Learning by Reading

Jornadas TIMM 2014

Horacio Rodríguez
TALP (UPC)

Guión

- Para qué
- Introducción
- Qué leer
- Tareas implicadas
- Conclusiones

Para qué

- Limitaciones de los sistemas de PLN que se basan únicamente en procesamiento superficial.
- Necesidad de procesamiento semántico
- Necesidad de conocimiento del mundo (World Knowledge, Common Sense Knowledge)
- Adquisición de este conocimiento.
- Un ejemplo. RTE

RTE

- Equivalence (**Paraphrase**): *expr1* $\Leftrightarrow$ *expr2*
- **Entailment:** *expr1* $\Rightarrow$ *expr2* – more general

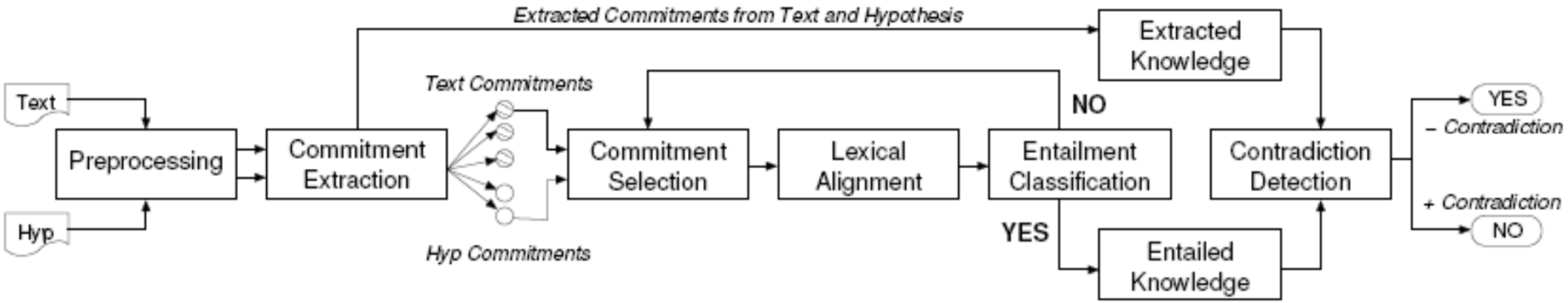
- Directional relation between two text fragments: *Text (t)* and *Hypothesis (h):*

t **entails** *h* (***t*$\Rightarrow$*h***) if, typically, a **human** reading *t* would infer that *h* is **most likely** true"

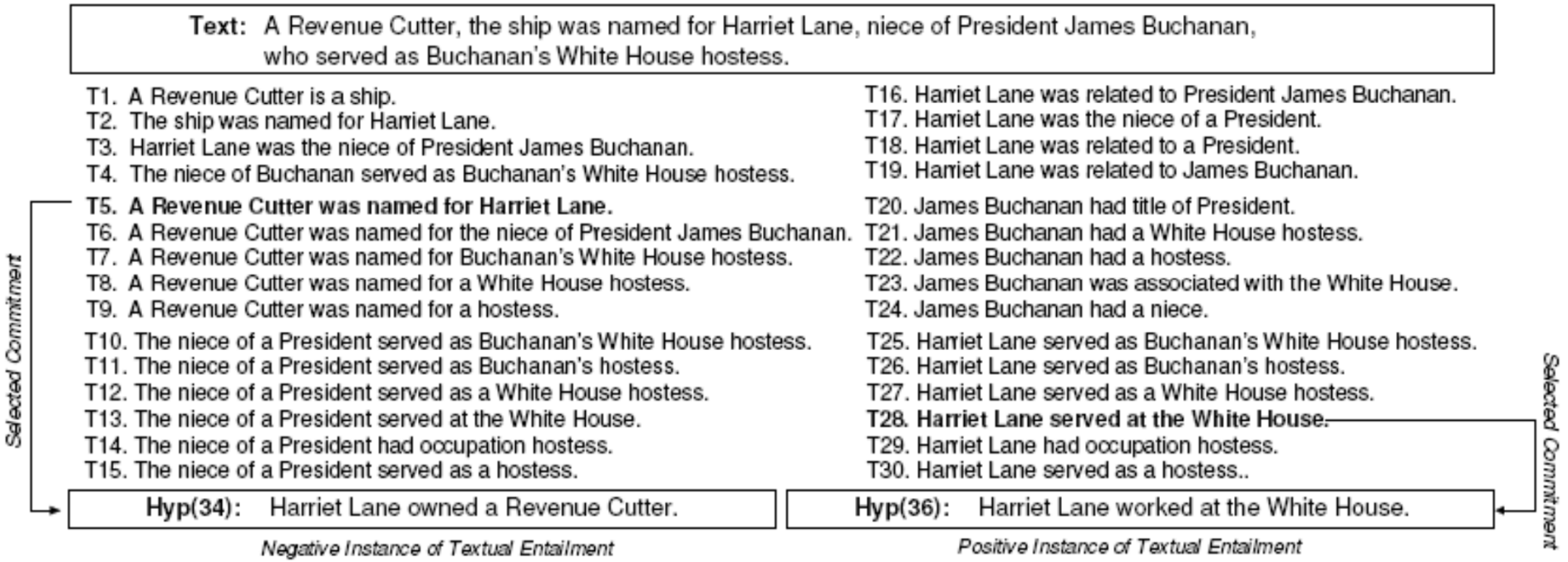
LCC Hickl

LCC, Hickl, 2007

- Discourse Commitment-based Framework



LCC Hickl


Text: A Revenue Cutter, the ship was named for Harriet Lane, niece of President James Buchanan, who served as Buchanan's White House hostess.
T1. A Revenue Cutter is a ship.
T2. The ship was named for Harriet Lane.
T3. Harriet Lane was the niece of President James Buchanan.
T4. The niece of Buchanan served as Buchanan's White House hostess.
T5. A Revenue Cutter was named for Harriet Lane.
T6. A Revenue Cutter was named for the niece of President James Buchanan.
T7. A Revenue Cutter was named for Buchanan's White House hostess.
T8. A Revenue Cutter was named for a White House hostess.
T9. A Revenue Cutter was named for a hostess.
T10. The niece of a President served as Buchanan's White House hostess.
T11. The niece of a President served as Buchanan's hostess.
T12. The niece of a President served as a White House hostess.
T13. The niece of a President served at the White House.
T14. The niece of a President had occupation hostess.
T15. The niece of a President served as a hostess.
T16. Harriet Lane was related to President James Buchanan.
T17. Harriet Lane was the niece of a President.
T18. Harriet Lane was related to a President.
T19. Harriet Lane was related to James Buchanan.
T20. James Buchanan had title of President.
T21. James Buchanan had a White House hostess.
T22. James Buchanan had a hostess.
T23. James Buchanan was associated with the White House.
T24. James Buchanan had a niece.
T25. Harriet Lane served as Buchanan's White House hostess.
T26. Harriet Lane served as Buchanan's hostess.
T27. Harriet Lane served as a White House hostess.
T28. Harriet Lane served at the White House.
T29. Harriet Lane had occupation hostess.
T30. Harriet Lane served as a hostess..
Selected Commitment
Hyp(34): Harriet Lane owned a Revenue Cutter.
Negative Instance of Textual Entailment
Hyp(36): Harriet Lane worked at the White House.
Positive Instance of Textual Entailment
Selected Commitment

