SemLink+: FrameNet, VerbNet, and Event Ontologies

Martha Palmer,

Claire Bonial, Diana McCarthy

University of Colorado

Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929 – 2014)

ACL Workshop

June 27, 2014

Computational Language and EducAtion Research

C L E A R

Outline

- Deep NLU?
- Where we are now
- Where we need to go
- More details about where we need to go
- The contributions and limitations of lexical resources to this process

Where we are now – shallow semantics

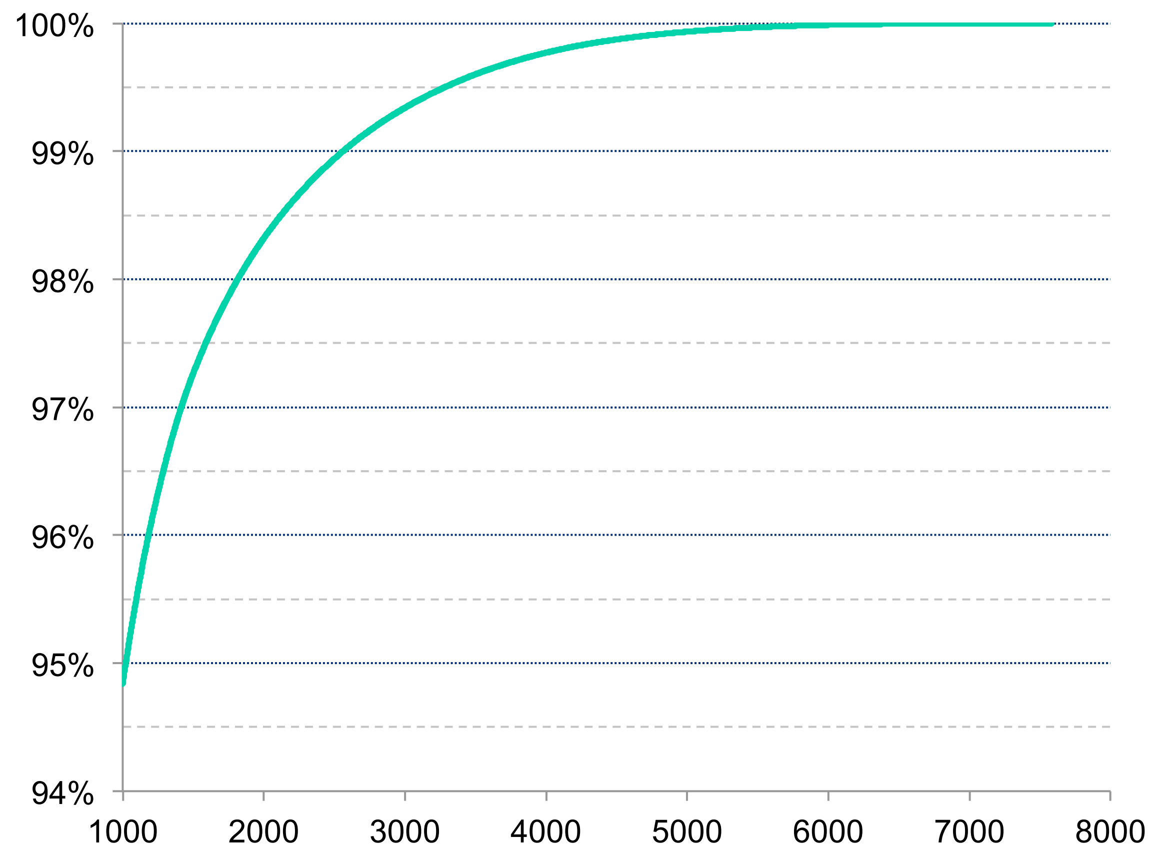
- Syntactic Structure – parse trees, Treebanks
- Semantic types – nominal entities [Person, Location, Organization], NE tagging
- Semantic roles – Agents, [PropBank FrameNet, VerbNet]
- Sense distinctions – *call me a taxi, call me an idiot,* WordNet, OntoNotes groups, FrameNet, VerbNet, vectors, etc.
- Coreference – *[President Obama: he]*

Where we are now - DETAILS

- DARPA-GALE, OntoNotes 5.0
  - BBN, Brandeis, Colorado, Penn
  - Multilayer structure: NE, TB, PB, WS, Coref
  - Three languages: English, Arabic, Chinese
  - Several Genres (@ ≥ 200K): NW, BN, BC, WT
    - Close to 2M words @ language (less PB for Arabic)
  - Parallel data, E/C, E/A
- DARPA BOLT – discussion forum, SMS
  - PropBank extensions: light verbs, function tags on core args, nominalizations, adjectives, constructions, ***often relying on FrameNet***

PropBank Verb Frames Coverage

- The set of verbs is open
- But the distribution is highly skewed
- For English, the 1000 most frequent lemmas cover 95% of the verbs in running text.
  - Graphs show counts over English Web data containing 150 M verbs.



FrameNet and VerbNet should have the same coverage, and we (or at least VerbNet)
desperately need help to do this semi-automatically!!

