UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Textual Entailment

Alicia Ageno
ageno@cs.upc.edu
Universitat Politècnica de Catalunya

Outline

- Introduction
- RTE in PASCAL and TAC
- Techniques and general structure
- Relevant Systems

Readings

- **Textual Entailment Community**:
  - The *RTE Resource Pool* can now be accessed from: http://aclweb.org/aclwiki/index.php?title=Textual_Entailment_Resource_Pool
  - The *Textual Entailment Subzone* can now be accessed from: http://aclweb.org/aclwiki/index.php?title=Textual_Entailment_Portal
- **PASCAL Challenges**
  - **RTE-1 2005**
  - **RTE-2 2006**
  - **RTE-3 2007**
- **TAC** has been proposed as a generic task that captures major semantic inference needs across many natural language processing applications.
- **TAC challenges**
  - **RTE-4 TAC 2008**
  - **RTE-5 TAC 2009**
  - **RTE-6 TAC 2010**
  - **RTE-7 TAC 2011**

Readings

- **Workshops**
  - ACL 2005 Workshop on Empirical Modeling of Semantic Equivalence and Entailment, 2005
  - First PASCAL Recognising Textual Entailment Challenge (RTE-1), 2005
  - Second PASCAL Recognising Textual Entailment Challenge (RTE-2), 2006
  - Third PASCAL Recognising Textual Entailment Challenge (RTE-3), 2007
  - Answer Validation Exercise at CLEF 2006 (AVE 2006)
  - Answer Validation Exercise at CLEF 2007 (AVE 2007)
  - TAC workshops since 2008
- **Tutorials**
  - Horacio Rodríguez, Tutorial on TE at University of Sevilla, 2008
    - http://www.lsi.upc.es/%7Ehoracio/varios/sevilla2008.zip
  - Horacio Rodríguez, talk on Paraphrasing at University of Basc Country, 2009
    - http://www.lsi.upc.es/%7Ehoracio/varios/paraphrasing.tar.gz
- **Surveys**
  - [Ghuge, Bhattacharya, 2013]

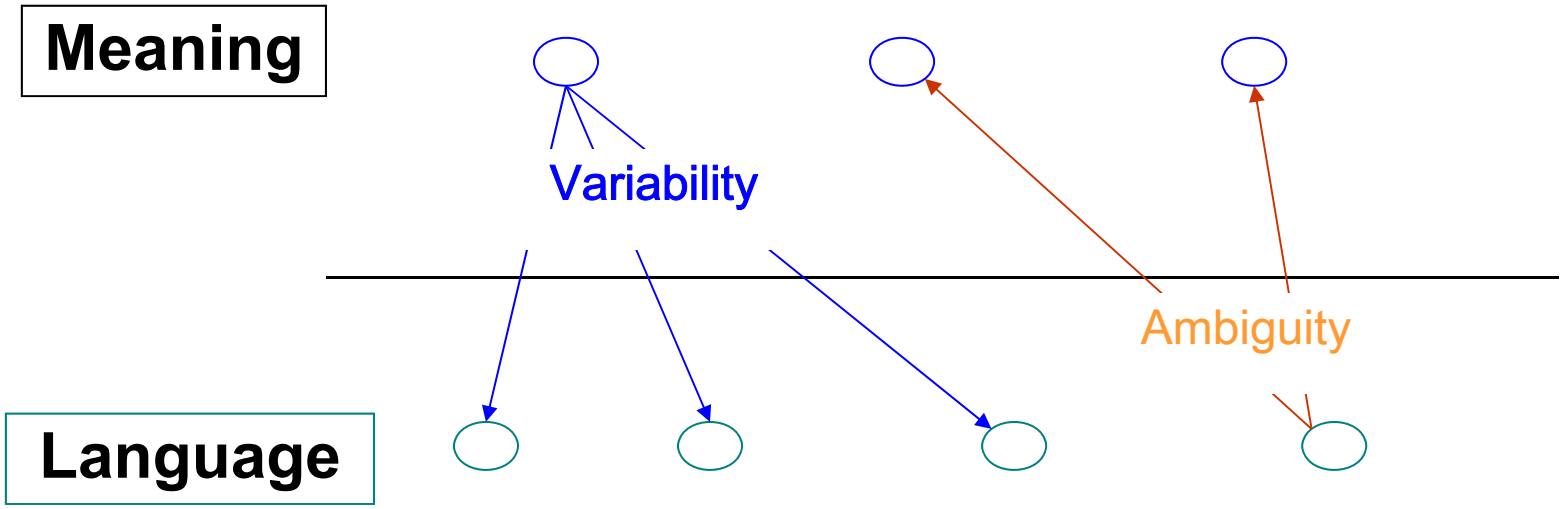
Readings

- **Thesis**
  - Oren Glickman (PHD, 2006)
  - Idan Szpecktor (MSC, 2005, PHD, 2009)
  - Milen Kouylekov (PHD, 2006)
  - Regina Barzilay (PHD, 2004)
  - Elena Cabrio (PHD, 2011)
  - Óscar Ferrández (PHD, 2009)
  - Prodromos Malakasiotis (PHD, 2011)
  - Annisa Ihsani (MSC, 2012)
  - Roy Bar Haim (PHD, 2010)
  - Shachar Mirkin (PHD, 2011)
- **Other material**
  - presentations of RTE 2 online
    - http://videolectures.net/pcw06_rus_trlsa/

Textual Entailment

Motivation:

- Text applications require *semantic* inference
- A common framework for applied semantics is needed, but still missing
- Textual entailment may provide such framework



Textual Entailment

The Dow Jones Industrial Average closed up 255

Dow ends up

Dow climbs 255

Dow gains 255 points

Stock market hits a record high

Model variability of semantic expression as relation between text expressions:

- Equivalence: *text1* ⇔ *text2* (paraphrasing)
- Entailment: *text1* ⇨ *text2* (the general case)

Textual Entailment

- Textual entailment recognition is the task of deciding, given two text fragments, whether the meaning of one text is entailed (can be inferred) from another text. This task captures generically a broad range of inferences that are relevant for multiple applications.

- For example, a QA system has to identify texts that entail the expected answer:

Question — *Expected answer form*

Who killed Kennedy? >> X killed Kennedy

The assassination of Kennedy by Oswald

text

⇒ *entails*

Oswald killed Kennedy

hypothesized answer

- Other QA ex: "Who owns Overture?"

"Yahoo acquired Overture" "Overture's acquisition by Yahoo", …

Textual Entailment

- In IE entailment holds between different text variants that express the same target relation: (X kill Y X acquire Y)
- In MDS a redundant sentence or expression, to be omitted from the summary, should be entailed from other expressions in the summary.
- In IR the concept denoted by a query expression should be entailed from relevant retrieved documents.
- In MT evaluation a correct translation should be semantically equivalent to the gold standard translation, and thus both translations have to entail each other.
- Thus, in a similar spirit to WSD and NER which are recognized as generic tasks, modeling textual entailment may consolidate and promote broad research on applied semantic inference.

Textual Entailment Applications

- Question-Answering
- Information Extraction
- Information Retrieval
- Multi-Document Summarization
- Named Entity Recognition
- Temporal and Spatial Normalization
- Semantic Parsing
- Natural Language Generation

Textual Entailment Applications

- Question-Answering
  - can often be reduced to a textual entailment problem by rephrasing the question as a declarative statement that is to be entailed by correct answers
- Information Retrieval
  - for certain IR queries the combination of semantic concepts and relations denoted by the query should be entailed from relevant retrieved documents
- Information Extraction
  - entailment holds between different text variants that express the same target relation
- Multi-Document Summarization
  - Identify a redundant sentence, to be omitted

Textual Entailment

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

- Classical entailment definition:
  A text t entails a hypothesis h if h is true in <u>every</u> circumstance (possible world) in which t is true

- Strict entailment - doesn't account for some uncertainty allowed in applications:

  t: The technological triumph known as GPS … was incubated in the mind of Ivan Getting.

  h: Ivan Getting invented the GPS.

