FROM INFORMATION EXTRACTION TO TEXT UNDERSTANDING

Marie-Francine Moens

Language Intelligence and Information Retrieval

Human Computer Interaction

Department of Computer Science

Katholieke Universiteit Leuven, Belgium

Overview

- Information extraction from text: some history
- **MUSE**: from information extraction to text understanding
- Information extraction: current NLP components
- Implementation : a machine learning framework
- **Challenges**:
  - Lack of training data for infrequent items
  - Learning of complex interrelated structures
  - Lack of world or domain knowledge
  - Text semantics vs. planning representation

[Wikipedia]

Information extraction?

"Information extraction is the identification, and consequent or concurrent classification and structuring into semantic classes, of specific information found in unstructured data sources, such as natural language text, providing additional aids to access and interpret the unstructured data by information systems."

[Moens 2006]

Early origin of information extraction

- ° end 1960s and 1970s: [Schank 1972, 1975]
  - Defines all natural language words in terms of elementary primitives or predicates in an attempt of capturing the semantic content of a sentence
  - **Conceptual dependency** representation specifies **semantic roles:** the **action** of the sentence (e.g., as reflected by the verbs of the text) and the **arguments** (e.g., agent, object) and **circumstances**
  - Main categories of concepts are PPs (i.e., picture producers, in other words, concrete nouns) and actions
  - Representations are ordered in a **script** or scenario which outlines sequences of events or actions

Script: human (X) taking the bus to go from LOC1 to LOC3

1. X **PTRANS** X **from** LOC1 **to** bus stop

2. bus driver **PTRANS** bus **from** LOC2 **to** bus stop

3. X **PTRANS** X **from** bus stop **to** bus

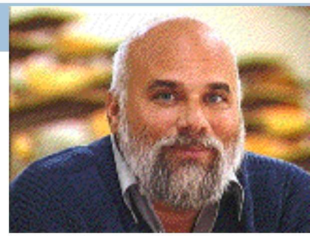
4. X **ATRANS** money **from** X **to** bus driver

5. bus driver **ATRANS** ticket **to** X

6. Various subscripts handling actions possible during the ride.

7. bus driver **PTRANS** bus **from** bus stop **to** LOC3

8. X **PTRANS** X **from** bus **to** LOC3



[Schank 1975]

X performs a physical transition expressed by PTRANS

X gives money to the bus driver. ATRANS is used to express a transfer of an abstract relationship, in this case the *possession* of money.

(3), (7), (8): mandatory

Frame-based approaches

- [Minsky 1975]: frame-based knowledge representations
  - Frames are often triggered by the occurrence of a certain word or phrase
  - Very **partial analysis** of the input text:
    - Algorithm tries to match natural language sentences with particular frames by simply filling out the slots in accordance with the constraints placed on them
    - Often top-down (expectation-driven): guided by the expected patterns to be found in the text
    - Robust: ignoring of irrelevant information
  - **Template frames** that outline the information can be used as **output**

Frame-based approaches

- Patterns to be identified can be encoded as regular expressions and recognized by finite state automaton
- Frames are often organized in a **script:**
  - because of their strict organization, scripts have good predictive ability useful in information extraction

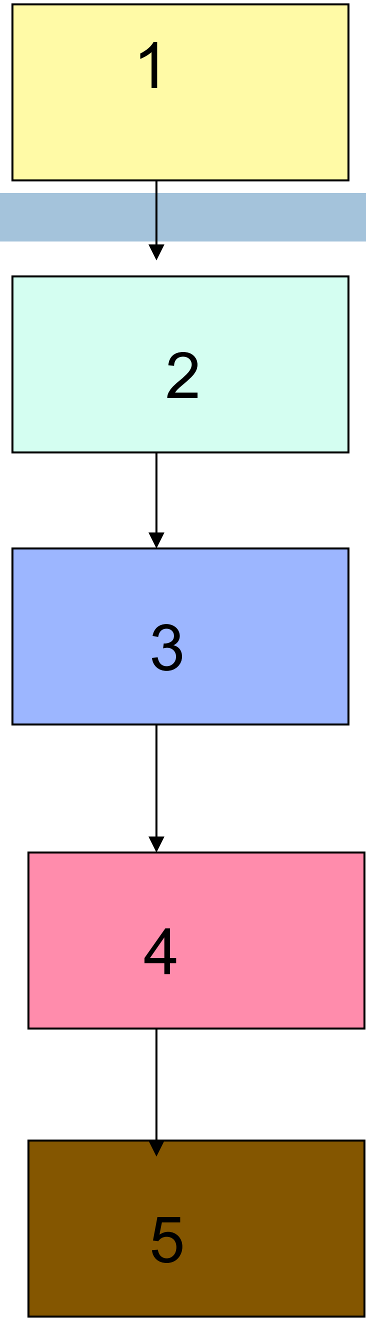
- Examples of some famous information extraction applications:
  - FRUMP: Yale University
  - FASTUS: Stanford Research Institute

FASTUS

- Finite state automaton implementation: **set of cascaded, non-deterministic finite-state transducers**
  - Application of symbolic rules in the form of hand-crafted regular expressions
  - Cascade: output of finite state transducer is input for next finite state transducer

[Hobbs et al. 1996] [Hobbs JBioInformatics 2002]

Cascade of finite state transducers



1. Recognition of compound words and named entities

2. Partial parse: recognition of verb, noun, prepositional phrases, actives, passives, gerunds

3. Recognition of complex noun groups

4. Resolution to active form, recognition of information to be extracted

5. Structure merging

- Example sentence:

  Bridgestone Sports Co. said Friday it has set up a joint venture in Taiwan with a local concern and a Japanese trading house to produce golf clubs to be shipped to Japan.