WordNet: - call, 28 senses, 9 groups

WN5, WN16,WN12
Loud cry

WN15 WN26
Bird or animal cry

WN3 WN19
WN1 WN22
Label

WN4 WN 7 WN8 WN9
Request

WN20 WN25
Call a loan/bond

WN18 WN27
Challenge

WN2 WN 13
Phone/radio
WN28

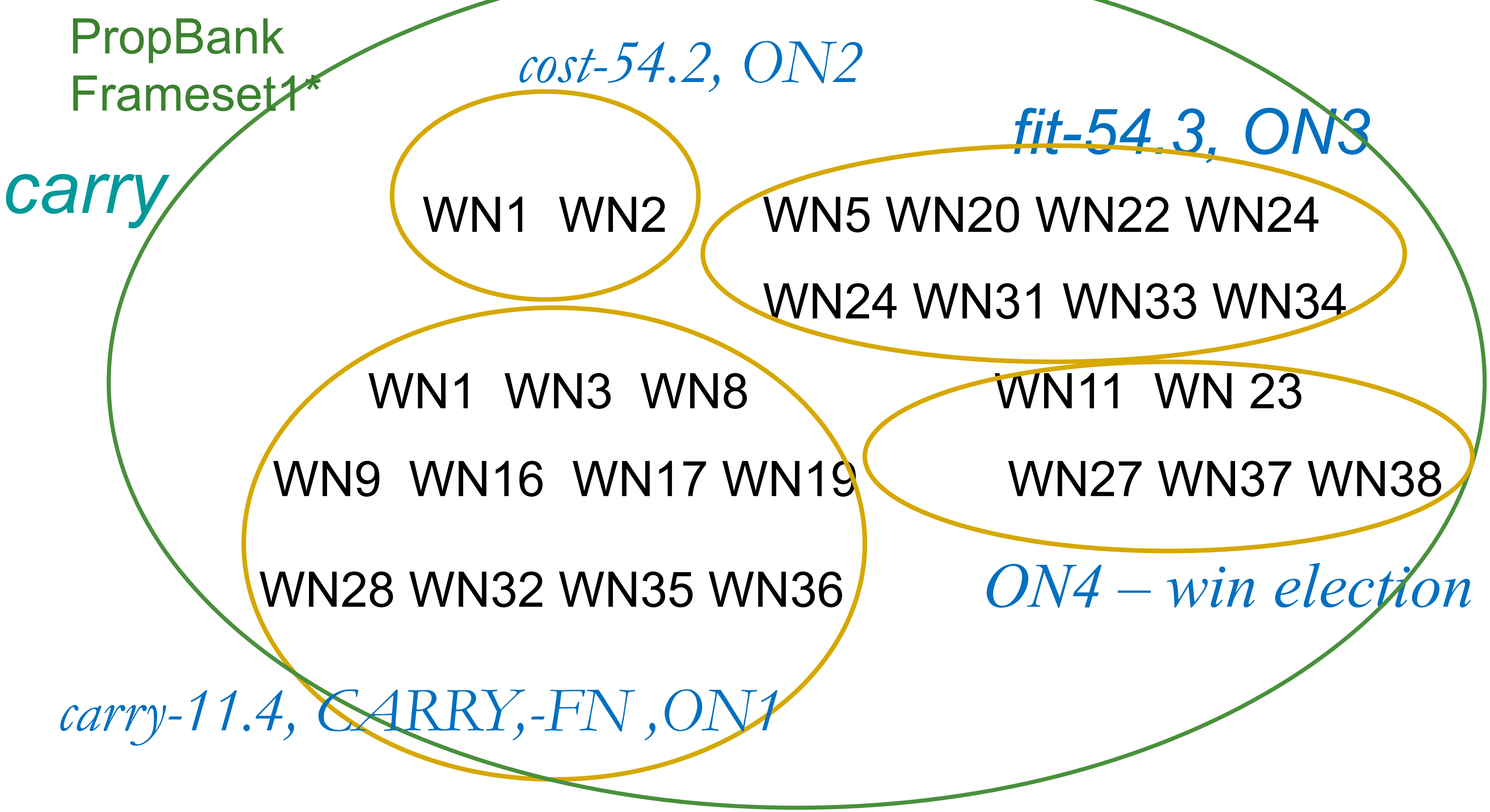
WN6 WN23
Visit

WN17 , WN 11
WN24,

WN10, WN14, WN21,
Bid

SEMLINK-PropBank, VerbNet, FrameNet, WordNet, OntoNotes

Palmer, Dang & Fellbaum, NLE 2007



*ON5-ON11 carry oneself,carried away/out/off, carry to term

Computational Language and EducAtion Research
C L E A R

Sense Hierarchy

- PropBank Framesets – ITA >90%
  coarse grained distinctions
  20 Senseval2 verbs w/ > 1 Frameset
  Maxent WSD system, 73.5% baseline, 90%
  - Sense Groups (Senseval-2/OntoNotes) - ITA 89%
    Intermediate level
    (includes Verbnet/some FrameNet) – SVM, 88+%

Dligach & Palmer, ACL2011

- WordNet – ITA 73%
  fine grained distinctions, 64%

SEMLINK

- Extended VerbNet: 6,340 senses
  - 92% PB tokens (8114 verb senses/12,646 all)
- Type-type mapping PB/VN, VN/FN, VN/WN
- Semi-automatic mapping of WSJ PropBank instances to VerbNet classes and thematic roles, hand-corrected. ***(now FrameNet also)***
- VerbNet class tagging as automatic WSD