Textual Entailment

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

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

- Operational (applied) definition:
  - Human gold standard – as in NLP applications
  - Assuming common background knowledge – which is indeed expected from applications

Textual Entailment

- Probabilistic interpretation:

> ***t* probabilistically entails *h* (*t* ⇒ *h*)** if
> P(*h is true* | *t*) > P(*h is true*)

  - *t* increases the likelihood of *h* being true
  - ≡ Positive PMI – *t* provides information on *h*'s truth

- P(*h is true* | *t*): *entailment confidence*
  - The relevant entailment score for applications
  - In practice: "most likely" entailment expected

Textual Entailment

- The role of knowledge:
  - For textual entailment to hold we require:

$$\textit{text} \text{ AND } \textit{knowledge} \Rightarrow h$$

  - But *knowledge* should not entail *h* alone
  - Systems are not supposed to validate *h*'s truth regardless of *t* (e.g. by searching *h* on the web)
- The knowledge sources available to the system are the most significant component of supporting TE

Textual Entailment ≈ Human Reading Comprehension

- From a children's English learning book (Sela and Greenberg):

Reference Text:

"…The Bermuda Triangle lies in the Atlantic Ocean, off the coast of Florida. …"

???

Hypothesis (True/False?):

The Bermuda Triangle is near the United States

- Common method for testing human reading comprehension: test the entailment capability
- Difficulties:
  - Variability between question and text
  - Knowledge needed - Florida in the US

Textual Entailment examples

TEXT	HYPOTHESIS	ENTAILMENT
Eyeing the huge market potential, currently led by Google, Yahoo took over search company Overture Services Inc last year.	*Yahoo bought Overture.*	**TRUE**
Microsoft's rival Sun Microsystems Inc. bought Star Office last month and plans to boost its development as a Web-based device running over the Net on personal computers and Internet appliances.	*Microsoft bought Star Office.*	**FALSE**
The National Institute for Psychobiology in Israel was established in May 1971 as the Israel Center for Psychobiology by Prof. Joel.	*Israel was established in May 1971.*	**FALSE**

Textual Entailment examples

TEXT	HYPOTHESIS	ENTAILMENT
Since its formation in 1948, Israel fought many wars with neighboring Arab countries.	*Israel was established in May 1948.*	**TRUE**
Putting hoods over prisoners' heads was also now banned, he said.	*Hoods will no longer be used to blindfold Iraqi prisoners.*	**FALSE**
The market value of u.s. overseas assets exceeds their book value.	*The market value of u.s. overseas assets equals their book value.*	**FALSE**

Textual Entailment

PASCAL Recognizing Textual Entailment (RTE) Challenges (*EU FP-6 Funded PASCAL Network of Excellence 2004-7)*

- 7 application settings in RTE-1, 4 in RTE-2/3
  - QA
  - IE
  - “Semantic” IR
  - Comparable documents / multi-doc summarization
  - MT evaluation
  - Reading comprehension
  - Paraphrase acquisition
- Most data created from actual applications output
- RTE-2/3: 800 examples in development and test sets
- 50-50% YES/NO split

Analysis

- Shallow techniques are promising when t-h pairs remain short, but their performance degrades as longer and more syntactically complex entailment pairs are considered

- For the first time: deep methods (semantic/ syntactic/ logical) clearly outperform shallow methods (lexical/n-gram)

- Still, most systems based on deep analysis did not score significantly better than the lexical baseline

Why?

- System reports point at two directions:
  - Lack of knowledge (syntactic transformation rules, paraphrases, lexical relations, etc.)
  - Lack of training data

- It seems that systems that coped better with these issues performed best:
  - LCC (Hickl et al.) - acquisition of large entailment corpora for training
  - COGEX (Tatu et al.) – large knowledge bases (linguistic and world knowledge)

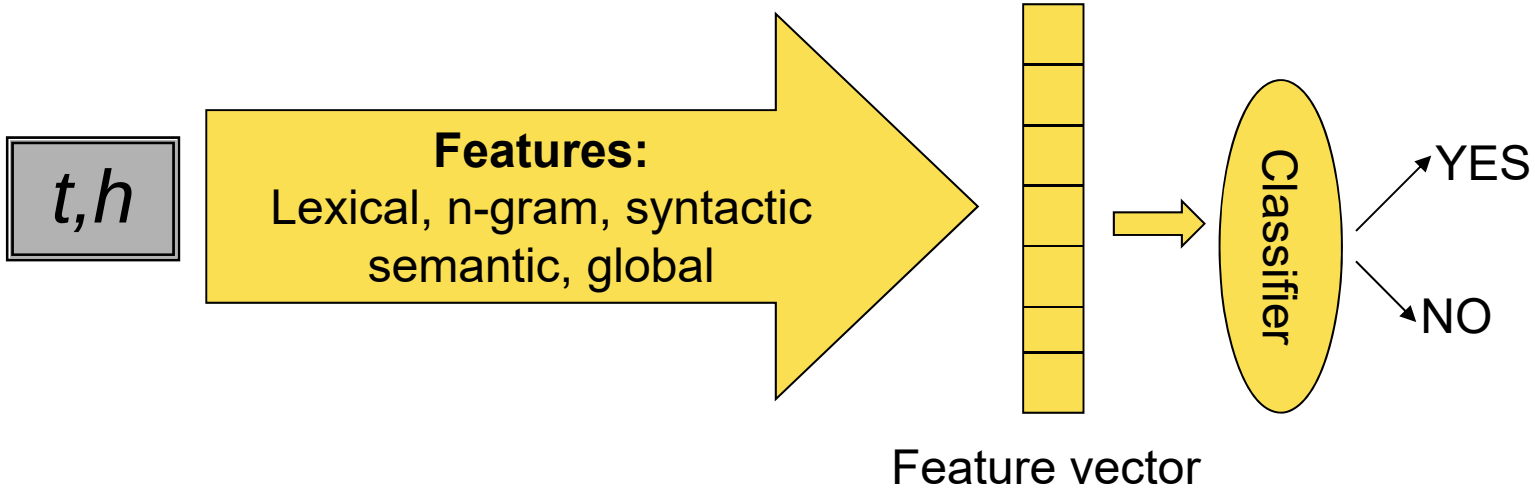
Methods and Approaches

- Word overlap
  - lexical, syntactic, and semantic
- Logical approaches
  - Raina et al, 2005
  - Bos et al, 2005, 2006
  - Moldovan et al, 2003
- Graph matching approaches
  - Haghighi et al, 2005
  - de Salvo et al, 2005
  - de Marneffe et al, 2005, 2006
- Paraphrases and Entailment Rules
  - Moldovan and Rus, 2001
  - Lin and Pantel, 2001 QA
  - Shinyama et al, 2002 IE

Methods and Approaches

- Measure similarity between *t* and *h* (*coverage* of *h* by *t*):
  - Lexical overlap (unigram, N-gram, subsequence)
  - Average Matched Word Displacement
  - Lexical substitution (WordNet, statistical)
  - Syntactic matching/transformations
  - Lexical-syntactic variations (“paraphrases”)
  - Semantic role labeling and matching
  - Global similarity parameters (e.g. negation, modality)
- Sentence Alignment
  - Exhaustive Sentence Alignment
  - Web-based Sentence Alignment
  - Bigrams
- Cross-pair similarity
  - Similarity measure between textual entailment examples
- Detect mismatch (for non-entailment)
- Interpretation to logic representation + logic inference

Dominant approach: Supervised Learning



- Features model both similarity and mismatch
- Classifier determines relative weights of information sources
- Train on development set and auxiliary *t-h* corpora

