

# General Lexical Representation for an Effect Predicate

Martha Palmer

*National University of Singapore,  
Department of Information Systems and Computer Science  
Lower Kent Ridge Road  
Singapore  
mpalmer%nusdiscs.bitnet*

## Abstract

This paper argues that there is no reason to distinguish between lexical information and real-world information on the basis of the formalisms used; that both types of knowledge can be expressed in the same formalism. However, it also argues that there is information that is uniquely lexical in content, and this information consists of verb-independent definitions for thematic relations such as *cause* and *effect* that alter the representation of a verb depending on the presence or absence of certain verb arguments.

## 1 Introduction

When considering the question of whether or not lexical information can be separated from real-world knowledge, it must first be recognized that most lexical information in a natural language processing system is expressed in terms of properties defined in a knowledge base representing the domain model - i.e., the "world," whether real or otherwise. For example, semantic class constraints consist of properties such as *animate* or *machine-part* which are associated with generic entities. Semantics is often described as the link between syntax and the knowledge base, so by definition it must have one foot in the knowledge base camp. In the absence of concrete examples of lexical facts or rules that are provably more difficult to represent than real-world knowledge, it also seems reasonable to assume that the knowledge representation formalism used for the domain model could serve to model the associated lexical information. This is not intended to imply that it will be possible to fully represent all of both types of information.. Formalisms for knowledge representation have been acknowledged to be inadequate for representing real-world knowledge, let alone lexical information [1]. It does make it especially burdensome to prove that a formalism that is inadequate for real world knowledge is even more inadequate for lexical information.

The question remains, is there lexical information which is uniquely "lexical" in that it can be distinguished from real-world knowledge albeit it may be expressed identically? This paper argues for the affirmative response. Given the approach to semantic representation implemented in PUNDIT and KERNEL [2], [3], [10] which is closely aligned to Jackendoff's lexical conceptual structures [4], [5], it is argued that the definitions of thematic relations such as *cause* and *effect* correspond to "uniquely lexical" items of information.

## 2 Reviewing the effect definition

It has already been proposed, [7], [9], that the thematic relation **effect** can be formalized as a predicate which takes two arguments, an Intermediary and another predicate. This is briefly summarized here. Examples of three verbs from a domain of pulley word problems that include this definition in their lexical conceptual representations are *connect*, *hang* and *suspend*.

### 2.1 Contact verbs

We will first look at *attach* and then contrast it with *connect*. An example sentence from the pulley domain is *A particle is attached [to/at] the end of a string*. The final semantic representation of this sentence needs to include a fixed, direct contact between the *particle* and the *end* of the string. This is accomplished by mapping the referent for the *particle*, **particle1**, onto one Patient role, and the referent for the *end*, **end1** onto the second Patient role in the following lexical conceptual structure: <sup>1</sup>

attach-:  
contact(Patient1,Patient2))

The semantic representation resulting from this lexical conceptual structure becomes **contact(particle1,end1)**. At first glance, *connect* seems to be quite similar to *attach*, but the possible inclusion of an Intermediary makes an important difference. Examples of *connect* as used in the pulley domain include: *A particle is connected to the end of a string*, *A string connects two particles*, *A particle is connected by a string to another particle...* In all of these phrases, the final semantic representation should include a fixed contact between location points on each of the two objects being connected to each other. In the first phrase, *A particle is connected to the end of a string*, the *particle* and the *end* are being attached directly to each other. In the second and third phrases, *A string connects two particles* and *A particle is connected by a string to another particle*, the two *particles* are being indirectly linked with each other by the use of an Intermediary, the *string*. There is still a fixed contact involved, in fact there are two fixed contacts, but they are between the *particles* and the ends of the *strings*, not between the *particles* themselves. This use of a *string* as a Intermediary to effect a *connection* between two objects is captured by the following general definition of *connect* which can produce distinct representations for the different usages of *connect* given in the examples here.

connect -:  
effect(Intermediary),  
contact(Patient1,Patient2))

Instantiating this definition by the referents to the noun phrases from *A string connects two particles* results in:

---

<sup>1</sup>The use of standard thematic role names in the verb definitions is not intended to mislead the reader into thinking that the arguments to the predicates correspond directly to traditional thematic roles. They are simply arguments to predicates, and could just have easily been labelled Arg1 and Arg2, and have been labelled as such in various incarnations. However, the author finds it difficult to keep Arg1 and Arg2 straight, and has resorted to the thematic role names in the hopes of preserving some clarity around which argument is which.

effect(string1,  
contact(particle1,particle2))

At this stage, the representation does not capture the contacts between the *particles* and the *string*. To achieve that it is necessary to apply the following linguistic definition of the **effect** thematic relation that is independent of the verb, i.e., it applies to more than one verb:

effect(Intermediary),  
Relation(Arg1,Arg2) :-  
    Relation(Arg1,Intermediary) AND  
    Relation(Intermediary,Arg2)

Matching this definition to our instantiated lexical conceptual structure from above replaces the general Relation predicate with **contact**, Intermediary with **string1**, Arg1 with **particle1** and Arg2 with **particle2**. It results in:

effect(string1),  
contact(particle1,particle2)) :-  
    contact(particle1,string1)  
    contact(string1,particle2)

This demonstrates that the general definition given above, a classic definition of logical transitivity (not linguistic transitivity!), can capture the essence of the meaning of the **effect** thematic relation in that the Intermediary **effects** an indirect contact between two objects by being in direct contact with each of them.