  Brown, Dligach, Palmer, IWCS 2011; Croce, et. al., ACL2012
- Run SRL, map Arg2 to VerbNet roles, Brown performance improves *Yi, Loper, Palmer, NAACL07*

Where we need to go – Richer Event Descriptions - RED

- *"Saucedo said that guerrillas in one car <u>opened fire</u> on police <u>standing guard</u>, while a second car <u>carrying</u> 88 pounds (40 kgs) of dynamite <u>parked</u> in front of the building, and a third car <u>rushed</u> the attackers away."*

- *Saucedo said* – reporting event, evidential

Computational Language and EducAtion Research
CLEAR

What we can do

- *that guerrillas in one car opened fire on police standing guard*
- *opened fire* = aspectual context,
  - fire(guerillas, police)
- *standing guard* = support verb construction/ aspectual?, reduced relative
  - guard(police, X)

What we can do, cont.

- *while a second car carrying 88 pounds (40 kgs) of dynamite parked in front of the building*

- *carrying* - reduced relative, correct head noun - pounds or dynamite?
  - carry(car2, dynamite)

- park(car2, front_of(building))

What we can do, cont.

- *and a third car* <u>*rushed*</u> *the attackers away*
- rush(car3, attackers, away)

Temporal & Causal ordering?

- *"Saucedo said that guerrillas in one car <u>opened fire</u> on police <u>standing</u> guard, while a second car <u>carrying</u> 88 pounds (40 kgs) of dynamite <u>parked</u> in front of the building, and a third car <u>rushed</u> the attackers away*
- guarding BEFORE/OVERLAP firing
- Narrative container – TimeX
  - *[firing, parking, rushing]* all overlap, all in the same temporal bucket?
  - [see Styler, et. al, ACL2014, Events Workshop & RED Guidelines]

Don’t mark the relations between EVENTs.

!

Instead, put EVENTs in temporal buckets and relate the buckets

Temporal & Causal ordering

- "*Saucedo said that guerrillas in one car <u>opened fire</u> on police <u>standing</u> guard, while a second car <u>carrying</u> 88 pounds (40 kgs) of dynamite <u>parked</u> in front of the building, and a third car <u>rushed</u> the attackers away*
  - guarding BEFORE/OVERLAP firing
  - X CONTAINS [firing, parking, rushing]
  - firing BEFORE parking
  - parking BEFORE rushed

Implicit arguments

- *that guerrillas in one car opened fire on police standing guard*

- *opened fire* = aspectual context,
  - fire(guerillas, police)

- *standing guard* = support verb construction or aspectual?, reduced relative
  - guard(police, X)

More compelling example
(thanks to Vivek Srikumar)

- *The bomb exploded in a crowded marketplace. Five civilians were killed, including two children. Al Qaeda claimed responsibility.*

- Killed by Whom?
- Responsibility for what?

- Need recovery of implicit arguments

VerbNet – based on Levin, B.,93

- Class entries: *Kipper, et. al., LRE08*
  - Capture generalizations about verb behavior
  - Organized hierarchically
  - Members have common semantic elements, semantic roles, syntactic frames, predicates
- Verb entries:
  - Refer to a set of classes (different senses)
  - each class member linked to WN synset(s), ON groupings, PB frame files, FrameNet frames,

VerbNet: *send-11.1* (Members: 11, Frames: 5)

includes "*ship*"

- Roles
  - Agent [+animate | +organization]
  - Theme [+concrete]
  - Source [+location]
  - Destination [+animate | [+location & -region]]
- Syntactic Frame:NP V NP PP.destination
  - example "*Nora sent the book to London*."
  - syntax Agent V Theme {to} Destination
  - semantics motion(during(E), Theme)
    location(end(E), Theme, Destination)
    cause(Agent, E)

*Recovering Implicit Arguments**

*[Palmer, et. al., 1986;
Gerber & Chai, 2010, 2012;
Ruppenhofer, Sporleder, Morante, Baker, Palmer. 2010.
SemevEval-2010 Task10:]*

** AKA definite null complements*

*[$_{Arg0}$ The two companies] [$_{REL1}$ produce] [$_{Arg1}$ market pulp, containerboard and white paper]. The goods could be manufactured closer to customers, saving [$_{REL2}$ **shipping**] costs.*

- *Used VerbNet for subcategorization frames*

Implicit arguments

- SYNTAX   Agent V Theme {to} Destination

[AGENT] shipped [THEME] to [DESTINATION]

- SEMANTICS
  - CAUSE(AGENT,E)
  - MOTION(DURING(E), THEME),
  - LOCATION(END(E), THEME, DESTINATION),

Implicit arguments instantiated using coreference

- *[AGENT] shipped [THEME] to [DESTINATION]*
- *[Companies] shipped [goods] to [customers].*

- SEMANTICS
  - CAUSE(Companies, E)
  - MOTION(DURING(E), goods),
  - LOCATION(END(E), goods, customers),

Can annotate, semi-automatically!

Another type of Implicit Relation

Example from Daniel Marcu, GALE Wrap-up Mtg

- Between Munich and LA you need less than 11 hours by plane.
- You can fly
  to Los Angeles from Munchen
  in no more than eleven hours.
- From Munich to Los Angeles,
  it does not take
  more than eleven hours by plane.