PASCAL RTE-3

- Analysis of the Knowledge Requirements
  - Clark et al, 2007
  - Lexico-Syntactic Matching
    - **Synonyms**
      - 648.T "...go through ... licencing procedures..."
      - 648.H "...go through the licencing processes."
    - **Generalizations (Hypernyms)**
      - 148.T "Beverly served...at WEDCOR"
      - 148.H "Beverly worked for WEDCOR."
    - **Noun Redundancy**
      - 607.T "single-run production process..."
      - 607.H "Single-run production..."
    - **Noun-Verb Relations**
      - 480 "Marquez is a winner..." → "Marquez won..."
    - **Compound Nouns**
      - 168 "Sirius CEO Karmazin" → "Karmazin is an executive of Sirius"
      - 583 "physicist Hawking" → "Hawking is a physicist"

PASCAL RTE-3

- Analysis of the Knowledge Requirements
  - Syntactic Matching
    - **Definitions**
      - 328 "sufferers of coeliac disease..." → "coeliacs..."
    - **World Knowledge: General**
      - 273 "bears kill people" → "bears attack people"
      - 499 "shot dead" → "murder"
      - 705 "under a contract with" → "cooperates with"
      - 721 "worked on the law" → "discussed the law"
      - 731 "cut tracts of forest" → "cut trees in the forest"
      - 732 "establishing tree farms" → "trees were planted"
      - 639 "*X*'s plans for reorganizing" → "*X* intends to reorganize"
      - 328 "the diets must be followed by <*person*>" → "the diets are for <*person*>"
      - 722 *X* and *Y* vote for *Z* → *X* and *Y* agree to *Z*.

PASCAL RTE-3

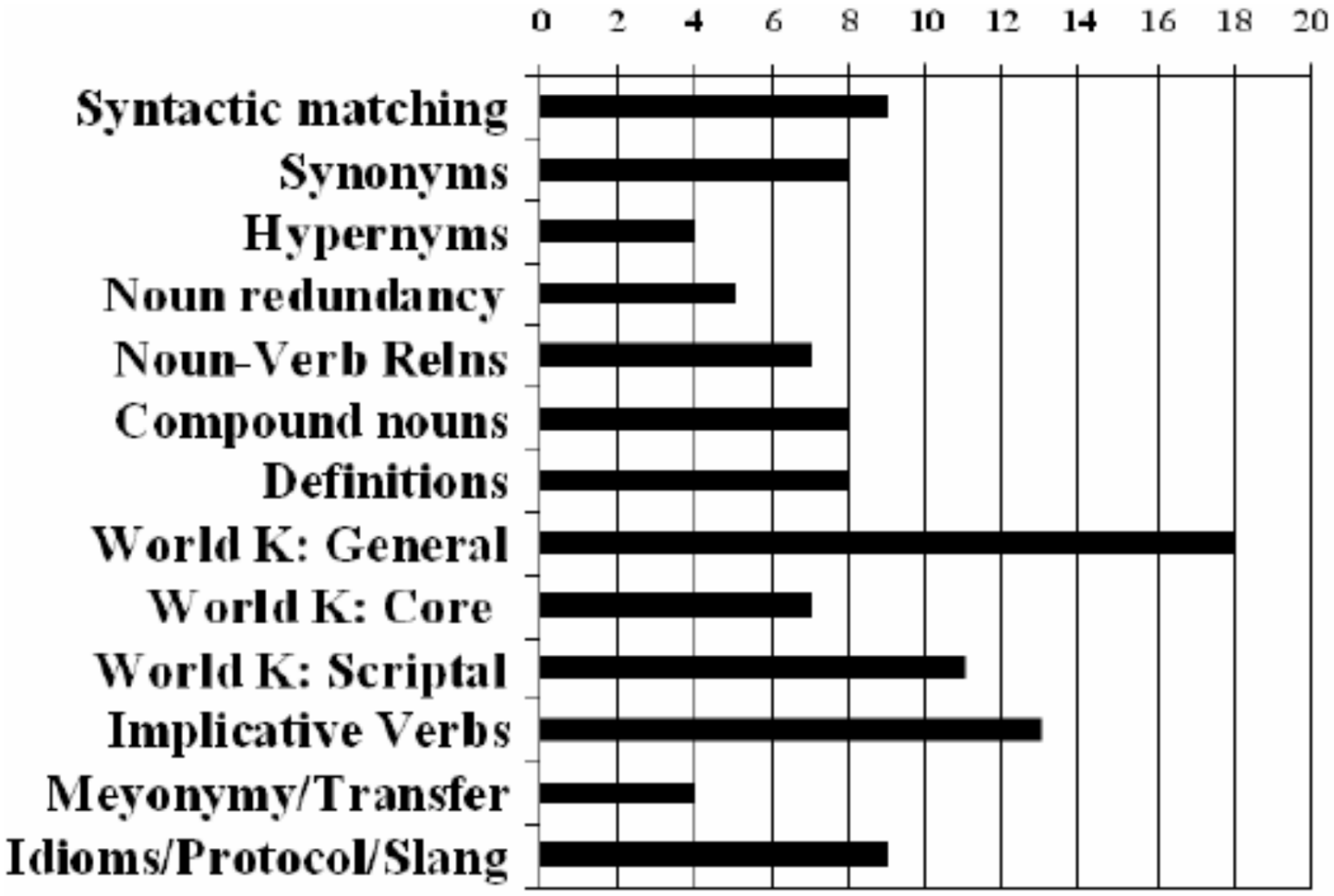
- Analysis of the Knowledge Requirements
  - World Knowledge
    - **Core Theories** (e.g., space, time, plans, goals)
      - 6.T "Yunupingu is one of the clan of..."
      - 6.H "Yunupingu is a member of..."
      - 491.T "...come from the high mountains of Nepal."
      - 491.H "...come from Nepal."
      - 178.T "...3 people in Saskatchewan succumbed to the storm."
      - 178.H "...a storm in Saskatchewan."

PASCAL RTE-3

- Analysis of the Knowledge Requirements
  - World Knowledge
    - **Frames and Scripts**
      - **Implicative Verbs**
        - » 668 "A survey shows that *X*..." → "*X*..."
        - » 657 "...*X* was seen..." → "...*X*..."
        - » 725 "...decided to *X*..." → "...*X*..."
        - » 716 "...have been unable to *X*..." → "...do not *X*"
      - **Metonymy/Transfer**
        - » 708.T "Revenue from stores funded..."
        - » 708.H "stores fund..."
      - **Idioms/Protocol/Slang**
        - » 12 "Drew served as Justice. Kennon returned to claim Drew's seat" → "Kennon served as Justice."
        - » 486 "*name*, 1890-1970" → "*name* died in 1970"
        - » 408 "takes the title of" → "is"
        - » 688 "art finds its way back" → "art gets returned"

PASCAL RTE-3

Frequency



Textual Entailment: General View

Entails

Subsumed by

Eyeing the huge market potential, currently led by Google, Yahoo took over search company Overture Services Inc. last year

Yahoo acquired Overture

Overture is a search company

Google is a search company

.....

