Proposition Bank (PropBank)

Daniel Zeman

April 15, 2020
The Proposition Bank

- Deep-syntactic annotation of the predicate-argument structure
- A deep layer over the constituents of the Penn Treebank (English)
- Goal: “consistent argument labels across different syntactic realizations of the same verb”
  - \([John]_{ARG0} \ [broke]_{\text{break.01}} \ [the \ window]_{ARG1}\)
  - \([The \ window]_{ARG1} \ [was \ broken]_{\text{break.01}} \ [by \ John]_{ARG0}\)
  - \([The \ window]_{ARG1} \ [broke]_{\text{break.01}}\)
Numbered Arguments

- Numbered arguments:
  - required for the valency of a predicate (e.g., agent, patient, benefactive)
  - or not required but occurring with high frequency in actual usage

- The semantic roles slightly differ given the usage of each predicate, but in general they correspond to the following:
  - ARG0 … agent
  - ARG1 … patient
  - ARG2 … instrument, benefactive, attribute
  - ARG3 … starting point, benefactive, attribute
  - ARG4 … ending point
  - ARGM … modifier
Semantic classification of modifiers of predicates:

- **ARGM-COM** ... comitative
- **ARGM-LOC** ... locative
- **ARGM-DIR** ... directional
- **ARGM-GOL** ... goal
- **ARGM-MNR** ... manner
- **ARGM-TMP** ... temporal
- **ARGM-EXT** ... extent
- **ARGM-REC** ... reciprocals
- **ARGM-PRD** ... secondary predication
- **ARGM-PRP** ... purpose
- **ARGM-CAU** ... cause
- **ARGM-DIS** ... discourse
- **ARGM-ADV** ... adverbials
- **ARGM-ADJ** ... adjectival
- **ARGM-MOD** ... modal
- **ARGM-NEG** ... negation
- **ARGM-DSP** ... direct speech
- **ARGM-LVB** ... light verb
- **ARGM-CXN** ... construction
Semantic Roles of Arguments

Semantic roles of arguments in the predicate dictionary
(sentence annotation refers to dictionary entry and arguments are just numbered):

- ARGn−PAG ... proto-agent
- ARGn−PPT ... proto-patient
- ARGn−PRD ... secondary predication
- ARGn−GOL ... goal, recipient, beneficiary
- ARGn−DIR ... direction
- ARGn−MNR ... manner, instrument
- ARGn−PRP ... purpose
- ARGn−VSP ... verb-specific
Verb Senses (Frames, Role Sets)

• Predicate: *break*
  - Roleset id: break.01 (*break, cause not to be whole*)
  - Roleset id: break.02 (*break in, gain entry*)
  - Roleset id: break.13 (*happen, come to light, onset, beginning*)
  - Roleset id: break.16 (*separate one entity from another*)
  - Roleset id: break.18 (*violate (the law, an oath, trust etc]*)
  - Roleset id: break.19 (*to suspend an action, take a break from something, pause, musical interlude*)
  - Roleset id: break.20 (*break into motion, for a goal or from a source location*)
  - Roleset id: break.21 (*notable occurance, point at which one’s luck turns (bad break, lucky break]*)

• Predicate: *break_away*
  - Roleset id: break_away.14 (*separate, withdraw, secede*)
  - ...
Verb Senses (Frames, Role Sets)

- **Predicate:** *break*
  - Roleset id: **break.01** *(break, cause not to be whole)*
    - ARG0-PAG: breaker
    - ARG1-PPT: thing broken
    - ARG2-MNR: instrument
    - ARG3-PRD: pieces
    - ARG4-DIR: arg1 broken away from what?
  - Roleset id: **break.02** *(break in, gain entry)*
  - Roleset id: **break.13** *(happen, come to light, onset, beginning)*
  - Roleset id: **break.16** *(separate one entity from another)*
  - Roleset id: **break.18** *(violate (the law, an oath, trust etc))*
  - Roleset id: **break.19** *(to suspend an action, take a break from something, pause, musical interlude)*
  - Roleset id: **break.20** *(break into motion, for a goal or from a source location)*
Verb Senses (Frames, Role Sets)

- **Predicate:** `break`
  - Roleset id: `break.01` *(break, cause not to be whole)*
  - Roleset id: `break.02` *(break in, gain entry)*
    - ARG0-PAG: enterer
    - ARG1-GOL: place/domain broken into
  - Roleset id: `break.13` *(happen, come to light, onset, beginning)*
  - Roleset id: `break.16` *(separate one entity from another)*
  - Roleset id: `break.18` *(violate (the law, an oath, trust etc))*
  - Roleset id: `break.19` *(to suspend an action, take a break from something, pause, musical interlude)*
  - Roleset id: `break.20` *(break into motion, for a goal or from a source location)*
Verb Senses (Frames, Role Sets)