The principle difference between *connect* and *attach* is the possible inclusion of the Intermediary. When the Intermediary in *connect* is not present, it is not assumed, and the desired representation is exactly the same as the representation for *attach*. This can be captured by a definition such as the one given below:

effect(absent,  
Relation(Arg1,Arg2) :-  
    Relation(Arg1,Arg2)

Given our first example phrase, *A particle is connected to the end of a string*, the first representation produced would be:

effect(absent,  
contact(particle1,end1))

The application of the **absent** definition strips off the **effect** predicate, and the resulting representation is identical to the *attach* representation given above, i.e., **contact(particle1,end1)**. This allows us to produce two quite different semantic representations for the different usages of *connect* in our example sentences. <sup>2</sup>

---

<sup>2</sup>This would also be consistent with the view that the different usages are actually separate lexical items, which should receive separate semantic representations, as in fact they do here. Then the principal

The verb-independence of the linguistic definition for **effect** is substantiated by applying it to *hang* and *suspend* as described in the following section.

## 2.2 Support verbs

Two principal **support** verbs in the pulley domain are *hang* and *suspend* which are used quite similarly as in *A particle is hung from a pulley by a string*, *Two particles are suspended from a pulley*.<sup>3</sup>

The lexical conceptual structure for *hang* and *suspend* is similar to *connect* and is given below:

```
hang -:  
  effect(Intermediary,  
        support(Patient, Source))
```

The first phrase, *A particle is hung from a pulley by a string* is treated almost identically to *A particle is connected to another particle by a string*. The *string* is the Intermediary, the *particle* is the Patient and the *pulley* is the Source, giving us:

```
hang -:  
  effect(string1,  
        support(particle1, pulley1))
```

The same linguistic definition for **effect** is applied, this time to **support** predicates instead of **contact** predicates, and the next stage of the representation is as follows:

```
effect(string1,  
       support(particle1, pulley1)) -:  
  support(particle1, string1),  
  support(string1, pulley1))
```

In other words, the **support** of the *particle* is the *string* and the **support** of the *string* is the *pulley*. The important point here is that the same verb-independent linguistic definition of the **effect** thematic relation as a logical transitive relation that applies to *connect* also applies to *hang*. In this case an Intermediary **effects** a support relationship between two objects by being supported by one of the objects and in turn supporting the other.

However, there are differences in how this definition is applied to the **support** verbs. With *connect*, if the Intermediary was **absent**, it was assumed it did not exist. With these **support** verbs, if it is absent it is assumed that it does exist. This is captured by defining the Intermediary in the context of a **support** verb as an *essential* role whose

---

role of the lexical conceptual structure given here becomes that of determining which lexical items is being used.

<sup>3</sup>In the pulley domain, the basic predicate for *hang* and *suspend* is actually location, and the presence of gravity allows the **support** relationship with the corresponding upwards and downwards forces to be inferred. This has been changed to **support** in order to focus the discussion on the transitivity issues. This is not intended to imply that location would not be equally central to a more general definition of *hang*.

filler can be hypothesized [8]. This prevents the **absent** definition from applying, and the same definition that applied in our first example also applies in the second example. So, in *Two particles are suspended from a pulley*, the following instantiations occur, where `string57` is an hypothetical string whose existence is assumed

```
hang -:  
    effect(string57,  
           support([particle1, particle2], pulley1))
```

This is expanded similarly to the expansion above, resulting in:

```
effect(string57,  
       support([particle1, particle2], pulley1)) -:  
    support([particle1, particle2], string57)),  
    support(string57,pulley1))
```

This is intended to represent the information that the two *particles* are supported by the *string*, and the *string* is in turn supported by the *pulley*.

### 2.3 Summary

This section has argued that the logical transitivity definition for the **effect** thematic relation represents knowledge that is uniquely lexical. It is derived from real world knowledge and is expressed in the same formalism that real world knowledge is expressed in, but is specifically concerned with the effect of the presence or absence of a particular type of verb argument.

## 3 An extension of the effect predicate

This section extends the previously defined notion of the **effect** thematic relation as summarized above by modifying it so that it can apply to the verb *shoot* as well. This provides further evidence for the argument that there exists a category of knowledge pertaining to the impact occasioned by the presence or absence of a particular thematic relation and that this category of knowledge is essentially lexical in nature.

Standard flat predicate argument structure representations have had difficulty with *shoot*:

John shot the turkey with a rifle.

John shot the turkey with a bullet.

John shot the turkey with a bullet from a rifle.