Step 2

Company nam e	**Bridgestone Sports Co.**
Verb group	**said**
Noun group	**Friday**
Noun group	**it**
Verb group	**had set up**
Noun group	**a joint venture**
Preposition	**in**
Location	**Taiwan**
Preposition	**with**
Noun group	**a local concern**
And	**and**
Noun group	**a Japanese trading house**
Verb group	**to produc e**
Noun group	**golf clubs**
Verb group	**to be shipped**
Preposition	**to**
Location	**Japan**

Step 4

Extraction rules:

<Company/ies> {Set-up} {Joint-Venture} {with} <Company/ies> {Produce} <Product>

Relation:	TIE-U P
Entities:	**Bridgestone Sports C o .** **a local concer n** **a Japanese trading house**
Joint Venture Company:	
Activity:	
Amount:	

Activity:	PRODUCTION
Company:	
Product:	**golf clubs**
Start Date :	

Symbolic techniques: results

- Successful systems, built and tested in many subject domains
- e.g., MUC-7 (1998): subject domain of air plane crashes:
  - Performance of individual systems: largely similar
  - Certain information much easier to extract than others
- **Problem**:
  - Infinite variety of subject domains: very difficult to exhaustively implement the symbolic knowledge
  - Very difficult to construct a script for every conceivable situation

Table 2: Maximum Results Reported in MUC-3 through MUC-7 by Task

Evaluation\Tasks	Named Entity	Coreference	Template Element	Template Relation	Scenario Template	Multilingual
MUC-3					R < 50% P < 70%	
MUC-4					F < 56%	
MUC-5					EJV F < 53% EME F < 50%	JJV F < 64% JME F < 57%
MUC-6	F < 97%	R < 63% P < 72%	F < 80%		F < 57%	
MUC-7	F < 94%	F < 62%	F < 87%	F < 76%	F < 51%	
Multilingual						
MET-1	C F < 85% J F < 93% S F < 94%					
MET-2	C F < 91% J F < 87%					

Legend: R = Recall P = Precision F = F-Measure with Recall and Precision Weighted Equally

E = English C = Chinese J = Japanese S = Spanish

JV = Joint Venture ME = Microelectronics

Information extraction from text

- Tested in ARPA's Tipster Text Program and in the
  - Past:
    - **Message Understanding Conferences (MUC)**
    - **Automatic Content Extraction (ACE)**
  - Current:
    - **Text Analysis Conference (TAC)** (National Institute of Standards and Technology, NIST)
    - **Computational Natural Language Learning (CoNLL)**
    - **SemEval** competitions

Today

- **Machine reading**
  - = The automated discovery of meaningful knowledge in free text

    => form of automated understanding of the text
- Can be evaluated in the translation to another modality

  => MUSE project : **translation of text into a virtual reality**

  => translation of text to planning language

  = translation to a meaning representation

MUSE project

- MUSE: Machine Understanding for interactive StorytElling
- Algorithms for translating text into virtual worlds, 9/2012-8/2015, EU FP7-296703 (FET-open call)

Teesside University

Institut "Jožef Stefan"

Universiteit Leiden

HAS
HAUTE AUTORITÉ DE SANTÉ

KU LEUVEN

- The ability of a computer or other machine to perform those activities that are normally thought to require intelligence:
  - **Automated reading or understanding of texts written in natural language**
  - Understanding of images
  - …

  - Beyond: story telling, game development, story generation, case based reasoning, ...

Current NLP Components

Information extraction: Where are we now?

- NLP components:
  - Named entity recognition
  - Noun phrase coreference resolution
  - Semantic role labeling
  - Event recognition
  - Temporal expression recognition
  - Temporal relation recognition
  - Spatial relation recognition
  - …

