In *PACLING-93, 1st Conf. of the Pacific Association for Computational Linguistics*, Vancouver: Apr. 93, 323–332.

An Image-Schematic System of Thematic Roles

Dekai Wu*
Department of Computer Science
The Hong Kong University of Science & Technology
Clear Water Bay, Hong Kong
dekai@cs.ust.hk

Abstract

We describe a system of thematic roles and frames designed to address a number of problems in semantic representations at the lexical semantic level. Our primary objective is broad expressiveness, so that real domains can practically be encoded. However, for both empirical and computational reasons we limit the number of role types to four, allocating this structure to the strongest associations. We show how the system incorporates image-schematic semantics to encode various schematization operations relating to scales and reification.

Keywords: semantics, cognitive modelling

1 Introduction

Anyone attempting to build a natural language processing model for any application beyond syntactic parsing faces a dilemma in choosing a meaning representation. On the one hand, traditional knowledge representation languages such as predicate logic, frames, or semantic networks permit arbitrary flexibility in choosing the predicates or types. This makes it difficult to compare or falsify theories since any small-scale domain can be accomodated by careful choice of predicates. On the other hand, linguistic theories of semantics are generally constructed under Occam's principle, using minimal pair methodology; the resulting semantic theories are overconstrained such that in nearly any application domain, the semantic primitives' expressiveness is rapidly exceeded.

The thematic role system described here is a part of our larger program to define a broad-scope set of ontological primitives for mental representation. The ontology is designed for AI and NLP applications and therefore expressiveness is of primary importance rather than Occam's razor; i.e., excessive representational flexibility is preferable to inexpressible semantic distinctions. Within this constraint, we maintain consistency as far as possible with the available linguistic, cognitive, and neuropsychological evidence. The hope behind this proposal is that by providing an ontology that is more constrained than logic or frames but still sufficiently expressive for language understanding models, larger-scale systems can be built on tighter, firmer foundations.

That said, it may be helpful to further delimit the approach within two well-known mileposts: CD and CYC. In a sense the proposal is reminiscent of that of CD or conceptual dependency (Schank 1973), in that the representation declares the set of permissible predicate (role) types. It diverges from CD in that there is no insistence on a canonical form, and there is a far greater range of predicates than the 13 or so permitted in CD. However, the approach does not go so far as CYC (Lenat & Guha 1988, 1990) in proliferating predicates. The CYC methodology attempts to encode much of human knowledge without distinguishing semantic levels of representation from the general conceptual system. In the proposed system there is a definite semantic level of representation that is constrained relatively strongly in the variety of roles, distinct from the general conceptual pool of frame types.

The role system has evolved through iterative refinement, while coding representations for hundreds of nominal compounds in the course of developing an interpretation model Wu (1992a, 1992b, 1993). Though of course it will continue to evolve, it has been our experience that the representation as it stands is versatile and quite practical.

*Preparation of this paper was partially supported by the Natural Sciences and Engineering Research Council of Canada while the author was a postdoctoral fellow at the University of Toronto. Much of this research was done at the Computer Science Division, University of California at Berkeley and was sponsored in part by the Defense Advanced Research Projects Agency (DoD), monitored by the Space and Naval Warfare Systems Command under N00039-88-C-0292, the Office of Naval Research under N00014-89-J-3205, the Sloan Foundation under grant 86-10-3, and the National Science Foundation under CDA-8722788.

2 Some Questions for a Semantic System

Subject to the goal of constructing a broadly expressive semantic system, we have attempted to address several recurrent representational issues in the design of the thematic role and frame system. The issues include:

- *How many role types?* Traditionally, AI permits a large range of role types reflecting conceptual relationships. In some representation languages, arbitrarily many types of roles are permitted. However, this can unnecessarily increase the size of the concept space, with corresponding growth in search times. Suppose there is some arbitrary number r of role types, and no restriction is placed on the combination of roles particular frames can have. If each frame can have up to b roles and frames can be nested to depth d, then the concept space includes $O(rcb^d)$ possible structures, where c is the number of different frame types (i.e., feature combinations). Clearly we do not want an excessively large rc that includes too many unused frame/role combinations.

  Moreover, as statistical language models gain prominence, the difficulty of representing roles in feature vectors becomes a more acute problem. It is a reasonable speculation that concept formation increases in difficulty more rapidly with role complexity than feature complexity. Thus, representation methods that reduce the complexity of roles stand to gain.

- *Image-schematic regularities.* A number of proposals for *image-schematic* semantics deal with grounding semantics in human perceptual experience (Talmy 1983, 1985, 1988; Lakoff 1987; Langacker 1987). Typically the fundamental roles in these systems are one or two *landmarks* (notated LM in this paper), a *trajector* (TR), and a *trajectory* (TRY). Though these studies must be regarded as extremely preliminary, they are attractive for the subregularities that can be captured, particular in explaining metaphoric usage patterns. However, image-schematic semantics has not been formalized for incorporation into a broad-coverage representation.

