FrameNet Annotation: Pointers & Pitfalls

FE annotation

Where to look for FEs

- Respect syntactic locality!
  - Usually the FE argument is an **immediate syntactic dependent** of the target. Exceptions are narrowly limited to situations like control, raising, and relative clauses (see "Syntactic locality" in The Book) where the target's subject is supplied syntactically by a higher constituent
  - Antecedents of **pronouns**/coreferring expressions are not annotated unless they are syntactically local to the target
    - Relative clauses: both relative pronoun and antecedent are annotated
- Null instantiation (INI, DNI, CNI) annotations indicate that a **core** FE is conceptually present but not stated overtly *in a syntactically local position*.
  - Core sets and Requires/Excludes relations formalize the notion that multiple core FEs may be considered facets of the same information, and thus if one is specified the other does not need to be null instantiated.
  - With **noun targets**, the use of null instantiation is more limited. Essentially, non-overt core FEs of a noun target should be considered null-instantiated only if the noun refers to a specific event instance.

Semantic criteria for FEs

- Some FE types (e.g. "Agent", "Protagonist") are for sentient/conscious participants. Abstract nouns or objects (such as *the life*, *the materials*) are unlikely to be Agents. Note that (i) for some targets the subject might a Cause, Action, Means, Instrument, etc., and (ii) for some targets sentient FEs may be null-instantiated.

What are the non-FE layers?

The GF and PT layers indicate the syntactic status and relationship of **overt frame elements with respect to the target**. This information is used to construct valence patterns for the LU, which can be seen in the Lexical Entry report (e.g. this one).

- **Grammatical Function (GF)** is simple—3 values account for most FEs: `Ext` for external argument (subject of the verb phrase whose predicate is the target, or of a higher clause with control/raising); `Obj` for direct and indirect object NPs, and `Dep` for most other dependents.
- **Phrase Type (PT)** has more values: regular `NP` , possessive NP `Poss`, modifier noun `N`; `PP` and some variants; `Sub` (subordinate clause); various types of `S` (clause) and `VP` (verb phrase, no subject); adjective, adverb, and quantifier phrases; and quotes.

The Book provides details—PT in ch. 4, GF in ch. 5.

Layers specific to the LU's POS (**Verb**, **Noun**, etc.) are for indicating attributes like regular or copular support (`Supp`, `Cop`). The **Other** layer is used to specify the antecedent (`Ant`) and relativizer (`Rel`) for an FE whose target is embedded in a relative clause, and `Null` for expletive *it* or *there*, which is not an FE. Consult the Book for details.

Issues with FE spans, GF, PT

- Be sure to select the entire phrase for the FE in question.
  - If the FE is introduced by a grammatical word such as a preposition, include it!
  - Coordinated phrases count as a single span
    - [Cats and dogs]$_{Self\_mover/Ext/NP}$ are **frolicking**
    - It is possible to **injure** someone [out of malicious intent or due to negligence]$_{Explanation/Dep/PP}$
- Usually there should be no more than one occurrence of a given FE per target. This is particularly true of core FEs.
- Occasionally the correct non-core FE may be missing from the frame.
- **Denoted FEs**: If the target also fills an FE, the FE span should just be the span of the target, and it should not receive GF or PT labels.
- If the FE is headed by an *-ing* verb with no subject, its PT is **VPing** (not `Sing`) and its GF is **Dep** (not `Obj`)
  - If we meet in public we **risk**$_{Run\_risk}$ [blowing our cover]$_{Bad\_outcome/Dep/VPing}$
- Pronouns and proper names have a PT of **NP** (not `N`, which here is reserved for noun modifiers within an NP)
  - If we meet in public [we]$_{Protagonist/Ext/NP}$ **risk**$_{Run\_risk}$ blowing our cover

Full-text annotation

Do your best to identify tokens that should be targets, and then see if an appropriate LU exists in an appropriate frame. (The lexical item may be missing entirely, or in some of its senses.)

Targets include, in principle:	The following are NOT targets:
Most content words (verbs, nouns, adjectives, adverbs)—it's OK to focus on these for the purposes of this course - Modal auxiliaries and *not* - Quantifiers - Some conjunctions (particularly causal and temporal ones; check the LU index) - Some prepositions	- Named entities - Pronouns, including relative and indefinite pronouns - Basic articles: *a(n), the* - Demonstratives: *this, that, these, those* *and, but, or* - Complementizer *that* - Infinitival *to* - Copulas (these are marked as `Cop` w.r.t. the predicate complement)

● Some "grammatical" adjectives and adverbs (*only, else, too, never*)	● Light/support verbs (these are marked as `Supp` w.r.t. the object) ● Non-modal auxiliaries ● Possessive *'s*

Numbers: There are frames Cardinal_numbers and Ordinal_numbers, but they do not contain an infinite number of LUs.

Some targets are **multiword expressions** (MWEs). To annotate a multiword target, select one of the words, choose the multiword LU, and then a dialog should appear in which you can check the other tokens in the sentence belonging to the target.