Computational Language and EducAtion Research
C L E A R

Constructions allow us to

- Recognize a path prepositional phrase, and that it necessarily goes with a “MOTION” event – Caused-motion constructions
  - *John sneezed the tissue off the table.*
  - *Mary blinked the snow off of her eyelashes.*
- If we detect a MOTION event we can associate the *plane* with it as a vehicle
- Just the *plane* itself can suggest a motion event…

Construction Grammar

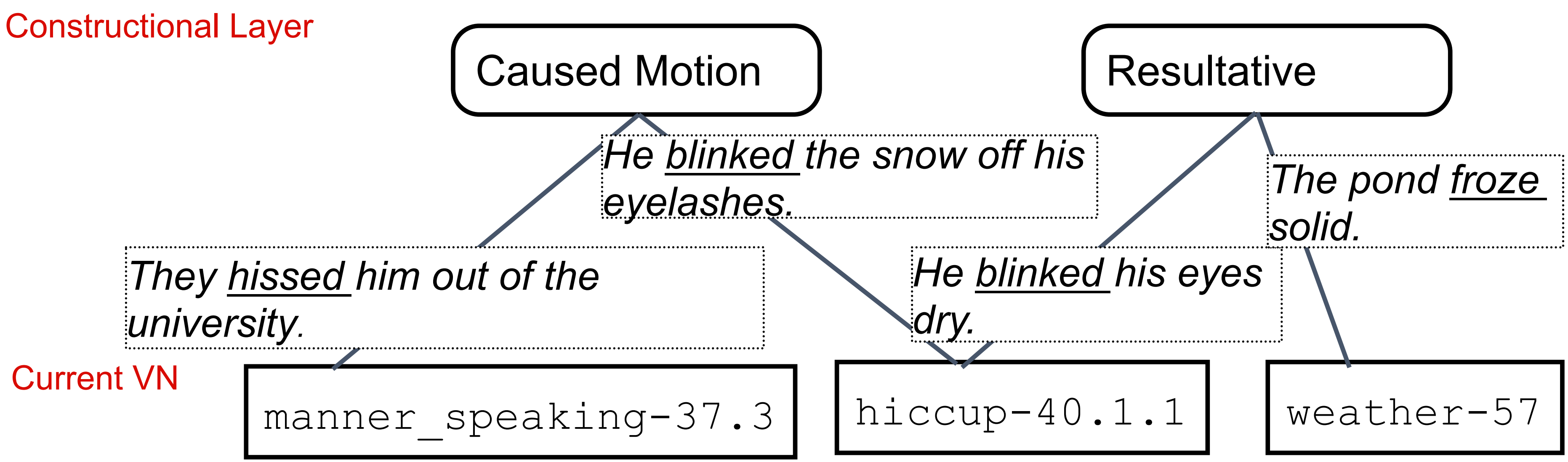
- In Construction Grammar
  (Fillmore, 1988, Goldberg, 1995, Kay and Fillmore, 1999, Michaelis, 2004, Goldberg, 2005)
  - ***constructions*** are carriers of meaning
  - ***constructions*** are assigned meaning in the same way that words are – via convention rather than composition.
- Invaluable resource –

  FrameNet Constructicon, Cxn Viewer

Introducing a Constructional Layer to VerbNet

Jena Hwang, LREC-2014

- Introduce a constructional ``layer'' to VerbNet, which attaches orthogonally to relevant VerbNet classes



VerbNet can also provide inferences – sometimes…

- *Every path from back door to yard was covered by a grape-arbor, and every yard had fruit trees.*
- *Where are the grape arbors located?*

VerbNet – *cover, fill-9.8 class*

- **Members:** fill, …, cover,…, staff, ….
- **Thematic Roles:** Agent
Theme
Destination
- **Syntactic Frames with Semantic Roles**

  "The employees staffed the store"

  " The grape arbors covered every path"

  Theme V Destination

location(E,Theme,Destination)
location(E,grape_arbor,path)

Computational Language and EducAtion Research
C L E A R

Inferences can inform Coreference?

- *"Saucedo said that guerrillas in one car opened fire on police standing guard, while a second car carrying 88 pounds (40 kgs) of dynamite parked in front of the building, and a third car rushed the attackers away."*

