5.1. Summary

In this thesis, we have addressed the problem of interpreting certain kinds of descriptive phrases, viz. noun phrases consisting of a single noun preceded by a sequence of adjectives. We found that in order to successfully interpret such phrases, it is necessary to take account of linguistic information about adjectives and nouns, as well as "real-world" knowledge pertaining to the properties of the kinds of objects being described.

The major linguistic problem we faced was that of determining the scope of each adjective in a sequence of them. An adjective in such a sequence could modify the noun exclusively or in combination with other adjectives to its right. We examined examples which suggested that the structure of modification in poly-adjectival noun phrases could be exceedingly complex, with certain adjectives "skipping over" others to combine with ones yet further removed.

Our approach to this problem was twofold. Firstly, we adopted a classification scheme for adjectives, which allowed us to distinguish between adjectives that could only modify the head of a noun phrase in isolation; adjectives that could modify the head of a noun phrase, possibly in combination with other adjectives (of the same class) to their right; and adjectives that always combined with others to their right in order to modify the noun. Classifying adjectives in this way allowed us to substantially reduce the number of candidate interpretations, in some cases to a single one.

The next step was to invoke other linguistic and non-linguistic information in an attempt to resolve any remaining ambiguities in modifier scope. We showed that "real-world" knowledge about objects and their properties was often essential in allowing us to select between competing interpretations. We also made use of punctuation cues and conversational principles to help resolve structural ambiguity in the noun phrase.

The major non-linguistic problem that we tackled was that of constructing representational structures that captured the descriptions embodied in interpreted noun phrases. This meant that our representations had to capture properties and attributes that could be said to be typical of the class of objects described by each noun phrase. Particular adjectives were taken to indicate restrictions on the values that objects could take on for associated properties.

Frame-like structures were used to represent the generic concepts that we took to be associated with noun phrases. Slots were used to represent stereotypical properties of generic objects, and interslot constraints enforced restrictions on the values taken on by inter-related properties. Our representation scheme was flexible enough to allow featural, dimensional, and functional properties to be associated with objects.

A single knowledge base was used to represent both the interpretations constructed for noun phrases and the "real-world" knowledge needed to resolve structural ambiguity in the noun phrase The contents of the knowledge base were organized along a classification axis and a generalization axis, as well as along an implicit aggregation axis. The use of these axes enabled us to capture the relationships that various concepts bore to one another, and enabled us to build up interpretations of noun phrases compositionally from the interpretations of their constituents.

5.2. Implementation

Our interpreter has been implemented in Franz Lisp on a VAX-11/780. The uncompiled code for the interpreter takes up about 64K bytes of storage. A small knowledge base containing the representations of concepts associated with 5 common nouns, and enough information to allow the interpretation of about 15 different adjectives, occupies about another 10K bytes. The knowledge base is currently being expanded so as to eventually permit the system to handle about fifty different adjectives. Appendix B of this thesis contains some examples that illustrate the operation of the system.

5.3. Evaluation

Hirst (1983) identifies the following qualities as being desirable in a semantic interpreter:

1. It should be provide semantic objects that are "amenable to computational manipulation, supporting inference and problem solving" (p.26).

2. There should be distinct semantic objects corresponding to different syntactically well-formed constituents, and the interpretation of a phrase should be built up compositionally from the interpretations of its syntactic constituents.

3. The interpreter should take into account the contribution of the syntactic structure of a phrase in determining its meaning.

4. The interpreter should be able to handle semantically complex phenomena, such as opaque contexts and generics.

5. The interpreter should be able to resolve instances of lexical ambiguity in the input.

6. The interpreter should be able to provide feedback to a parser.

In this section, we shall argue that our interpreter meets the first three criteria, partially satisfies the fourth one, and could easily be extended to meet the last two requirements as well. We also discuss how it could be modified to construct multiple interpretations and handle at least some intensional adjectives.

Our interpreter uses PSN classes as its semantic objects. These are frame-like objects that are structured into a semantic network along a number of organizational axes. The representation formalism provides mechanisms that allow these objects to be computationally manipulated, and the organizational axes permit flexible control strategies and inference schemes to be easily implemented.[1] Thus, our system clearly satisfies the first requirement listed above.

Each semantic object in our knowledge base corresponds either to a common noun, or else is constructed to represent the modification of a common noun by one or more adjectives. The interpretation of the latter is constructed in a principled, compositional manner from the interpretation of the common noun and the interpretations of subphrasal constituents, each consisting of the head noun preceded by a subsequence of the original sequence of adjectives. Each of these constituents is syntactically well-formed, and each contributes its interpretation to the interpretation of the whole noun phrase. We thus claim that our interpreter also satisfies the second requirement listed above, compositionality.

With regard to the desirability of taking account of syntactic structure, it should be remarked that parsers generally do not assign any internal structure to the prenominal modifier sequence of a noun phrase. The assignment of structure to such sequences has traditionally been viewed as a problem that lies in the domain of semantic interpretation. We have shown, in this thesis, that it is necessary to take account of both linguistic and non-linguistic knowledge in order to uncover the structure of modification within noun phrases; we have presented a system that makes use of such knowledge in order to assign structure to a noun phrase, and interprets it on the basis of the structure inferred. One source of syntactic information which this system takes account of is word order. Additional information is available when when punctuation (or conjunction) is present in the noun phrase, and this is made use of also. In that our interpreter takes punctuation and word order into account when interpreting noun phrases, it may be said to satisfy the third requirement listed above, also.