- Predicate: *break*
  - Roleset id: **break.01** (*break, cause not to be whole*)
  - Roleset id: **break.02** (*break in, gain entry*)
  - Roleset id: **break.13** (*happen, come to light, onset, beginning*)
    - ARG0-PAG: *causal agent*
    - ARG1-PPT: *thing coming to light*
  - Roleset id: **break.16** (*separate one entity from another*)
  - Roleset id: **break.18** (*violate (the law, an oath, trust etc]*)
  - Roleset id: **break.19** (*to suspend an action, take a break from something, pause, musical interlude*)
  - Roleset id: **break.20** (*break into motion, for a goal or from a source location*)
Verb Senses (Frames, Role Sets)

- **Predicate:** break
  - Roleset id: **break.01** *(break, cause not to be whole)*
  - Roleset id: **break.02** *(break in, gain entry)*
  - Roleset id: **break.13** *(happen, come to light, onset, beginning)*
  - Roleset id: **break.16** *(separate one entity from another)*
    - ARG0-PPT: *entity in motion*
    - ARG1-DIR: *movement away from what*
    - ARG2-PRD: *secondary predication on arg0 (free, loose etc)*
  - Roleset id: **break.18** *(violate (the law, an oath, trust etc))*
  - Roleset id: **break.19** *(to suspend an action, take a break from something, pause, musical interlude)*
  - Roleset id: **break.20** *(break into motion, for a goal or from a source location)*
Verb Senses (Frames, Role Sets)

- **Predicate**: break
  - **Roleset id**: break.01: **break, cause not to be whole**
  - **Roleset id**: break.02: **break in, gain entry**
  - **Roleset id**: break.13: **happen, come to light, onset, beginning**
  - **Roleset id**: break.16: **separate one entity from another**
  - **Roleset id**: break.18: **violate (the law, an oath, trust etc))**
    - ARG0–PAG: scofflaw, violator
    - ARG1–PPT: law, oath, trust, thing violated
    - ARG2–MNR: violation committed
  - **Roleset id**: break.19: **to suspend an action, take a break from something, pause, musical interlude**
  - **Roleset id**: break.20: **break into motion, for a goal or from a source location**
Verb Senses (Frames, Role Sets)

- Predicate: **break**
  - Roleset id: **break.01** (break, cause not to be whole)
  - Roleset id: **break.02** (break in, gain entry)
  - Roleset id: **break.13** (happen, come to light, onset, beginning)
  - Roleset id: **break.16** (separate one entity from another)
  - Roleset id: **break.18** (violate (the law, an oath, trust etc))
  - Roleset id: **break.19** (to suspend an action, take a break from something, pause, musical interlude)
    - ARG0-PAG: person taking/causing break/pause
    - ARG1-PPT: activity interrupted
    - ARG2-PRP: purpose of pause, interrupted for or with what
  - Roleset id: **break.20** (break into motion, for a goal or from a source location)
Verb Senses (Frames, Role Sets)

- **Predicate:** *break*
  - Roleset id: **break.01** *(break, cause not to be whole)*
  - Roleset id: **break.02** *(break in, gain entry)*
  - Roleset id: **break.13** *(happen, come to light, onset, beginning)*
  - Roleset id: **break.16** *(separate one entity from another)*
  - Roleset id: **break.18** *(violate (the law, an oath, trust etc))*
  - Roleset id: **break.19** *(to suspend an action, take a break from something, pause, musical interlude)*
  - Roleset id: **break.20** *(break into motion, for a goal or from a source location)*
    - ARG0-PAG: mover
    - ARG1-DIR: place left, broke from
    - ARG2-GOL: destination
Grammatical Coreference

- Find antecedents for empty arguments of the verbs.
- Missing or moved constituents are represented as empty categories [*] in the Penn Treebank.
- Empty categories that can be coreferred with a NP within the same sentence are linked in coreference chains in PropBank.
  - *I made a decision [*PRO*] to leave.*
    - REL: leave
    - ARG0: [*PRO*] * [I]

- Note: textual coreference is not resolved.
“The thrift holding company said it expects to obtain regulatory approval and complete the transaction by year-end.”

Note that the textual coreference between *it* and *the thrift holding company* is not annotated.
The thrift holding company said it expects to obtain regulatory approval and complete the transaction by year-end.

Note that the textual coreference between *it* and *the thrift holding company* is not annotated.
The thrift holding company said -NONE-.

The thrift holding company obtained regulatory approval to complete the transaction by year-end.

Proposition Bank (PropBank)
PropBanks for Other Languages

- Use of the English Propbank on other languages:
  - IBM Universal Proposition Banks
- Propbank annotations with language-specific lexicons:
  - Hindi Propbank
  - Urdu Propbank
  - Persian Propbank
  - Arabic Propbank
  - Chinese Propbank
  - Finnish Propbank
  - Portuguese Propbank
  - Basque Verb-Index
  - Turkish Propbank
- Some of them are annotated over dependency rather than constituency trees