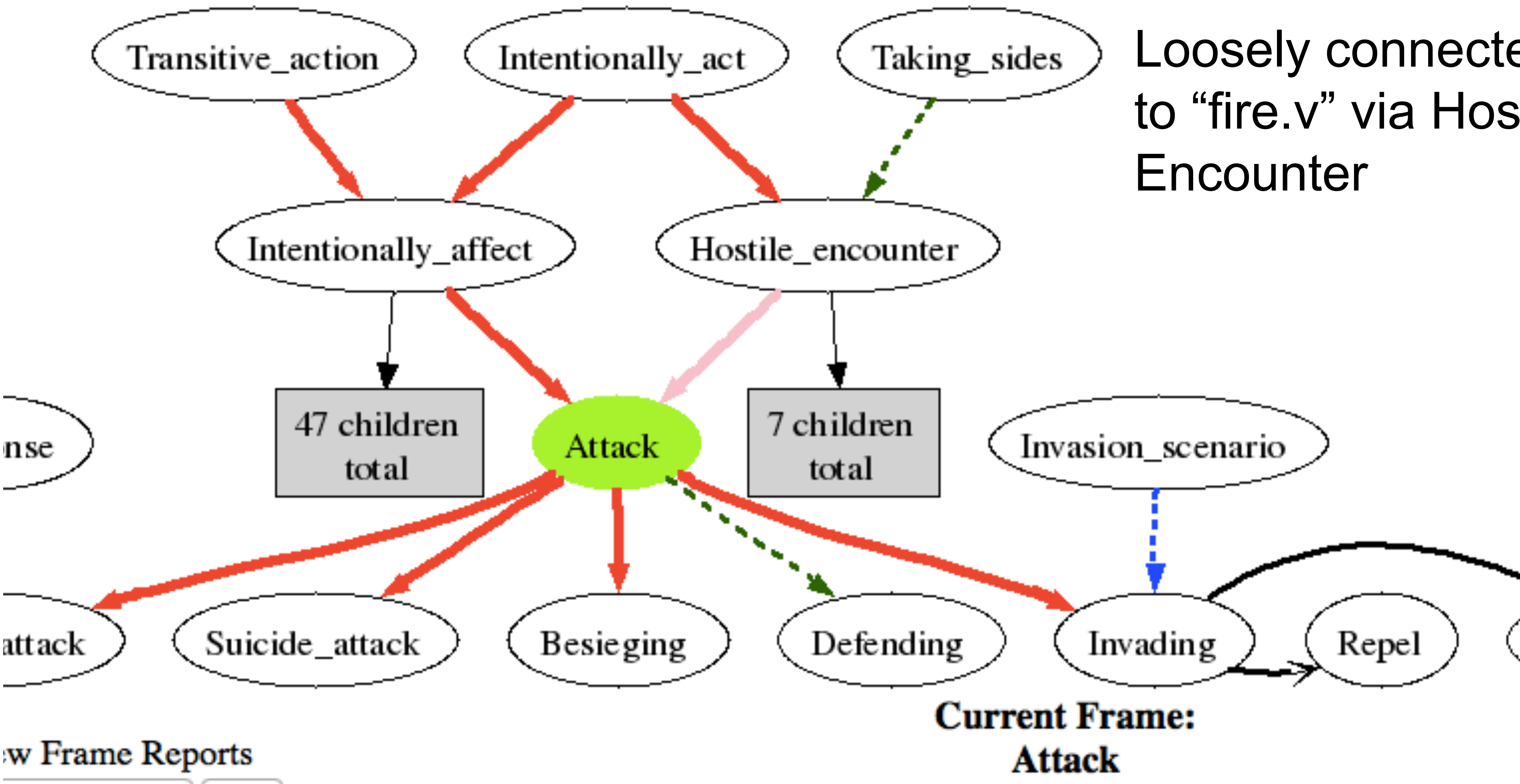
- *AMR*
  - *guerilla – Arg0 of fire.01*
  - *attacker – person – Arg0-of attack.01*