On the subject of being able to handle semantic complexities, we claim that our analysis provides us with the ability to resolve complex instances of modification. We can give principled accounts for why *tall man* is interpreted as "tall for a man", whereas a *tall waiter* is normally taken to be tall for an adult. We can successfully predict that a *tall red-headed young child* is probably tall for a young child, whereas a *tall handsome young professor* is likely to be tall for an adult. We can assert that a *pleased woman* is one who serves the role of PATIENT in a particular kind of action, whereas a *pleased voice* is one that belongs to a pleased person. While our interpreter presently cannot handle opaque contexts, such as the ones induced by intensional adjectives, it does handle generic (readings of) noun phrases. Thus, we claim that it goes some way towards meeting the fourth desideratum among those listed above.

One of the weaknesses of our system is its inability to deal with lexical ambiguity. In its current incarnation, the system requires that each common noun map to a unique

[1] For examples of the use of complex control regimens in reasoning over PSN knowledge bases, see Tsotsos (1983, 1984).

object in the knowledge base, and that each adjective refer to a single, fixed property. This restriction is unreasonable, since many adjectives and nouns are ambiguous, some highly so. For example, the noun *hand* may be used to variously refer to different kinds of objects, as shown by the following:[2]

(1) She shook my hand and invited me in.

(2) The hand on the clock was pointing at 3.

(3) I dealt another hand, and the game continued.

(4) They hired a new hand to help out on the farm.

(5) Give me a hand with these boxes!

Similarly, an adjective may refer to a number of different kinds of properties. For example, the adjective *dead* refers to different properties in each of the following:

dead cow; dead engine; dead issue; dead stop; dead language

In some instances, an ambiguous adjective can be correctly interpreted simply because only one of its associated properties is relevant to the kind of object that its complement refers to. For example, with *dead engine* we could rule out the "deprived of life" reading, since engines are not alive in the first place; the other candidate interpretations could probably also be ruled out in the same fashion. Similarly, an ambiguous noun may be constrained to one meaning by the adjectives that modify it. For example, in a literal reading, *young hand* would only make sense under the "employee" reading of *hand*.

In some cases, a noun phrase may be successfully interpreted even when both the noun as well as the prenominal adjectives are ambiguous, provided that they are mutually disambiguating. An example due to Small (1980) illustrates the point nicely: both the adjective *deep* and the noun *pit* are ambiguous, since *deep* can mean both "extending far beneath the surface" and "profound", whereas *pit* can mean "a deep hole in the earth" or "the stone of a drupaceous fruit"; however, when the two words are combined in the noun phrase *deep pit*, only one of the four possible interpretations makes any sense. Unfortunately, words are not always mutually disambiguating in this way. For example, the noun phrase *slow hand* could be interpreted in any one of the following ways, depending upon the context:

a slow-moving bodily appendage;
the hand of a slow-running clock;

---

[2] Of course, the word *hand* is also categorially ambiguous, since it can be used both as a noun and as a verb. However, as Hirst (1983, p.6) points out, categorial ambiguity has traditionally been regarded as a problem in parsing, and we will therefore ignore it in this work.

an abnormally time-consuming (or boring) round of cards;
a dim-witted employee

Hirst (1983) points out that word-sense disambiguation requires, in the general case, contextual knowledge, the ability to find associations between "nearby" words, sensitivity to syntactic cues, the use of selectional restrictions on case-slot fillers, and inference "as a last resort" (p. 86).[3] He presents a lexical disambiguation mechanism called POLAROID WORDS which works in parallel with a parser and a semantic interpreter, "permitting them to deal with ambiguous words as if their semantic object were assigned immediately" (p. 137). The Polaroid Words mechanism is designed to operate on a knowledge base consisting of a network of frames, not unlike that used by our interpreter. It works by searching for connections between concepts in such a network, and attempts to find a set of mutually satisfactory word meanings corresponding to these concepts. It should be relatively straightforward to modify our interpreter to make use of a lexical disambiguation system such as Hirst's.

In cases where it is not possible to resolve ambiguity in the noun phrase, it would be desirable to construct all the possible interpretations, rather than just picking one of them. This problem does not just arise with lexically ambiguous noun phrases, but also with structurally ambiguous ones. For example, a noun phrase like *small powerful engine* is structurally ambiguous, since the two measure adjectives *small* and *powerful* could either be taken to modify the noun separately or else in combination, corresponding to a flat or a nested interpretation, respectively. In Chapter 4 we indicated that when the interpreter is faced with such a choice it always selects the nested interpretation over the flat one. This yields the preferred interpretation in examples like *large powerful engine*, where the flat interpretation carries redundant information. However, in genuinely ambiguous examples like *small powerful engine*, all the candidate interpretations should probably be constructed.

