FrameNet for NLP

Getting the **meaningful** roles right

Collin F. Baker, ICSI

FrameNet Tutorial, NAACL-HLT Denver, 2015

Overview

- **Frame semantics for NLP**
- **Comparison with three other Semantic Resources for NLP**
  - WordNet : FrameNet
  - PropBank : FrameNet
  - AMR : FrameNet
- **Frame Semantics across languages**

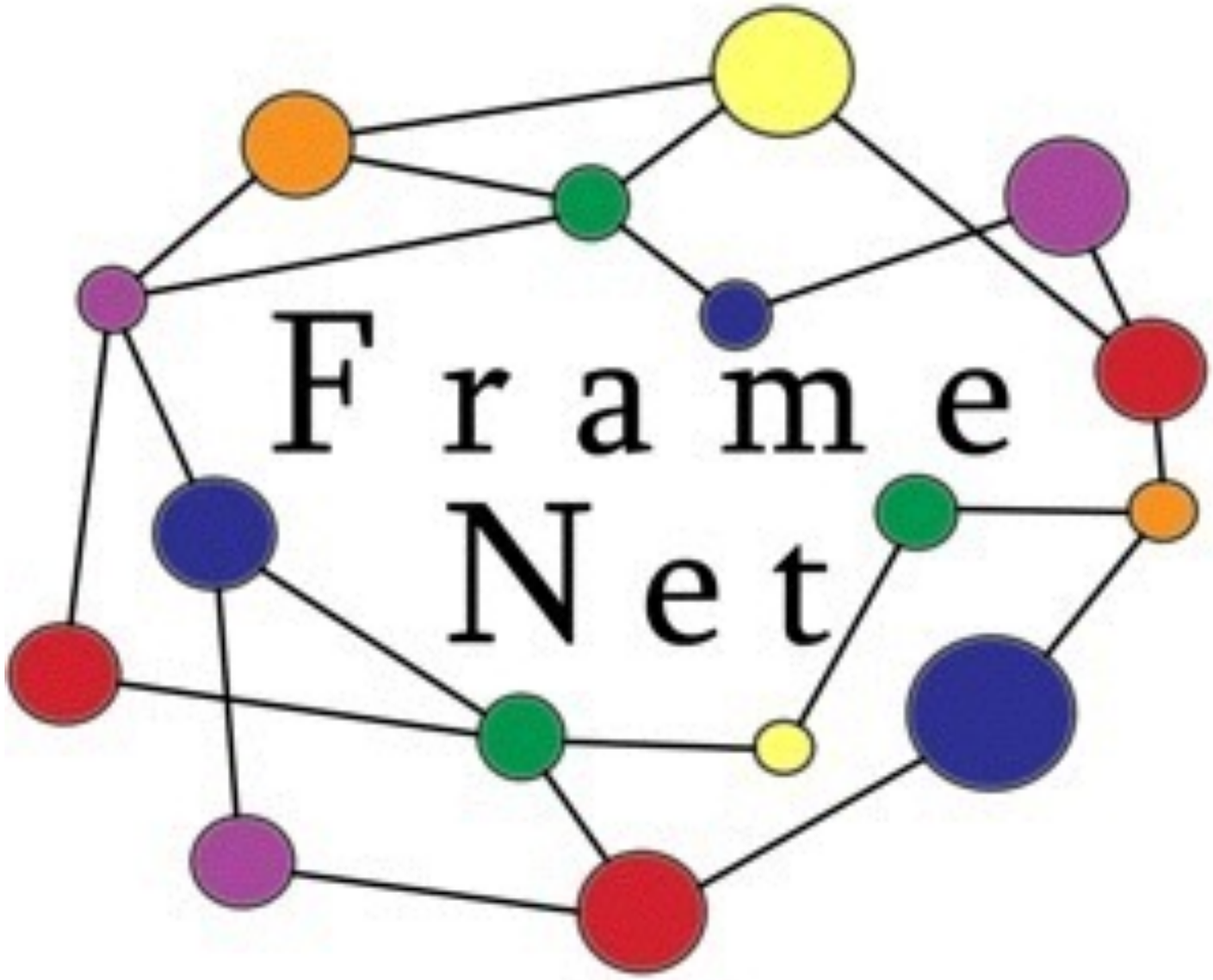
Frame Semantics

- The central idea of Frame Semantics is that word meanings must be described in relation to semantic frames – schematic representations of the conceptual structures and patterns of beliefs, practices, institutions, images, etc. that provide a foundation for meaningful interaction in a given speech community. (Fillmore *et al.* 2003)

Frame Semantics

- The central idea of Frame Semantics is that word meanings must be described in relation to semantic frames – schematic representations of the conceptual structures and patterns of beliefs, practices, institutions, images, etc. that provide a foundation for meaningful interaction in a given speech community. (Fillmore 2003)

- Meanings are relativized to frames. (Fillmore 1977)



Frame Semantics in Practice

https://framenet.icsi.berkeley.edu

The Scope of Frame Semantics

The Scope of Frame Semantics

- **Events**
  - **Being born**: LUs: *born.v, come into the world.v*
  - **Giving birth**: *bear.v, beget.v, birth.n, birth.v, bring forth.v,, carry to term.v, have.v .*
  - **Death**: c*roak.v, death.n, demise.n, die.v, end.n, expire.v, kick the bucket.v . . .*

The Scope of Frame Semantics

- **Events**
  - **Being born**: LUs: *born.v, come into the world.v*
  - **Giving birth**: *bear.v, beget.v, birth.n, birth.v, bring forth.v,, carry to term.v, have.v .*
  - **Death**: c*roak.v, death.n, demise.n, die.v, end.n, expire.v, kick the bucket.v . . .*
- **Relations**
  - **Being relevant**: *irrelevant.a, pertinent.a, play (into).v, relevant.a*
  - **Personal Relationship**: *adultery.n, affair.n, affianced.a, amigo.n, bachelor.n, beau.n*