Semantic labeling

Named entity recognition

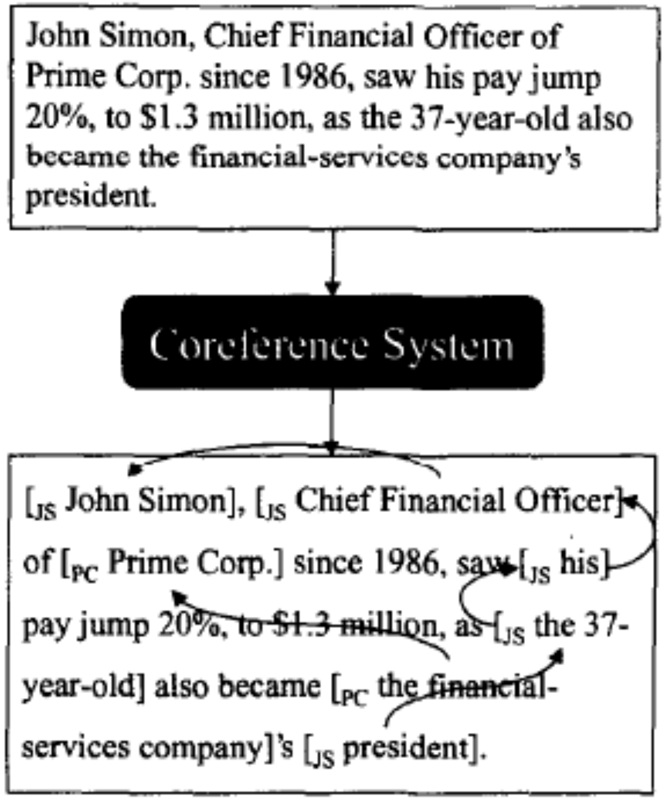
- Recognition of classes of entities (persons, locations, companies, organizations, …)

Over the past month the BBC has had rare access to International Monetary Fund boss Christine Lagarde. Now, as European leaders meet in Brussels, she will be at the centre of the fight to avert another financial crisis.

(BBC News 1-3-2012)

Noun phrase coreference resolution

Identify all noun phrases (mentions) that refer to the same entity



[Ng & Cardie ACL 2002]

Figure 1: Coreference System

Semantic role labeling

Recognizing the basic event structure of a sentence ("**who**" "**does what**" "**to whom/what**" "**when**" "**where**", ...)

fall.01

Arg1: Logical subject, patient, thing falling

Arg2: Extent, amount fallen

Arg3: Start point

Arg4: End point, end state of Arg1

Ex1: [$_{\text{Arg1}}$ Sales] *fell* [$_{\text{Arg4}}$ to $251.2 million] [$_{\text{Arg3}}$ from $278.7 million].

Ex2: [$_{\text{Arg1}}$ The average junk bond] *fell* [$_{\text{Arg2}}$ by 3.7%].

(PropBank roles)

Event recognition

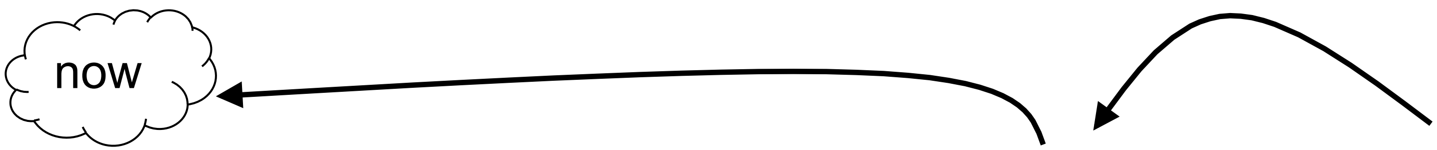
Events = situations that happen or occur

Events can be punctual (1-2) or last for a period of time (3-4), or describing states or circumstances in which something obtains or holds true (5)

1. *Ferdinand Magellan, a Portuguese explorer, first [event **reached**] the islands in search of spices.*
2. *A fresh flow of lava, gas and debris [event **erupted**] there Saturday.*
3. *11,024 people, including local Aeta aborigines, were [event **evacuated**] to 18 disaster relief centers.*
4. *"We're [event **expecting**] a major eruption," he said in a telephone interview early today.*
5. *Israel has been scrambling to buy more masks abroad, after a [event **shortage**] of several hundred thousand gas masks.* (Sauri et al. 2005)

Temporal information recognition

- Recognition: identify which phrases are temporal and which are not
- Normalization: translate the temporal phrase to a standard time expression



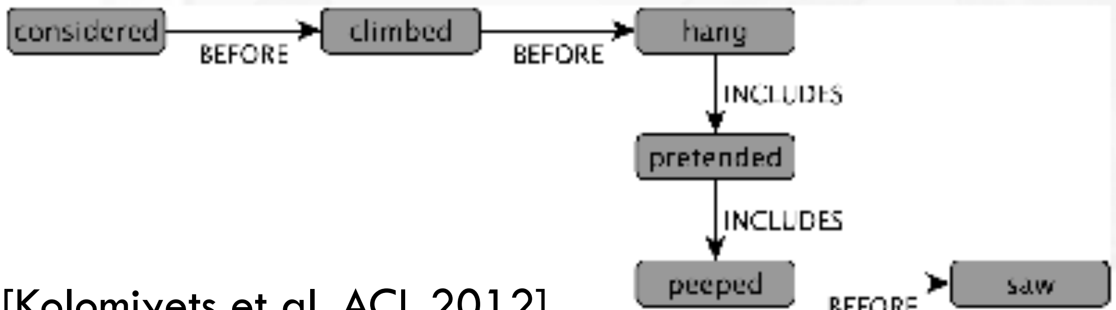
Show me platform six five days ago at 17:20.