Computational Language and EducAtion Research
C L E A R

FrameNet Attack frame

attacker.n, fire.n

Loosely connected to “fire.v” via Hostile Encounter

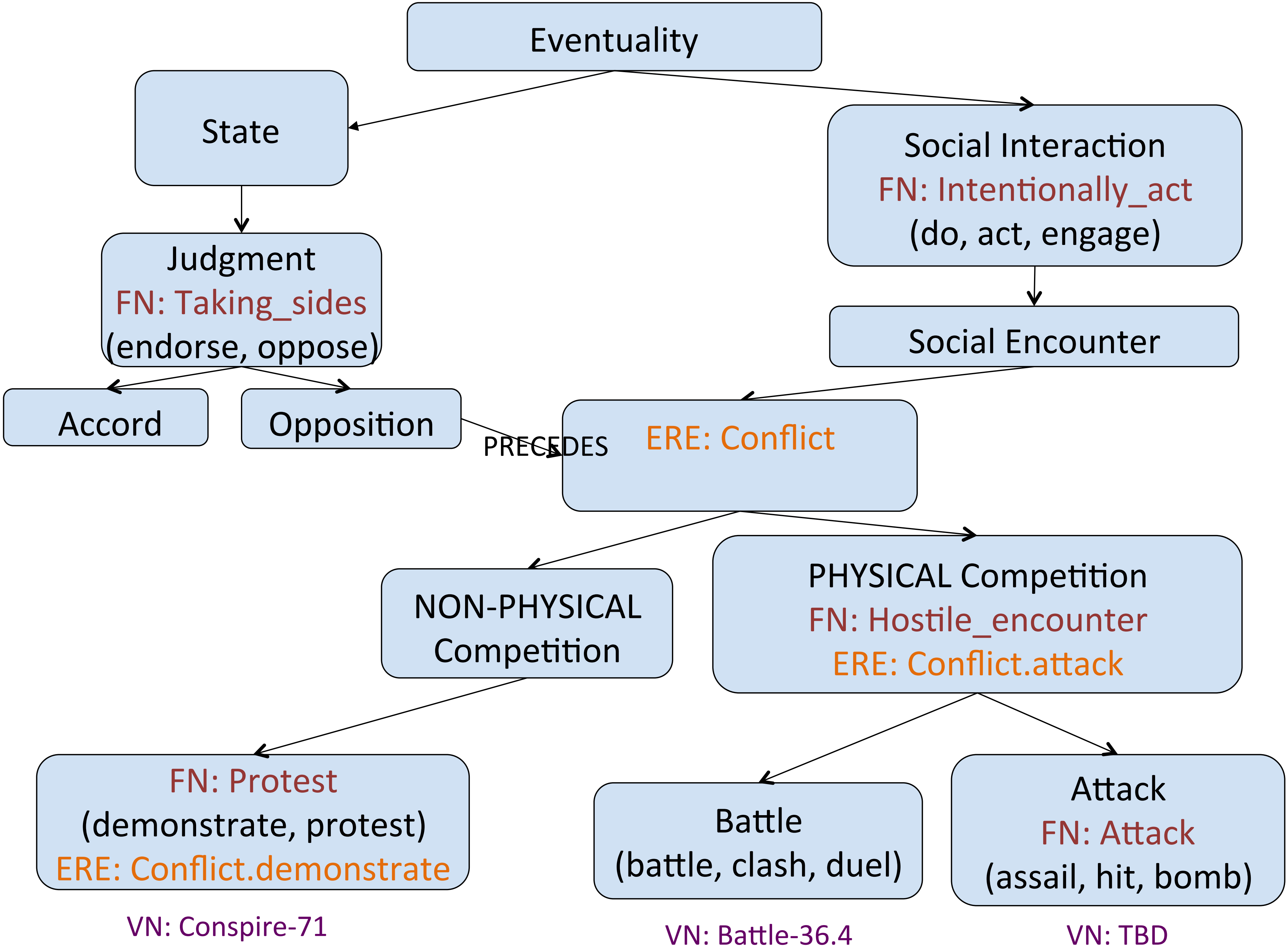


Needed – An Event Ontology

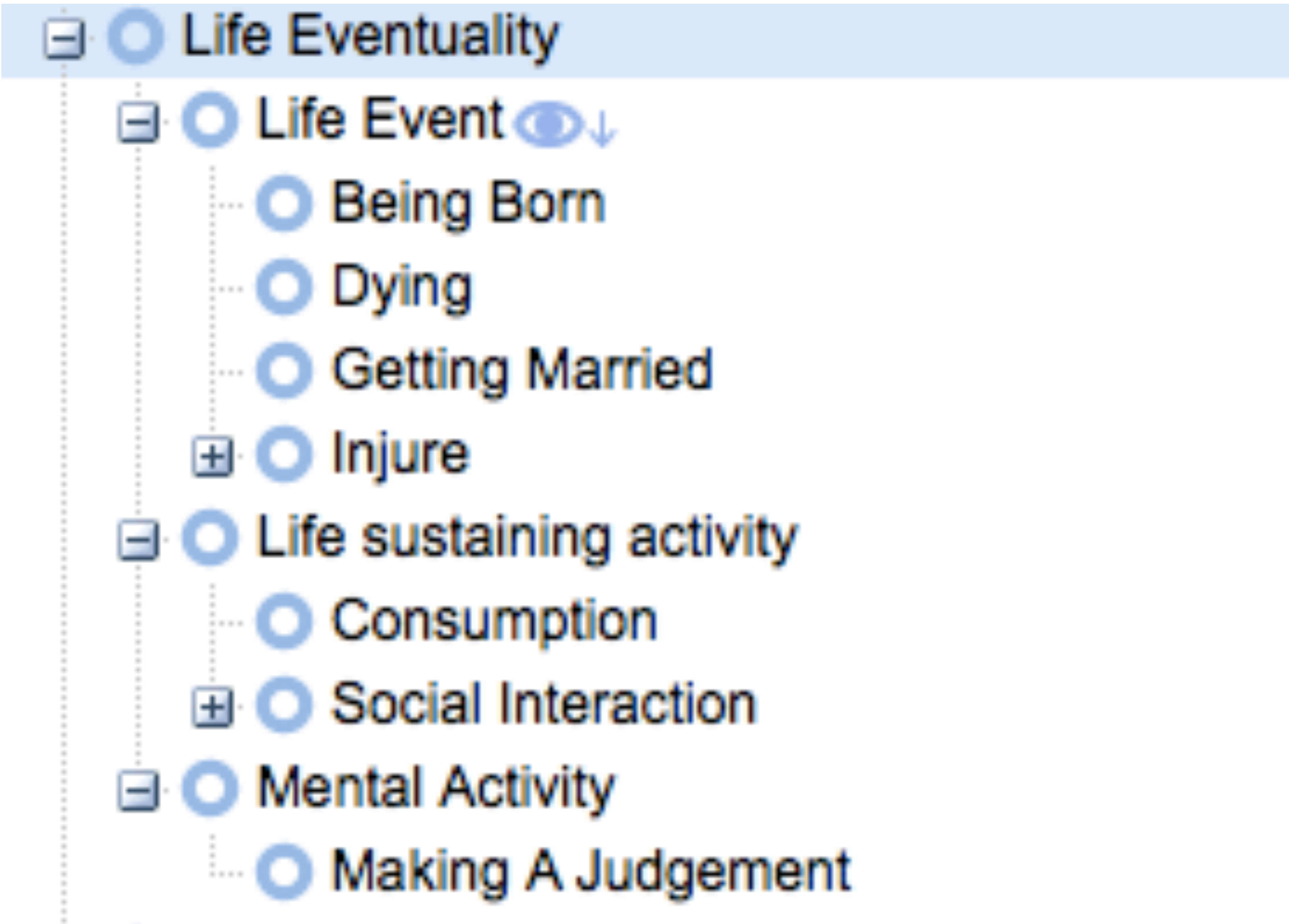
- That can provide appropriate levels of generalization
- DEFT - Event Ontology conference calls
  - *Martha Palmer, James Pustejovsky, Annie Zaenen, Diana McCarthy, Teruko Mitamura, German Rigau, Ann Bies, Kira Griffit, Julie Fitzgerald, Claire Bonial, Derek Palmer*
- Map ERE event types to FrameNet?
- Develop an upper level event ontology that ERE and FN can both be mapped to?