The Scope of Frame Semantics

- **Events**
  - **Being born**: LUs: *born.v, come into the world.v*
  - **Giving birth**: *bear.v, beget.v, birth.n, birth.v, bring forth.v,, carry to term.v, have.v .*
  - **Death**: c*roak.v, death.n, demise.n, die.v, end.n, expire.v, kick the bucket.v . . .*
- **Relations**
  - **Being relevant**: *irrelevant.a, pertinent.a, play (into).v, relevant.a*
  - **Personal Relationship**: *adultery.n, affair.n, affianced.a, amigo.n, bachelor.n, beau.n*
- **States**
  - **Being in operation**: *off.prep, on.prep, operate.v, operational.a*
  - **Being located**: *find.v, lie.v, located.a, sit.v, situated.a, (ten)-twenty.n, whereabouts.n*

The Scope of Frame Semantics

- **Events**
  - **Being born**: LUs: *born.v, come into the world.v*
  - **Giving birth**: *bear.v, beget.v, birth.n, birth.v, bring forth.v,, carry to term.v, have.v .*
  - **Death**: c*roak.v, death.n, demise.n, die.v, end.n, expire.v, kick the bucket.v . . .*
- **Relations**
  - **Being relevant**: *irrelevant.a, pertinent.a, play (into).v, relevant.a*
  - **Personal Relationship**: *adultery.n, affair.n, affianced.a, amigo.n, bachelor.n, beau.n*
- **States**
  - **Being in operation**: *off.prep, on.prep, operate.v, operational.a*
  - **Being located**: *find.v, lie.v, located.a, sit.v, situated.a, (ten)-twenty.n, whereabouts.n*
- **Entities**
  - **Gizmo**:*appliance.n, centrifuge.n, contraption.n, device.n, gear.n, machine.n,*

FrameNet in two slides (1)

- 1,195 Semantic Frames
- Frame Elements (FEs) (roles) ~12/frame, some ``core'' FEs
- 12,989 Lexical Units (LUs), connections between one lemma+POS and one frame
- 198,932 Manual annotations of corpus examples
- 1,774 Frame-to-frame relations: Inheritance, Using, Perspective on, ...
- (and closely-related FE-to-FE relations)

FrameNet in two slides (2)

- Multiple inheritance, forms a lattice (Valverde-Albacete 2008)
- Constructed bottom-up, quasi-ontology
- Semantic type labels: Animate, Human, Positive_judgement, ...
- Assumes Construction Grammar (“the Constructicon”)
- Metaphor and metonymy: sometimes even marked

Is FrameNet an Ontology?

- Not intended as a formal ontology
- Linguistically motivated, bottom up
- We have made non-lexical frames as needed to connect the nodes in certain cases (e.g. Placing, Filling)
- Some non-lexical frames could be eliminated if we had a new frame-frame relation, Entailment.
- Cross-linguistic differences are to be expected: e.g. Commercial_transaction vs. Criminal_process

NLP needs frame semantics

- Frames provide generalizations about lexical units at a useful level of abstraction, e.g. **Operate vehicle** covers *drive.v, fly.v, paddle.v, sail.v,* etc. useful for paraphrase

- Roles (Frame Elements) are also more meaningful than traditional semantic role labels, e.g. **Driver** in **Operate vehicle** for all the types of vehicle tells us more than just **Agent.**

- Frames represent conceptual gestalts--more than just the sum of their parts

NLP and FrameNet

- Automatic Semantic Role Labeling (ASRL)-- long history, beginning with Gildea and Jurafsky 2002, through Dipanjan Das et al. 2010 and ff.

- Semantic parsing, joint inference (Das et al. ACL 2013)

- Automatic frame induction (Hermann et al. ACL 2014)

- Controlled crowd-sourcing of annotation (current work with Google, in Nancy Chang *et al.* LAW 2015)

Comparison of Resources

Comparison of NLP Resources

- WordNet : FrameNet
- PropBank : FrameNet
- AMR : FrameNet

WordNet : FrameNet

POS	WordNet	FrameNet
Noun	146,312	5,177
Verb	25,047	4,879
Adj	30,002	2,270
Adv	5,580	(other) 387
Total word senses	206,941	12,713

- Q: Why would you use FrameNet when you have WordNet?
- A: The lexical information each contains is different, and in many ways complementary:
- FN has little to say about common nouns; WN noun hierarchies are usually good
- WN has little syntagmatic information, FN has a lot

WordNet : FrameNet

Structural differences:

- WN: separate hierarchies for N, V, Adj, and Adv., (only)
- Each FN frame can cover words of any POS
- WN: hyper-/hyponymy relations between synsets (i.e. between LUs that are roughly synonymous)
- FN: no **LU** relations *per se*, but several types of **frame** relations
- **Content differences:**
- FN: annotated examples showing syntax and semantics of each LU.
- FN describes roles (Frame Elements) for each frame
- FN frame hierarchy often provides better generalizations than WN synset hierarchy

PropBank : FrameNet

- Proposition Bank (Palmer *et al.* 2005 *CL*) began with labeling verbs and their arguments (and adjuncts) in *WSJ*
- Uses Penn POS tags, Penn TreeBank parse structures
- Later added nouns and the roles from associated verbs
- Substantial work on Chinese, Korean, Arabic, Hindi, biomedical domain
- Efficient semantic role labeling systems available
- No equivalent of frames: two levels of role labels:
    1. completely general, ARG0-ARG5, ARGM-Time, ARGM-Loc, etc.
    2. specific to lexical unit (word sense)

Comparing annotation PB:FN

Text	PB Arg	PB Verb Specific	FN FE name
The internal investigation	Arg0	critic	Communicator
also	ArgM-dis	-	-
CRITICIZED	(Rel)	-	(Target)
MiniScribe's auditors, Coopers & Lybrand,	Arg1	entity being criticized	Evaluee
for allegedly ignoring numerous red flags	Arg2	on what grounds?	Reason