Phrasal verb paraphrasing

Entity matching

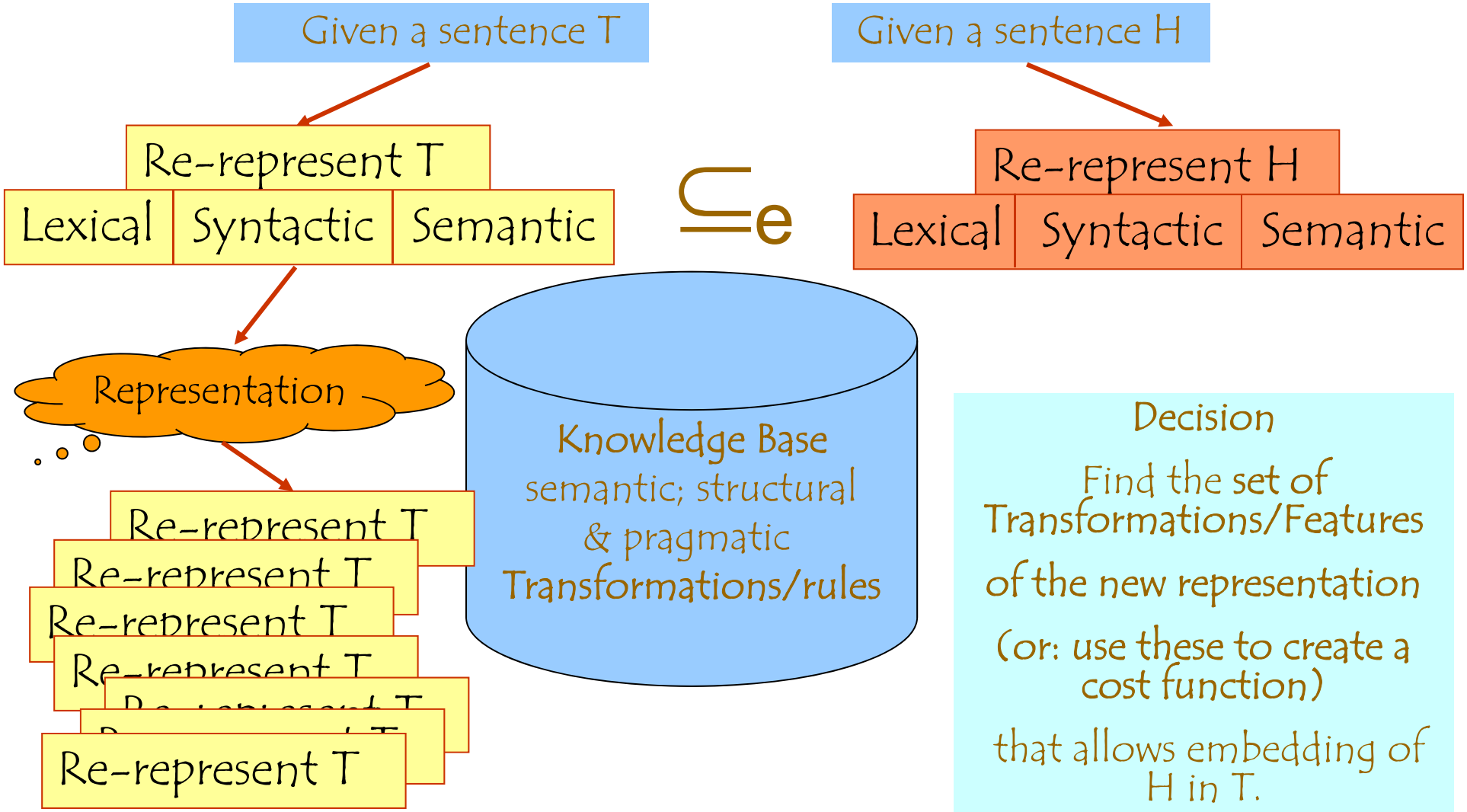
Alignment

How?

Semantic Role Labeling

Integration

A general Strategy for Textual Entailment



Details of the entailment strategy

1. Preprocessing
   - Multiple levels of lexical pre-processing
   - Syntactic Parsing
   - Shallow semantic parsing
   - Annotating semantic phenomena
2. Representation
   - Bag of words, n-grams through tree/graph based representation
   - Logical representations
3. Knowledge sources
   - Syntactic mapping rules
   - Lexical resources
   - RTE specific knowledge sources
   - Additional Corpora/Web resources
4. Control Strategy & Decision Making
   - Single pass/ iterative processing
   - Strict vs. Parameter based

Preprocessing

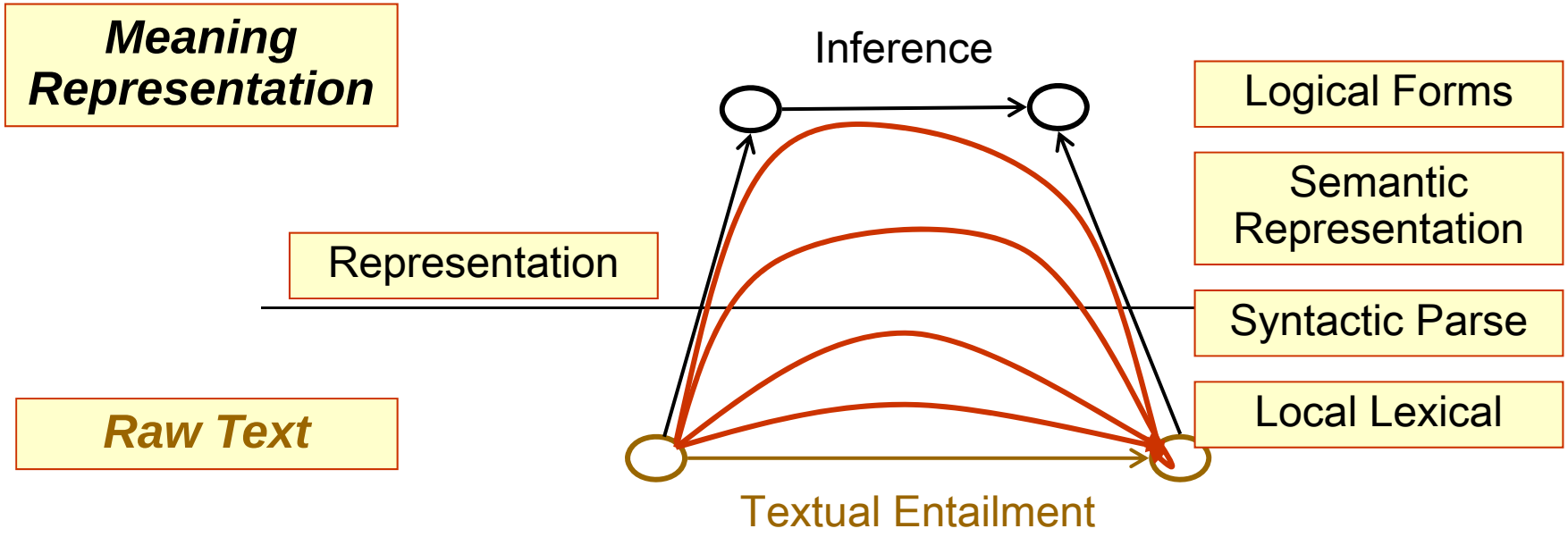
- Syntactic Processing:
  - Syntactic Parsing (Collins; Charniak; CCG)
  - Dependency Parsing (+types)
- Lexical Processing
  - Tokenization; lemmatization
  - For each word in Hypothesis, Text
  - Phrasal verbs
  - Idiom processing
  - Named Entities + Normalization
  - Date/Time arguments + Normalization
- Semantic Processing
  - Semantic Role Labeling
  - Nominalization
  - Modality/Polarity/Factive
  - Co-reference

Only a few systems

often used only during decision making

often used only during decision making

Representation



- Most approaches augment the basic structure defined by the processing level with additional annotation and make use of a tree/graph/frame-based system.

Knowledge Sources

- The knowledge sources available to the system are the most significant component of supporting TE.
- Different systems draw differently the line between preprocessing capabilities and knowledge resources.
- The way resources are handled is also different across different approaches.

Control Strategy and Decision Making

- Single Iteration
  - Strict Logical approaches are, in principle, a single stage computation.
  - The pair is processed and transformed into the logic form.
  - Existing Theorem Provers act on the pair along with the KB.
- Multiple iterations
  - Graph based algorithms are typically iterative.
  - Transformations are applied and entailment test is done after each transformation is applied.
  - Transformations can be chained, but sometimes the order makes a difference. The algorithm can be a greedy algorithm or can be more exhaustive, and search for the best path found [Braz et. al'05;Bar-Haim et.al 07]
  - Possible transformations: phrasal verb rewriter, nominalization promoter, apposition rewriter, use of Wordnet, ...