- *Scales; discretization and scalarization.* Existing representations do not generally deal with scales. A *scale* is a concept against which relative positions or intervals can be defined, and is the basis for representing degrees such as weight or temperature. Scales may be bounded at both ends (fullness; sometimes health, temperature), unbounded at one end (weight, cost, speed; sometimes health, temperature), or unbounded at both ends (beauty, poverty). A number of operations relate to scales.

  Discretization is an operation whereby a new discrete property is defined, relative to some interval on a scale.

  Scalarization occurs when a scale is defined by turning an ordinarily discrete property or object into a matter of degree (Wilensky 1989). For example, normally someone is either pregnant or not, but in the sentence

  (1) The cover photo of an unclad and very pregnant actress caused a media overreaction.

  the property *pregnant* is converted into a degree.

- *State and property reification* In knowledge representation terminology, to reify is to view an abstract or implicit relation as a concrete concept, node, or object. Several kinds of reification are generally overlooked but are quite useful. *State reification* turns a concept such as *red apple* into a state frame *being a red apple* (we discuss below what a state frame is). In contrast, *property reification* turns the *red apple* concept into *apple having a red color*.

The following section describes a semantic system that proposes some solutions to these issues.

3 A Thematic Role and Frame System

To help control the growth of the rc factor in concept space size, the proposed thematic role system accomodates only four basic role types. Because the roles may serve different purposes from one frame to another, this is actually another way of saying that any frame can have a maximum of four *different* types of roles. Thus the effective type of a role is context-sensitive.

The four role types are also motivated by empirical considerations. We wish to use the each basic role type to capture the strongest possible generalizations. Since it is difficult to capture generalizations across all of knowledge using only four roles—even less than in an average case role theory—the core semantics of these roles must be quite abstract. The four core thematic roles used are SOURCE, PATH, GOAL, and LOCAL, after Somers's (1987) case grid theory. Each of these roles takes on specialized purposes depending on the frame type; the indentation of the frame types indicates the hierarchical specialization. As shown in figure 1, when a role's purpose is sufficiently specialized we assign it an alternate, more mnemonic name.

Thus, the assignment of specialized roles to one of SOURCE, PATH, GOAL, or LOCAL is intended to reflect patterns of usage across different frame types. This is the only significance of the core thematic roles' names, which otherwise could be left as 1, 2, 3, and 4: the labels are intuitively suggestive of the core semantics common to all the specializations. That four is a good number is purely an

Frame type	**Role 1**	**Role 2**	**Role 3**	**Role 4**
eventuality	SOURCE	PATH	GOAL	LOCAL
state				
lm-tr	-	TRY	LM	TR
containment	-	BOUNDARY	CONTAINER	CONTENT
part-whole	-		WHOLE	PART
isa-category	-		CATEGORY	MEMBER
peripheral	-	BOUNDARY	CENTER	PERIPHERY
possession	-	-	POSSESSION	POSSESSOR
has-property	-	-	PROP	TR
lm-lm	LMONE	-	LMTWO	-
link	LMONE	LINK	LMTWO	-
lm-tr-lm	LMONE	TRY	LMTWO	TR
linear-order	LMONE		LMTWO	
locative				
temporal				
value				
front-back	BACK		FRONT	ENTITY
flank	LEFT		RIGHT	ENTITY
above-below	BELOW		ABOVE	ENTITY
state-schema				
opposition	OPPST:state		ST:state	TR
nonbinary	STSET		ST:state	TR
discretization		MIDPT		TR
scalarization				
bounded	BEGPT		ENDPT	
locative				
vert	DOWNPT		UPPT	
horiz	LEFTPT		RIGHTPT	
quant	LEAST		MOST	
positive	BEGPT		ENDLM	
...				
negative	BEGLM		ENDPT	
...				
unbounded	BEGLM		ENDLM	
...				
quant	LESS		MORE	
event				
accomplishment				
locative-acc	ORIGIN:loc	TRY	DESTINATION:loc	PATIENT
achievement				
locative-ach	ORIGIN:loc		DESTINATION:loc	PATIENT
*acc/ach**				
causal	CAUSER	MEANS	EFFECT:s/e	PATIENT
volitive	AGENT	PLAN	INTENT:s/e	PATIENT
instrumental	INSTRUMENT	MEANS	EFFECT:s/e	PATIENT
psychological	STIMULUS	CAUSE	EXPERIENCER	EXPERIENCE
possessive	GIVER	CAUSE	RECIPIENT	PATIENT
value-ev	INITVAL:val		FINALVAL:val	PATIENT
state-change	INITST:state		FINALST:state	PATIENT
loc-st-chg	INITST:loc-st		FINALST:loc-st	PATIENT
val-st-chg	INITST:val-st		FINALST:val-st	PATIENT
process				
activity				EV:event

Figure 1: Overview of thematic role system. TR stands for trajector, TRY for trajectory, and LM for landmark (schemas with two landmarks have LMONE and LMTWO). A dash indicates that a role cannot meaningfully be filled in that frame; in contrast, blanks leave open the option of merging two frames providing their roles do not overlap (see figure 3 for example). Where specified, the label to the right of a colon indicates the expected type of filler. *An acc/ach frame is an accomplishment if PATH is a process, and an achievement if PATH is an event.

empirically-motivated claim, since theoretically one could minimize the number of role types to just 1 and 2, and then reify all roles to binary frames; but this would not capture any subregularities between specialized roles.