Conflict events

- ERE
  - Attack events
  - Protest/Demonstration events
- FrameNet
  - Attack events - See previous slide
  - Protest, not present
  - Demonstration – Reasoning frame
- VerbNet
  - Attack - Judgment
  - Protest - Conspire
  - Demonstrate – Transfer_message


Eventuality
State
Social Interaction
FN: Intentionally_act
(do, act, engage)
Judgment
FN: Taking_sides
(endorse, oppose)
Social Encounter
Accord
Opposition
PRECEDES
ERE: Conflict
NON-PHYSICAL
Competition
PHYSICAL Competition
FN: Hostile_encounter
ERE: Conflict.attack
FN: Protest
(demonstrate, protest)
ERE: Conflict.demonstrate
Battle
(battle, clash, duel)
Attack
FN: Attack
(assail, hit, bomb)
VN: Conspire-71
VN: Battle-36.4
VN: TBD


Classes
Create
Delete
Search:
Type search string
owl:Thing
Entity
abstract entity
Community
Human Community
concrete entity
Biological Entity
Eventuality
Features
protégé
WebProtege
Verbnet O

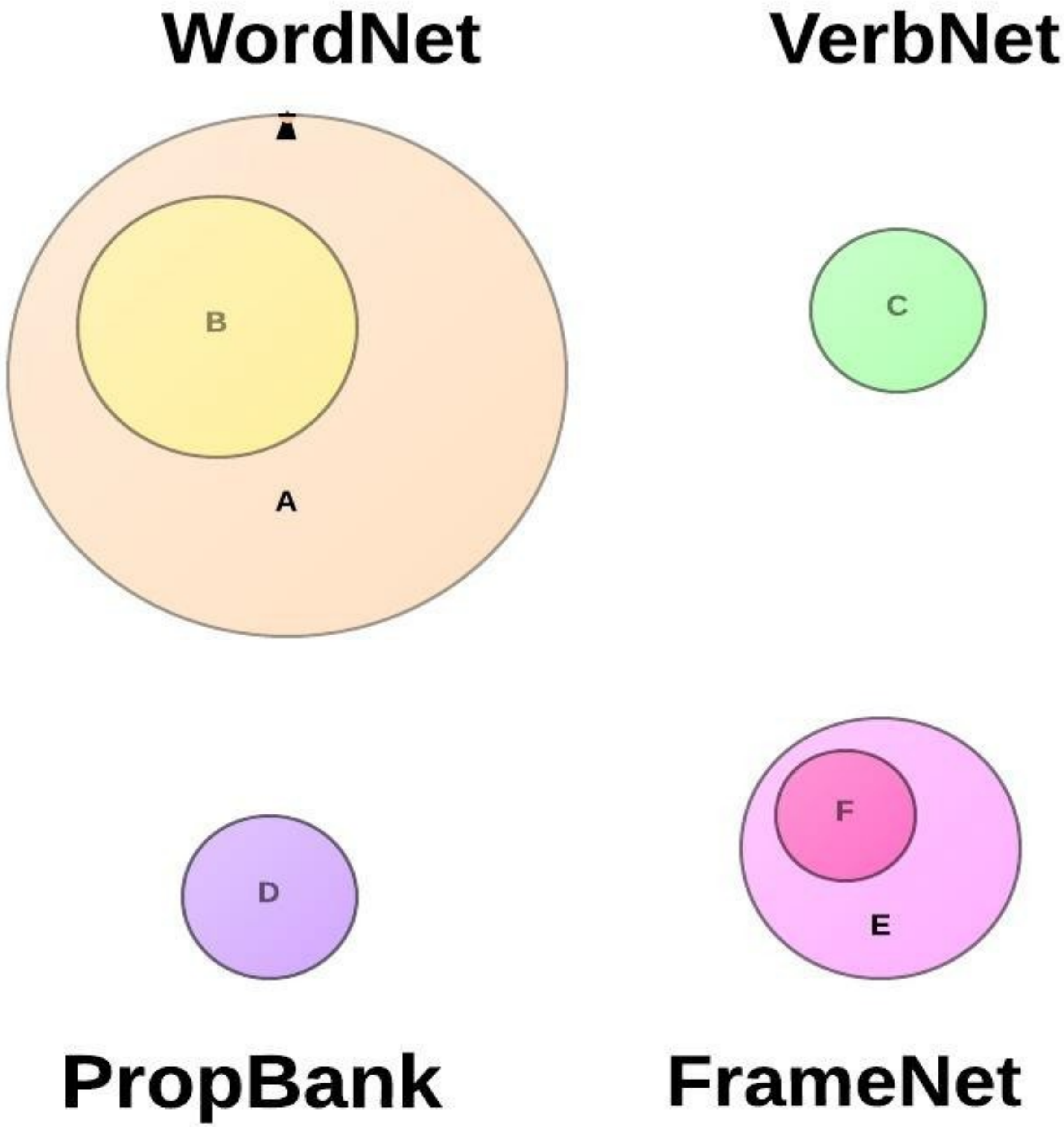

Life Eventuality
Life Event
Being Born
Dying
Getting Married
Injure
Life sustaining activity
Consumption
Social Interaction
Mental Activity
Making A Judgement