Comparing Role Names PB:FN

Verb	Arg0	Arg1	Arg2
criticize	critic	entity being criticized	on what grounds?
disparage	talker	victim	-
denigrate	speaker	subject	grounds, reason
acclaim	acclaimer	acclaimed	cause, acclaimed for what?
commend	entity giving praise	entity being praised	praised for what?
FN FE Name	Communicator	Evaluee	Reason

Abstract Meaning Representation

(Banarescu et al., LAW 2013)

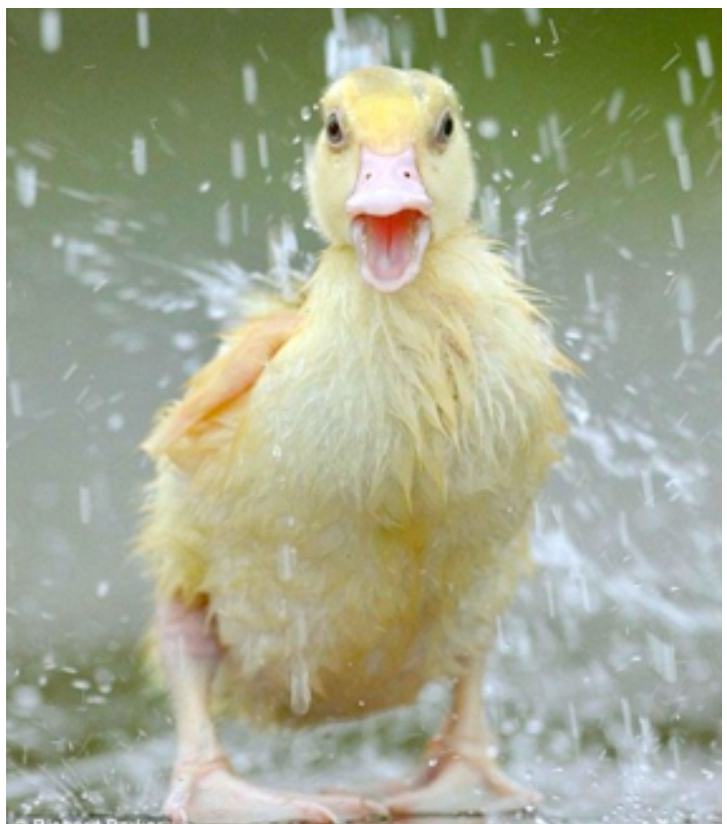
A graph-based representation of lexical concepts and typed relations between those concepts that are denoted by an English sentence.

AMR integrates several aspects of lexical/relational meaning—abstracting away from the grammatical details—in a single structure designed to support rapid corpus annotation and data-driven NLP.

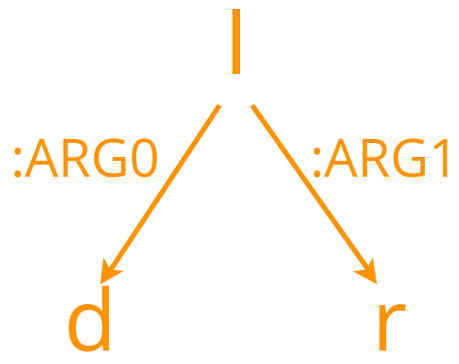
AMRs

```
(l / like-01
   :ARG0 (d / duck)
   :ARG1 (r / rain-01))
```