As examples of role subregularities across frame types, consider the following:

- The EXPERIENCE (*grief*) and PATIENT (*roses*) roles both appear frequently in the accusative as in *He caused her grief* and in *for giving him roses of the wrong color.*
- The fact that EXPERIENCER and RECIPIENT are both GOAL roles surfaces in grammatical subregularities like their use of dative case—*She caused him grief* and *She gave him roses*—or the corresponding *to-PP*s—*She caused grief to him* and *She gave roses to him.*[1]
- That both EXPERIENCER and LMTWO (the second landmark in image schemas with two landmarks, like *West Coast* in the locative *They moved from the East Coast to the West Coast*) are both GOAL roles captures the shared use of *to-PP*s, though the locative does not permit the dative.
- The PATIENT in a *location state change* frame is semantically very close to the ENTITY role in a *flank* frame.
- Polarization tendencies like the fact that *down* is typically considered *less* are captured by subsumption under the SOURCE role.

In other words, we are attempting to allocate the limited amount of structure we permit ourselves to the strongest possible associations. The correspondence is by no means perfect and there are certainly many pairs of specific SOURCE role types that by themselves predict no surface correlations. However, the fact that they share surface grammatical predictions with other related roles constitutes evidence for an indirect "family resemblance" similarity in the way the roles are cognitively represented. The labels are merely intended to convey some notion close to a core significance.

Following Bach's (1983, 1986) work on aspect, every frame is an *eventuality* and falls into three major classes: *state*, *event*, and *process*. States describe non-changing situations. They do not intrinsically include any information on duration, and can hold for arbitrary lengths of time. A state has no culmination or intrinsic termination.[2]

In contrast, all events have intrinsic termination points, so one can sensibly ask whether an event is "finished". Usually events also have intrinsic, definite culmination points. Events are broken down into two major classes: the *accomplishment* and the *achievement*. An achievement is a happening that is conceived of as an instantaneous event such as *hit*, *stop*, *win*, or *stumble*. One cannot meaningfully ask for the duration of an achievement. In contrast, an accomplishment is an event that is conceived of as an event whose progress spans some temporal interval.

Many types of event frames can be used for either accomplishments or achievements. In all event frames except locative events, the PATH role specifies the means by which a state change is brought about. If the semantic substructure that fills the PATH role is a process (see below) then the event frame is taken to be an accomplishment. On the other hand, if the substructure is another event, the event frame is taken to be an achievement.

Locative events are statically schematized versions of spatiotemporal (motion) images. The representation currently has no provision for spatiotemporal images. If a representation were developed, the method of schematization would be what Langacker (1987, p. 145) calls "summary scanning". The spatial path taken by the trajector/patient over time is converted to a static, directed curve containing all and only those points passed through. Such a mechanism has been implemented in a connectionist model to handle motion prepositions in the L_0 project (Regier 1992).

The notion of agency is accounted for by wrapping a causal event around some effect (state or event). This can be glossed by saying that *skater* is designated the agent in

(2) The skater rolled down the hill.

by using the semantic interpretation *the skater caused the skater to roll down the hill.* This is a different resolution of the problem that Jackendoff (1972) noted with sentences whose subject functions both as theme and agent. Jackendoff argued that thematic relations, unlike case systems, do not constrain each noun to only one role, and thus more elegantly capture the similarity of the above sentence to

(3) The tire rolled down the hill.

---

[1] Many people find *He caused grief to her* strange, yet do find the construct acceptable with heavy NPs as in *He caused grief to anyone foolish enough to get involved with him* or *He caused great and severe grief to her* (Charles Fillmore, personal communication).

[2] This may be confusing since one might argue that *owning a 57 Chevy from 1963 to 1966* is a state. Note that this phrase might mean two different things, however, one of which is an event and one of which is a state with no termination. First, in the terminated sense it is an event, a sense we might reinforce by changing the tense—*owned a 57 Chevy from 1963 to 1966*. Second, viewed as a state, someone who can be predicated *now* as *owning a 57 Chevy from 1963 to 1966* will *always* be predicated as such. In other words, the state sense has no termination, even though the event from which the state is reified contains a termination point.

In Jackendoff's proposal *skater* functions as both the theme and the agent of the same sentence frame. In contrast, here we distinguish the causal frame, with which the agent role is associated, from the effect frame, with which the theme role is associated. Both the causal frame's agent role and the effect frame's theme role are source roles in the case grid. This captures the generalization between the previous two sentences more cleanly, as well as for

(4) The skater rolled the tire down the hill.