It should be reasonably straightforward to modify the interpreter to allow it to construct more than one interpretation for a noun phrase. Since the design of the interpreter is quite modular, the change could be achieved at the cost of complicating the top-level of the control structure somewhat; however, most of the code would remain unaltered. Moreover, since the knowledge base stores partial interpretations, much like the well-formed substring table of a chart parser stores parses of subconstituents, these could be shared among multiple interpretations, avoiding inefficiency. The main expense would be in the increased running time required to construct all the interpretations of a noun phrase. In principle, it should be possible to speed up the program by having it construct the different interpretations in parallel. Provided that some care is taken to disallow simultaneous access to the knowledge base, running the modified interpreter on a machine with a parallel architecture should allow multiple interpretations to be constructed fairly efficiently.

[3] Inference should be used only as a last resort because it is usually a computationally expensive operation.

Our interpreter is capable of running in parallel with a parser, and may be used to provide feedback to the syntactic component of a natural language understanding system. In the present implementation, the interpreter takes as input the head noun of a noun phrase, and a list of prenominal adjectives. The noun phrase is processed from right to left, starting with the head noun and working back through the list of adjectives. Partial interpretations are constructed incrementally, as each adjective is processed, and added to the knowledge base. Finally, the semantic object that serves as the interpretation of the complete noun phrase is returned.

The design of the interpreter permits its operation to be either loosely or tightly coupled with that of a parser. For instance, the interpreter could be invoked after the parse of a noun phrase, or even of a sentence, had been completed. At the other extreme, it could repeatedly be called upon to judge the semantic acceptability of the modification of a noun by each successive adjective in a sequence. In the latter case, a parser might pause after processing the sequence of prenominal adjectives and the head noun, and then repeatedly call the interpreter, each time supplying it with the head noun and successively larger subsequences of adjectives, starting with the rightmost one and adding from the left.[4] If any of these calls failed to produce an acceptable interpretation, the system could either terminate processing of the phrase entirely, or else invoke higher-level modules to resolve the problem, perhaps by modifying the interpreter's knowledge base.

The operation of our interpreter is fairly similar to that of the descriptive semantic interpretation component of the PSI-KLONE natural language understanding system (Bobrow and Webber 1980a, 1980b). PSI-KLONE consists of an ATN parser that is "cascaded" with a semantic interpreter which, in turn, feeds its output to a discourse module. The parser processes syntactic constituents, producing representations of their structure in KL-ONE (Brachman 1978), a semantic network formalism. The parser can interact with the PSI-KLONE interpreter by TRANSMITTING to it a proposal for a new syntactic constituent to be added to the phrase that the system is currently processing. The interpreter can either reject the proposal, causing the parser to back up, or else it can indicate that the constituent is semantically acceptable, allowing the parse to proceed. PSI-KLONE's interpreter is flexible enough to be able to verify that constituents are semantically acceptable, without necessarily constructing incremental interpretations of phrases. Instead, the actual construction of interpretations can be postponed until after syntactic analysis has succeeded. In contrast, our interpreter presently constructs interpretations "on the fly" for each constituent, and its subconstituents, that it receives as input. However, we may easily modify it to permit more flexible behaviour. For instance, we could allow it to be called with a flag whose value would determine whether an actual interpretation was to be returned, or whether it should merely create a data structure from which an interpretation could later be constructed.[5]

---

[4] This would be a reasonably efficient scheme, since partial interpretations are added to the knowledge base and are thus available for subsequent use if the same subphrases are encountered again.

[5] Such data structures are already used by the interpreter. The only difference is that, at present, interpretations are constructed as soon as their data structures have been "filled in". However, there is no reason why the construction of interpretations could not be postponed, provided that the data structures were saved.

Finally, it would be desirable to extend the coverage of our interpreter to at least some intensional adjectives. There are a number of difficulties with doing this. One is that the meanings of such adjectives tend to be radically context-dependent. We treat adjectives as referring to, and restricting the values of, fixed properties of objects. However, many intensional adjectives do not seem to refer to a single property, or even a small number of properties, in a determinate fashion. For example, it would be extremely difficult to spell out in advance which properties the intensional adjective *good* refers to. When modifying a noun, *good* seems to pick out some salient property of the noun's referent, relative to the context in which the phrase is used. For example, the interpretation of *good performance* in

(6) Joan gave a good performance.

would be very different, depending on whether we were talking about her performance in a play or her performance in a baseball game. Similarly, the term *good engine* could be variously used to describe engines that are powerful, fuel-efficient, easy to maintain, etc. We could consider applying techniques used to resolve lexical ambiguity to adjectives such as *good*. If we provided our interpreter with knowledge about many of the different senses taken on by a word like *good*, it should be able to do a better job of interpreting it, by picking the sense that is most appropriate to the noun that the adjective modifies. However, it would still be extremely difficult to ensure that the most appropriate sense of the word is discerned in each example, just because almost any conceivable property could be referred to, given a suitable context.

An additional difficulty is presented by intensional adjectives such as *fake*, which do not preserve the subproperty assumption, and thus cannot be interpreted compositionally (see section 4.2). Since a *fake gun* is not a gun, strictly speaking we should not treat *fake gun* as a subconcept of *gun*. Indeed, since PSN enforces strict typing, meaning that IS-A subclasses cannot violate any of the properties of their superclasses, making FAKE-GUN, the putative interpretation of *fake gun*, a subclass of GUN would mean that instances of the former would also be taken as instances of the latter, which is clearly incorrect. On the other hand, if FAKE-GUN was not made a subclass of GUN, it would not be able to inherit any of the properties of GUN. However, as Lakoff and Johnson (1980) point out, the adjective *fake* "preserves certain *kinds* of the properties of GUNS and negates others" (p.121). In particular, they argue that *fake* preserves perceptual properties (how an object looks), motor-activity properties (how an object is handled), and some of the purposive properties (what an object is used for) of objects that it applies to, while negating their functional properties (what an object does) and history of function (what an object was designed to do). Thus, under their account, an object would be a *fake gun* if it looked and handled something like a gun, and could be displayed or used to threaten, as could a real gun, but couldn't be used to shoot anyone, and was not originally designed to shoot anyone (i.e., a gun that is merely broken or otherwise inoperable cannot be described as a fake).