- ducks like rain
- the duck liked that it was raining

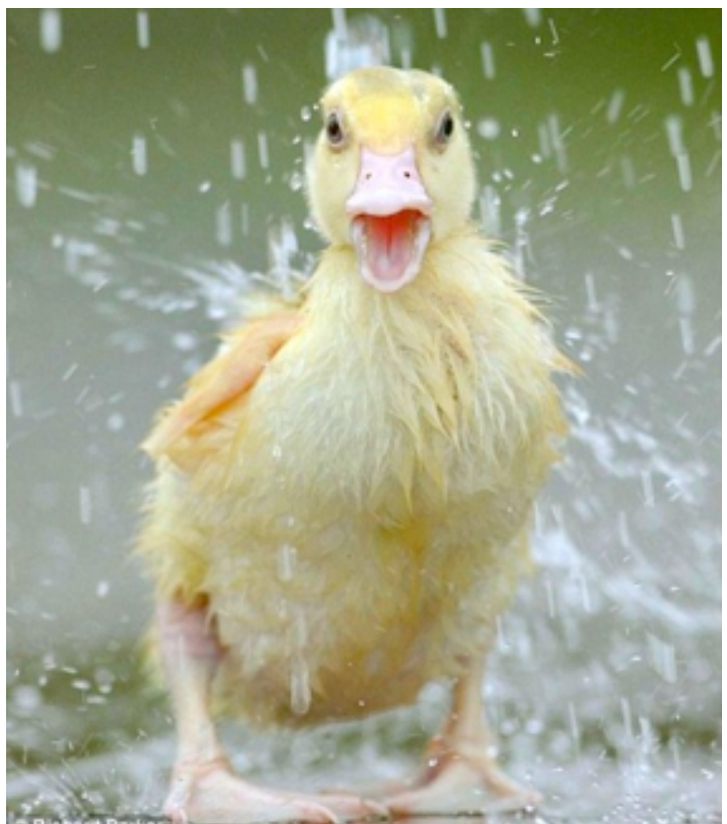


AMRs

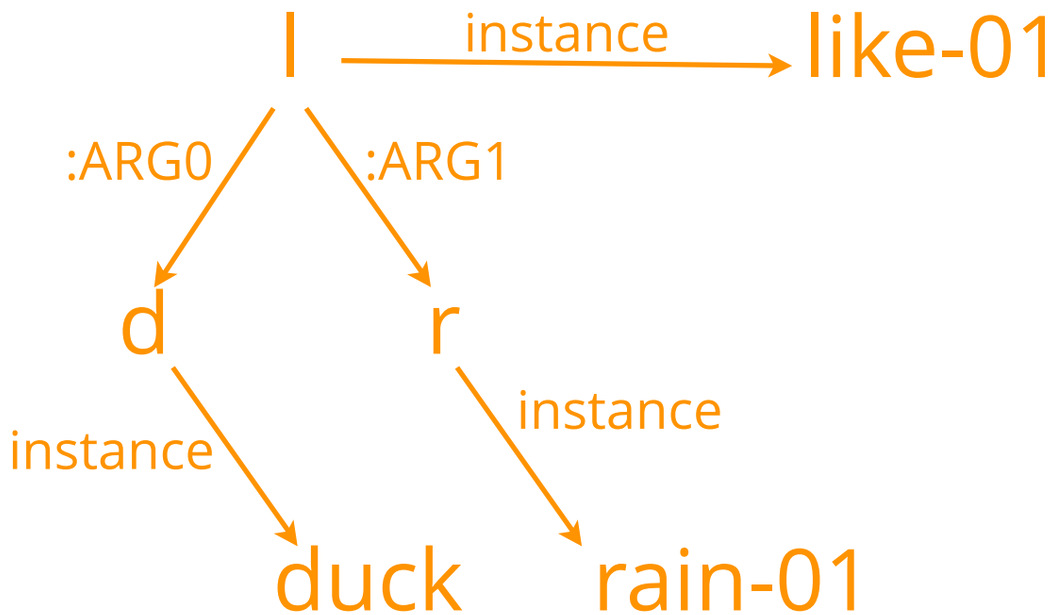



```
(l / like-01
   :ARG0 (d / duck)
   :ARG1 (r / rain-01))
```

- ducks like rain
- the duck liked that it was raining

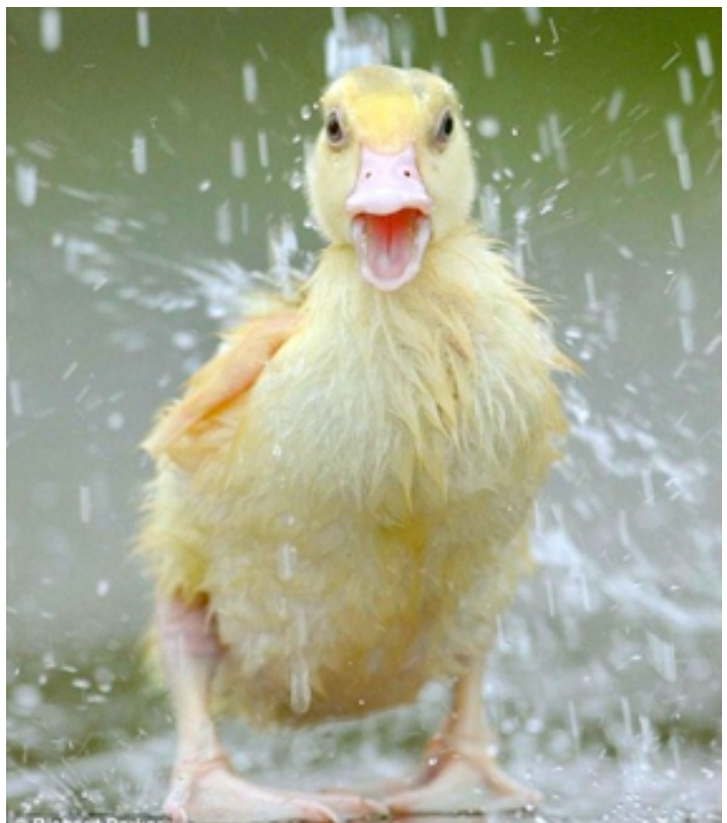


AMRs



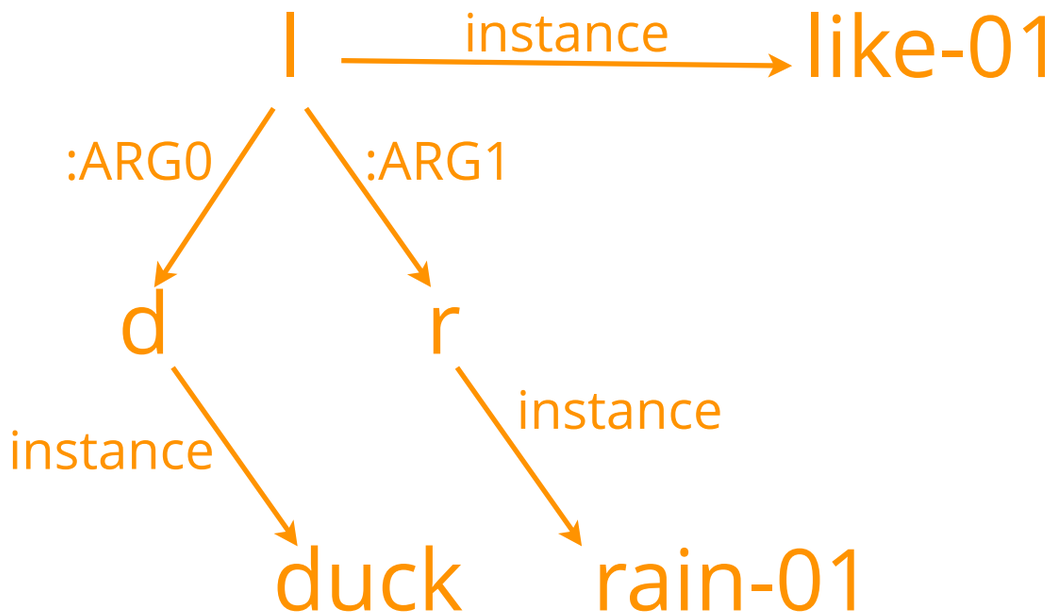
(l / like-01
:ARG0 (d / duck)
:ARG1 (r / rain-01))