Another innovation here relates to the distinction between events and processes: the *discrete/mass* feature is used to mark this. Though the original motivation for the feature is primarily for *concrete* concepts, we have generalized it to events and processes which are *abstract* entities. Processes resemble events in being dynamic, changing situations. However, unlike either accomplishments or achievements, they are not intrinsically terminating. Whereas events have culminations giving them the *discrete* property, processes are indefinitely bounded entities, giving them the *mass* property.

The literature does not usually differentiate the terms "process" and "activity". We define *activity* as a special case of *process* that is created by schematizing an indefinite sequence of event repetitions as a process. For example, the nominal compound *hurricane season*[3] requires schematizing an indefinite set of *hurricane* events as an activity that occurs during hurricane season. The schematization of events into activities is represented as a special case of *particle mass* constituency, further details of which can be found in Wu (1992b).

Differences from Case Theories The proposed representation depends more heavily on nested structures than ordinary case systems, thereby reducing the number of thematic roles. A side effect of this is that surface cases do not map straightforwardly onto thematic roles in a one-to-one fashion. For instance, the sentence

(5) John moved his furniture from San Francisco to Berkeley.

is analyzed as an outer *volitive* frame whose EFFECT role is filled by an inner *locative*. Thus we control proliferation of cases by defining the less immediate cases as internal roles. Similarly,

(6) John broke the window with a hammer.

is analyzed as an outer *volitive* frame whose PATIENT is the *window*, EFFECT is a *window breaking state change*, and PLAN is an inner *instrumental* frame, whose EFFECT role is in turn filled by the same *window breaking state change*. No single frame can have both AGENT and INSTRUMENT roles; a verb frame involving both must use nested semantic structures. For convenience we can use a chain notation, so that the traditional instrument case role becomes PLAN.INSTRUMENT. Similarly, the beneficiary case role is equivalent to INTENT.PATIENT.

Note that INSTRUMENTs are considered to be CAUSERs, just like AGENTs except for being nonvolitive. Thus in our analysis the hammer in sentence (6) is an instrumental CAUSER internal to the PLAN of the breaking event, and the internal frame filling the PLAN role is itself the semantic structure for

(7) A hammer broke the window.

where the hammer is the non-volitive cause. Omitted here is John's precise action upon the hammer; to represent this one would have to add a third *volitive* frame in which the AGENT is John and the PATIENT is the hammer. Of course we would expect a high degree of statistical association between this *volitive use of tool* frame and the *instrumental* frame.

We considered motivations for restricting frame roles earlier; we now add another *peripheral semantic role* motivation. There are analogous precedents in syntactic theory for the philosophy of narrowing the number of direct roles of a frame. Various theories of grammatical cases treat clauses as layered structures with a central and peripheral argument structure. In these theories the clause consists of one or two core arguments surrounding the predicate, typically labelled "subjects", "direct objects", "actors", "undergoers", and so forth. In addition a larger number of peripheral arguments express spatiotemporal information or secondary participants such as "beneficiaries". The distinction has variously been termed "core" versus "peripheral" (Silverstein 1976; Foley & van Valin 1984; Foley & Olson 1985), "inside the vp" versus "outside the vp" (Fillmore 1968), "nuclear" versus "satellite" (Dik 1978), "inner" versus "outer" (Halliday 1970; Platt 1971; Somers 1987), "nuclear" versus "peripheral" (Longacre 1976), and "propositional" versus "modal" (Cook 1972). Though some of the predicate argument classes are characterized using semantic notions, the theories are syntactic and do not extend the core-periphery claims to the semantic representation. The proposed representation does exactly that: instrumental roles, for example, are postulated to be semantically more peripheral than agent roles if the verb expresses a volitive rather than non-volitive causal event. Note that this does not present any verb valency problems since nested structures are still available for specifying the linking between verb clauses and their underlying semantic frame(s). The core-periphery

[3]Our study employs data from the Brown Corpus (Kučera & Francis 1967; Warren 1978); *hurricane season* is from Warren (1978, p. 246).

	Source	*Path*	*Goal*	*Local*
Active	instigator of action ±volitive	instrument or means	intended result (−animate) active recipient (+animate)	non-passive patient
Objective	original state (−concrete) material (+concrete)	counter-instrument passive means	result state (−concrete) factitive (+concrete)	undergoing change-of-state
Dative				
Psychological	stimulus	medium	experiencer (±dynamic)	content
Possessive	original owner	medium/price	recipient	thing transferred
Locative	place from where	space traversed	final destination	static position
Temporal	time since	duration	time until	time at which
Ambient	reason	manner	aim (+volitive) consequence (−volitive)	condition

Figure 2: Somer's (1987) case grid proposal.

contrast here applies to semantic frames—Parsons (1990) might say "subatomic"—rather than verb frames. Across many languages, including English, certain roles are more readily marked by word order while others retain case or prepositional markers.[4] A semantic explanation of why this pattern should evolve would be more satisfying than a purely grammatical one.