LCC Hickl

- Algunas de estos compromisos se pueden deducir con facilidad del texto:
  - Harriet Lane es sobrina de James Buchanan
  - James Buchanan (es/fue) presidente
  - Un Revenue Cutter es un barco
  - …
- **Otros no aparecen explícitamente en el texto y deben ser adquiridos como conocimiento del mundo**:
  - James Buchanan (es/fue) presidente de USA
  - Los presidentes de USA residen en la Casa Blanca
  - La Casa Blanca está en Washington
  - Azafata es una profesión
  - …

Nutcracker

Nutcracker, Roma (La Sapienza)

- Johan Bos, 2007
- Components of Nutcracker:
  - The C&C parser for CCG
  - Boxer
  - Vampire, a FOL theorem prover
  - Paradox and Mace, FOL model builders
- Background knowledge
  - WordNet [hyponyms, synonyms]
  - NomLex [nominalisations]

Nutcracker

- Given a textual entailment pair T/H withtext T and hypothesis H:
  - Produce DRSs for T and H
  - Translate these DRSs into FOL
  - **Generate Background Knowledge in FOL**
- Use ATPs to determine the likelyhood of entailment

Nutcracker

- ...
  - Generate Background Knowledge in FOL
- ...

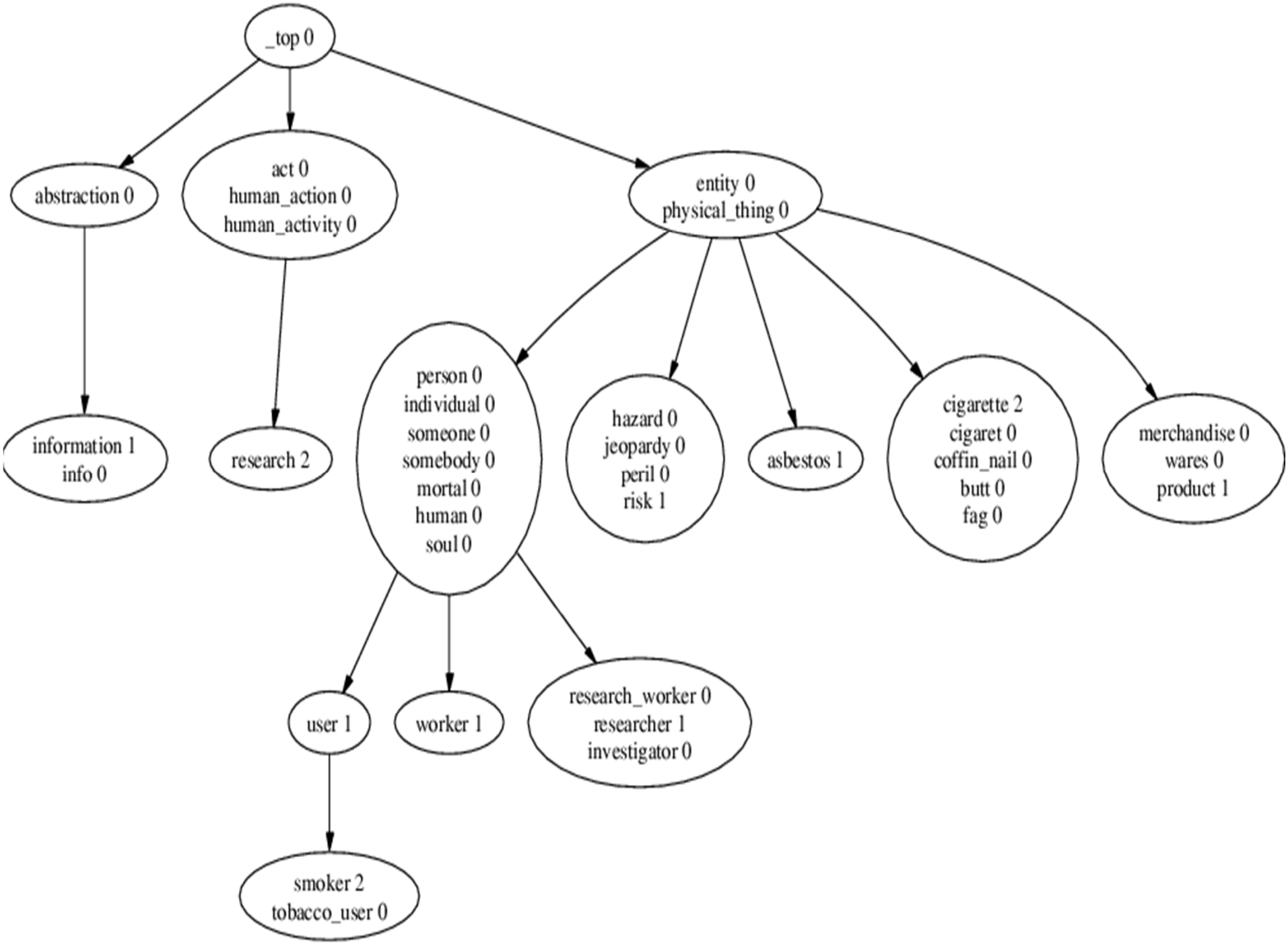
  - MiniWordNets
    - Use hyponym relations from WordNet to build an ontology
    - Do this only for the relevant symbols
    - Convert the ontology into first-order axioms