- ducks like rain
- the duck liked that it was raining

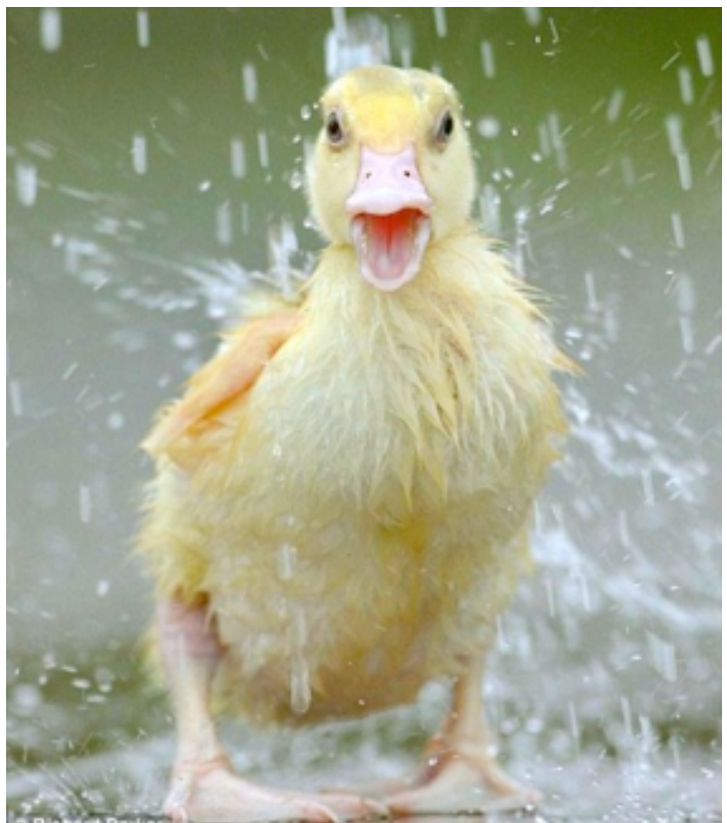


(l / like-01
  :ARG0 (d / duck)
  :ARG1 (r / rain-01))

AMRs



- ducks like rain
- the duck liked that it was raining

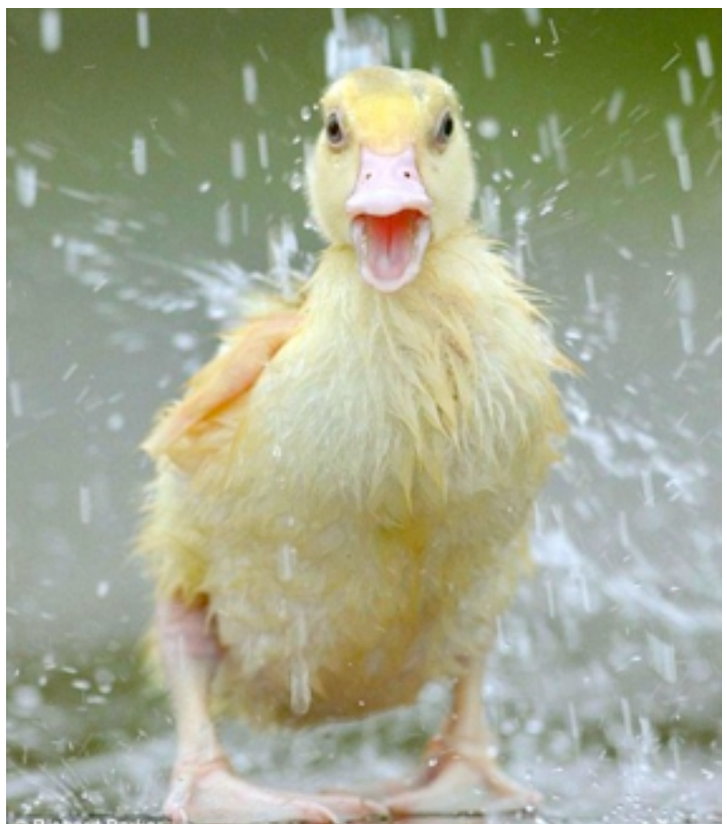


AMRs

```
(l / like-01
  :ARG0 (d / duck)
  :ARG1 (r / rain-01))
```

```
(s2 / see-01
  :ARG0 (i / i)
  :ARG1 (d / duck
          :poss (s / she)))
```

- I saw her duck

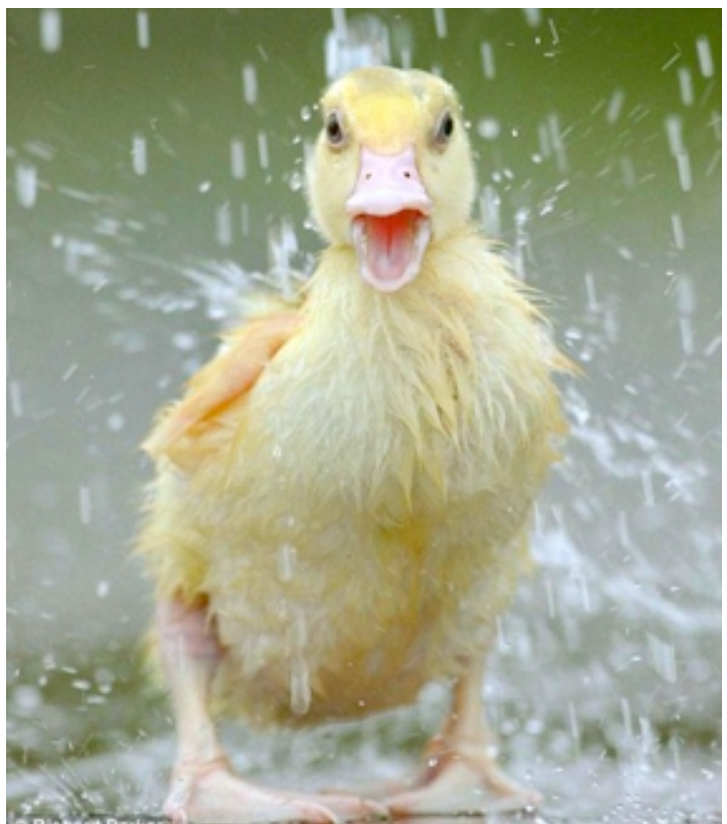


(l / like-01
  :ARG0 (d / duck)
  :ARG1 (r / rain-01))