Temporal relation recognition

- Recognition of temporal relations between events:

e.g., BEFORE, AFTER, INCLUDES, IS_INCLUDED, DURING, SIMULTANEOUS, BEGINS, … (TimeML)

So she [$_{event}$ **considered**] a while, and then [$_{event}$ **climbed**] up the wall and let herself [$_{event}$ **hang**] down by her hind legs from a peg, and [$_{event}$ **pretended**] to be dead. By and by a Mouse [$_{event}$ **peeped**] out and [$_{mover}$ **saw**] the Cat hanging there.



[Kolomiyets et al. ACL 2012]

Spatial relation recognition

- Recognition of:
  - **Trajector**: entity (person, object or event) whose location is described
  - **Landmark**: the reference entity in relation to which the location or the motion of the trajector is specified
- Recognition of more fine-grained spatial meanings: e.g., direction, motion

[$_{\text{Trajector}}$ She]
[$_{\text{Motion-Indicator}}$ went]
[$_{\text{Spatial-Indicator(type=REGION/RCC8,value=TPP)}}$ to]
[$_{\text{Landmark(path=END)}}$ school].

[Kordamshidi et al. TSLP 2011]

28

Implementation of NLP Components

Classifier approach to extracting semantics

- Method:
  - The text fragments (tokens, chunk, parse nodes) in training and test data are **segmented**
  - Special attention goes to **feature engineering**: use of NLP resources to extract features (e.g., part-of-speech tagging, syntactic dependencies, signaling words, etc.)
  - A **classifier** is trained on manually annotated examples (e.g. maximum entropy) possibly integrating constraints in structured output classifiers (Markov random fields, structured support vector machines)
  - The classifier is applied on previously unseen test examples

Classifier approach to extracting semantics

- = assignment of controlled language descriptors to words or phrases in a text
- Semantic classes or labels:

  $C = \{C_1, C_2, \ldots, C_m\}$

- Although not a necessary condition, classes are usually mutually exclusive:
  - Often seen as a single-label multi-class learning problem
  - But classes can be organized in a hierarchy, ontological structure

What are machines capable of?

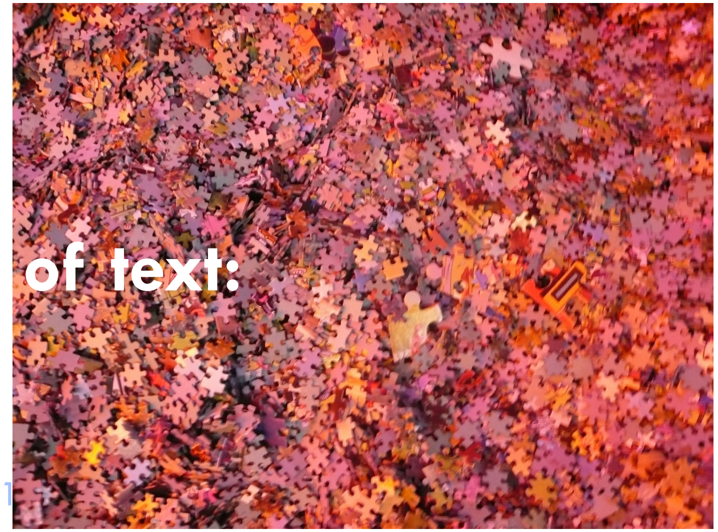
- (1) If sufficient training examples and if the classifier is tested on texts of a same domain as the training examples
  - Roughly: F1-values of the recognition maximum ca. 80% (lower for noun phrase coreference resolution, temporal expression normalization and temporal relation extraction)
  - **Room for improvement !**

http://www.etoon.com/cartoon-store/Vlad-Kolarov/

What are machines capable of?

- (2) The machine recognizes fragmentary pieces (e.g., names, facts) and the recognition of related fragments of text are often limited to the sentence level
  - Emerging recognition of discourse understanding: e.g. noun-phrase coreference resolution and temporal relation recognition

Human understanding of text:
inferencing,
connecting content



[Wikipedia]

What are machines capable of?

- (3) The machine only uses information resided in the texts

[Wikipedia]

- **Human understanding of text: humans connect to their world/domain knowledge**

What are machine capable of in restricted domains?