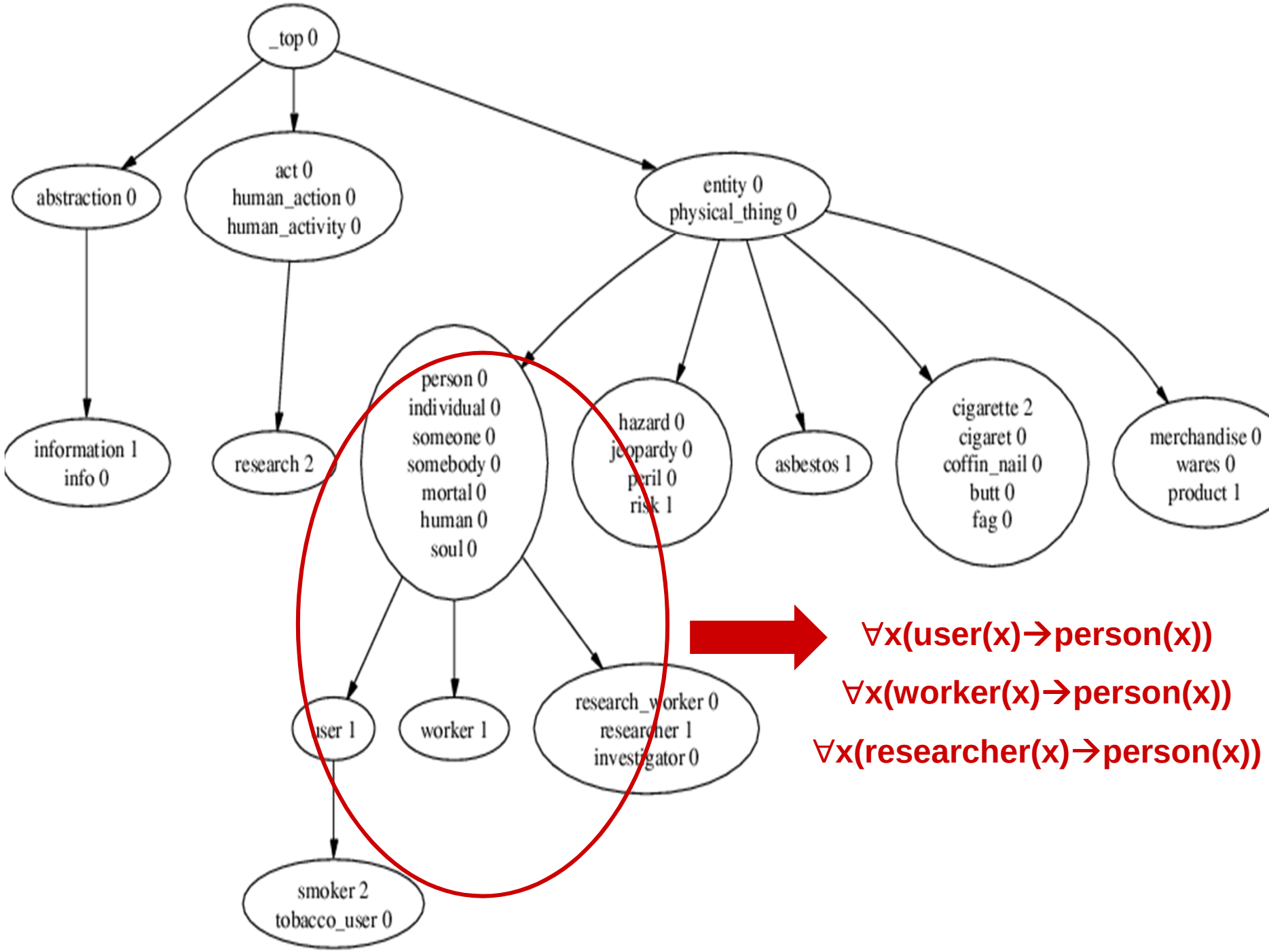
Nutcracker

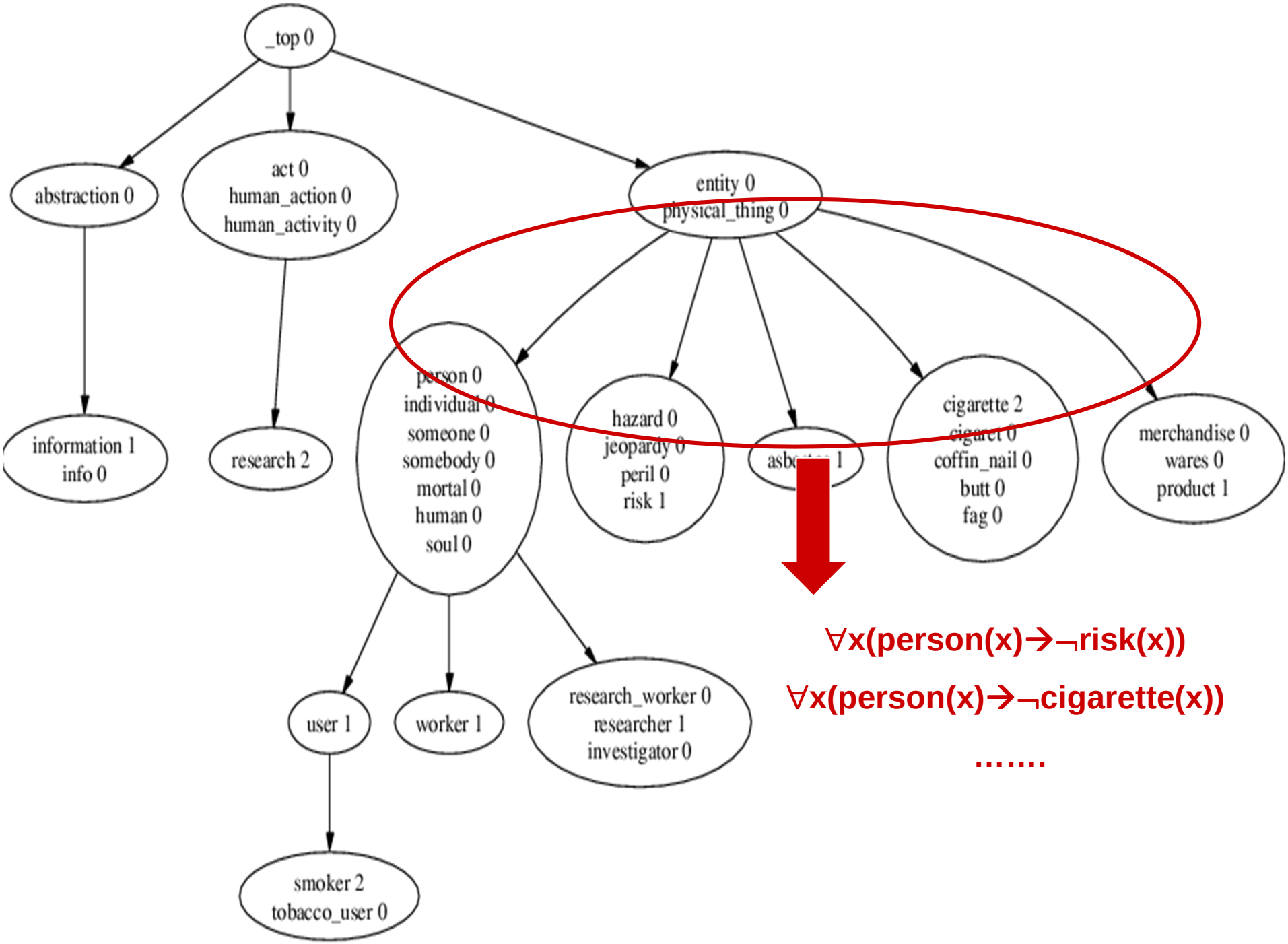
- ...
  - MiniWordNets

    - Example text:

      There is no asbestos in our products now. Neither Lorillard nor the researchers who studied the workers were aware of any research on smokers of the Kent cigarettes.


_top 0
abstraction 0
act 0
human_action 0
human_activity 0
entity 0
physical_thing 0
information 1
info 0
research 2
person 0
individual 0
someone 0
somebody 0
mortal 0
human 0
soul 0
hazard 0
jeopardy 0
peril 0
risk 1
asbestos 1
cigarette 2
cigaret 0
coffin_nail 0
butt 0
fag 0
merchandise 0
wares 0
product 1
user 1
worker 1
research_worker 0
researcher 1
investigator 0
smoker 2
tobacco_user 0


_top 0
abstraction 0
act 0
human_action 0
human_activity 0
entity 0
physical_thing 0
information 1
info 0
research 2
person 0
individual 0
someone 0
somebody 0
mortal 0
human 0
soul 0
hazard 0
jeopardy 0
peril 0
risk 1
asbestos 1
cigarette 2
cigaret 0
coffin_nail 0
butt 0
fag 0
merchandise 0
wares 0
product 1
user 1
worker 1
research_worker 0
researcher 1
investigator 0
smoker 2
tobacco_user 0
∀x(user(x)→person(x))
∀x(worker(x)→person(x))
∀x(researcher(x)→person(x))


_top 0
abstraction 0
act 0
human_action 0
human_activity 0
entity 0
physical_thing 0
information 1
info 0
research 2
person 0
individual 0
someone 0
somebody 0
mortal 0
human 0
soul 0
hazard 0
jeopardy 0
peril 0
risk 1
cigarette 2
coffin_nail 0
butt 0
fag 0
merchandise 0
wares 0
product 1
user 1
worker 1
research_worker 0
researcher 1
investigator 0
smoker 2
tobacco_user 0
∀x(person(x)→¬risk(x))
∀x(person(x)→¬cigarette(x))
.....