Using protégé for Ontologies

- Will coordinate with SUMO, WN also

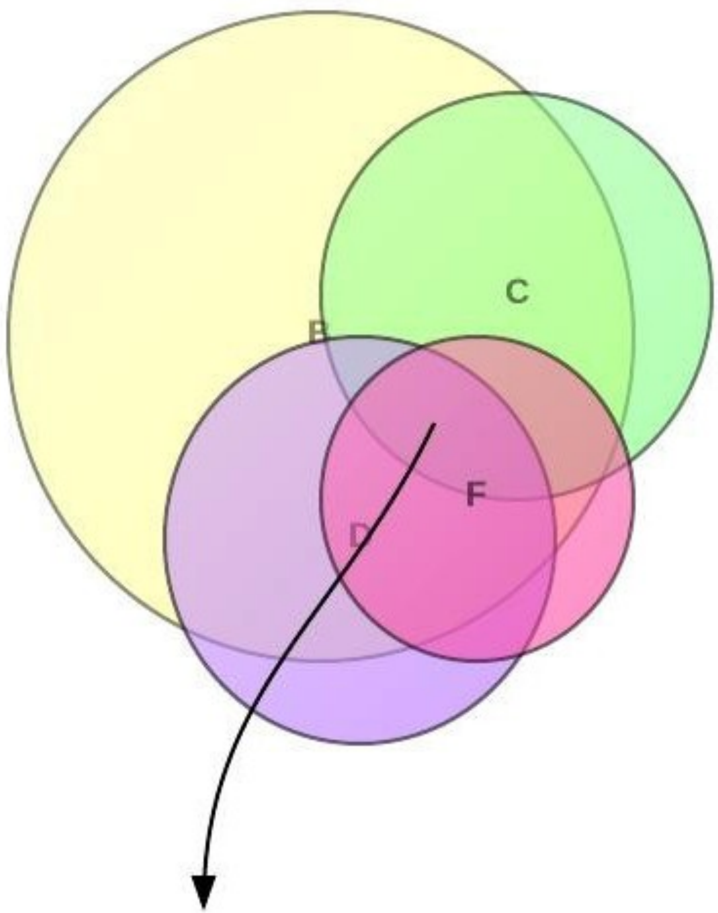
- Any advice?
- Could really use some help from an experienced user
- Could also REALLY use input from all of these "clustering" techniques!

Predicate Matrix - Lacalle, Laparra, Rigau, LREC 2014 -



A: WN Senses (206,941)
B: WN Verb senses (25,041)
C: VN predicates(6293)
D: PB predicates (6,181)
E: FN lexical-units (10,195)
F: FN verb lexical-units (4095)

Predicate Matrix



B: WN Verb senses (25,041)
C: VN predicates(6,293)
D: PB predicates (6,181)
F: FN verb lexical-units (4,095)

SemLink < UNION (B,C,D,E)

Predicate Matrix

- First version 1.0 (GWC 2014)
  - SemLink +
  - Monosemous verbs from VN +
  - Synonyms from WN

- Second version 1.1 (LREC 2014)
  - SemLink +
  - Automatic mappings between predicates +
    - WN-VN and WN-FN (**new mappings!**)
  - Project VN roles to FN roles (**complete gaps!**) +
  - Synonyms from WN

Where we need to go

- Recovery of implicit arguments
- Recovery of implicit relations
- Better Entity coreference
- Event coreference
- Temporal and Causal ordering of events
- Generalizations over event types

Lexical resources can provide

- Generalizations about subcat frames & roles
- Backoff classes for OOV items for portability
- Semantic similarities/”types” for verbs
- Event type hierarchies for inferencing
- Need to be unified and empirically validated and extended: Semlink+
  - VN & FN need PB like coverage, and techniques for extension and automatic domain adaptation
- ***Hybrid lexicons – symbolic and statistical lexical entries?***

Acknowledgments

- We gratefully acknowledge the support of the National Science Foundation Grants for Consistent Criteria for Word Sense Disambiguation, Robust Semantic Parsing, Richer Representations for Machine Translation, A Bayesian Approach to Dynamic Lexical Resources for Flexible Language Processing, DARPA-GALE via a subcontract from BBN, DARPA-BOLT & DEFT via a subcontract from LDC, and DTRA: SemLink+ via a subcontract from BBN, and NIH THYME.
- Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, DARPA or NIH.

And thanks to

- Postdocs: Paul Kingsbury, Dan Gildea, Nianwen Xue, Jinying Chen
- Students: Joseph Rosenzweig, Hoa Dang, Tom Morton, Karin Kipper Schuler, Jinying Chen, Szu-Ting Yi, Edward Loper, Susan Brown, Dmitriy Dligach, Jena Hwang, Will Corvey, Claire Bonial, Jinho Choi, Lee Becker, Shumin Wu, Kevin Stowe
- Collaborators: Christiane Fellbaum, Suzanne Stevenson, Annie Zaenen, Orin Hargraves, James Pustejovsky