Notable TE Systems

- TEASE
- COGEX
- LCC Hickl

Notable Systems

TEASE and improvements

- Idan Szpektor (2005) Scaling Web-based Acquisition of Entailment Relations (Ms. thesis)
- Idan Szpektor et al (2004)
- Idan Szpektor and Ido Dagan (2007)
- Lorenza Romano et al (2007)
- Ido Dagan et al (2008)

TEASE

Input Verb	Learned Templates	
X approach *Y*	*X* go to *Y* *X* step to *Y* *X* walk to *Y*	*X* near *Y* *X* stride to *Y* *X* pass a note to *Y*
X defeat *Y*	*Y* lose to *X* *X* beat *Y* *X* victory over *Y*	*X* destroy *Y* *X* win *Y* *X* conquer *Y*
X host *Y*	bring *Y* to *X* *Y* is held in *X* *Y* come to *X*	*Y* is played in *X* *X* venue of *Y* *X* play host to *Y*
X preclude *Y*	*X* prevent *Y* *X* bar *Y* *X* prohibit *Y*	*X* exclude *Y* *X* deny *Y* *X* forbid *Y*
X plant *Y*	*X* grow *Y* *X* produce *Y* *X* cultivate *Y*	*X* raise *Y* *X* sow *Y* *X* farm *Y*

- First direct algorithm for extracting rules

TEASE

- This system proposes a principled semantic inference mechanism over parse-based representations. That is, a proof system that operates over syntactic trees.

- Control Strategy: it is an iterative system. The prover applies entailment rules that aim to transform *T* into *H* through a sequence of intermediate parse trees

- Uses MINIPAR
  - Broad-coverage dependency parser for English
  - It generates dependency trees, each link between 2 words in a dependency tree represents a direct semantic relationship

TEASE

- Scalable unsupervised web-based method for extracting candidate entailment relations
- Web as main textual resource
- Acquired template structures as general as possible
- Feature-Vector Similarity Approach
  - The algorithm constructs for each possible template two feature vectors, representing its co-occurrence statistics with the two anchors. Two templates with similar vectors are suggested as paraphrases (termed inference rule)
- Unsupervised manner
  - Input is a verb lexicon (verb-based expressions), each element describing a specific relation
  - For each verb, it searches the Web for related syntactic entailment templates

TEASE

- A *template*, T_i, is a connected parse graph fragment (or dependency parse-tree fragment) with optional variables at some nodes.
  - example

$$X \xleftarrow{subj} prevent \xrightarrow{obj} Y$$

- A *pair of templates* T_1 and T_2 is denoted as $<T_1, T_2>$.
- A pair of templates is called an *entailment relation* if T_1 and T_2 contain the same variables and the meaning of T_2 can be inferred from the meaning of T_1, or vice versa, in some contexts, under the same variable instantiation.

TEASE

- example $X \xleftarrow{subj} prevent \xrightarrow{obj} Y$ *entails* $X \xleftarrow{subj} reduce \xrightarrow{obj} Y\ risk$

  "*aspirin reduces heart attacks risk*" can be inferred from "*aspirin prevents heart attacks*"

- An entailment relation does not need to hold under all possible variable instantiations, i.e. the correctness/validity of an entailment relation depends on specific variable instantiations
- A *pivot P* is a lexical phrase, such as a verb, a phrasal verb or a noun phrase (typically a nominalization) that expresses a semantic relation.

  Ex: aquire, fall to, prevent, victory over, near, ...
- A *pivot template*, denoted $\{P, T_P\}$, is a pivot P with its syntactic template T_P in the form of a parse graph fragment including at least two variable slots (nodes).

TEASE

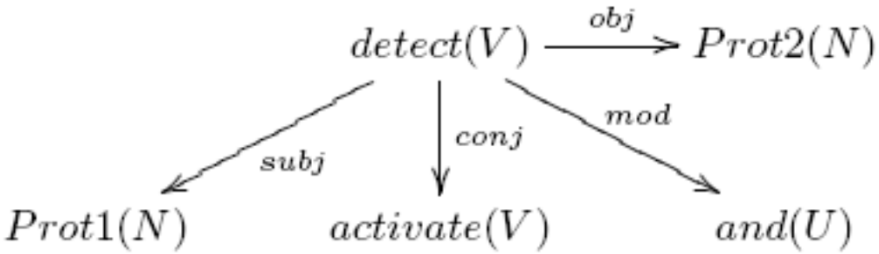


Figure 1: The dependency parse graph of the sentence "*Prot1 detected and activated Prot2*".

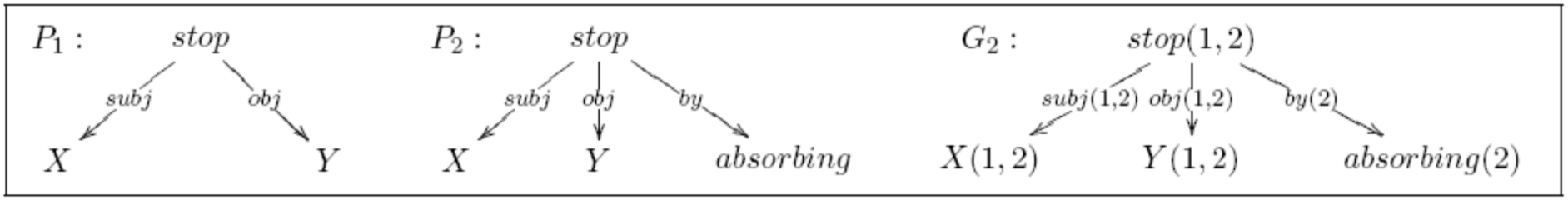


Figure 2: Two parse trees and their compact representation (sentence sets are shown in parentheses).

TEASE

First step: create a complete template T_P for the input pivot P

- Variable slots are added for the major types of syntactic relations that interact with P, based on its syntactic type
- input template for a transitive verb V

$$X \overset{subj}{\leftarrow} V \overset{obj}{\rightarrow} Y$$

- phrasal verb consisting of a verb V and a preposition

$$X \overset{subj}{\leftarrow} V \overset{prep}{\rightarrow} Preposition \overset{\text{mod}}{\rightarrow} Y$$

- The output of the TEASE algorithm is a ranked list of templates $\{T_i\}$ $<T_P, T_i>$ is an entailment relation candidate

TEASE

- An *anchor-set* is a set of words (or terms). Each member of an anchor-set is termed an *anchor*
- An anchor in an anchor-set that is designated to be an instantiation of a template variable in a sentence is termed a *slot anchor*
- An anchor in an anchor-set that is not designated to be an instantiation of a template variable in a sentence is termed a *context anchor*.
- A *matching* of a template *T* in a sentence *s* is the embedding of the parse-graph of *T* as a sub-graph in the parse-graph of *s*
- An *instantiation* of a template *T* by an anchor-set *AS* in a sentence *s* is a matching of *T* in *s* where each variable of *T* is instantiated with the corresponding value of a slot anchor in *AS*, and all the values of the context anchors in *AS* appear elsewhere in *s*

"Antibiotics in pregnancy prevent miscarriage"

Slot anchors: $\{antibiotics \overset{subj}{\leftarrow}, miscarriage \overset{obj}{\leftarrow}\}$

Context anchor: *pregnancy*

$$X \overset{subj}{\leftarrow} prevent \overset{obj}{\rightarrow} Y$$

TEASE

A good *AS* should satisfy a proper balance between specificity and generality

- A *minimal anchor-set* is an anchor-set that contains only slot anchors
- *AS* is termed a *characteristic anchor-set* of a template *T* if for every template T_i that is instantiated by *AS* in some sentence, the entailment relation $<T, T_i>$ holds between *T* and T_i
- A *diverse anchor-set* is an anchor-set that instantiates more than one template in sentences
- An anchor-set that is both characteristic of a template *T* and diverse is termed a *productive anchor-set* for a template *T*
- Contex anchors provide for specificity

TEASE

- Algorithm
  - For each input pivot template T_P:
    - Extract productive anchor-sets for the pivot template (ASE phase)
      - Construct a sample corpus for the pivot template by retrieving sentences containing the pivot template from the Web.
      - Extract candidate anchor-sets from the sentences in the sample corpus.
      - Filter out candidate anchor-sets that fail certain criteria
    - Extract templates (TE phase)
      - Construct a sample corpus by retrieving sentences containing the anchor-sets extracted in phase 1 from the Web.
      - Extract repeated sub-structures in the sample corpus to be template candidates T_i.
    - Rank each extracted template T_i according to the confidence level in the correctness of the entailment relation $<T_P, T_i>$.

