9 Constraining the Interpretation of Nominal Compounds in a Limited Context

Timothy W. Finin
Computer and Information Science
University of Pennsylvania

Nominal compounds are notoriously ambiguous. Syntactically, the number of parse trees of a nominal compound is exponential in the number of terms making up the compound. Worse, the semantic relationship between each term and its modifier has been deleted from the surface form and must be inferred. This makes nominal compounds even more highly ambiguous from a semantic point of view. The seemingly inherent ambiguity of nominal compounds might suggest that their utility is limited. On the contrary, they are very heavily used, especially in the context of a *sublanguage*.

In earlier work, we described a system that synthesized candidate semantic interpretations of nominal compounds from possible interpretations of their constituents. The candidates were partially ordered with respect to an appropriateness measure based on a number of syntactic, semantic and pragmatic factors. More recent work has examined the additional constraints that arise in the context of an extended dialogue. In addition, we discuss the use of nominal compounding as a short-term naming device in these two situations.

INTRODUCTION

In earlier work (Finin, 1980a, 1980b), we described UNCLE, a system that built semantic interpretations of nominal compounds – sequences of two or more nouns (or nominal adjectives) that function as a unit. Some examples of the kinds of nominal compounds dealt with are: *aircraft engine repairs, fuel pump float valve adjustments* and *January F105 maintenance data*. In the simplest case, where a compound contains only two nominals, the problem reduces to understanding the intended relationship between the concepts the nouns denote. For example, *engine repairs* might refer to events in which someone repaired an engine; *screwdriver repairs,* to events in which someone repaired something using a screwdriver; and *mechanic repairs,* to events in which someone who is a kind of mechanic repaired something.

As these examples show, there are typically more than one potential relationship between the modifier and modified nouns. Our initial system took the representations of the concepts making up the compound and synthesized a number of concepts which represented possible interpretations. Each candidate interpretation was then assigned an *appropriateness* measure based on a variety of factors.

One source of knowledge that was not used was the discourse context. Nominal compounding is often used as a naming device to refer to a concept already introduced, either directly or indirectly, into the discourse. For example, if one is fixing a broken screwdriver, then it is perfectly proper to refer to the activity as a *screwdriver repair.* If we are discussing the distant future in which we have robot automoble mechanics, then we may want to use the compound *mechanic repairs* to refer to the action of fixing a broken automobile repair robot.

Our current work in this area centers around three questions: (1) How can we effectively use discourse context as a primary source of evidence for both hypothesizing candidate interpretations (roughly in order of plausibility) and choosing a preferred interpretation from among them? (2) How much analysis is really required to do an adequate job of understanding nominal compounds used in discourse? We believe that many compound strings may require only a relatively shallow analysis. (3) When is it appropriate to use nominal compounds? We would like to formulate the conditions when a particular nominal compound is likely to be accurately and efficiently interpreted by a hearer.

In this chapter, we first describe the model developed in earlier work for understanding nominal compounds and then describe how it is being extended to account for the effects of discourse context.

BACKGROUND

The semantics of nominal compounds have been studied, either directly or indirectly by linguists and AI researchers. In an early study, Lees (1960) developed an impressive taxonomy of the forms. More recently, Levi (1979) and Downing (1977) attempted to capture the linguistic regularities evidenced by nominal compounding. Rhyne (1976) explored the problem of generating compounds from an underlying representation. Brachman (1978) used the problem of interpreting and representing nominal compounds as example domain in the development of his SI-Net representational formalism. Gershman (1977) and McDonald and Hayes-Roth (1978) attempt to handle noun-noun modification in the context of more general semantic systems.

Our own work on nominal compounds was done in the context of the natural language data base accessing system JETS (1979). UNCLE was designed to act as a specialist that, given a nominal compound, would produce its best guess at the most appropriate interpretation. In UNCLE, the interpretation of nominal compounds was divided into three intertwined subproblems: *lexical interpretation* (mapping words into concepts), *modifier parsing* (discovering the structure of compounds with more than two nominals), and *concept modification* (assigning an interpretation to the modification of one concept by another). The essential feature of this form of modification is that the underlying semantic relationship between the two concepts is not explicit. Moreover, a large number of relationships might, in principal, exist between the two concepts. The selection of the most appropriate one can depend, in general, on a host of semantic, pragmatic and contextual factors.