Differences from Case Grid Theories The proposed system is distinguished from previous case grid theories (Ostler 1980; Somers 1987) in the following ways:

1. The grid's vertical dimension designates frame types; for each frame used in a semantic representation, roles may not be chosen from other rows.
2. The categories in the vertical dimension are proliferated. Previous proposals employ only a small number of categories, typically less than ten.
3. The categories of the vertical dimension are specialized hierarchically.
4. A major class of *state* frames is defined. Case theories are typically only concerned with what is called here the *event* class of frames.
5. Image schemas are strongly incorporated into the case grid.

Certain distinctions can be made in more detail. First, Somers is unsure whether to interpret his grid as an elaborate model for inner cases only, or a sparse model of both inner and outer roles, but he leans towards the former. As we have just seen, here we propose to distinguish outer roles by hierarchical nesting, rendering the question moot. Second, Somers' grid, shown in figure 2, employs additional role features (although he qualifies them as "optional"), whereas the restriction to four role types is absolute in the proposed system. Third, in contrast to the case grids, activeness/passiveness is a distinction made orthogonally to the case system. There is no "eventive/agentive" (Ostler 1980) or "active/objective" (Somers 1987) separation within the role system. Somers sees Objectives as "processes" (a different usage than the eventuality sense) as opposed to Actives which are actions. In the proposed model, the *volitive* frame type distinguishes actions with volitional agents and all others are considered passive. There is no need to distinguish active and passive patients; Somers suggests handling reflexives by using Active Local as an "active patient", but reflexives can be handled equally well by coindexing the AGENT and PATIENT. Fourth, the function of a role cannot be ambiguous, as opposed to cells like Somer's Objective Path, which can indeterminately function either as a "counter-instrument" (enabler or non-tool-like instruments), or as the instrument in an agent-less event as in sentence (7).

4 Examples

All examples in this section use standard typed feature-structure notation, as described for example by Shieber (1986), except for the $\emptyset$ extension explained later; the notation is fully compatible with frame representations.[5]

The discretization mechanism operates by using the MIDPT and TR (trajector) roles of a predefined *scale* such as *automobile speed*, as shown in figure 3(a). The *reference property* is a predefined point or interval on the scale such as *high*. The *greater degree property* is then defined to be the

[4] The fact that other languages like Latin or Russian use case markers for a wider set of roles does not contradict the hypothesis that not all the cases encode semantic roles of the same status.

[5] Details are given in Wu (1992b).

$$\text{(a)}\quad \begin{bmatrix} \text{TYPE:} & \textit{scale} \\ \text{MIDPT:} & \textit{reference-property} \\ \text{TR:} & \textit{greater-degree-property} \end{bmatrix} \quad \begin{bmatrix} \text{TYPE:} & \textit{automobile-speed} \\ \text{MIDPT:} & \textit{high} \\ \text{TR:} & \textit{high-automobile-speed} \end{bmatrix}$$

$$\text{(b)}\quad \begin{bmatrix} \text{TYPE:} & \textit{scale} \\ \text{BEGPT:} & \textit{opposition-object} \\ \text{ENDPT:} & \textit{ideal-object} \end{bmatrix} \quad \begin{bmatrix} \text{TYPE:} & \textit{calmness} \\ \text{BEGPT:} & \textit{anxiety} \\ \text{ENDPT:} & \textit{calm} \end{bmatrix}$$

$$\text{(c)}\quad \begin{bmatrix} \text{TYPE:} & \textit{scale} \\ \text{BEGPT:} & \textit{opposition-object} \\ \text{MIDPT:} & \textit{reference-object} \\ \text{ENDPT:} & \textit{ideal-object} \\ \text{TR:} & \textit{greater-degree-property} \end{bmatrix} \quad \begin{bmatrix} \text{TYPE:} & \textit{calmness} \\ \text{BEGPT:} & \textit{anxiety} \\ \text{MIDPT:} & \textit{inverse-low} \\ \text{ENDPT:} & \textit{calm} \\ \text{TR:} & \textit{low-anxiety} \end{bmatrix}$$

Figure 3: (a) Discretization, (b) scalarization, and (c) similarity schemas, using the *scale* type and with examples.

property ranging over the part of the scale greater than the reference property. The resulting *high automobile speed* property can be used in the semantic representation of, for example, *high-speed buses*.[6]

Figure 3(b) shows the scalarization mechanism's use of the BEGPT and ENDPT roles of a *scale* to specify the endpoints or landmarks of the scale. The *ideal object* is an object that serves as the reference point for a greater degree on the new scale, for example *calm* on the *calmness* scale, and the *opposition object* serves as the reference point for the lesser degree, for example *anxious*. Like primitive scales, scalarized scales can be bounded on both ends, as in the example, or they can be unbounded on either or both ends. For an unbounded end, specifying a landmark indicates the directional point of reference, rather than an absolute endpoint of the scale.