While it is not clear that Lakoff and Johnson's analysis of the adjective *fake* is completely correct,[6] [7] it seems reasonable that some account along these lines might indeed be accurate. There does not appear to be any entirely satisfactory way to capture such a treatment within the PSN framework. One approach might be to make use of a semantic network formalism that allowed exceptions in the inheritance of properties (see, for instance, Etherington 1983). Then, we might conceive of constructing an interpretation of *fake gun* as a subconcept of *gun* subject to certain exceptions in the inheritance of properties from its superclass. However, this would still violate the principle that a *fake gun* is not, under any circumstances, a *gun*.

5.4. Extensions

There are a number of linguistic phenomena which are not accounted for by our analysis, yet which bear similarities to the data that we do handle. In this section, we point out some of these cases, and briefly comment on their nature.

Recall that, according to our characterization, the attribute referred to by an adjective could be directly ascribed to the objects denoted by its complement, or it could be ascribed to a more general class of objects than denoted by the complement, or it could be ascribed to objects having physical or logical subparts that were denoted by the complement. However, there seem to be cases in which, rather than ascribing the attribute associated with an adjective to some other objects than those denoted by its complement, one is forced to regard the adjective as referring to an attribute that it would not normally denote. We classify these as being metaphorical uses of adjectives. Some examples:

loud dress; wooden smile; cruel sea; tall story; cold look

The problem of recognizing and resolving metaphor is extremely difficult in general, and it is one that we have avoided tackling in this thesis. The best that we could do, given our present system, is to guess that an adjective is being used metaphorically if we are unable to find any reasonable (literal) interpretation under our analysis. However, in order to actually resolve metaphors, a lot of high-level inference seems to be necessary. For a good overview of various approaches to the study of metaphor, the interested reader is referred to Ortony (1979); Hayes (1977b) and Hobbs (1979) discuss metaphor from an AI perspective. Lakoff and Johnson (1980) argue, contrary to the received view, that metaphor is "pervasive in everyday language and thought" and that it should therefore be regarded as "a matter of central concern, perhaps the key to giving an adequate account of understanding" (Preface, p. ix).

Another phenomenon which bears upon our work is that of noun-noun modification. As mentioned in Chapter 1, nouns, like adjectives, can appear as premodifiers of the heads of noun phrases. Since, in this work, we were concerned with analyzing the nature of

---

[6] For instance, is it the case that a *fake coin* cannot possess, or have been constructed to possess, the functional properties of a coin? What would be characterized as the functional properties of a Picasso painting or a Beethoven symphony?

[7] Graeme Hirst (personal communication) points out that the adjective *fake* does not appear to negate the functional properties of all its complements. In particular, it does not seem to do this in the case of *fake furs*.

adjectival modification, we ignored the possibility that more than one noun might be present in a noun phrase. However, a more complete treatment of modification in the noun phrase would seek to integrate work like ours with that done by Brachman (1978), Finin (1980), D.B. McDonald (1982), and others on noun-noun modification. We believe that such integration is feasible, since the approach that we have adopted to our problem is quite similar to those taken by the above-mentioned researchers in tackling noun-noun modification. In particular, the use of a store of frame-based "real-world" knowledge about the nature of the concepts represented by nouns is something that is common to our system and their systems. Similarly, the form of the interpretations that each of the systems construct is quite similar.

We should point out that a complete noun phrase interpreter could not be constructed simply by linking up our system with one of the noun-noun modification systems mentioned above. The reason this would not work is that there are examples in which adjective-noun modification and noun-noun modification interact, and any interpreter for these must be able to simultaneously bring to bear techniques for dealing with both phenomena. Consider the following examples:

heavy water consumption;
new computer journal;
large oil company earnings;
expensive car dealership;
small powerful engine repair shop

In each of these noun phrases, there are conflicting interpretations based upon whether an adjective is taken to modify a noun compound, or whether an adjective-noun combination is taken to modify another noun. Thus, *heavy water consumption* could be either a heavy consumption of water, or the consumption of heavy water; a *new computer journal* could either be a new journal about computers, or it could be a journal about new computers; *large oil company earnings* could be taken to refer to the excessive profits of oil companies in general, or only to the earnings of the large oil companies; an *expensive car dealership* could be a company that sells expensive cars, or an expensive company that sells cars; a *small powerful engine repair shop* could be a repair shop for small powerful engines (recall that the subphrase *small powerful engine* is itself ambiguous), or it could be a small shop that specializes in repairing powerful engines, or it could be a shop that performs small repairs on powerful engines. It should be clear that such complex cases could only be handled by an interpreter that fully integrated the analysis of adjective-noun and noun-noun modification.