Nutcracker

- …
  - Use ATPs to determine the likelyhood of entailment

    - Create Background Knowledge for T&H
    - Give this to the theorem prover:
      - (BK & T') → H'
    - If the theorem prover finds a proof, then we predict that T entails H

Nutcracker

- El problema básico de la aproximación es el uso de BK
  - Los resultados son excelentes en cuanto a precisión pero muy limitados en cuanto a cobertura.
  - WN es claramente insuficiente para representar el BK necesario.
  - Se necesitan otras fuentes para la obtención de BK.

Introducción

- Formas de adquisición
  - Manual
  - A partir de fuentes (semi-)estructuradas preexistentes
    - Wikipedia
    - Linked data
      - Dbpedia
      - Yago
      - Freebase
    - Ontologías de dominio
    - Terminologías
    - Glosarios
  - **A partir de textos (LbR)**

Introducción

- To build a formal representation of a specific, coherent topic through deep processing of concise texts focused on that topic
- Natural Language Understanding (**NLU**) + Knowledge Integration (**KI**)
- Formas varias de representación del conocimiento:
  - RDF triples
    - <CONCEPT, RELATION, CONCEPT>
  - Frames
- Tipo de conocimiento:
  - Episódico
    - Podemos ha obtenido 5 diputados en las elecciones al Parlamento europeo del 25 de mayo de 2014
  - Genérico
    - Los perros ladran

Introducción

- Ejemplo tomado de Barker et al, 2007
  - Dominio CORAZÓN
- Relaciones
  - EVENT-to-ENTITY: agent, donor, instrument, etc.
  - ENTITY-to-ENTITY: has-part, location, material, etc.
  - EVENT-to-EVENT: causes, defeats, enables, etc.
  - EVENT-to-VALUE: rate, duration, manner, etc.
  - ENTITY-to-VALUE: size, color, age, etc.
- Component Library
  - 700 conceptos generales
  - events
    - TRANSFER, COMMUNICATE, ENTER
  - entities
    - PLACE, ORGANISM, CONTAINER.
- Seed concepts
  - Possitive (10)
    - PUMP, MUSCLE
  - Confusers (20)
    - MUSICAL-INSTRUMENT, SHOE

Introducción

- *1. The heart is a pump that works together with the lungs.*
- *2. The heart consists of 4 chambers.*
- *3. The upper chambers are called atria, and the lower chambers are called ventricles.*
- *4. The right atrium and ventricle receive blood from the body*
- *through the veins and then pump the blood to the lungs.*
- *5. It pumps blood in 2 ways.*
- *6. It pumps blood from the heart to the lungs to pick up oxygen.*
- *7. The oxygenated blood returns to the heart.*
- *8. It then pumps blood out into the circulatory system of blood vessels that carry blood through the body.*

Introducción

- *NLU*
  - *Parsing: CONTEX*
  - *Logical Form Generation: LF toolkit*
  - *Abductive Expansion and Reformulation: TACITUS*
- *KI*
  - *Word to Concept Mapping*
  - *Concept Creation*
  - *Instance Unification*
  - *SRL*
  - *Constraint Assertion*
  - *Adjective Elaboration*
  - *KB matching*

Introducción

- *Output*
  - *Concepts*
    - *(a subclass of CHAMBER), BLOOD (LIQUID-SUBSTANCE), HEART (PUMPING-DEVICE), LUNG (INTERNAL-ORGAN), OXYGEN (GAS-SUBSTANCE), VEIN (BODY-PART), VENTRICLE (CHAMBER) and VESSEL (BODY-PART).*
  - *Axioms*
    - *<Pumping*
      - *object Blood*
      - *destination Lung>*
    - *<Heart*
      - *has-part (exactly 4 Chamber)>*
    - *<Receive*
      - *origin Body*
      - *path Vein*
      - *recipient Ventricle*
      - *object Blood>*

Qué leer

- LbR
  - Mobius
    - Barker et al, 2007
  - Learning Reader
    - Forbus, et al, 2007
  - Explanation Agent
    - Forbus, et al, 2009
    - McFate, Forbus, Hinrichs, 2013
- Commonsense
  - Eslick, 2006
  - Clark, Harrison, 2009 (**DART**)
  - Havasi et al, 2009 (**ConceptNet**)
  - Singh, 2005 (**EM-ONE**, PHD , **Open Mind Common Sense**, **LifeNet**, **StoryNet**, **ShapeNet**)
- EL, SF
  - TAC KBP Proceedings 2010, 2011, 2012, 2013
- Event
  - Chambers, Jurafsky, 2007, 2008, …
  - Riloff et al, 2007, …
  - Filatova, Prager, 2005, 2012

Tareas implicadas

- Entity Linking (EL)
- Slot Filling (SF)
- Event Detection
  - Event enrichment
  - Event Generalization
  - Scenario induction
- Relation Extraction
  - Domain restricted
  - Unrestricted