- Method:
  - Use of handcrafted grammar that translates to primitives in form of prolog commands that steer a virtual world [Rochefort et al. AAAI 1997]
  - Learn alignments between words/phrases and semantics
    - Grounded language learning in form of PCFG [Börschinger et al. EMNLP 2011], [Kim & Mooney EMNLP 2012], CCG [Zettlemoyer & Collins EMNLP 2007], or logical form [Liang et al. HLT 2011]
    - Mapping from a dependency parsing tree to a planning tree [cf. Jones et al. ACL 2012]
    - Often EM like reinforcements [e.g., Chen & Mooney 2008; Goldwasser et al. 2011] or EM/Gibbs sampling trained generative models [Liang et al. ACL-IJCNLP 2009], though some interest in spectral models [Dhillon et al. EMNLP 2012]
- Results:
  - Some success on restricted domains
    - Robot directions, weather reports, sports descriptions

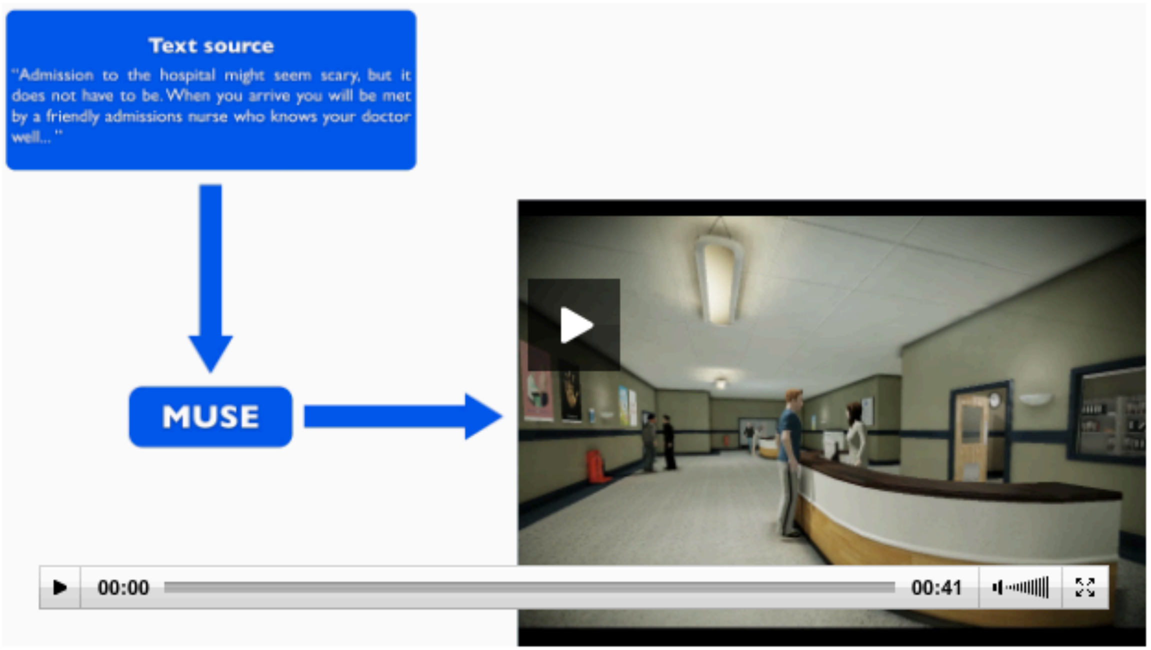
NLP in MUSE

MUSE

- = bringing text to life
  - Children's stories and patient education guidelines

render these as 3D-virtual worlds

: http://www.muse-project.eu/

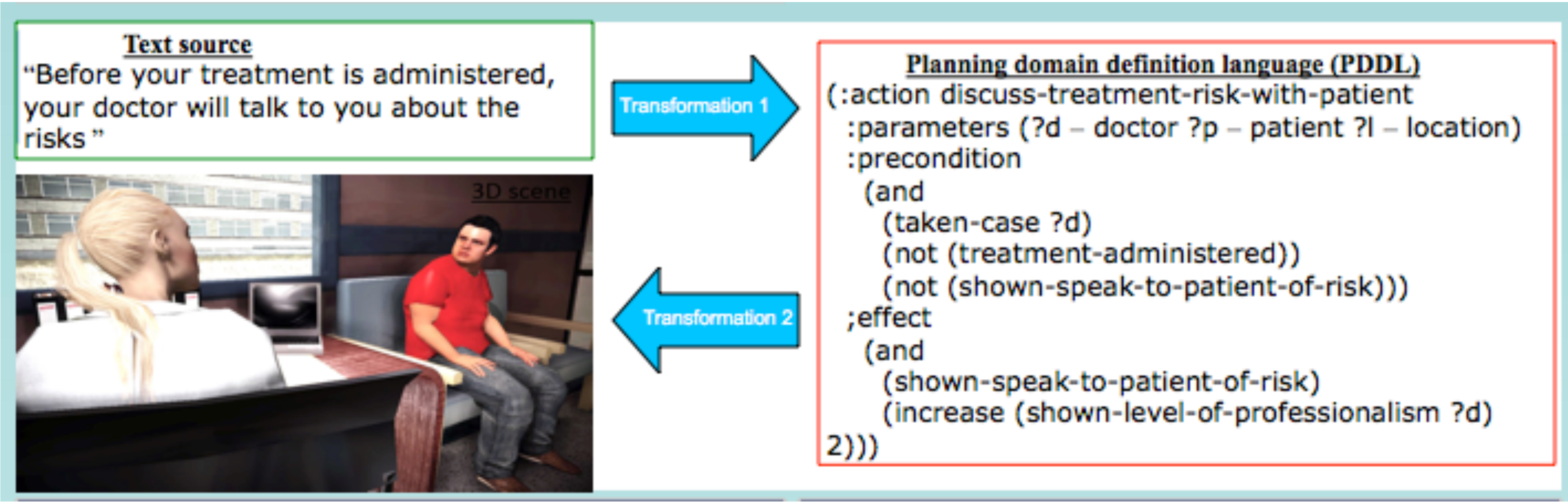


MUSE: What do we recognize in the texts?