The discretization and scalarization operations combine to form the *similarity* mechanism, which places discrete properties or objects on a scale to support a new discretization. In figure 3(c) the *calmness* scale from (b) is combined with a discretization operation, producing the new property *low anxiety* which is closer to the ideal end, *calm*. This property is used in the semantic representation for *low-anxiety child*[7], as shown in figure 4 which requires several points of explanation. First, the SEM label identifies the conceptual referent of the phrase, which in this case has been co-indexed to the *child*, while FRM holds the surrounding contextual structure. Second, the $\emptyset$ labels simply hold other contextual structures that are necessary but play no specified global role. Also note the polarity of the scale has been chosen to produce the desired property; were the endpoints reversed, the scale would be *anxiousness* and the property *low anxiety* could not be defined. This is not a limitation on expressiveness because scales are always permitted to be inverted.

[6]From Warren (1978, p. 148).

[7]The compound *low-anxiety children* is from Warren (1978, p. 148).

The similarity mechanism can also be used to derive "peripheralized" (Wilensky 1989) concepts like *reddish* which apply to objects that fall in the periphery of the category described by *red*.[8] As shown in figure 5, a scale is created between the *any red* and ideal *prototype red* concepts. The *reddish* property is then defined by the interval between *any red* and the reference point, which is a default or average point corresponding to *-ish* forms.[9]

Comparatives like *redder* are handled using the similarity mechanism; several variations are found in the reification examples in figure 6. In (a) a *redder red* property is defined to characterize the color of *a redder apple [than apple/2]*, where the *reference red* property is contextually defined in $\emptyset_1$ to be the color of some other apple *apple/2*. (The slash / notation differentiates extension.) An alternate interpretation is shown in (b), where instead of defining a redness scale, we define a "redness-of-apple" scale with a slightly different connotation.

Representing state and property reification
Figure 6(c) shows how the *redder apple* from (a) is turned into *being a redder apple* by embedding it in a *state schema*. The difference between a mental image state and a state eventuality cannot be

[8]Wilensky's use of the term "peripheralized" extends the periphery of a concept beyond its normal applicable range. The term is used here instead to mean the peripheral area of the applicable range, and a term like "peripheral extension" is suggested for the former.

[9]Consider, however, that sometimes *reddish* is extended to describe objects one would never call red, like *tangerines* in

(8) The study showed that consumers shun reddish tangerines.

To handle such cases it is necessary to instead relativize to a scale whose endpoints are *prototype tangerine* and *reddest tangerine*.

$$\begin{bmatrix} \text{SEM: } {}^{1} & \\ \text{FRM:} & \begin{bmatrix} \text{TYPE:} & \textit{causal} \\ \text{CAUSER:} & {}^{1} \\ \text{MEANS:} & \begin{bmatrix} \text{TYPE:} & \textit{habitual-volitive} \\ \text{AGENT:} & {}^{1}\textit{child} \end{bmatrix} \\ \text{EFFECT:} & \begin{bmatrix} \text{TYPE:} & \textit{psychological} \\ \text{EXPERIENCER:} & {}^{2} \\ \text{EXPERIENCE:} & {}^{3} \end{bmatrix} \\ \text{PATIENT:} & {}^{2}\textit{parent} \end{bmatrix} \\ \emptyset_1: & \begin{bmatrix} \text{TYPE:} & \textit{parental-link} \\ \text{LM1:} & {}^{1} \\ \text{LM2:} & {}^{2} \end{bmatrix} \\ \emptyset_2: & \begin{bmatrix} \text{TYPE:} & \textit{calmness} \\ \text{BEGPT:} & \textit{anxiety} \\ \text{MIDPT:} & \textit{inverse-low} \\ \text{ENDPT:} & \textit{calm} \\ \text{TR:} & {}^{3}\textit{low-anxiety} \end{bmatrix} \end{bmatrix}$$

Figure 4: Use of the defined *low anxiety* concept in the semantic representation for *low-anxiety child.*

represented in the simpler single-level *+/–dynamic* distinction between events and states proposed by Pinker (1989, p. 195) which is in turn based on Jackendoff.

Property reification is implemented by making a proper concept out of a (discrete) property role in a feature structure. Putting ^{1} or ^{2} in the SEM role in figure 6(d) determines whether it represents *apple having a redder color [than apple/2]* versus *apple which has a redder color [than apple/2]*. In either case, the property of having a red color is reified by the *has color schema* which is a specialized frame of the *has property* type.

5 Conclusion

We have found the thematic role and frame system's expressive coverage at this point to be quite broad. As discussed above, it incorporates image-schematic, aspectual, and case grid semantics to handle a number of subtler semantic distinctions. In our mental representation model, the case grid organization also serves the secondary purpose of providing primitive roles for the general conceptual system; it defines a combinatorically restricted AI-style role hierarchy. We believe the generalization performance of statistical language learning models should also be improved by the thematic role system's language bias. The system as it stands is practical and useful, and has substantially decreased the time and effort required to encode the semantics of new examples, while increasing consistency amongst representations for examples across a broad range of domains.

