

Introduction to Semantic Role Labelling

Behrang QasemiZadeh

zadeh@phil.hhu.de

Computational Linguistics Department, HHU – DRAFT

October 2018–January 2019

Recap

We discussed Jackendoff's LCS system last week.

Jackendoff's LCS focuses on clarifying the “nature” of semantic roles and their systematic mappings to syntactic forms (e.g., distinguishing Agent and Theme, adjunct and arguments, etc.).

LCS try to achieve this goal by forming the conceptual predicates behind grammatical constructions in which verbs are realized.

By contrast, Dowty (a prominent semanticist) chooses a different approach, and proposes his **PROTO-ROLES** theory using **prototype theory**.

That's in our agenda!

Tigers...

How many tigers do you see? How do you define a tiger? Are they all target with an equal weight?!



Tigers... (contd.)

Is Tigger a tiger?!



PROTO-ROLES (Background)

To describe PROTO-ROLES, Palmer et al. (2010) start with **semantic categories**:

*Psychologists, philosophers and linguists had long recognized the difficulty of defining **axioms** that unequivocally identify members of individual semantic **categories** such as tigers, so they had eventually resorted to defining **prototypes** for the categories instead.*

But what does it mean?

PROTO-ROLES (Background) (contd.)

Axiom is a statement that is taken to be **true**, and is used as a **premise** (the ground) for additional reasoning and arguments.

In other words, axioms are assumed true without being proven. Subsequently, one or several axioms are used to prove/reject a **Theorem**.

For example, the axioms of *Boolean algebra* are: $0 \wedge 0 = 0$,
 $0 \vee 0 = 0$, $1 \wedge 1 = 1$, $1 \vee 1 = 1$, $0 \wedge 1 = 1 \wedge 0 = 0$,
 $0 \vee 1 = 1 \vee 0 = 1$, $x = 0 \rightarrow \neg x = 1$, $x = 1 \rightarrow \neg x = 0$.

From these, we derive Identity, Commutative, etc. properties. Eventually these simple axioms give us computers!

PROTO-ROLES (Background) (contd.)

Defining axioms for natural things such as language is difficult.

Sometimes, thinkers define axioms for natural things based on their observations, in particular at the absence sufficient means and instruments for studying/explaining them.

All I can see is that the earth is flat, thus

Axiom: The earth is flat because. ✗

By contrast, the apple fell, Isaac Newton observed it, and eventually came up with the axiom of gravitation. ✓

You can imagine what will happen when we attempt to define axioms for language and linguistic phenomenon.

PROTO-ROLES (Background) (contd.)

NB: Disclaimer, what mentioned here regarding axioms was an overt over-simplification of the matter – maybe inaccurate or even incorrect.

PROTO-ROLES (Background) (contd.)

Philosophers, psychologists, and linguists, etc. admit that defining axioms for *unequivocally* identifying members of individual **semantic categories** is difficult. (maybe even impossible)

For instance, different people have different ideas regarding what can be counted as the semantic category of Tiger.

Put it coarsely, the ontology and structure of concepts is a major theme in several disciplines. We can try to classify lexical concepts in several ways (e.g., the members of the class Tiger), two of them being, the **classical theory** and **prototype theory**.

PROTO-ROLES (Background) (contd.)

Classical theory and prototype theory for Categorization:

Some (the classical theory) believe a lexical concept carries a **definitional** structure. And, this definitional structure contains necessary and sufficient information (conditions) to determine members of it.

Prototype theory is the alternative to the classical theory. It states that concepts have **probabilistic** structure and definitional in the sense that something belongs to a category C as long as it satisfies a sufficient number of **properties** from C 's elements.

Kind of rationalist vs empiricist way to do the task!

Between, there are many more theories; and again, above is an oversimplification! . . .

PROTO-ROLES (Background) (contd.)

Getting back to our example, not every tiger

has whiskers and gold and black stripes;

lives in the jungle;

hunts voraciously;

etc.

But most people would recognize an animal with the above mentioned characteristics as *a prototypical tiger*.

PROTO-ROLES (Background) (contd.)

Dowty used the same argument as psychologist used for Tiger.
Although every **Agent** may not be

a **sentient** being (able to feel/observe things)
that volitionally causes a change of state in another
participant

most people recognize a being with the above mentioned
characteristics as a prototypical Agent ?.

For verb predicates, Dowty proposed only two **thematic-role-like
concepts**: the **proto-agent** and **proto-patient**.

Respectively, Dowty defined a set of properties which are likely to
be associated with a **Proto-Agent** and **Proto-Patient**.

Proto-Agent, Proto-Patient Properties

Proto-Agent Properties:

Volitional involvement in event or state;

Sentience (and/or perception);

Causing an event of change of state in another participant;

Movement (relative to position of another participant);

Exists independently of event named;

Agent, also Experiencer, Instrument, Actor, Causer, etc..

Proto-Agent, Proto-Patient Properties (contd.)

Proto-Patient Properties:

Undergoes change of state

Incremental theme

Causally affected by another participant

Stationary relative to movement of another participant

Does not exist independently of the event, or at all;

Patient, also Theme, Precepts, etc.

Dowty's Proto-Role System

Instead of using conceptual predicate to capture/present the notion of agenthood (e.g., as in LCS), Dowty focuses on characteristics/features of agenthood and patienthood (i.e., definitional vs probabilistic structures).

No predicate or logical statement but a collection of properties that let us decide the category of an argument (Proto-Agent, Prot-Patient).

Properties are implied by the participant being associated with the action/event described by the verb in this role.

Dowty's Proto-Role System (contd.)

For a participant of an Agent-like role, at least one of the Proto-Agent properties must be implied.