TEASE

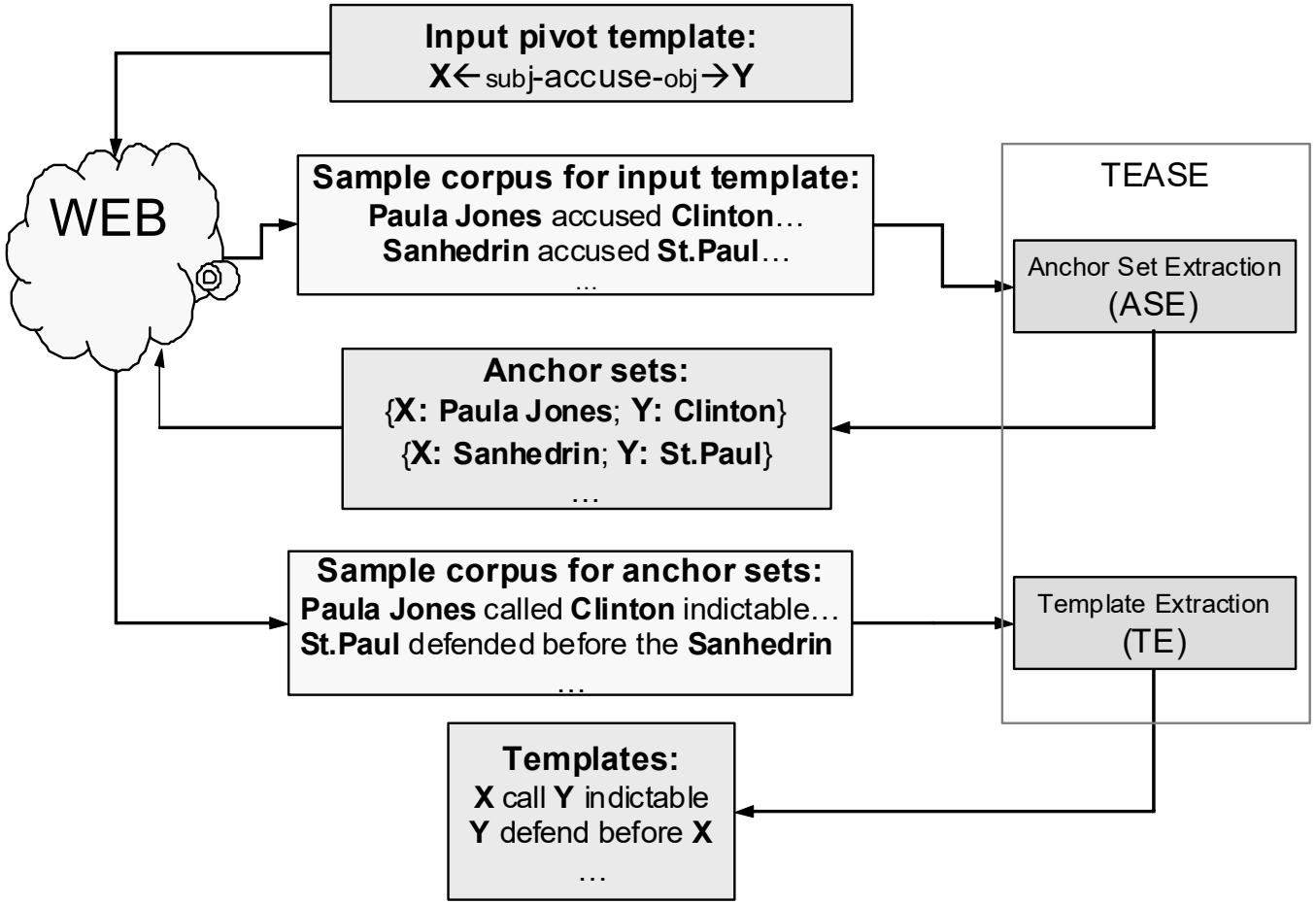


Figure 3: An example of the TEASE algorithm flow for the input pivot template '*X* accuse *Y*'.

TEASE

- ASE Algorithm
  - For each input pivot template T_P:
    - Construct a sample corpus that consists of sentences containing T_P.
      - Retrieve sentences the Web using a query containing the template's words
      - Retrieve more sentences from the Web using refined queries, based on the sentences retrieved at step
    - Extract productive anchor-set candidates from the constructed corpus.
      - Extract one minimal anchor-set, containing only the slot anchors, from each sentence in the sample corpus.
      - Extract one more anchor-set from each sentence, containing one context anchor in addition to the slot anchors, if possible.
    - Filter out candidates that fail certain filtering criteria:
      - Applying thresholds over individual anchor-set statistics.
      - Filtering anchor-sets that are redundant or inconsistent relative to other anchor-sets.

TEASE – ASE phase

Pivot Template	Learned Anchor Sets
X establish *Y*	*X* = epa, *Y* = national emission standard, C_1 = asbestos *X* = canada agricultural products act, *Y* = review tribunal *X* = school district, *Y* = breakfast program *X* = federal government, *Y* = conservation corps *X* = erisa, *Y* = minimum standards *X* = constantine, *Y* = new rome
X write *Y*	*X* = laurie, *Y* = numerous songs *X* = lewis carrol, *Y* = alice's adventures *X* = plato, *Y* = detailed account, C_1 = atlantis *X* = mendelssohn, *Y* = incidental music *X* = shakespeare, *Y* = great tragedies *X* = thomas malthus, *Y* = essay
X calculate *Y*	*X* = katz equation, *Y* = membrane potential *X* = eratosthenes, *Y* = circumference *X* = nernst equation, *Y* = equilibrium potential *X* = language model, *Y* = probabilities *X* = following table, *Y* = annual cost *X* = acos, *Y* = arc cosine

Table 4: Examples of anchor-sets learned for several pivot templates. *X* and *Y* are the slot anchors. C_1 is a context anchor.

TEASE – TE phase

Pivot Template	Learned Templates	
X establish *Y*	*X* set *Y* *X* develop *Y* *X* create *Y* *X* found *Y* *X* enforce *Y* *X* form *Y* *X* offer *Y* *X* release *Y*	*X* promulgate *Y* *X* issue *Y* *X* implement *Y* *X* provide *Y* *X* make *Y* *X* launch *Y* *X* institute *Y* *X* for the establishment of *Y*
X write *Y*	*X* who write *Y* *X* publish *Y* *X* compose *Y* read *Y* by *X* *Y* attributed to *X* perform *Y* by *X* *X* writer of *Y* selected *Y* of *X*	*X* produce *Y* *X* pen *Y* *X* create *Y* *X*'s *Y* *X* complete *Y* *X* book of *Y* *X* say in *Y* *X* work include *Y*
X calculate *Y*	*X* determine *Y* *X* compute *Y* *X* give estimate of *Y* *X* return *Y* *X* assess *Y* *X* generate *Y* *X* recalculate *Y* *X* work out *Y*	*X* measure *Y* *X* calculation of *Y* *X* yield *Y* *X* get *Y* *X* produce *Y* *Y* according to *X* *Y* obtained from *X* *X* evaluate *Y*

Table 5: Examples of templates learned for several pivot templates.