- At the **sentence level**:
  - Actions/events and their semantic roles (actor, patient, instrument …)
  - Scope of negation, modality
- At the **discourse level**:
  - Conditional, pre/postconditions of actions/scenario, causal (local, multicausal)
  - Coreferent noun phrases
  - Temporal relations between actions
  - (Spatial relations)
  - Intention (of actors), goals
  - Abstract attitudes, personality treats

MUSE: Beyond information extraction

- But **mapping** needed to PDDL **knowledge representation**
  - Possibility of consistency feedback with setting specific knowledge



MUSE

- Challenges still remain
  - Lack of training data for infrequent items
  - Learning of complex interrelated structures
  - Lack of world or domain knowledge
- Additional challenge:
  - Text is incomplete or not right level of detail is given to translate to planning language

Challenge 1:
Lack of training data for infrequent items

Lack of training data

- Lexical items (words) are very important: they carry an important part of the semantic meaning
- Many, many different words in a language ! Many of which are never seen in the training examples !

Leveraging unlabeled data

- Semi-supervised learning (some labeled data)
  - Self-learning: iterative retraining after labeling of data points for which the current model is most confident
  - Transductive inference: labels of the unlabeled examples are predicted according to a most likely model that explains the labeled and unlabeled examples
- Unsupervised learning (no labeled data)
  - Try to find meaningful clusters
  - Clusters can be used as features for supervised models

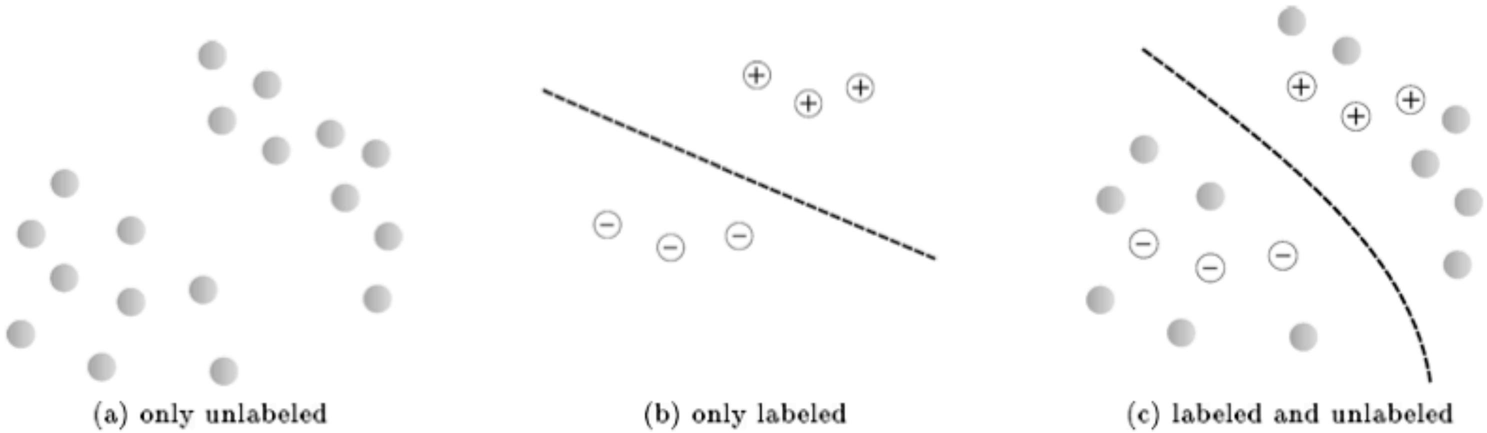
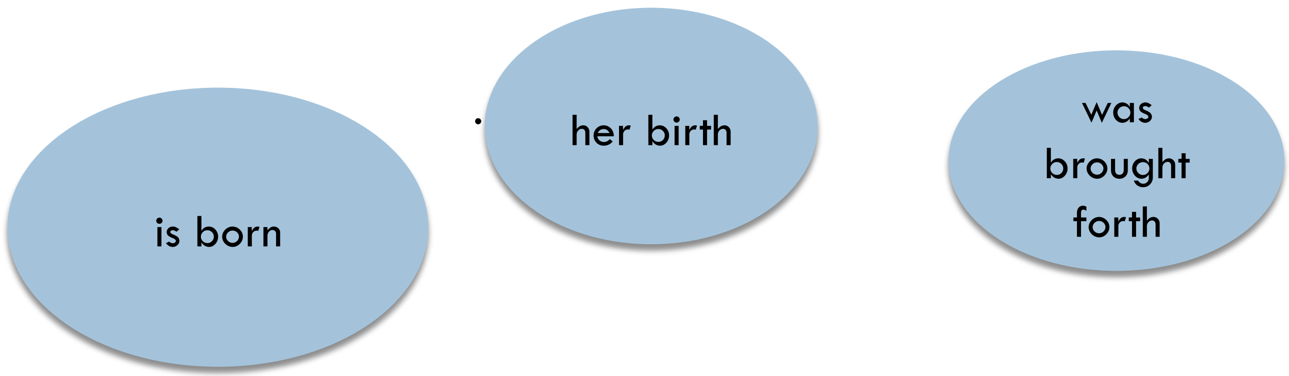