Let's restrict our attention for a moment to the simplest of compounds—those made up of just two nouns, both of which unambiguously refer to objects that we know and understand. What is the fundamental problem in interpreting the modificaton of the second noun by the first? The problem is to find the underlying relationship that the utterer intends to hold between the two concepts that the nouns denote. for example, in the compound "aircraft engine" the relationship is *part of*; in "meeting room" it is *location*; in "salt water" it is *dissolved in*.

There are several aspects that make this problem difficult. First, the relationship is not always evident in the surface form of the compound. What is it about the compound *GM cars* that suggests the relationship *made by*? The correct interpretation of this compound depends on our knowledge of several facts. We must know that GM is the name of an organization that manufactures things, in particular, automobiles. Another fact that strengthens this interpretation is that the identity of an artifact's manufacturer is a salient fact. It is even more important when the artifact is an automobile (as opposed to, say, a pencil).

A second source of difficulty is the general lack of syntactic clues to guide the interpretation process. The interpretation of clauses involves discovering and making explicit the relationships between the verb and its "arguments," for example, the subject, direct object, tense marker, aspect, etc. Clauses have well developed systems of syntactic clues and markers to guide interpretation. These include word order (e.g. the agent is usually expressed as the subject, which comes before an active verb); prepositions, which suggest case roles; and morphemic markers. None of these clues exists in the case of nominal compounds.

Third, even when the constituents are unambiguous, the result of compounding them may be multiply ambiguous. For example, *a woman doctor* may be a doctor who is a woman or a doctor whose patients are women. Similarly, *Chicago flights* may be those bound for Chicago, coming from Chicago, or even making a stop in Chicago.

A fourth aspect is that compounds exhibit a variable degree of lexicalization and idiomaticity. In general, the same compound form is used for lexical items (e.g. duck soup, hanger queen) and completely productive expression (e.g. engine maintenance, faculty meeting).

Finally, it is possible for any two nouns to be combined as a compound and be meaningful in some context. In fact, there can arbitrarily be many possible relationships between the two nouns, each relationship appropriate for a particular context.

INTERPRETATION RULES FOR NOMINAL COMPOUNDS

The UNCLE system uses three general classes of interpretation rules in the interpretation of nominal compounds. The first class contains *idiomatic rules*—rules in which the relationship created is totally dependent on the identity of the rule's constitutents. These rules typically match surface lexical items directly. Often, the compounds have an idiomatic or exocentric meaning.

The second class consists of *productive rules.* These rules attempt to capture forms of modification that are productive in the sense of defining a general pattern that can produce many instantiations. They are characterized by the semantic reltionships they create between the modifying and modified concepts. That is, the nature of the relationship is a property of the rule and not of the constituent concepts. The nature of the concepts determines only whether or not the rule applies and, perhaps, how strong the resulting interpretation is. For example, a rule for *dissolved in* could build interpretations of such compounds as "salt water" and "sugar water" and be triggered by compounds in which the head noun is a kind of liquid and the modifier is a kind chemical compound.

The third class contains the *structural rules.* These rules are characterized by the structural relationships they create between the modifying and modified concepts. The semantic nature of the relationship that a structural rule creates is a function of the concepts involved in the modification. Many of these rules are particularly useful for analyzing compounds which contain nominalized verbs.

Structural Rules

This class contains the most general semantic interpretation rules, those that help to achieve a degree of closure with respect to semantic coverage. Similar structural rules form the basis of the approaches of Brachman (1978) and McDonald and Hayes-Roth (1978). This section presents some of the structural rules we have catalogued. Each rule handles a compound with two constituents.

Role Value Modifies a Concept The first structural rule is the most common. It interprets the modifying concept as specifying or filling one of the roles of the modified concept.Some examples of compounds that can be successfully interpreted by this rule are:

engine repair	(a to-repair with object = (an engine))
January flight	(a to-fly with time = (a January))
F4 flight	(a to-fly with vehicle = (an F4))
engine housing	(a housing with superpart = (an engine))
iron wheel	(a wheel with raw-material = (a iron))

Note that when the compound fits the form "subject + verb" or "object + verb," this works very nicely. The applicability of this rule is not limited to such compounds, however, as the last two examples demonstrate.