TEASE

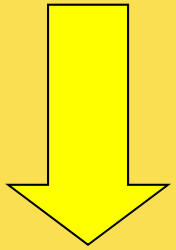
Pivot template	Learned template	Reason
$X \xleftarrow{subj} \text{write} \xrightarrow{obj} Y$	$X \xleftarrow{subj} \text{present} \xrightarrow{obj} Y$	Presenting a cheque is synonymous to writing a cheque.
$X \xleftarrow{subj} \text{write} \xrightarrow{obj} Y$	$X \xleftarrow{subj} \text{issue} \xrightarrow{obj} Y$	Issuing a report (e.g. by a policeman) is synonymous to writing a report.
$X \xleftarrow{subj} \text{produce} \xrightarrow{obj} Y$	$X \xleftarrow{subj} \text{obtain} \xrightarrow{obj} Y$	A telescope producing images is synonymous to a telescope obtaining images.

Table 6: Examples of learned templates that participate in non-trivial entailment relations with the related pivot template.

TEASE

TEASE and improvements

- Learning Canonical Forms of Entailment Rules
  - Idan Szpektor and Ido Dagan (2007)
  - Morpho-Syntactic Template Variations
    - there is still no sound solution for addressing it at learning time.



  - **A Modular Approach for Entailment Rule Learning**
    - generic morphological and syntactic regularities.
    - transform lexical-syntactic template variations that occur in a text into their *canonical form*

TEASE's improvements

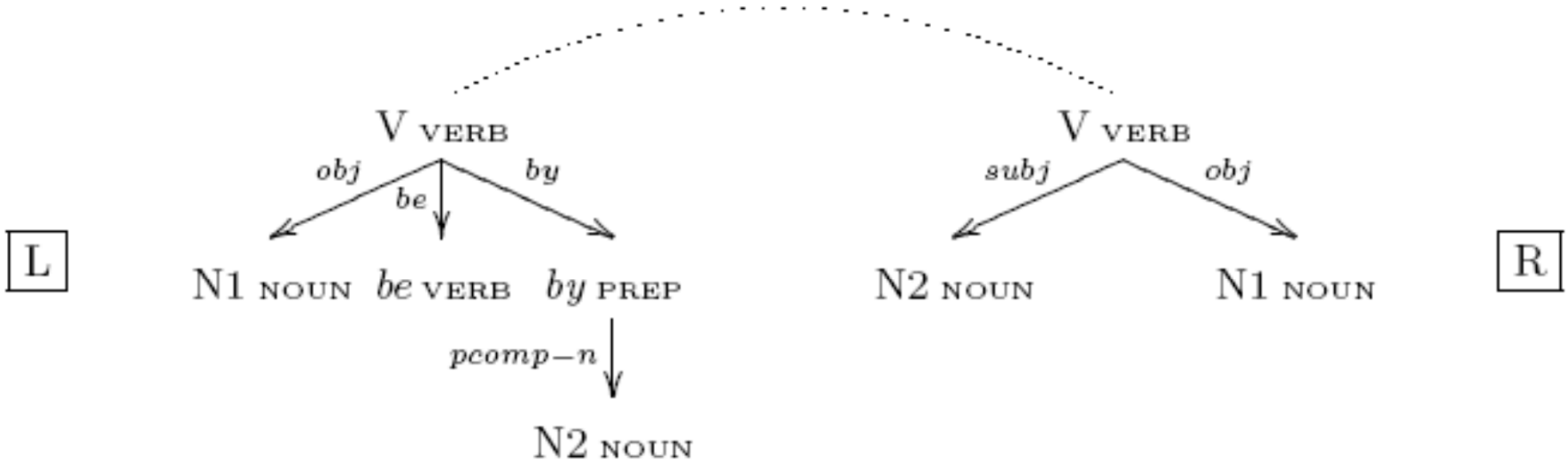
- Morpho-Syntactic Canonization Module
  - Syntactic-based Rules

Rule	Description	Original Template	Simplified Template
passive to active	$\xleftarrow{pcomp-n}\ \xleftarrow{by-subj}$ $\Longrightarrow \xleftarrow{subj}$	$\xleftarrow{pcomp-n}\ \xleftarrow{by-subj}\ \xrightarrow{obj}$	$\xleftarrow{subj}\ \xrightarrow{obj}$
conjunction	$\xrightarrow{conj}\ \Longrightarrow$	$\xleftarrow{subj}\ \xrightarrow{obj}\ \xrightarrow{conj}$	$\xleftarrow{subj}\ \xrightarrow{obj}$
apposition	$\xrightarrow{appo}\ \Longrightarrow$	$\xleftarrow{subj}\ \xrightarrow{obj}\ \xrightarrow{appo}$	$\xleftarrow{subj}\ \xrightarrow{obj}$
abbreviation	$\xrightarrow{spellout}\ \Longrightarrow$	$\xleftarrow{subj}\ \xrightarrow{obj}\ \xrightarrow{spellout}$	$\xleftarrow{subj}\ \xrightarrow{obj}$

Table 3: *Some of the syntactic rules used in our implementation, together with usage examples (the application of the second rule and the third rule is demonstrated in Figure 1).*

TEASE's improvements

- Passive to active substitution rule



(b) Passive to active substitution rule. The dotted arc represents alignment.

TEASE's improvements

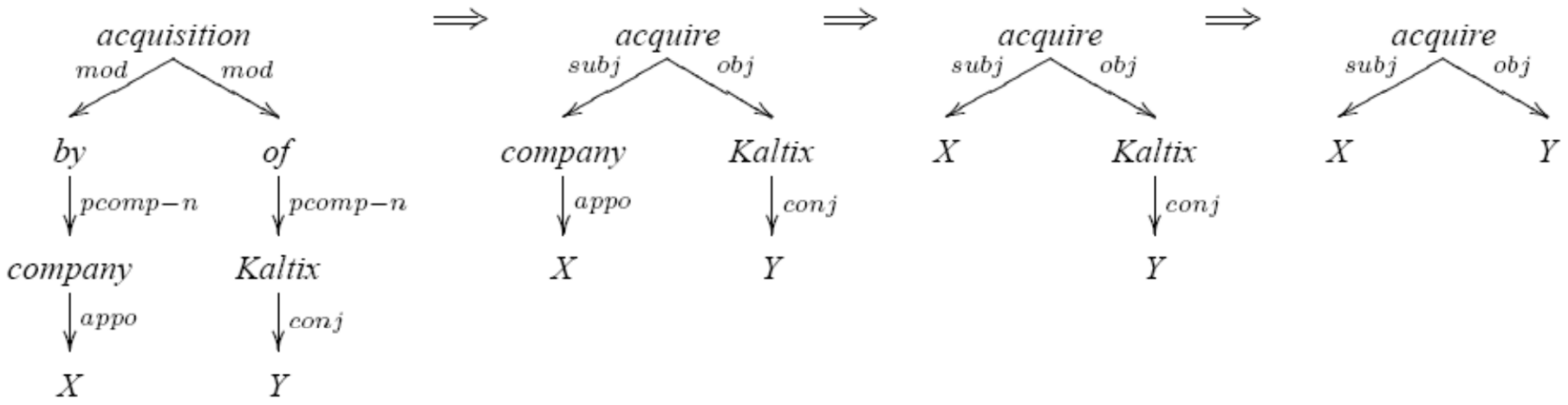


Fig. 1: *Chaining of canonization rules that transforms the path template between the arguments {X='Google';Y='Sprinks'}, which occurs in the sentence "*We witnessed the acquisition of Kaltix and Sprinks by another growing company, Google*", into a canonized template form. The first rule applied is a nominalization rule, followed by removal of apposition and removal of conjunction (as described in Table 3). As can be seen, applying the rules in any order will result in the same final canonized form.*

TEASE's improvements

examples of canonization rules

#	Category	Example: source	Example: derived
1	Conjunctions	Helena's very experienced and has played a long time on the tour.	⇒ Helena has played a long time on the tour.
2	Clausal modifiers	But celebrations were muted as many Iranians observed a Shi'ite mourning month.	⇒ Many Iranians observed a Shi'ite mourning month.
3	Relative clauses	The assailants fired six bullets at the car, which carried Vladimir Skobtsov.	⇒ The car carried Vladimir Skobtsov.
4	Appositives	Frank Robinson, a one-time manager of the Indians, has the distinction for the NL.	⇒ Frank Robinson is a one-time manager of the Indians.
5	Determiners	The plaintiffs filed **their** lawsuit last year in U.S. District Court in Miami.	⇒ The plaintiffs filed **a** lawsuit last year in U.S. District Court in Miami.
6	Passive	We have been approached by the investment banker.	⇒ The investment banker approached us.
7	Genitive modifier	Malaysia's crude palm oil output is estimated to have risen by up to six percent.	⇒ The crude palm oil output of Malasia is estimated to have risen by up to six percent.
8	Polarity	Yadav was forced to resign.	⇒ Yadav resigned.
9	Negation, modality	What we've **never** seen is actual costs come down.	What we've never $\overline{\text{seen}}$ is actual costs come down. (⇏ What we've seen is actual costs come down.)