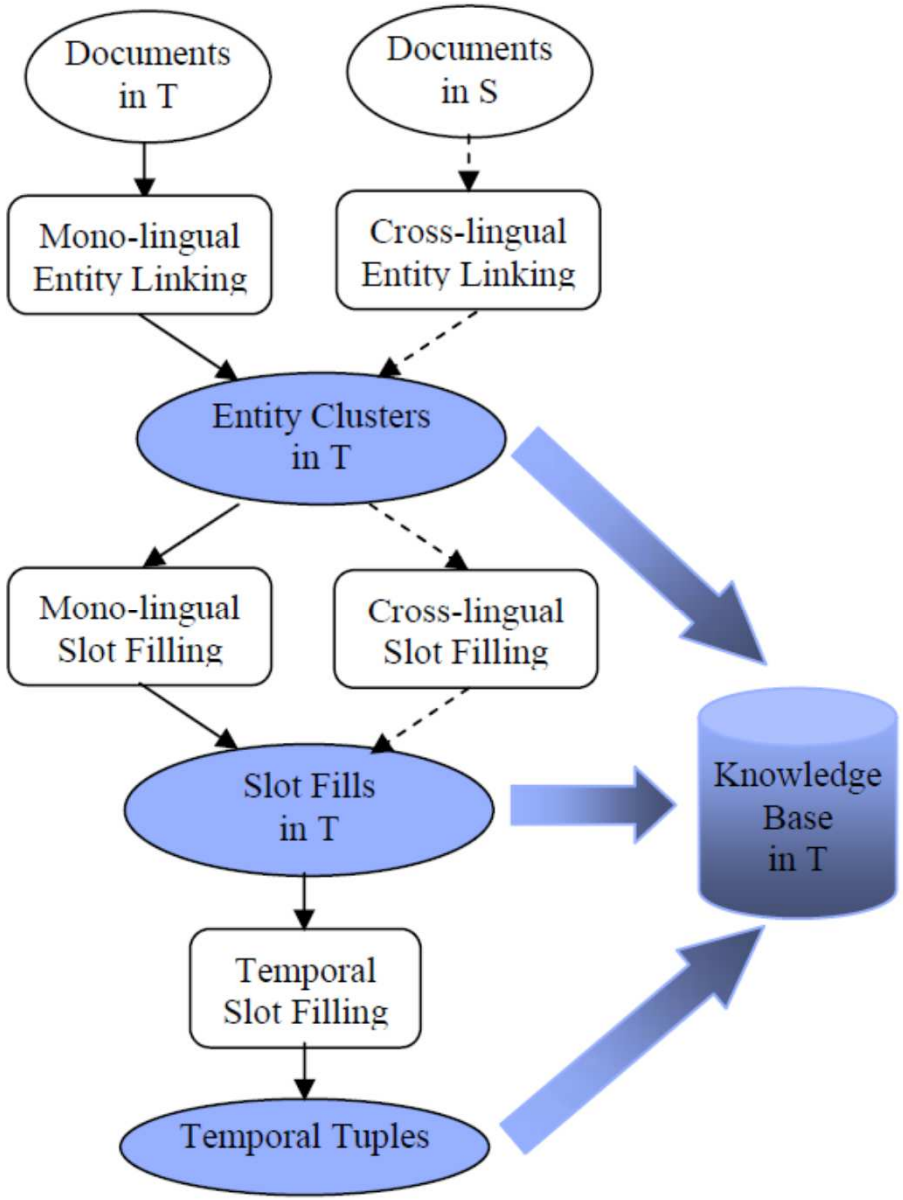
TAC KBP

- KB
  - 818,741 nodes
  - From English WP 2008
  - For each node:
    - Facts + Reference document
- Traks
  - Cold Start
  - SF
  - EL
  - Event
  - Sentiment

TAC KBP SF

- Tasks
  - English SF
  - Chinese SF
  - Temporal SF
  - Sentiment SF
  - SF Validation
  - Cold start

SF


Documents in T
Documents in S
Mono-lingual Entity Linking
Cross-lingual Entity Linking
Entity Clusters in T
Mono-lingual Slot Filling
Cross-lingual Slot Filling
Slot Fills in T
Knowledge Base in T
Temporal Slot Filling
Temporal Tuples

SF

Person Slots

Name	Type	List?
per:alternate_names	Name	Yes
per:date_of_birth	Value	
per:age	Value	
per:country_of_birth	Name	
per:stateorprovince_of_birth	Name	
per:city_of_birth	Name	
per:origin	Name	Yes
per:date_of_death	Value	
per:country_of_death	Name	
per:stateorprovince_of_death	Name	
per:city_of_death	Name	
per:cause_of_death	String	
per:countries_of_residence	Name	Yes
per:statesorprovinces_of_residence	Name	Yes
per:cities_of_residence	Name	Yes
per:schools_attended	Name	Yes
per:title	String	Yes
per:employee_or_member_of	Name	Yes
per:religion	String	Yes
per:spouse	Name	Yes
per:children	Name	Yes
per:parents	Name	Yes
per:siblings	Name	Yes
per:other_family	Name	Yes
per:charges	String	Yes

SF

Organization Slots

Name	Type	List?
org:alternate_names	Name	Yes
org:political_religious_affiliation	Name	Yes
org:top_members_employees	Name	Yes
org:number_of_employees_members	Value	
org:members	Name	Yes
org:member_of	Name	Yes
org:subsidiaries	Name	Yes
org:parents	Name	Yes
org:founded_by	Name	Yes
org:date_founded	Value	
org:date_dissolved	Value	
org:country_of_headquarters	Name	
org:stateorprovince_of_headquarters	Name	
org:city_of_headquarters	Name	
org:shareholders	Name	Yes
org:website	String	