To apply this rule, we must be able to answer two questions. First, which of the modified concept's roles can the modifier fill? Obviously, some roles of the modified concept may be inappropriate. The concept for the to-repair event has many roles, such as an agent doing the repairing, an object being repaired, an instrument, a location, a time, etc. The concept representing an engine is clearly inappropriate as the filler for the agent and time roles, probably inappropriate as a filler for the location and instrument roles, and highly appropriate as the object's filler.

Second, given that we have found a set of roles that the modifier may fill, how do we select the best one? Moreover, is there a way to measure how well the modifier fits a role? Having such a figure of merit allows one to rate the overall interpretation. The process of determining which roles of a concept another may fill and assigning scores to the alternatives is called *role fitting.* This process returns a list of the roles that the modifier can fill and, for each, a measure of how "good" the fit is. Each possibility in this list represents one possible interpretation. Not all of the possibilities are worthy of becoming interpretations, however. A selection process is applied that takes into account the number of possible interpretations, their absolute scores, and their scores relative to each other. Making a role fit into an interpretations involves making a new instantiation of the modified concept and filling the appropriate role with modifier. Details of this process are presented in the next section.

Concept Modifies a Role Value This rule is similar to the first, except that the concepts change places. In interpretations produced by this rule, the modified concept is seen as filling a role in the modifier concept. Note that the object referred to by the compound is still an instance of the modified concept. Some examples where this rule yields the most appropriate interpretation are:

drinking water	(a water which is (an object of (a to-drink)))
washing machine	(a machine which is (an instrument of (a to-wash)))
maintenance crew	(a crew which is (an agent of (a to-maintain)))

Again, the application of this rule is mediated by the *role fitting* process.

Concept Modifies a Role Nominal This rule is applicable when the modified concept is in the class I call *role nominals,* nouns that refer to roles of other underlying concepts. English has but one productive system for naming role nominals: the agent of a verb can commonly be referenced by adding the -er or -or suffix to the verb stem. This should not hide the possibility of interpreting many concepts as referring to a *role* in another related concept. Some examples are:

a student is the *recipient* of a teaching,
a pipe is the *conduit* of a flowing,
a pump is the *instrument* of a pumping, and
a book is the *object* of a reading.

This rule tries to find an interpretation in which the modifier actually modifies the underlying concept to which the role nominal refers. For example, given *F4 Pilot*, the rule notes that *pilot* is a role nominal referring to the agent role of the to-fly event and attempts to find an interpretation in which *F4* modifies that to-fly event. The result is something like "an F4 pilot is the agent of a to-fly event in which the vehicle is an F4." Some other examples are:

cat food	(an object of (a to-eat with agent = (a cat)))
oil pump	(an instrument of (a to-pump with object = (an oil)))
dog house	(a location of (a to-dwell with agent = (a dog)))

Viewing a concept as a *role nominal* (e.g. food as the object of eating) ties the concept to a characteristic activity in which it participates. It is very much like a relative clause, except that the characteristic or habitual nature of the relationship is emphasized.

Role Nominal Modifies a Concept This rule is similar to the previous one, except that it applies when the modifying concept is a role nominal. The action is to attempt an interpretation in which the modification is done, not by the first conept, but by the underlying concept to which it refers. For example, given the compound *pilot school*, we can derive the concept for "an organization that teaches people to fly." This is done by noting that pilot refers to the agent of a to-fly event and then trying to modify *school* by the *to-fly*. This, in turn, can be interpreted by the *Concept + Role Nominal* rule if school is defined as "an organization which is the agent of a to-teach." This leads to an attempt to interpret to-fly as modifying to-teach. The *Role Value + Concept* rule interprets to-fly as filling the object (or discipline) role of to-teach.

Some other examples of compounds that benefit from this interpretation rule are newspaper glasses (glasses used to read a newspaper), driver education (teaching people to drive), food bowl (a bowl used to eat food out of).

Other Structural Rules Other structural interpretation rules that I have identified include *Specific + Generic*, which applies when the modifier is a specialization of the modified concept (e.g. F4 planes, boy child); *Generic + Specific*, which applies when the modifier is a generalization of the modified concept (e.g. Building NE43, the integer three); *Equivalence*, in which the resulting concept is descendant from both the modifier and modified concepts (e.g. woman doctor); and *Attribute Transfer*, in which a salient attribute of the modifier is transferred to the modified concept (e.g. iron will, crescent wrench.)