AMRs

(s2 / see-01
  :ARG0 (i / i)
  :ARG1 (d / duck
    :poss (s / she)))

- I saw her duck



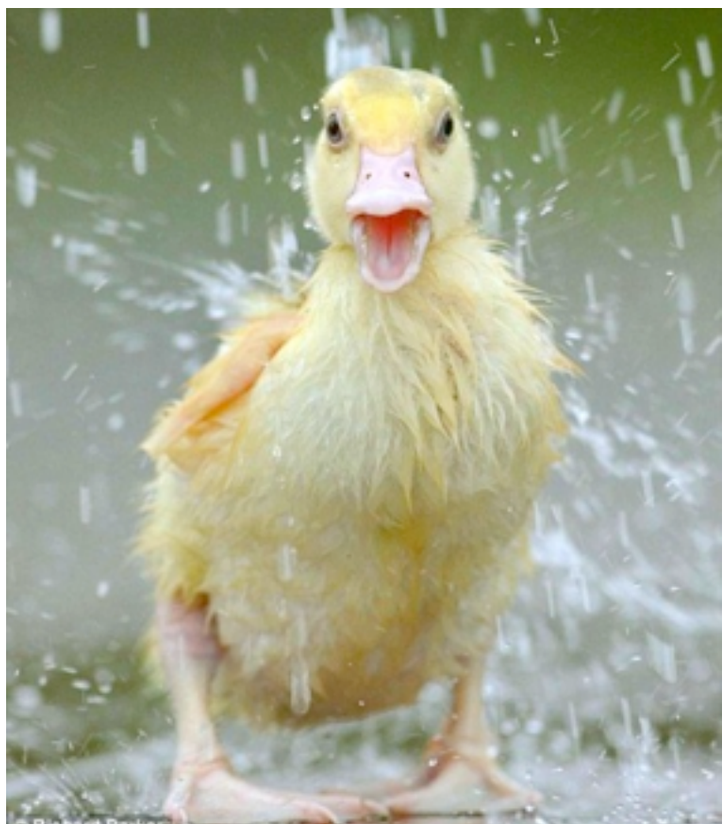
AMRs

```
(l / like-01
  :ARG0 (d / duck)
  :ARG1 (r / rain-01))
```

```
(s2 / see-01
  :ARG0 (i / i)
  :ARG1 (d / duck
        :poss (s / she)))
```

```
(s2 / see-01
  :ARG0 (i / i)
  :ARG1 (d / duck-01
        :ARG0 (s / she)))
```

- I saw her duck **(alternate interpretation)**



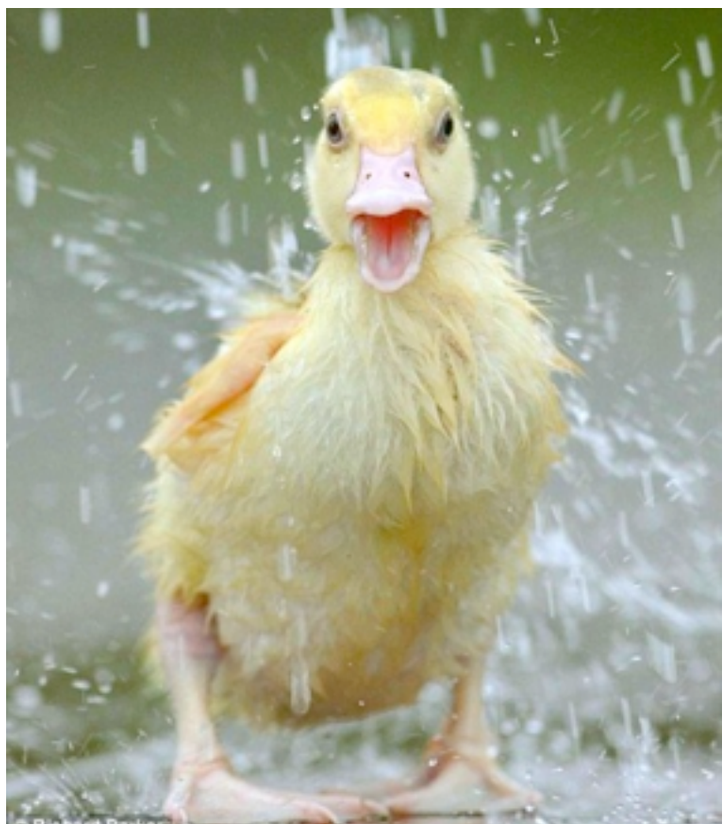
AMRs

```
(l / like-01
  :ARG0 (d / duck)
  :ARG1 (r / rain-01))
```

```
(s2 / see-01
   :ARG0 (i / i)
   :ARG1 (d / duck
            :poss (s / she)))
```

```
(s2 / see-01
   :ARG0 (s / she)
   :ARG1 (d / duck
            :poss s))
```