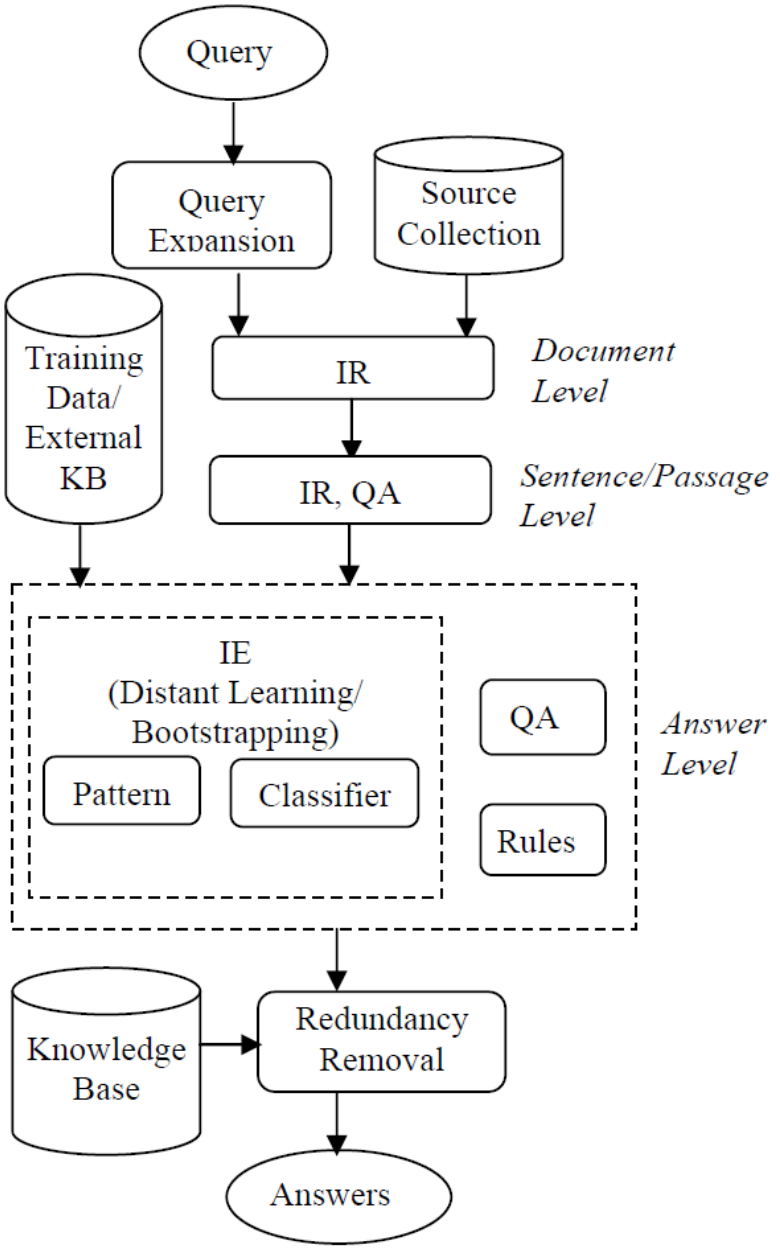
SF

	Diagnostic Scores			Official Scores		
	Recall	**Precision**	**F1**	**Recall**	**Precision**	**F1**
lsv	**32.93**	38.50	35.50	**33.17**	42.53	**37.28**
ARPANI*	29.10	47.83	**36.18**	27.45	50.38	35.54
RPI-BLENDER	30.62	38.19	33.98	29.02	40.73	33.89
PRIS2013	27.82	35.33	31.13	27.59	38.87	32.27
BIT	22.06	57.86	31.94	21.73	61.35	32.09
Stanford	28.46	32.30	30.26	28.41	35.86	31.70
NYU	17.35	50.70	25.85	16.76	53.83	25.56
UWashington	10.31	**59.72**	17.59	10.29	**63.45**	17.70
CMUML	10.63	28.79	15.53	10.69	32.30	16.07
SAFT_KRes	13.43	12.43	12.91	14.99	15.67	15.32
UMass_IESL	18.47	9.43	12.48	18.46	10.88	13.69
utaustin	7.91	21.85	11.62	8.11	25.16	12.26
UNED	9.11	15.08	11.36	9.33	17.59	12.19
Compreno	13.19	8.69	10.48	12.74	9.74	11.04
TALP_UPC	9.67	6.54	7.81	9.81	7.69	8.62
IIRG	3.20	7.38	4.46	2.86	7.72	4.17
SINDI	2.80	7.26	4.04	2.59	7.84	3.89
CohenCMU	3.68	1.69	2.32	3.68	1.98	2.57
LDC	58.35	83.81	68.80	57.08	85.60	68.49

SF

	Entity Count	Value Count (Pct)
per:title	33	142 (10.8%)
org:top_members_employees	41	116 (8.8%)
org:alternate_names	45	82 (6.2%)
per:employee_or_member_of	28	72 (5.5%)
per:children	23	52 (3.9%)
per:cities_of_residence	30	51 (3.9%)
per:age	31	51 (3.9%)
per:date_of_death	36	48 (3.6%)
per:cause_of_death	33	47 (3.5%)
per:charges	13	45 (3.4%)
per:alternate_names	24	45 (3.4%)
per:countries_of_residence	25	36 (2.7%)
per:city_of_death	32	35 (2.6%)
org:country_of_headquarters	34	34 (2.6%)
org:website	32	32 (2.4%)
per:origin	28	32 (2.4%)
per:spouse	23	28 (2.1%)
per:statesorprovinces_of_residence	23	28 (2.1%)
per:schools_attended	16	27 (2.0%)
org:subsidiaries	13	25 (1.9%)
per:parents	18	25 (1.9%)
org:city_of_headquarters	23	24 (1.8%)
org:members	4	22 (1.6%)
org:founded_by	11	21 (1.6%)
org:stateorprovince_of_headquarters	20	20 (1.5%)

SF


Query
Query Expansion
Source Collection
Training Data/ External KB
IR
Document Level
IR, QA
Sentence/Passage Level
IE (Distant Learning/ Bootstrapping)
Pattern
Classifier
QA
Rules
Answer Level
Knowledge Base
Redundancy Removal
Answers

SF

Alternate Names (query expansion)

Query	Name	Alternate names	#
SF558	Barbara Boxer	1.0 Barbara Levy Boxer 0.8 Barbara L. Boxer 0.64 B. L. Boxer 0.56 B. Boxer 0.49 Boxer ...	37
SF520	Hong Kong Disneyland	1.0 Hong Kong Disneyland 1.0 HKDL 0.8 H. Kong Disneyland 0.8 Hong K. Disneyland 0.7 Hong Disneyland 0.7 Kong Disneyland 0.64 Disneyland ...	30

TAC KBP SF

- distant supervision
  - A partir de Freebase y Wikipedia
  - Filtrado para la reducción del ruido
  - Wikificación, normalización de expresiones temporales,
  - Parsing de dependencias
- QA
- hand-coded rules
- IE
- Bootstrapping
- Aprendizaje no supervisado

TAC KBP EL

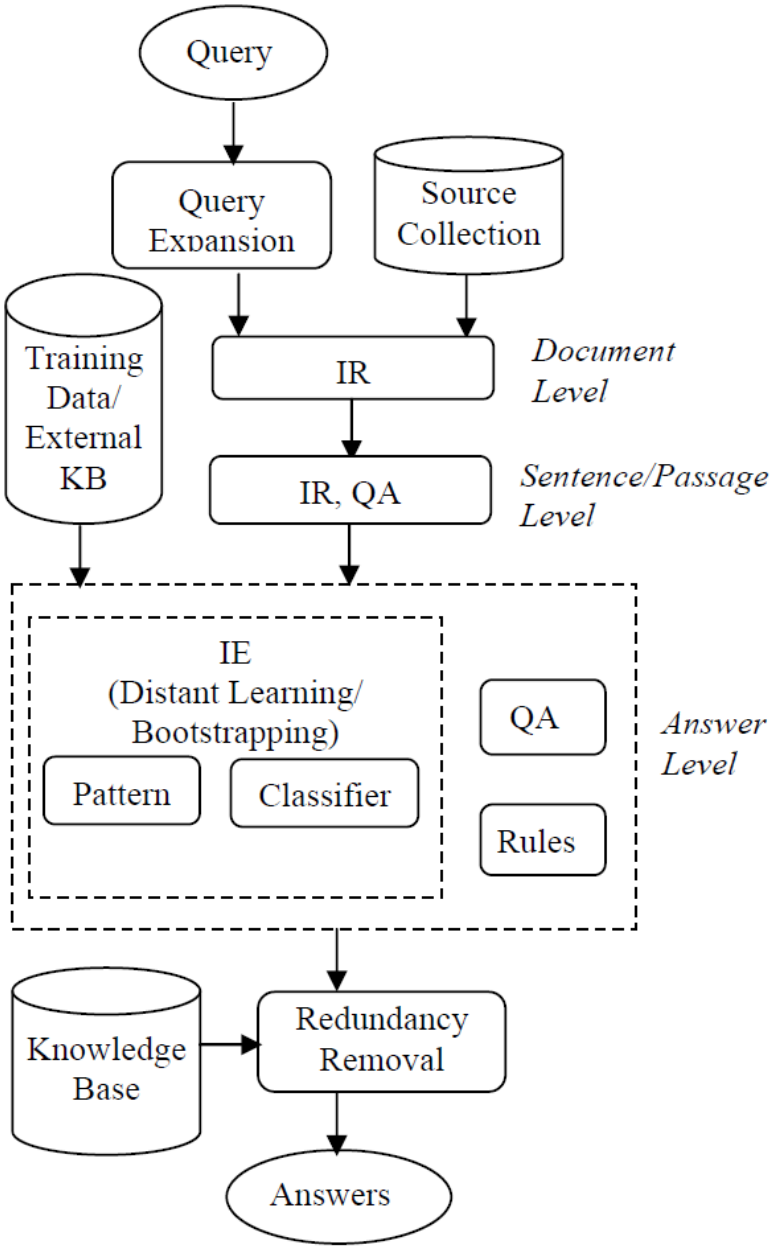
- Query
  - Name + reference document
- Desambiguación
  - KB
  - NIL
- Preproceso
  - Query classification (PER, ORG, GPE)
  - Query expansion
  - Proceso lingüístico del documento de Referencia
    - Segmentación en oraciones, tokenización, POS tagging, NERC, Correferencia, WSD, SRL, parsing de dependencias,