Figure 2: Schematic figure illustrating how unlabeled data might improve a supervised classifier. Grey dots are unlabeled data, white dots labeled data and the dotted line the classification boundary.

Leveraging unlabeled data

- The semantic classes are expressed with many different words that when used as features do not cluster

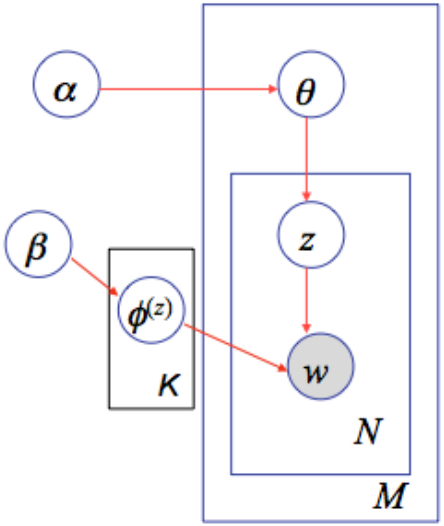


- Even if we normalize syntactic constructions clustering is difficult
- Many words have different meanings

Unsupervised learning

- Learn classes of exchangeable phrases, words and syntactic constructs

- Potential of latent class models



Learn from large textual data sets !

[Blei et al. ML 2003]

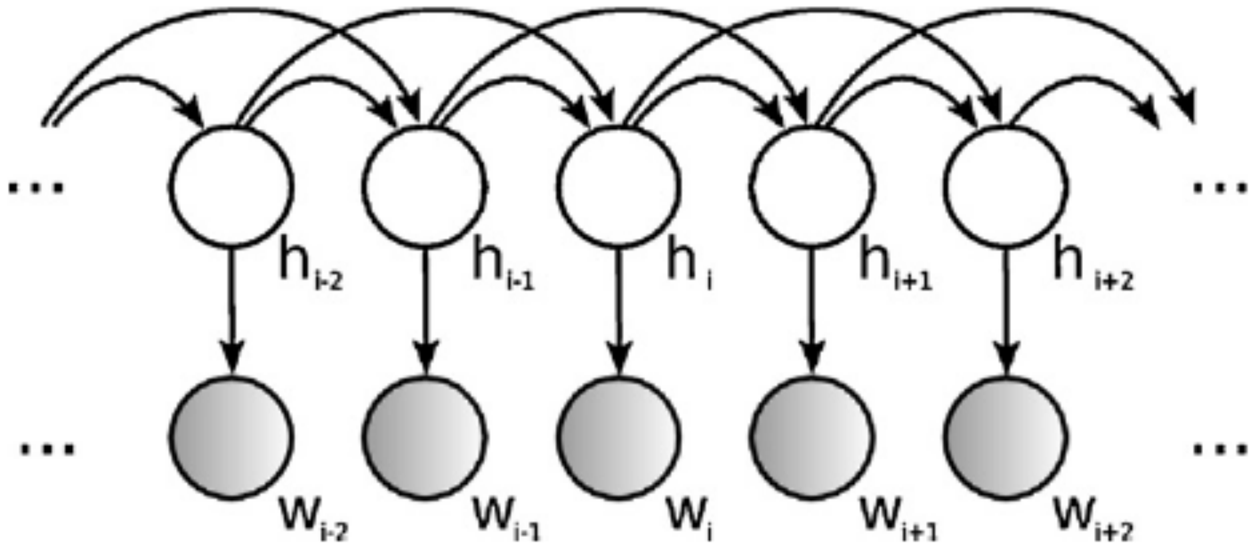
Latent Words Language Model

Compuserve	**corp**	**said**	**Tuesday**	**it**	**anticipates**	**a**	**loss**
Microsoft	inc	told	Friday	they	expects	the	profit
Crysler	corp.	reported	Thursday	he	expected	some	gain
Oracle	ltd	added	Monday	she	assumes	an	deficit
Software	co	say	Wednesday	this	doubts	another	earnings

A	**Japanese**	**electronics**	**executive**	**was**	**kidnapped**	**in**	**Mexico**
the	U.S.	tobacco	director	is	abducted	on	Usa
its	German	sales	manager	we	killed	at	UK
an	British	consulting	economist	are	found	of	Australia
one	Russian	electric	spokesman	be	abduction	into	Canada