Role Fitting

The process of role fittng is one in which we are given two concepts, a *role Value* and a *Host*, and attempt to find appropriate roles in the Host concept in which the Role Value concept can be placed. Briefly, the steps carried out by the program are: (a) collect the local and inherited roles of the Host concept; (b) filter out any inappropriate ones (e.g. structural ones); (c) for each remaining role, compute a score for accepting the Role Value concept; (d) select the most appropriate role(s).

In the third step, the goodness-of-fit score is represented by a signed integer. Each role of a concept is divided into an arbitrary number of facets, each one representing a different aspect of the role. In computing the goodness of fit measure, each facet contributes to the overall score via a characteristic scoring function. The facets which participate include the following:

Requirements	descriptions candidate value *must* match.
Preferences	descriptions candidate value *should* match.
Default Value	a default value.
Typical Value	other very common values for this role.
Modality	one of Optional, Mandatory, Dependent or Prohibited.

Multiplicity	maximum and minimum number of values.
Salience	a measure of the role's importance with respect to the concept.

For example, the scoring function for the *requirements* facet yields a score increment of +1 for each requirement that the candidate value matches and a negative infinity for any mismatch. For the *preferences* facet, we get a +4 for each matching preference description and a −1 for each mismatching description. The *salience* facet holds a value from a 5-point scale (i.e. Very Low, Low, Medium, High, Very High). Its scoring function maps these into the integers −1, 0, 2, 4, 8.

DISCOURSE CONTEXT

The problem with this analysis is that it has ignored the discourse context as a source of evidence for selecting appropriate interpretations of nominal compounds. One can take a radical view that discourse should be the primary source of evidence and treat most nominal compounds as referring expressions. Under this view, a compound's function is to select an object that has been previously introduced into the discourse. Determining the correct interpretation of a nominal compound, then, is primarily a matter of finding the proper referent.

Nominal Compounds as Referring Expressions

We can generalize our earlier approach to discover compounds whose interpretation is to be understood referentially. Consider a system that works in a domain of discourse in which there are flights made by aircraft and which have an origin, a destination, and sometimes a stopover location. The compound *Chicago flights* might well be used to describe flights coming from, going to, or stopping at Chicago. Viewed in isolation, there is little evidence in the compound to support one interpretation over another. In the context of a particular discourse, there may be strong evidence for selecting one or another of the candidate interpretations. Consider interpreting the compound *Chicago flights* in each of the following discourse fragments.

> All of the flights coming from the midwest are delayed by the weather.
> The *Chicago flights* are a full hour late.

> I usually take a flight going through Chicago or St. Louis.
> I prefer *Chicago flights*, sinnce they are usually shorter.

> Last week I made trips to our Illinois and California offices.
> The food on the *Chicago flight* was so bad I got sick.

As an example, consider a discourse containing the following text:

> All flights scheduled to stop at Chicago are being rerouted through Milwaukee. Flights scheduled to stop at St. Louis are being rerouted through Indianapolis. The *Chicago flights* will experience delays of up to one hour while the St. Louis flights will suffer little or no delays.

We want to account for the fact that, in this text, the compound *Chicago flights* is unambiguously interpreted as meaning "the flights that were to have stopped in Chicago."

The concept lattice can be seen as encoding all possible relationships between concepts. Finding candidate interpretations for nominal compounds composed of the nouns denoting CONCEPT1 and CONCEPT2 can then be described as:

> Candidate interpretations for a nominal compound with constituents denoting CONCEPT1 and CONCEPT2 can be found by considering all potential relationships between CONCEPT1 *and its generalizations* and CONCEPT2 *and its generalizations*.

For our example, the concept *flight* would be defined to include the source, destination, and stopovers roles:

```
(a movement-event is (an event) with
   . . .
   source matching (a location)
   modality = 1
   destination matching (a location)
   modality = 1
   . . .)

(a flight is (a movement-event) with
   . . .
   source typically (a city)
   destination typically (a city)
   stopovers typically (a city)
   modality > = 0
   . . .)
```

Chicago would be defined to be an instance of the concept CITY which, in turn, would be defined to be a subconcept of LOCATION. Thus, the strongest relationships are found between the concepts FLIGHT and CITY—those in which Chicago is seen as filling one of the three roles *source, destination,* and *stopovers.* The fact that the earlier text specifically mentioned flights that stopped in Chicago does not play a part in selecting or ranking the candidate interpretations.