John shot the turkey with his rifle (of?) a bullet.

because of the complexity of the event (or events) being represented. It has been argued that both a **launch** event and a **contact** event take place with the rifle being the Instrument of the **launch** of the bullet and the bullet being an Instrument by which John **effects a contact** with the turkey [6]. A general definition for this type of verb, which might also provide a core representation for other meanings or uses of *shoot* as well as similar verbs such as *fire* is suggested below. This definition makes use of an **effect** thematic relation very closely aligned to the notion of transitivity defined above, but this time used with a Propellant, the object being set into motion, as well as an Intermediary.. The central notion is once again one of logical transitivity, this time applied twice.

shoot -:

```
cause(Agent,
      effect_launch(Intermediary, Propellant,
                    apply_force(Agent, Goal)))
```

Given the sentence *John shot the turkey with a bullet from a rifle*, John is the Agent and the turkey is the Goal. The rifle is the Intermediary and the bullet is the Propellant. This would result in:

shoot -:

```
cause(John,
      effect_launch(rifle, bullet,
                    apply_force(John, turkey)))
```

This is intended to represent John as the Agent of an event in which he uses a rifle (which propels a bullet) to apply-force against a turkey. The representation of this event can be further expanded by the application of the following definition:

```
cause(Agent,
      effect_launch(Intermediary, Propellant,
                    apply_force(Agent, Goal))) -:
      apply_force(Agent,Intermediary) AND
      apply_force(Intermediary, Propellant) AND
      move(Propellant,Intermediary,Goal) AND
      apply_force(Propellant, Goal)

      apply_force(John,rifle) AND
      apply_force(rifle, bullet) AND
      move(bullet,rifle,turkey) AND
      apply_force(bullet,turkey)
```

These predicates represent John's application of force to the rifle, (by pulling the trigger), which then applies force to the bullet. The bullet moves from the rifle to the turkey, and applies force to the turkey. Similarly to *hang*, even if the Intermediary is not linguistically realized, it still needs to be hypothesized, as does the Propellant. If the minimal sentence of John shot the turkey is used, then the specifics of the Intermediary

and the Propellant have to be left undefined, but it is clear that some Intermediary and some Propellant (perhaps a bow and arrow?) have been used. This is certainly not intended to be a complete representation of the event - the trigger is never mentioned explicitly - and may need to be modified substantially before it could be used. But it is intended to support the hypothesis that logical transitivity, as an inherent component of the meaning of the effect thematic relationship, adds a uniquely lexical component to verb representations.

## 4 Conclusion

This paper argues that the definitions of the effect predicates presented here, as well as the optionality or essentialness of the roles, is a necessary part of the lexical semantics of the verb, which may rely heavily on real-world knowledge about launching Propellants, yet is still essentially lexical as evidenced by its being independent of the context. This in turn supports the claim that, while lexical information and real-world knowledge may share the same representation formalism, a distinction can still be maintained between information which is fundamentally lexical in nature, and information which is not.

**Acknowledgements** I would like to thank Tara Mohanan for the insightful comments she made with respect to the representation of *shoot* and her comments on a draft of the paper. I would also like to thank Alain Polguere for his comments.

## References

- [1] James Allen, "Natural Language, Knowledge Representation and Logical Form", *BBN symposium, Natural Language Processing: Language and Action in the World*, Nov, 1989.
- [2] Deborah A. Dahl and Martha S. Palmer and Rebecca J. Passonneau, "Nominalizations in PUNDIT", *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, Stanford University, Stanford, CA, Jul, 1987
- [3] L. Hirschman and M. Palmer and J. Dowding and D. Dahl and M. Linebarger and R. Passonneau and F-M. Lang and C. Ball and C. Weir, "The PUNDIT Natural-Language Processing System", *AI Systems in Government Conf*, Computer Society of the IEEE, Mar, 1989
- [4] R. S. Jackendoff, *Semantics and cognition*, MIT Press, Cambridge, Mass. 1983
- [5] R. S. Jackendoff, *Semantics Structures*, MIT Press, Cambridge, Mass. 1990
- [6] B. Levin, *Instrumental with and the control relation in English*. MIT AI Memo 552, MIT, MIT Master's Thesis, 1979.
- [7] M. Palmer, "A case for rule-driven semantic analysis", *Proceedings of the 19th Annual Meeting of the Association for Computational Linguistics*, Stanford, CA, Jun, 1981

- [8] Martha S. Palmer and Deborah A. Dahl and Rebecca J. [Schiffman] Passonneau and Lynette Hirschman and Marcia Linebarger and John Dowding, "Recovering Implicit Information", *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, Columbia University, New York, Aug,1986
- [9] M. Palmer, *Semantic Processing for Finite Domains*, Cambridge University Press, Cambridge, England,1990
- [10] Rebecca Passonneau and Carl Weir and Tim Finin and Martha Palmer, "Integrating Natural Language Processing and Knowledge Based Processing", *Proceedings of Eighth National Conference on Artificial Intelligence, AAAI 90*, 1990