Bibliography

BACHE, Carl (1978). *The Order of Premodifying Adjectives in Present-Day English.* Odense University Press, 1978.

BARWISE, Jon (1981). "Some Computational Aspects of Situation Semantics". *Proceedings, 19th Annual Meeting of the Association for Computational Linguistics*, Stanford, CA, June 1981, pp.109-111.

BARWISE, Jon and PERRY, John (1983). *Situation Semantics.* MIT Press, Cambridge, MA, 1983.

BARWISE, Jon and PERRY, John (1984). "Shifting Situations and Shaken Attitudes: An Interview with Barwise and Perry". CSLI-Report No. 14, Center for the Study of Language and Information, Leland Stanford Junior University, Palo Alto, CA (1984).

BOBROW, Robert J and WEBBER, Bonnie Lynn (1980a). "PSI-KLONE: Parsing and Semantic Interpretation in the BBN natural language understanding system". *Proceedings of the Third Biennial Conference of the Canadian Society for Computational Studies of Intelligence*, Victoria, BC, May 1980, pp. 131-142.

BOBROW, Robert J and WEBBER, Bonnie Lynn (1980b). "Knowledge Representation for Syntactic/Semantic Processing". *Proceedings, National Conference on Artificial Intelligence*, Stanford, CA, August 1980, pp. 316-323.

BRACHMAN, Ronald J (1978). "A Structural Paradigm for Representing Knowledge".
[1] Report 3605, Bolt, Beranek and Newman, Cambridge, MA, May 1978.
[2] Ablex Publishing: Norwood, NJ, 1984.
[3] revised from Doctoral dissertation, Division of Engineering and Applied Physics, Harvard University, May 1977.

BRACHMAN, Ronald J and LEVESQUE, Hector (1982). "Competence in Knowledge Representation". *Proceedings, National Conference on Artificial Intelligence*, Pittsburgh, PA, August 1982, pp. 189-192.

CAREY, Susan (1982). "Semantic Development: the state of the art". In: *Language Acquisition: the state of the art*, Eric Wanner and Lila R. Gleitman (editors), Cambridge University Press, New York, NY, 1982.

DOWNING, Pamela (1977). "On the Creation and Use of English Compound Nouns". *Language*, 53(4), December 1977, pp. 131-147.

DOWTY, David R; WALL, Robert Eugene and PETERS, Stanley (1981). *Introduction to Montague Semantics*.
(Synthese Language Library 11),
D. Reidel, Dordrecht, 1981.

DRETSKE, Fred I (1981). *Knowledge and the Flow of Information*.
Bradford Books/MIT Press: Cambridge, MA (1981).

DRETSKE, Fred I (1983). "The Epistemology of Belief".
Synthese, 55 (1983), pp. 3-19.

ETHERINGTON, David W (1983). "Formalizing Non-Monotonic Reasoning Systems".
Technical Report 83-1, Department of Computer Science, University of British Columbia, 1983.

FAHLMAN, Scott Elliot (1979). *NETL: A System for Representing and Using Real-World Knowledge*.
[1] The MIT Press, Cambridge, MA, 1979.
[2] an earlier version: Doctoral dissertation, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, September 1977.

FININ, Timothy Wilking (1980). *The Semantic Interpretation of Compound Nominals*.
Doctoral dissertation, Department of Computer Science [available as report T-96, Coordinated Science Laboratory], University of Illinois at Urbana-Champaign.

GIL, David (1983). "Stacked Adjectives and Configurationality".
Linguistic Analysis, 12(2), 1983.

GOODMAN, Bradley A (1983). "Repairing Miscommunication: Relaxation in Reference".
Proceedings, National Conference on Artificial Intelligence,
Washington, D.C., August 1983, pp. 134-138.

GRICE, H Paul (1975). "Logic and Conversation".
In: *Syntax and Semantics 3: Speech Acts*, Peter Cole and Jerry L Morgan (editors),
Academic Press, New York, NY, 1975.

HAYES, Philip J (1977a). *Some Association-based Techniques for Lexical Disambiguation by Machine*.
[1] Doctoral dissertation, Département de Mathématiques, Ecole polytechnique fédérale de Lausanne.
[2] Technical report 25, Department of Computer Science, University of Rochester, June 1977.

HAYES, Philip J (1977b). "On Semantic Nets, Frames and Associations".
Proceedings of the 5th International Joint Conference on Artificial Intelligence,
Cambridge, MA, August 1977, pp. 99-107.

HELLAN, L (1981). *Towards an Integrated Analysis of Comparatives.*
Narr, Tübingen, 1981.

HIRST, Graeme John (1983). "Semantic Interpretation Against Ambiguity".
Doctoral dissertation [available as technical report CS-83-25], Department of Computer Science, Brown University, 1983.

HOBBS, Jerry Robert (1979). "Metaphor, Metaphor Schemata, and Selective Inferencing".
Technical Note 203, Artificial Intelligence Center, SRI International, December 1979.

KAMP, J A W (1975). "Two Theories About Adjectives".
In: *Formal Semantics of Natural Language.*
E.L. Keenan (editor),
Cambridge University Press, Cambridge, 1975.

KIEFER, Ferenc (1978). "Adjectives and Presuppositions".
Theoretical Linguistics, 5(2/3), 1978, pp.135-173.