Acknowledgements

I am indebted to Robert Wilensky, Charles Fillmore, George Lakoff, Jerome Feldman, Graeme Hirst, Nigel Ward, and the members of the BAIR, L_0, and University of Toronto Computational Linguistics groups for comments and many valuable discussions.

References

BACH, EMMON. 1983. On time, tense, and aspect: An essay in English metaphysics. In *Meaning, use, and interpretation*, ed. by R. Bauerle, C. Schwarze, & A. von Stechnow. New York: de Gruyter.

BACH, EMMON. 1986. The algebra of events. *Linguistics and Philosophy*, 9:5–16.

COOK, W. A. 1972. A set of postulates for case grammar analysis. In *Languages and linguistics working papers 4*. Georgetown University. Reprinted in Cook 1979.

DIK, SIMON C. 1978. *Functional grammar*. Amsterdam: North-Holland.

FILLMORE, CHARLES J. 1968. The case for case. In *Universals in linguistic theory*, ed. by Emmon W. Bach & Robert T. Harms, 1–88. New York: Holt, Rinehart, and Winston.

FOLEY, WILLIAM A. & MICHAEL L. OLSON. 1985. Clausehood and verb serialization. In *Grammar inside and outside the clause*, ed. by Johanna Nichols & Anthony Woodbury. Cambridge: Cambridge University Press. Cited in Foley & van Valin 1984.

FOLEY, WILLIAM A. & ROBERT D. VAN VALIN, JR. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

HALLIDAY, M. A. K. 1970. Language structure and language function. In *New horizons in linguistics*, ed. by John Lyons, 140–165. Harmondsworth, Middx: Penguin.

JACKENDOFF, RAY S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

KUČERA, HENRY & W. NELSON FRANCIS. 1967. *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

$$\begin{bmatrix}\text{TYPE:} & \textit{scale}\\ \text{BEGPT:} & \textit{ideal-object}\\ \text{MIDPT:} & \textit{reference-object}\\ \text{ENDPT:} & \textit{object}\\ \text{TR:} & \textit{greater-degree-property}\end{bmatrix} \qquad \begin{bmatrix}\text{TYPE:} & \textit{reddishness}\\ \text{BEGPT:} & \textit{proto-red}\\ \text{MIDPT:} & \textit{default-ish-point}\\ \text{ENDPT:} & \textit{any-red}\\ \text{TR:} & \textit{reddish}\end{bmatrix}$$

Figure 5: Use of the similarity mechanism for peripheralization.

$$\text{(a)}\ \begin{bmatrix}\text{SEM:} & {}^{1}\\ \text{FRM:} & {}^{1}\begin{bmatrix}\text{TYPE:} & \textit{apple/1}\\ \text{COLOR:} & \textit{redder-red}\end{bmatrix}\\ \emptyset_1: & \begin{bmatrix}\text{TYPE:} & \textit{apple/2}\\ \text{COLOR:} & \textit{reference-red}\end{bmatrix}\\ \emptyset_2: & \begin{bmatrix}\text{TYPE:} & \textit{redness}\\ \text{MIDPT:} & \textit{reference-red}\\ \text{ENDPT:} & \textit{proto-red}\\ \text{TR:} & \textit{redder-red}\end{bmatrix}\end{bmatrix} \qquad \text{(b)}\ \begin{bmatrix}\text{SEM:} & {}^{1}\\ \text{FRM:} & \begin{bmatrix}\text{TYPE:} & \textit{redness}\\ \text{MIDPT:} & \begin{bmatrix}\text{TYPE:} & \textit{apple/1}\\ \text{COLOR:} & \textit{reference-red}\end{bmatrix}\\ \text{ENDPT:} & \begin{bmatrix}\text{TYPE:} & \textit{apple/1}\\ \text{COLOR:} & \textit{proto-red}\end{bmatrix}\\ \text{TR:} & {}^{1}\begin{bmatrix}\text{TYPE:} & \textit{apple/1}\\ \text{COLOR:} & \textit{redder-red}\end{bmatrix}\end{bmatrix}\\ \emptyset_1: & \begin{bmatrix}\text{TYPE:} & \textit{apple/2}\\ \text{COLOR:} & \textit{reference-red}\end{bmatrix}\end{bmatrix}$$