EL

<DOCID> eng-NG-31-108519-8977045 </DOCID>
<DOCTYPE SOURCE="usenet"> USENET TEXT </
DOCTYPE>
<DATETIME> 2007-10-10T22:25:00 </DATETIME>
<BODY>
<HEADLINE>
Dollars for Death
</HEADLINE>
<TEXT>
<POST>
<POSTER> Anybody <anybod...@canada.com> </
POSTER>
<POSTDATE> 2007-10-10T22:25:00 </POSTDATE>
Dieticians

The **American Dietetic Association (ADA)** has 67,000 members. Their motto is "Everything in moderation." That includes McDonald's, other fast food restaurants, dairy products, NutraPoison, and sugar-rich soda. Of course, the one concept that they do not limit is donations by various industry groups who delight in seeing the ADA's continuing ……. i4crob(at)earthlink.net
</POST></TEXT></BODY></DOC>

EL


Query
Query Expansion
Source Collection
Training Data/ External KB
IR
Document Level
IR, QA
Sentence/Passage Level
IE (Distant Learning/ Bootstrapping)
Pattern
Classifier
QA
Rules
Answer Level
Knowledge Base
Redundancy Removal
Answers

TAC KBP EL

- Query expansion
  - AN generation
    - PER
      - Person name grammars
    - ORG
      - Acronyms, suffixes
    - GPE
      - Gazetteers
        - » Geonames, GNIS, DBPEDIA, YAGO, YAGO 2
  - WP hyperlinks
  - Coreference
  - Statistical models

TAC KBP EL

- Candidate generation
  - IR
  - Document semantic analysis and context modeling
  - Collaborative clustering
    - Go beyond single query and single KB entry
    - All entities in the reference document
    - Graph-based clustering

TAC KBP EL

- Candidate ranking
  - VSM (unsupervised similarity between the vector representing the query and the vectors representing the candidates)
  - Supervised classification and ranking
  - Global graph-based ranking
  - Rule based
  - Ranking algorithms
    - SVMRank, MaxEnt, Random Forests, ListNet

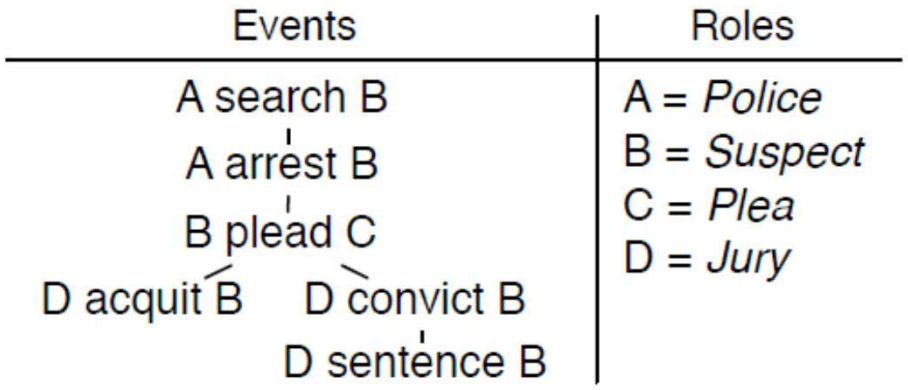
TAC KBP EL

- NIL clustering
  - Name string matching
  - Longest mention with within-document coreference
  - HA clustering
  - Global graph-based clustering
  - Linking to larger KB and mapping down
  - Topic modeling, LDA, LSA, WP categories, SUMO nodes

Event

Chambers & Jurafsky

- Chambers, Chambers, Jurafsky, 2007, ..., 2013, PHD thesis, 2011
- Narrative Event Chains
  - Partially ordered sets of events centered around a common **Protagonist**.
- Typed Narrative Events

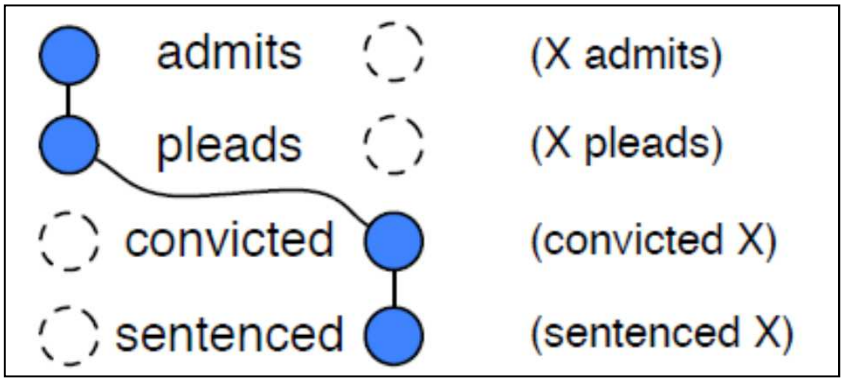


Event

Chambers & Jurafsky

L = (X pleads), (X admits), (convicted X), (sentenced X)

O = {(pleads, convicted), (convicted, sentenced), ...}



A *detain* B
A *confiscate* B
A *seize* B
A *raid* B
A *search* B
A *arrest* B

A ∈ {police, agent, officer, authorities, troops, official, investigator, ... }

B ∈ {suspect, government, journalist, *monday*, member, citizen, client, ... }

Event

Chambers & Jurafsky

- A database of Narrative Schemas (LREC 2010)
  - http://www.usna.edu/Users/cs/nchamber/data/schemas/acl09/
  - Narrative Schemas (unordered)
    - Various sizes of schemas (6, 8, 10, 12)
    - 1813 base verbs
    - 69 documents
    - 740 events
  - Temporal Orderings
    - Pairs of verbs
    - Counts of before and after relations

Authorship

Events: publish sell write translate distribute edit produce read

Role 1: { translate-s produce-s sell-s write-s distribute-s publish-s read-s edit-s
company author group year microsoft magazine my time firm writer government }

Role 2: { produce-o edit-o sell-o translate-o publish-o read-o write-o distribute-o
book report novel article story letter magazine film letters movie show }