KLEIN, Ewan (1980). "A Semantics for Positive and Comparative Adjectives".
Linguistics and Philosophy, 4(1), 1980, pp. 1-45.

KLEIN, Ewan (1982). "The Interpretation of Adjectival Comparatives".
Journal of Linguistics, 18, 1982, pp. 113-136.

LAKOFF, George and JOHNSON, Mark (1980). *Metaphors We Live By.*
The University of Chicago Press, Chicago, 1980.

LEES, Robert B (1960). *The Grammar of English Nominalizations.*
Mouton & Co.: The Hague, 1960.

LEVESQUE, Hector (1977). "A Procedural Approach to Semantic Networks".
Technical Report 105, Department of Computer Science, University of Toronto, 1977.

LEVESQUE, Hector (1984). "A Fundamental Tradeoff in Knowledge Representation and Reasoning".
Proceedings of the Fifth Biennial Conference of the Canadian Society for Computational Studies of Intelligence,
London, ONT, May 1984, pp. 141-152.

LEVESQUE, Hector and MYLOPOULOS, John (1979). "A Procedural Semantics for Semantic Networks".
In: *Associative Networks: Representation and Use of Knowledge by Computer*, Nicholas V. Findler (editor),
Academic Press, 1979.

LEVI, Judith N (1978). *The Syntax and Semantic of Complex Nominals.*
Academic Press: New York, 1978.

MATTHEI, Edward H (1982). "The Acquisition of Prenominal Modifier Sequences".
Cognition, 11(3), 1982.

McCAWLEY, James D (1981). *Everything that Linguists Have Always Wanted to Know About Logic* *
* *but were ashamed to ask.*
The University of Chicago Press, Chicago, 1981.

McDONALD, David Blair (1982). *Understanding Noun Compounds.*
Doctoral dissertation [available as technical report CMU-CS-82-102], Department of Computer Science, Carnegie-Mellon University, January 1982.

McDONALD, David D and CONKLIN, E J (1982). "Salience As a Simplifying Metaphor for Natural Language Generation".
Proceedings, National Conference on Artificial Intelligence,
Pittsburgh, PA, August 1982, pp. 75-78.

MONTAGUE, Richard (1970). "English as a Formal Language".
[1] In: Visentini, Bruno et al. *Linguaggi Nella Società e Nella Tecnica.*
Milan, Edizioni Di Comunità, 1970, pp. 189-224.
[2] In: Thomason (1974a), pp. 188-221.

MYLOPOULOS, John and LEVESQUE, Hector (1984). "An Overview of Knowledge Representation".
[1] Technical Note #36, Computer Systems Research Group, University of Toronto, May 1984.
[2] In: *On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages*, Michael L. Brodie, John Mylopoulos, and Joachim V. Schmidt (editors),
Springer-Verlag, New York, 1983.

NISHIDA, Toyoaki (1983). *Studies on the Application of Formal Semantics to English-Japanese Machine Translation.*
Doctoral dissertation, Department of Information Science, Kyoto University, Japan.

ORTONY, Andrew A (editor) (1979). *Metaphor and Thought.*
Cambridge University Press: New York, 1979.

PUTNAM, Hilary (1970). "Is Semantics Possible?".
[1] In: *Language, Belief, and Metaphysics*, H.E. Kiefer and M.K. Munitz (editors),
State University of New York Press, NY (1970).
[2] In: *Naming, Necessity, and Natural Kinds*, Stephen P. Schwartz (editor),
Cornell University Press, Ithaca, NY (1977).

PUTNAM, Hilary (1973). "Meaning and Reference".
[1] *Journal of Philosophy*, LXX, 1973, pp.699-711.
[2] In: *Naming, Necessity, and Natural Kinds*, Stephen P. Schwartz (editor),
Cornell University Press, Ithaca, NY (1977).

QUILLIAN, M Ross (1968). "Semantic Memory".
[1] In: *Semantic Information Processing*, Marvin Lee Minsky (editor),
The MIT Press, Cambridge, MA (1968).

[2] Doctoral Dissertation, Carnegie Institute of Technology [Carnegie-Mellon University], October 1966.
Published as Report 2, project 8668, Bolt, Beranek and Newman Inc., October 1966.

QUIRK, Randolph; GREENBAUM, Sidney; LEECH, Geoffrey and SVARTVIK, Jan (1972). *A Grammar of Contemporary English.*
Longman Group Limited, London, 1972.

SIDNER, Candace L (1983). "The Pragmatics of Non-Anaphoric Noun Phrases".
In: "Research in Knowledge Representation for Natural Language Understanding: Annual Report, 1 September 1982 to 31 August 1983".
Report No. 5421, Bolt, Beranek and Newman Inc., Cambridge, MA, October 1983.

SIEGEL, Muffy (1976). *Capturing the Adjective.*
Doctoral dissertation, University of Massachusetts, Amherst, MA, 1976.

SIEGEL, Muffy (1979). "Measure Adjectives in Montague Grammar".
In: *Linguistics, Philosophy, and Montague Grammar*, S. Davis and M. Mithun (editors), University of Texas Press, 1979.

SMALL, Steven Lawrence (1980). *Word Expert Parsing: A Theory of Distributed Word-Based Natural Language Understanding.*
Doctoral dissertation [available as technical report 954], Department of Computer Science, University of Maryland, 1980.