- She saw her (own) duck

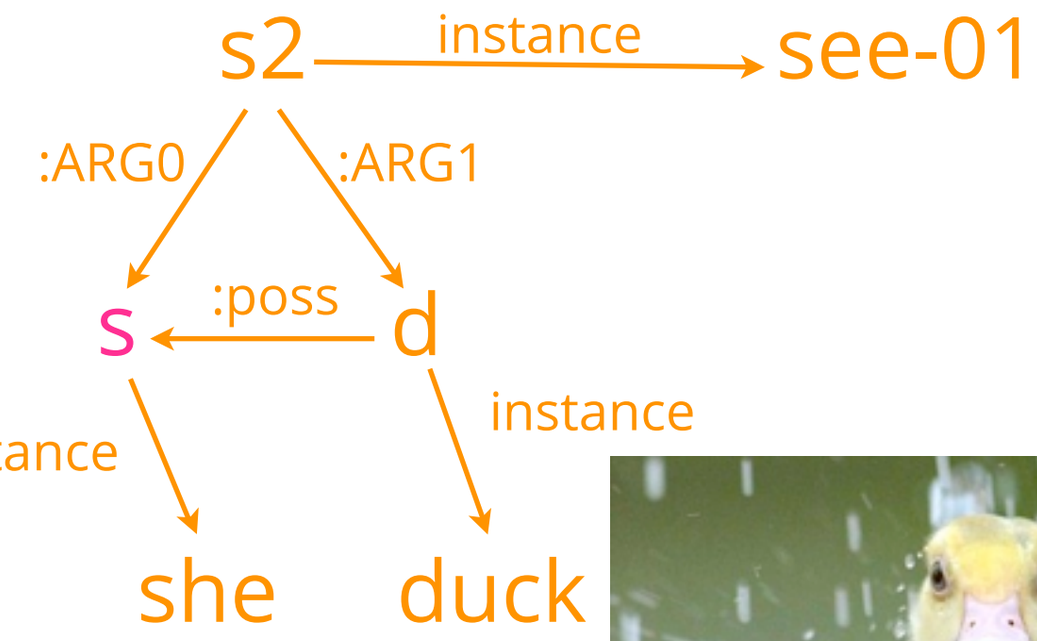


AMRs

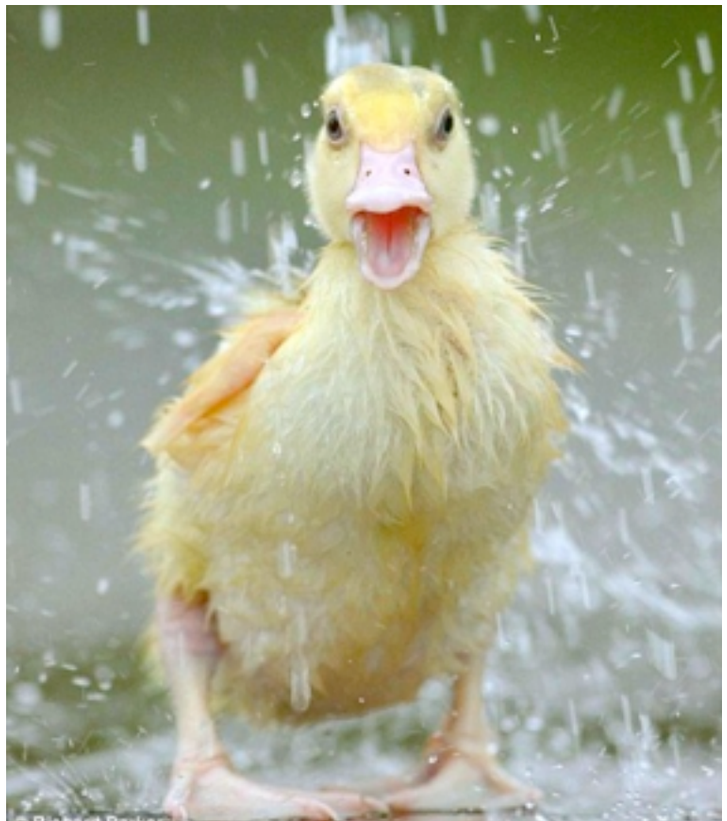
(l / like-01
  :ARG0 (d / duck)
  :ARG1 (r / rain-01))

(s2 / see-01
  :ARG0 (i / i)
  :ARG1 (d / duck
    :poss (s / she)))

(s2 / see-01
  :ARG0 (s / she)
  :ARG1 (d / duck
    :poss s))



- She saw her (own) duck



AMR Features

AMR Features

- PropBank predicate-argument semantics

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference entities & events

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference
- modality, negation, questions

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference
- modality, negation, questions
- relations between nominals

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference
- modality, negation, questions
- relations between nominals
- canonicalization of content words (remove inflectional morphology, convert adv → adj → noun → verb where possible)

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference
- modality, negation, questions
- relations between nominals
- canonicalization of content words (remove inflectional morphology, convert adv → adj → noun → verb where possible)

his trial → (t / try-02 :ARG1 (h / he))

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference
- modality, negation, questions

  history teacher → (p / person
- relations between nominals :ARG0-of (t / teach-01
  :ARG1 (h / history)))
- canonicalization of content words (remove inflectional morphology, convert adv → adj → noun → verb where possible)

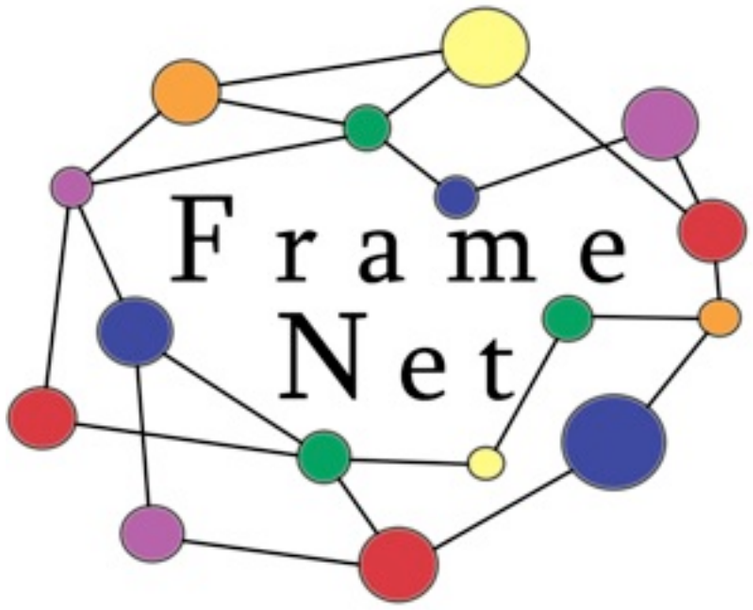
  his trial → (t / try-02 :ARG1 (h / he))

AMR Features

- PropBank predicate-argument semantics
- name & value entities; entity linking (wikification)
- coreference
- modality, negation, questions
- relations between nominals
- canonicalization of content words (remove inflectional morphology, convert adv → adj → noun → verb where possible)
- …*all in a single graph!*

AMR Assets

- Snazzy annotation tool
- Evaluation method (smatch)
- Extensive documentation (guidelines, help pages in tool, heuristics in tool)
- Close coordination with PropBank
- Annotation sites: CU, ISI, SDL, LDC
- Data: 15,000 AMRs (270k words) **released**, another 5,000 AMRs (150k words) **annotated** and in the pipeline



vs.