For a participant with a Patient-like role, at least one of the Proto-Patient properties must be met.

If an event participant have one or more Proto-Agent properties (or Proto-Patient properties) it can be labeled as an Agent (or as a Patient).

Not all the properties are applicable for a specific participant and a specific verb.

Dowty's Proto-Role System (contd.)

Dowty states that his Proto-role system were intended for the purpose of argument selection, e.g. assignment of grammatical functions to arguments, as described in his **Argument Selection Principle**:

The argument of a predicate having the greatest number of proto-agent properties entailed by the meaning of the predicate will, all else being equal, be lexicalised as the subject of the predicate; the argument having the greatest number of Proto-patient properties will, all else being equal, be lexicalised as the direct object of the predicate.

Dowty's Proto-Role System (contd.)

Simplified version: Arguments that have more of Proto-Agent properties tend to be syntactic subject; Arguments that have higher number of Proto-Patient properties, will be direct object.

For example, for the following subcat of the verb break:

Subject **break** Object with Oblique

most examples are

subject filler: `cause+exist = 2+0`;

object filler: `exist+ stationary = 1+1`;

oblique filler: `exist = 1+0`

which somehow also works in a similar way for the pattern

Subject **break** Object

Dowty's Proto-Role System (contd.)

There are also two corollaries:

- 1) If two arguments have an equal number of implied prot-role properties (agent–patient), then both have equal chance to fill subject and object (e.g., many communication verbs, soon we will be able to compile an an approximate list).
- 2) If there are three arguments:
 - a) argument of the highest Proto-patient properties will be the direct object;
 - b) the non-subject argument with the fewer proto-patient properties will fill the oblique/prepositional-object place.

Dowty's Proto-Role System (contd.)

Argument Selection Principle (ASP) has been often used when studying diathesis and verb alternation patterns.

According to ASP, alternation patterns occurs depending on the rankings of arguments wrt. to proto-agent-patient properties (in particular, equal number of properties for different arguments).

See examples from EAGLES (1996 but still good as new) recommendations on subcategorization frames:

<http://www.ilc.cnr.it/EAGLES96/synlex/synlex.html>.

Dowty's Proto-Role System (contd.)

Palmer et al. (2010) points out certain interesting aspects of Dowty.

Dowty (like Jackendoff) proposed his own hierarchy of roles:

Agent/Instrument/Benefactor > Patient > Source/Goal

For the objects in motion, Dowty uses **Agent** rather than **Theme** (see the SRL table from 2nd session).

Additionally, Dowty introduces a new role named **Incremental Theme**.

Dowty's Proto-Role System (contd.)

An **Incremental Theme** is an event participant that is *undergoing* a change of state and the completion of the change of state signals the termination of the event.

(1) Mary drank a glass of **water**.

How could **Incremental Theme** help?

(2) Mary loaded the hay onto the truck.

(3) Mary loaded the truck with hay.

Are (2) and (3) different (different senses of load)? If they are different, how?

Dowty's Proto-Role System (contd.)

- If the Incremental Theme is the *hay*, the change of state is its change of location to the truck.
- If the *truck* is the Incremental Theme, it is changing state from being empty to being filled.
- In either case, the event only terminates when all of the hay is moved or when the truck is completely filled.
- According to Dowty, it is the shift from one type of Incremental Theme to another that accounts for the difference in meaning, but the same sense of the verb is being referenced.

Dowty's Proto-Role System (contd.)



To get to know Dowty's work requires time beyond what we can afford now!

Disregarding the examples discussed here, the novelty and importance of Dowty's proto-roles is inarguable and still an inspiration for researchers of several disciplines.

A noteworthy project is Decomp (<http://decomp.io/>); see Reisinger et al. (2015) to see how they exploit Dowty's proto-role.¹

¹Or go through these slides:

Last week homework

Bring a lexical conceptualization representation for the following sentences:

John opened the door.

The door is open.

The ball rolled to the other end of the hall.

Jim gave the book to the professor.

Mary ejected Jim from the room.

Laura quickly whispered the latest rumour about the Queen to her cousin.

Bill cut his hair with a razor.

Gina crashed the car into the embankment.

Last week homework (contd.)

- * John opened the door.
CAUSE(John, [GO_{IDENT}(door, \neg open, open)])
- * The door is open.
BE(door, open)
- * The ball rolled to the other end of the hall.
GO_{POSIT}(ball, x, end-of-hall)
- * Jim gave the book to the professor.
CAUSE(Jim, [GO_{POSS}(book, Jim, professor)])
- * Mary ejected Jim from the room.
CAUSE(Mary, [GO_{POSIT}(Jim, room, y)])
- * Laura quickly whispered the latest rumour about the Queen to her cousin.
CAUSE(Laura, [GO(rumor, Laura, cousin)]), MANNER(whisperingly)
- * Bill cut his hair with a razor.
CAUSE(Bill, [GO_{IDENT}(hair, x, cut)])
- * Gina crashed the car into the embankment.
CAUSE(Gina, [GO_{POSIT}(car, x, kaput)])

Homework

Compile the list of proto-role properties for of the arguments of the following sentences.

John opened the door.

~~The door is open.~~

The ball rolled to the other end of the hall.

Jim gave the book to the professor.

Mary ejected Jim from the room.

Laura quickly whispered the latest rumour about the Queen to her cousin.

Bill cut his hair with a razor.

Gina crashed the car into the embankment.

Bibliography

- Palmer, M., Gildea, D., and Xue, N. (2010). *Semantic Role Labeling*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.
- Reisinger, D., Rudinger, R., Ferraro, F., Harman, C., Rawlins, K., and Durme, B. V. (2015). Semantic proto-roles. *Transactions of the Association for Computational Linguistics*, 3:475–488.