SMITH, Brian Cantwell (1982). "Linguistic and Computational Semantics".
Proceedings, 20th Annual Meeting of the Association for Computational Linguistics, Toronto, ONT, June 1982, pp. 9-15.

THOMASON, Richmond Hunt (1974a) (editor). *Formal Philosophy: Selected Papers by Richard Montague.*
Yale University Press, New Haven, 1974.

THOMASON, Richmond Hunt (1974b). "Introduction".
In: Thomason (1974a), pp. 1-69.

TSOTSOS, John Konstantine (1983). "Knowledge Organization: Its Role in Representation, Decision-making and Explanation Schemes for Expert Systems".
Technical Report LCM-TR83-3, Laboratory for Computational Medicine, Department of Computer Science, University of Toronto, December 1983.

TSOTSOS, John Konstantine (1984). "Representational Axes and Temporal Cooperative Processes".
Technical Report RCBV-TR-84-2 on Research in Biological and Computational Vision, Department of Computer Science, University of Toronto, April 1984.

ZIFF, Paul (1972). *Understanding Understanding.*
Cornell University Press, Ithaca, NY, 1972.

APPENDIX A

Sources of Selected Examples from Bache's Study

This appendix contains a list of sources used by Bache in his study of poly-adjectival noun phrases (see section 2.2). The list in this appendix only includes sources for examples actually mentioned in this thesis.

TAG	NAME OF SOURCE
CMMQ	Agatha Christie, *The Mysterious Mr. Quinn* (Dell 6246).
CSC	Agatha Christie, *Sad Cypress* (Fontana 1358).
DE	*Daily Express*, January 22, 1976.
KOFOCN	Ken Kesey, *One Flew Over the Cuckoo's Nest* (Picador 0 330 23564 8).
MHFF	F.L. Marcuse, *Hypnosis, Fact and Fiction* (Pelican 14 120446 6).
Ms	*Ms*, January 1976.
PBASS	James Cochrane (editor), *The Penguin Book of American Short Stories* (Penguin 14002919 2).
PBESS	Christopher Dolley (editor), *The Penguin Book of English Short Stories* (Penguin 2617).
SCR	J.D. Salinger, *The Catcher in the Rye* (Penguin 1400 1248 6).
Td	*Time*, September 29, 1975.
VS5	Kurt Vonnegut, *Slaughterhouse Five* (Panther 586 13328 9).

APPENDIX B

The Performance of the Implemented System

This appendix contains some examples illustrating the operation of the interpreter. The top-level function is named *interpret*, and it takes as arguments a list of adjectives and a single common noun. The adjectives in the list may be separated by the special symbol COMMA, which corresponds to the presence of a comma in that position in the noun phrase. Invocation of the interpreter on a given noun phrase usually results in the creation of new knowledge base structure, as discussed in Chapter 4. If an interpretation is found for the noun phrase, the interpreter prints out a pseudo-English description of this interpretation. Otherwise, it prints out a diagnostic message saying why it was unable to construct an interpretation.

I have appended brief explanatory remarks to some of the examples below. Where such remarks appear, they have been set in italics, in order to prevent any confusion with the output of the interpreter.

The running times printed below are for interpreted code, and include time used up in garbage collection.

```
-> [interpret (blonde) person]

  blonde-person is a person
      that has golden or auburn hair-colour


  CPU TIME = 1.22 SECONDS
```

Comment: Direct modification by an absolute adjective.

```
-> [interpret (redheaded) person]

  redheaded-person is a person
      that has red or orange hair-colour


  CPU TIME = 2.7 SECONDS
```

```
-> [interpret (blonde) man]

  blonde-man is a man
      and a blonde-adult
  blonde-adult is a adult
      and a blonde-person
  blonde-person is a person
      that has golden or auburn hair-colour


  CPU TIME = 26.7 SECONDS
```

Comment: Subclass modification by an absolute adjective

```
-> [interpret (blonde redheaded) man]

  *** Unable to find an interpretation:
     redheaded-man, blonde-man conflict with one another
     in the values that they can take on
     for the hair-colour attribute.


  CPU TIME = 50.52 SECONDS
```

Comment: Here the two absolute adjectives require disjoint values to be taken on by the same attribute.

```
-> [interpret (tall) person]

  tall-person is a person
      that is tall for a person


  CPU TIME = 5.65 SECONDS
```

Comment: Direct modification by a measure adjective

```
-> [interpret (tall) adult]

   tall-adult is a adult
       that is tall for a adult


   CPU TIME = 3.28 SECONDS
```

Comment: Direct modification, since height is salient to adulthood.

```
-> [interpret (tall) man]

   tall-man is a man
       and a tall-adult
   tall-adult is a adult
       that is tall for a adult


   CPU TIME = 18.08 SECONDS
```

Comment: Subclass modification, since in the knowledge base the class MAN does not restrict height to a greater degree than the class ADULT does.

```
-> [interpret (tall) basketballplayer]

   tall-basketballplayer is a basketballplayer
       that is tall for a basketballplayer


   CPU TIME = 5.42 SECONDS
```

Comment: Direct modification, since BASKETBALLPLAYER restricts height to a greater degree than does ADULT.

```
-> [interpret (tall young) person]

   tall-young-person is a tall-person
       and a young-person
```

tall-person is a person
that is tall for a person
young-person is a person
that is young for a person

CPU TIME = 24.43 SECONDS

Comment: The new class YOUNG-PERSON does not restrict height to a greater degree than its is-a parent PERSON (see the discussion in section 4.2.1.3).