$$\text{(c)}\ \begin{bmatrix}\text{SEM:} & {}^{1}\\ \text{FRM:} & {}^{1}\begin{bmatrix}\text{TYPE:} & \textit{state-schema}\\ \text{TR:} & {}^{2}\\ \text{ST:} & {}^{2}\begin{bmatrix}\text{TYPE:} & \textit{apple/1}\\ \text{COLOR:} & \textit{redder-red}\end{bmatrix}\end{bmatrix}\\ \emptyset_1: & \begin{bmatrix}\text{TYPE:} & \textit{apple/2}\\ \text{COLOR:} & \textit{reference-red}\end{bmatrix}\\ \emptyset_2: & \begin{bmatrix}\text{TYPE:} & \textit{redness}\\ \text{MIDPT:} & \textit{reference-red}\\ \text{ENDPT:} & \textit{proto-red}\\ \text{TR:} & \textit{redder-red}\end{bmatrix}\end{bmatrix} \qquad \text{(d)}\ \begin{bmatrix}\text{SEM:} & {}^{1\text{ or }2\text{ (see text)}}\\ \text{FRM:} & {}^{1}\begin{bmatrix}\text{TYPE:} & \textit{has-color-schema}\\ \text{PROP:} & {}^{3}\\ \text{TR:} & {}^{2}\begin{bmatrix}\text{TYPE:} & \textit{apple/1}\\ \text{COLOR:} & {}^{3}\textit{redder-red}\end{bmatrix}\end{bmatrix}\\ \emptyset_1: & \begin{bmatrix}\text{TYPE:} & \textit{apple/2}\\ \text{COLOR:} & \textit{reference-red}\end{bmatrix}\\ \emptyset_2: & \begin{bmatrix}\text{TYPE:} & \textit{redness}\\ \text{MIDPT:} & \textit{reference-red}\\ \text{ENDPT:} & \textit{proto-red}\\ \text{TR:} & \textit{redder-red}\end{bmatrix}\end{bmatrix}$$

Figure 6: Schematizations of *redder apple* with alternate scalarizations and reifications.

LAKOFF, GEORGE. 1987. *Women, fire, and dangerous things*. Chicago: University of Chicago Press.

LANGACKER, RONALD W. 1987. *Foundations of cognitive grammar*, volume 1. Stanford, CA: Stanford University Press.

LENAT, DOUG & R. V. GUHA. 1988. The world according to CYC. Technical Report ACA-AI-300-88, MCC Artificial Intelligence Laboratory, Austin, TX.

LENAT, DOUGLAS B. & R. V. GUHA. 1990. *Building large knowledge-based systems: Representation and inference in the CYC project*. Reading, MA: Addison-Wesley.

LONGACRE, R. E. 1976. *An anatomy of speech notions*. Lisse: Peter de Ridder Press.

OSTLER, NICHOLAS. 1980. *A theory of case linking and agreement*. Bloomington, IN: Indiana University Linguistics Club.

PARSONS, TERENCE. 1990. *Events in the semantics of english: A study in subatomic semantics*. Cambridge, MA: MIT Press.

PINKER, STEVEN. 1989. *Learnability and cognition*. Cambridge, MA: MIT Press.

PLATT, J. T. 1971. *Grammatical form and grammatical meaning: A tagmemic view of Fillmore's Deep Structure Case concepts*. Amsterdam: North-Holland.

REGIER, TERRY, 1992. *The acquisition of lexical semantics for spatial terms: A connectionist model of perceptual categorization*. University of California at Berkeley dissertation.

SCHANK, ROGER C. 1973. Identification of conceptualizations underlying natural language. In *Computer models of thought and language*, ed. by Roger C. Schank & Kenneth Mark Colby. San Francisco: W. H. Freeman.

SHIEBER, STUART M. 1986. *An introduction to unification-based approaches to grammar*. Stanford, CA: Center for the Study of Language and Information.

SILVERSTEIN, MICHAEL. 1976. Hierarchy of features and ergativity. In *Grammatical categories in Australian languages*, ed. by R. M. W. Dixon. Canberra: Australian Institute of Aboriginal Studies. Cited in Foley & van Valin 1984.

SOMERS, H. L. 1987. *Valency and case in computational linguistics*, volume 3 of *Edinburgh Information Technology Series*. Edinburgh: Edinburgh University Press.

TALMY, LEONARD. 1983. Spatial orientation: Theory, research, and application. In *How lan-*

guage structures space, ed. by Herbert Pick & Linda Acredolo. New York: Plenum Press.

TALMY, LEONARD. 1985. Lexicalization patterns: Semantic structure in lexical forms. In *Language typology and syntactic description*, ed. by T. Shopen, volume 3: Grammatical Categories and the Lexicon. Cambridge: Cambridge University Press.

TALMY, LEONARD. 1988. Force dynamics in language and cognition. *Cognitive Science*, 12:49–100.

WARREN, BEATRICE. 1978. *Semantic patterns of noun-noun compounds*. Gothenburg, Sweden: Acta Universitatis Gothoburgensis.

WILENSKY, ROBERT, 1989. (untitled). Unpublished draft, University of California at Berkeley.

WU, DEKAI. 1992a. Approximate maximum-entropy integration of syntactic and semantic constraints. In *AAAI-92 Workshop on Statistically-Based NLP Techniques*, 98–108, San Jose, CA.

WU, DEKAI, 1992b. *Automatic inference: A probabilistic basis for natural language interpretation.* University of California at Berkeley dissertation. Available as UC Berkeley Computer Science Division Technical Report UCB/CSD 92/692.

WU, DEKAI. 1993. Estimating probability distributions over hypotheses with variable unification. In *Proceedings of AAAI-93, Eleventh National Conference on Artificial Intelligence*, Washington, D.C. To appear.