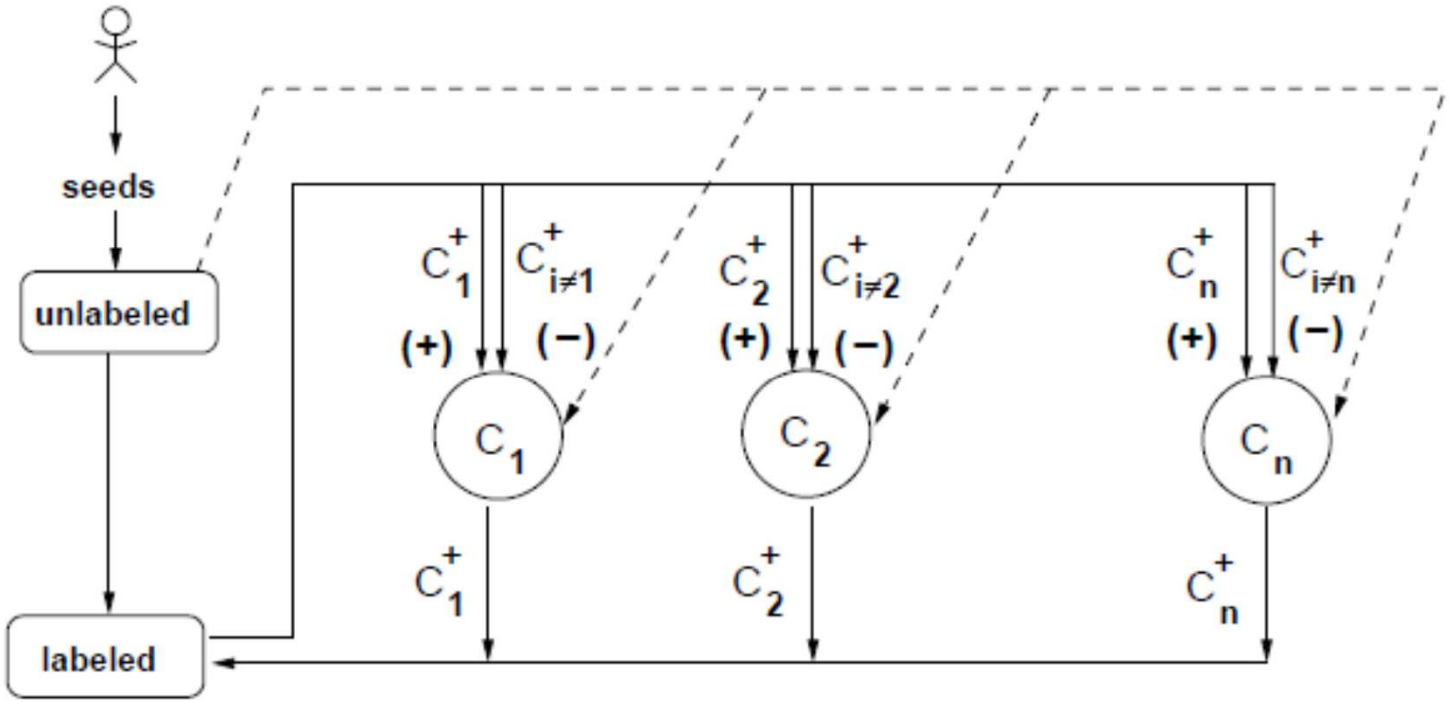
Event

Riloff et al

- Huang & Riloff, 2010
- Semantic tagging
- Bootstrapping approach
  - Seed words
  - Example in veterinary medicine domain
    - [A 14yo doxy]ANIMAL owned by [a reputable breeder]HUMAN is being treated for [IBD]DISEASE with [pred]DRUG.
  - Two steps:
    - **Inducing a Contextual Classifier**
    - **Cross-Category Bootstrapping**

Event

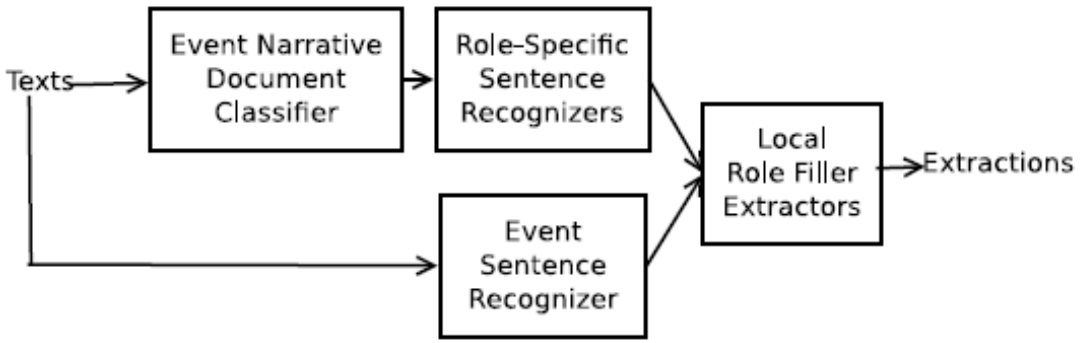
Riloff et al



Event

Riloff et al

- Huang & Riloff, 2012
- Event extraction classifiers
- **TIER** Event Extraction Model
- Bootstrapping approach
  - 4 classifiers
    - Document level
    - Sentence level
    - Noun phrase level



Event

Filatova et al

- Domain independent Detection, Extraction and Labeling of Atomic Events

THING: China Airlines Flight 676 from Bali to Taipei crashes
PLACE: Taipei, Taiwan
WHEN: February 16, 1998
TOPIC EXPLICATION: The flight was from Bali to Taipei. It crashed several yards short of the runway and all 196 on board were believed dead. China Airlines had an already sketchy safety record. This crash also killed many people who lived in the residential neighborhood where the plane hit the ground. Stories on topic include any investigation into the accident, stories about the victims/their families/the survivors. Also on topic are stories about the ramifications for the airline.

Event

Filatova et al

- Domain independent Detection, Extraction and Labeling of Atomic Events
  - NE pairs

Relation Frequency	First Element	Second Element
0.0212	China Airlines	Taiwan
0.0191	China Airlines	Taipei
0.0170	China Airlines	Monday
0.0170	Taiwan	Monday
0.0170	Bali	Taipei
0.0148	Taipei	Taiwan
0.0148	Bali	Taiwan
0.0148	Taipei	Monday
0.0127	Bali	Monday
0.0127	International Airport	Taiwan

Event

Filatova et al

- Domain independent Detection, Extraction and Labeling of Atomic Events
  - Top connectors

Relation	Connector	Connector Frequency
China Airlines – Taiwan	crashed/VBD	0.0312
	trying/VBG	0.0312
	burst/VBP	0.0267
	land/VB	0.0267
China Airlines – Taipei	burst/VBP	0.0331
	crashed/VBD	0.0331
	crashed/VBN	0.0198

Event

Filatova et al

- Domain independent Detection, Extraction and Labeling of Atomic Events
  - Output Event

First named entity	Second named entity	Connectors
China Airlines	Taiwan; Taipei	crashed/VBD trying/VBG burst/VBP land/VB killing/VBG

Event

Filatova et al

- Learning occupation related activities for Biographies

Artist	Dancer	Physicist	Singer
Born/VBN	Made/VBD	Born/VBN	Said/VBD
Painted/VBD	Died/VBD	Died/VBD	Born/VBN
Painted/VBN	Appeared/VBD	Announced/VBD	Died/VBD
Including/VBG	Been/VBN	Discovered/VBD	Join/VB
Be/VB	Founded/VBD	Be/VB	Singing/VBG
Became/VBD	Became/VBD	Including/VBG	Sang/VBD
Died/VBD	Born/VBN	Became/VBD	Has/VBZ
Been/VBN	Danced/VBD	Wrote/VBD	Conducting/VBG
Showed/VBD	Blessed/VBN	Helped/VBD	Made/VBD
Had/VBD	Perform/VB	Named/VBN	Became/VBD

Conclusiones

- **LbR** es una tarea necesaria y difícil
- Implica la realización de varias subtareas (**SF**, **EL**, **Event Detection**, **Event Enrichment**, **Scenario Induction**, **Relation Extraction**) que también son difíciles y despiertan el interés de los investigadores en dos áreas de interés: **NLU** y **KI**
- Las técnicas empleadas son muy variadas
- Se trata de un campo sumamente activo de investigación