COGEX

COGEX, Tatu 2006, 2007

- Combination of LEX, COGEXd, COGEXc
- Textual entailment as a logical implication between meanings:
  1. Transforms *T* and *H* into three-layered semantically-rich logic form representations
  2. Generates an abundant set of lexical, syntactic, semantic and world knowledge axioms
  3. Iteratively searches for a proof for the entailment between *T* and (a possibly relaxed version of) *H*
  4. A pair *T*-*H* is labeled as positive if the score of the found proof is above a threshold learned on the training data

COGEX

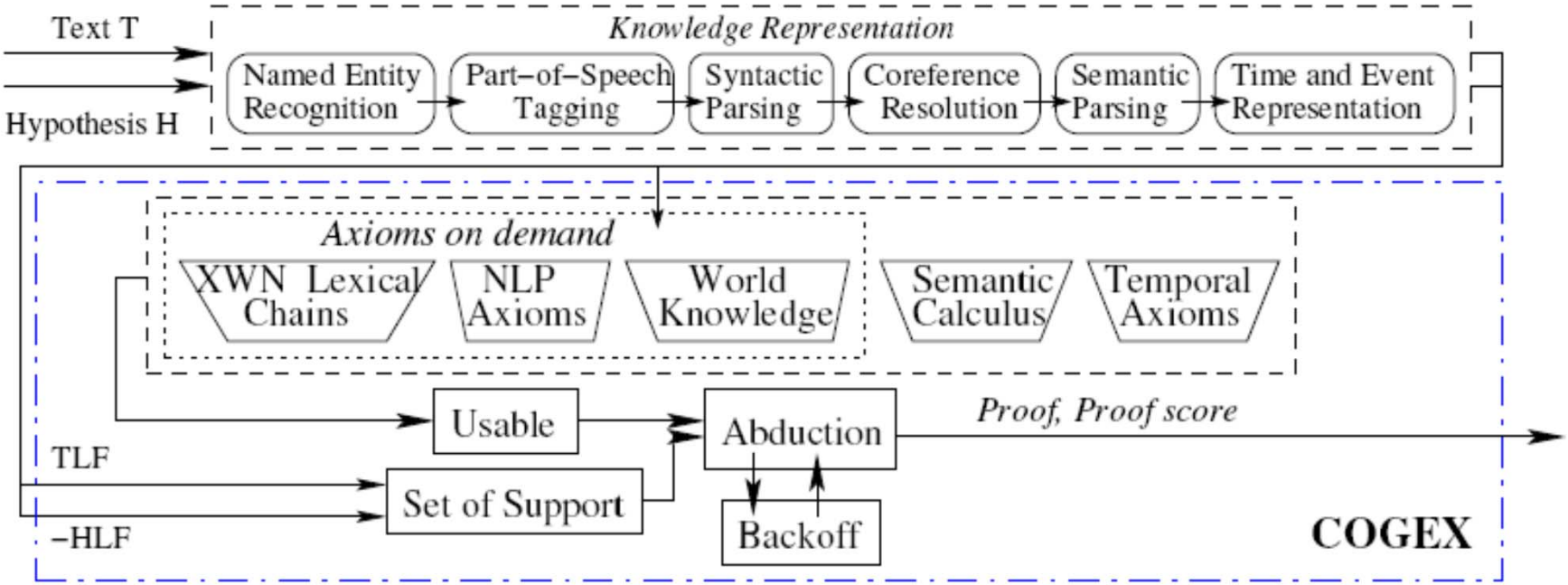


Figure 1: Cogex's Architecture

COGEX

The system is a natural language prover that performs proof by refutation: if a refutation is found then the proof is complete

- Usable list: axioms used by the system to generate inferences
- Set of support: list of clauses used to initiate the search for inferences. It is initially loaded with the negated form of *H* and the predicates that make up the logic form of *T*.
- Backoff: relaxation of predicate arguments that is carried out if a refutation cannot be found. If argument relaxation fails to produce a refutation, predicates are dropped from the negated *H* until a refutation is found.
- Scoring: once a proof found, a score for it is calculated by starting with an initial perfect score and deducting points for axioms that are used in the proof, arguments that are relaxed, and predicates that are dropped

COGEX

- EXtended WordNet Knowledge Base (XWN-KB)
  - XWN Lexical Chains: for each relation in the best lexical chain found between one of T's constituents and one of H's constituents, an axiom is created
  - Captures and stores the rich world knowledge encoded in WN's glosses into a KB: glosses are transformed into a set of semantic relations using a semantic parser

Nobelist, Nobel_laureate; synsetId: 06822770
gloss: *winner of a Nobel_prize*
ISA(Nobel_laureate,winner)
THEME(Nobel_prize,winner)

- NLP Axioms
  - Break down complex syntactic structures (complex nominals, coordinating conjunctions)
  - Links a NE to its set of aliases

T: A leading human rights group on Wednesday identified Poland and Romania as the likely locations in eastern Europe of secret prisons where al-Qaeda suspects are interrogated by the Central Intelligence Agency.
H: CIA secret prisons were located in Eastern Europe.

- World Knowledge
  - For ex., it allows to infer "nearly half a million dollars" from "$480.000"

COGEX

- Named Entity Check
  - In the proof scoring phase, deducts points for each pair whose *H* contains at least one *named entity* not-derivable from *T*

T: Thus, China's President repeatedly sent letters and envoys to the Dalai Lama and to the Tibetan Government asking that Tibet "join" the Republic of China.
H: Dalai Lama and the government of the People's Republic of China are in dispute over Panchen Lama's reincarnation.

- TARSQI Toolkit
  - Temporal Awareness and Reasoning System for Question Interpretation
  - Modular system which detects, resolves and normalizes time expressions
- Logic Representation of Events
  - Logic representation of the event's describing concept augmented with a special predicate *event_EV(e1)*
- Negation
  - Use polarity info attached to the identified events (TARSQUI output) and negate the event's predicate
- Coreference Resolution
  - Dedicated pronominal coreference resolution module to cope with long text pairs

T: The Kinston Indians are a minor league baseball team in Kinston, North Carolina. The team, a Class A affiliate of the Cleveland Indians, plays in the Carolina League.
H: Kinston Indians participate in the Carolina League.

Textual Entailment: what are we missing?

- It is completely clear that the key resource missing is knowledge.
  - Better resources translate immediately to better results.
  - At this point existing resources seem to be lacking in coverage and accuracy.
  - Not enough high quality public resources; no quantification.
- Some Examples
  - Lexical Knowledge: Some cases are difficult to acquire systematically.
    - A bought Y → A has/owns Y
    - Many of the current lexical resources are very noisy.
  - Numbers, quantitative reasoning
  - Time and Date; Temporal Reasoning.
  - Robust event based reasoning and information integration