A natural language understanding system using a representation system such as this would, in the course of processing the text of our example, add to the lattice a concept representing "flights stopping in Chicago." This concept might look something like the following:

(a flight32 is (a flight) with
stopOver = Chicago)

Adding this new concept to the lattice adds another connection between the concept FLIGHT and the concept CITY, namely, that there is a particular flight (FLIGHT32) that has a particular city (Chicago) as the filler for its *stopover* role. By changing the strategy for finding candidate interpretations slightly we can find this new connection:

> Candidate interpretations for a nominal compound with constituents denoting CONCEPT1 and CONCEPT2 can be found by considering all potential relatinships between CONCEPT1 and its generalizations *and specializations* and CONCEPT2 and its generalization *and specializations*.

In order to work this into the scheme developed previously we need to cause such a match to be recognized and provide a method to assign a score to the match. This can be easily accomplished by adding a characteristic scoring function for the *value* facet. Furthermore, we can adjust this function to give more or less weight to the "discourse-bases" interpretations.

Generating Nominal Compounds

The system described in the foregoing can be adapted for use by a natural language generation system. We can use such a system as a critic for proposed nominal compounds. If the language generation system proposes to realize a constituent as a compound, then the proposed compound can be analyzed by the interpretation system. The result would be a list of all interpretations whose strength lay above the threshold. Criteria can be developed for deciding whether or not the proposed compound is appropriate. For example, we might choose to use it only if the meaning it as intended to convey is selected by the interpreter as the most likely one. Furthermore, we might require that the distance between its strength and the strength of the next best interpretation be greater than a certain threshold.

SUMMARY

This chapter discusses one approach to the task of interpreting nominal compounds. A nominal compound is a sequence of two or more nouns or nominal adjectives (i.e. nonpredicating) related through modification. The concepts denoted by the nouns (and the compound) are expressed in a frame-based representation system. The knowledge that drives the interpretation comes from the knowledge of the concepts themselves and from three classes of interpretation rules. Examples of the most general class of interpretation rules have been given. The basic approach can be extended slightly to take the discourse context into account.

REFERENCES

Brachman, R. A. (1978). *A structured paradigm for representing knowledge* (Tech. Rep. No. 3605). Cambridge, MA: Bolt Beranek & Newman.

Downing, P. (1977, December). On the creation and use of English compound nouns. *Language, 53* (4).

Finin, T., Goodman, B., & Tennant, H. (1979). JETS: Achieving completeness through coverage and closure. *Proceedings of the Sixth International Joint Conference on Artificial Intelligence.* Tokyo, Japan.

Finin, T. (1980a). *The semantic interpretation of compound nominals* (Tech. Rep. No. T–96). Champaign-Urbana: University of Illinois, Coordinated Science Laboratory.

Finin, T. (1980b). *The semantic interpretation of compound nominals.* Proceedings of the First National Conference on Artificial Intelligence. Stanford, CA: American Association of Artificial Intelligence.

Gershman, A. V. (1977). Conceptual analysis of noun groups in English. *Proceedings of the Fifth International Joint conference on Artificial Intelligence,* Cambridge, MA.

Grosz, B. (1982). Focusing and description in natural language dialogues. In Joshi, Webber (Eds.), *Elements of discourse understanding.* Cambrige, UK: Cambridge University Press.

Lees, R. B. (1960). *The grammar of English nominalizations.* Bloomington: Indiana University Press.

Levi, J. N. (1979). *The syntax and semantics of complex nominals.* New York: Academic Press.

McDonald, D. & Hayes-Roth, F. (1978). Inferential searches of knowledge networks as an approach to extensible language understanding systems. In Waterman & Hayes-Roth (Eds.), *Pattern-directed inference systems.* New York: Academic Press.

Rhyne, J. R. A lexical process model of nominal compounding in English (1976). *Computational Linguistics.*

Roberts, B. R. & Goldstein, I. P. (1977). *The FRL manual* (Tech. Rep.No. MIT-AI-409). Cambridge, MA: M. I. T., Artificial Intelligence Laboratory.

Waltz, D. (1978). An English language question answering system for a large relational data base. *Communications of the Association for Computational Machinery, 21,* 526–539.