Berkeley FrameNet
https://framenet.icsi.berkeley.edu/

http://amr.isi.edu/

AMR vs. FrameNet

AMR vs. FrameNet

- *Scope:* AMR also describes coreference, named entities & value expressions, etc.—roughly, accounts for **all content words** and their relations. But **lexical semantics is much shallower** than FN.

AMR vs. FrameNet

- *Scope:* AMR also describes coreference, named entities & value expressions, etc.—roughly, accounts for **all content words** and their relations. But **lexical semantics is much shallower** than FN.

- *Original goals:* Cost-effective **annotation + NLP** vs. **lexicography**

AMR vs. FrameNet

- *Scope:* AMR also describes coreference, named entities & value expressions, etc.—roughly, accounts for **all content words** and their relations. But **lexical semantics is much shallower** than FN.

- *Original goals:* Cost-effective **annotation + NLP** vs. **lexicography**

- *Event predicates:* **PropBank rolesets** (lexicalized, coarsely disambiguated) vs. **FrameNet frames** (rich semantic groupings)

AMR vs. FrameNet

- *Scope:* AMR also describes coreference, named entities & value expressions, etc.—roughly, accounts for **all content words** and their relations. But **lexical semantics is much shallower** than FN.

- *Original goals:* Cost-effective **annotation + NLP** vs. **lexicography**

- *Event predicates:* **PropBank rolesets** (lexicalized, coarsely disambiguated) vs. **FrameNet frames** (rich semantic groupings)

- **FN:** annotations are labeled spans in the sentence;
  AMR: parts of graph not explicitly aligned to the sentence, and not all concepts are words in the sentence

AMR vs. FrameNet

- *Scope:* AMR also describes coreference, named entities & value expressions, etc.—roughly, accounts for **all content words** and their relations. But **lexical semantics is much shallower** than FN.

- *Original goals:* Cost-effective **annotation + NLP** vs. **lexicography**

- *Event predicates:* **PropBank rolesets** (lexicalized, coarsely disambiguated) vs. **FrameNet frames** (rich semantic groupings)

- **FN:** annotations are labeled spans in the sentence;
  AMR: parts of graph not explicitly aligned to the sentence, and not all concepts are words in the sentence

- **FN:** for a sentence, no explicit relationship across frame annotations;
  AMR: composition of predicates, shared arguments are explicit

Frame Semantics across languages

Example FrameNets

- Spanish FrameNet http://sfn.uab.es (UA Barcelona, Carlos Subirats)

- Swedish FrameNet http://spraakbanken.gu.se/eng/swefn (U Gothenburg, Lars Borin)

- Japanese FrameNet http://jfn.st.hc.keio.ac.jp (Keio U/U Tokyo, Kyoko Ohara)

Spanish FrameNet

- Created a new balanced corpus, mainly New World Spanish, and used their own POS tagger
- Manual annotation, following Berkeley closely,
- Generally English frames were OK, differences re: verbs of motion (verb-framed vs. satellite framed)

Swedish FrameNet++

Objectives:

- to link a number of existing free lexical resources, both in-house and external, both modern and historical, into an integrated lexical macro-resource

- to create a full-scale Swedish FrameNet with at least 50,000 lexical units and fully integrated into the macro-resource

- to develop methodologies and workflows which make maximal use of LT tools and large text corpora in order to minimize the human effort needed in the work.

Japanese FrameNet

- Annotating texts from "Balanced Corpus of Contemporary Written Japanese" core data
- Created an web-based annotation tool with multilingual support
- Lots of work on the "Constructicon" for Japanese
- Publications on Japanese-English differences

Counts by POS

	Spanish	Swedish	Japanese	English
Nouns	271	28,891	2,043	5,299
Verbs	856	5,398	908	5,141
Adjectives	99	3,293	134	2,347
Adverbs	16	322	89	214
Other	26	124	231	420
Total LUs	1,268	38,028	3,405	13,421
Annotations	11 k	9 k	73 k	200 k

Towards a Multilingual FrameNet

acercarse.v	ankomma	たどりつく.v	appear.v
acudir.v	ankomst	くる.v	approach.n
adentrarse.v	anlända	つく.v	approach.v
aproximarse.v	anlöpa	なる.v	arrival.n
arrimarse.v	båtankomst	はいる.v	arrive.v
avanzar.v	bussankomst	いたる.v	come.v
encaminarse.v	dyka_upp	せまりくる.v	crest.v
entrar.v	ensamkomma	届く.v	descend_(on).v
inmigrar.v	färjeankomst	帰る.v	enter.v
irrumpir.v	flygankomst	到着.n	entrance.n
llegada.n	fram	到来.n	entry.n
llegar.v	framkomst	もどる.v	find.v
peregrinar.v	framme	やっ て くる.v	get.v
replegarse.v	hamna	入る.v	hit.v
venir.v	hemkommen	入港.n	influx.n
	hemkomst	着く.v	make it.v
	infinna_sig	達する.v	make.v
	inlöpa	迫りくる.v	reach.v
	inresa	近づく.v	return.n
	inställa_sig	至る.v	return.v
		来る.v	visit.n
			visit.v