-> [interpret (tall young) man]

tall-young-man is a tall-man
and a young-man
tall-man is a man
and a tall-adult
tall-adult is a adult
that is tall for a adult
young-man is a man
and a young-adult
young-adult is a adult
that is young for a adult

CPU TIME = 47.85 SECONDS

-> [interpret (tall young) child]

tall-young-child is a young-child
that is tall for a young-child
young-child is a child
that is young for a child

CPU TIME = 12.28 SECONDS

Comment: Here, the new class YOUNG-CHILD restricts height to a greater degree than its is-a parent CHILD does (see section 4.2.1.3). Thus, the adjective 'tall' modifies 'young child' directly.

```
-> [interpret (tall COMMA young) child]

  tall-COMMA-young-child is a young-child
      and a tall-child
  young-child is a child
      that is young for a child
  tall-child is a child
      that is tall for a child


  CPU TIME = 23.82 SECONDS
```

Comment: Placing a comma between the two adjectives forces a non-nested interpretation; compare with the previous example.

```
-> [interpret (tall COMMA short) person]

  *** Unable to find an interpretation:
     tall-person, short-person conflict with one another
     in the values that they can take on
     for the height-range attribute.


  CPU TIME = 21.08 SECONDS
```

Comment: The comma forces a non-nested interpretation, resulting in the antonymous pair of measure adjectives selecting disjoint subranges of PERSON-HEIGHT-RANGE, as described in section 4.2.1.3.

```
-> [interpret (tall short) person]

  tall-short-person is a short-person
      that is tall for a short-person
  short-person is a person
      that is short for a person


  CPU TIME = 12.3 SECONDS
```

Comment: A non-nested interpretation is not possible, but nesting the adjectives does allow the noun phrase to be interpreted. Unlike the previous example, there is no comma to prevent a nested interpretation.

-> [interpret (pleased) person]

pleased-person is a person
that is the patient of please actions

CPU TIME = 0.35 SECONDS

Comment: Direct modification by a role adjective.

-> [interpret (pleased) man]

pleased-man is a man
that is the patient of please actions

CPU TIME = 0.37 SECONDS

-> [interpret (pleasing) smile]

pleasing-smile is a smile
that is the instrument of please actions

CPU TIME = 0.28 SECONDS

-> [interpret (pleased) smile]

pleased-smile is a smile
that has as agent a pleased-person

```
  pleased-person is a person
      that is the patient  of please actions

  CPU TIME = 0.72 SECONDS
```

Comment: Subpart modification by a role adjective.

```
    ----------------------

-> [interpret (annoying) look]

  annoying-look is a look
      that is the instrument  of annoy actions

  CPU TIME =  0.3 SECONDS

    ----------------------

-> [interpret (annoyed) look]

  annoyed-look is a look
      that has as agent a annoyed-person
  annoyed-person is a person
      that is the patient  of annoy actions

  CPU TIME = 0.77 SECONDS
```

Comment: Subpart modification by a role adjective.

```
    ----------------------

-> [interpret (tall redheaded young white) man]

  tall-redheaded-young-white-man is a tall-man
      and a young-man
      and a redheaded-man
      and a white-man
  tall-man is a man
      and a tall-adult
  tall-adult is a adult
      that is tall for a adult
  redheaded-man is a man
```

```
    and a redheaded-adult
redheaded-adult is a adult
    and a redheaded-person
redheaded-person is a person
    that has red or orange hair-colour
young-man is a man
    and a young-adult
young-adult is a adult
    that is young for a adult
white-man is a man
    and a white-adult
white-adult is a adult
    and a white-person
white-person is a person
    that has white skin-colour

CPU TIME = 107.72 SECONDS
```

Comment: A flat structure is deduced for the noun phrase.

```
-> [interpret (tall redheaded young white) child]

tall-redheaded-young-white-child is a tall-young-child
    and a redheaded-child
    and a white-child
tall-young-child is a young-child
    that is tall for a young-child
redheaded-child is a child
    and a redheaded-person
redheaded-person is a person
    that has red or orange hair-colour
young-child is a child
    that is young for a child
white-child is a child
    and a white-person
white-person is a person
    that has white skin-colour

CPU TIME = 54.45 SECONDS
```

Comment: Here, a hybrid structure is deduced, with the measure adjectives 'tall' and 'young' nesting with one another. Notice that these two adjectives are not even contiguous in the noun phrase.

-> [interpret (tall) smile]

*** The adjective tall cannot modify smile, nor any of its is-a ancestors.

CPU TIME = 0.07 SECONDS

Comment: The property referred to by 'tall' is not possessed by the complement 'smile'.

-> [interpret (black) person]

black-person is a person
    that has black skin-colour

CPU TIME = 0.85 SECONDS

-> [interpret (black) look]

*** The adjective black cannot modify look, nor any of its is-a ancestors.

CPU TIME = 0.07 SECONDS

Comment: The interpreter cannot handle metaphor, or even lexical ambiguity.