E.g., learning of synonyms and related words from a representative corpus
[Deschacht, De Belder & Moens CSL 2012] [Kolomiyets et al. ACL 2011]
Cf. use of language modeling in information extraction:: [Deschacht, De Belder & Moens CSL 2012] [Yates et al. Comp. Ling. 2013]

Latent Words Language Model



Generative model: Bayesian network with observed (grey nodes) and hidden variables (white nodes). The hidden nodes represent the probability distribution of each word of the vocabulary being present in the specific context. In our implementation a second order Markov dependency of the hidden words left and right of the target hidden word is modeled. The model is trained with a forward-forward beam search or with Markov Chain Monte Carlo sampling.

Latent class model challenges

- Latent class models are computationally expensive
  - Essentially have to enumerate all possibilities for all latent class assignments
  - Opportunities for new optimization techniques, sampling techniques
  - Opportunities for building contextual vectors [Vulic & Moens NAACL 2013]
- In MUSE, the domain is constrained
  - So computational expense may be acceptable given potential performance improvements

Challenge 2:
Learning complex interrelated structures

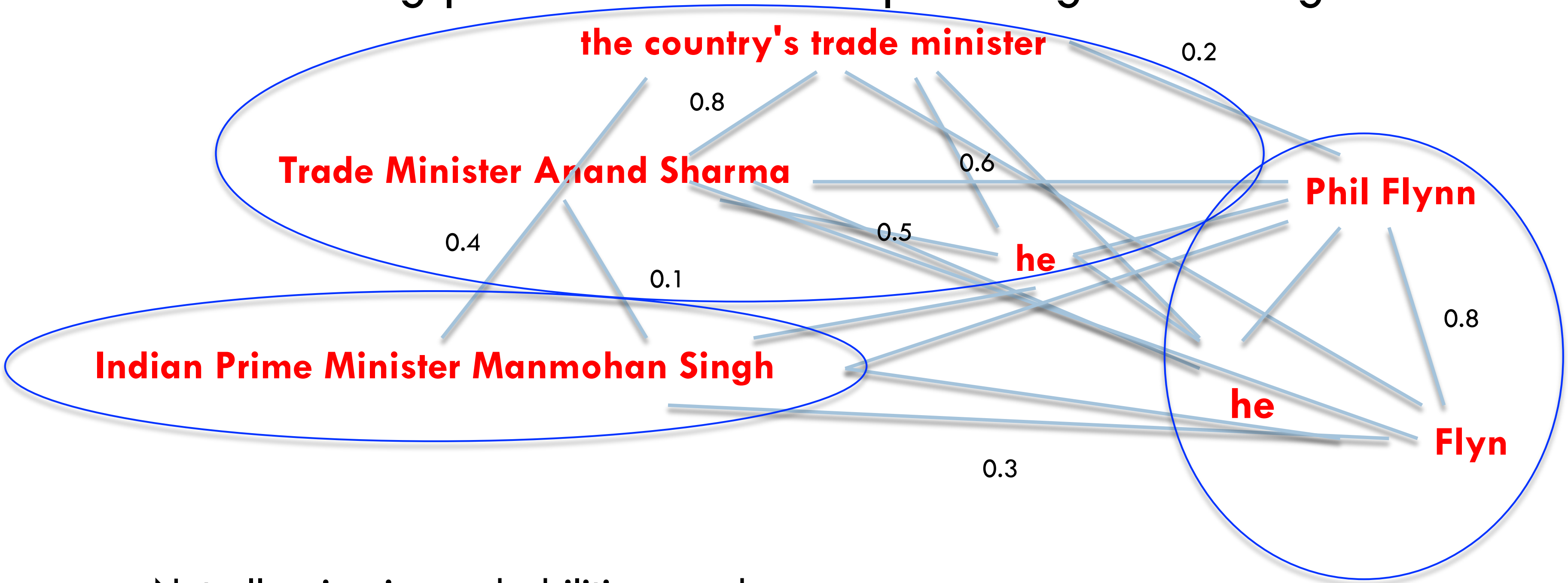
Complex structures and combining evidence

- Already done on a limited scale in the literature:
  - Several passes through the data (output of one pass serves as input for other)
  - Output of local extractors is combined (e.g., template filling) possibly by using additional constraints
  - Prediction of a structured output (parse tree, predicate logic, or graph) (e.g., learning of semantic parser, jointly learning the ontological classes)
- Uncertain recognition can be reinforced by other evidences:

Supporting and combining evidence !

Several passes through the data

- Noun phrase coreference resolution:
  - First step: train and apply classifier with local context
  - Reinforcing pair-wise relationships through clustering



Not all pair-wise probabilities are shown

Cf. [Culotta et al. ACL 2007]

[CNN 13-3-2012]

Several passes through the data

- Named entity recognition:
  - First step: train classifier with local context
  - Second step: retrain with additional features which are the output of the first classifier

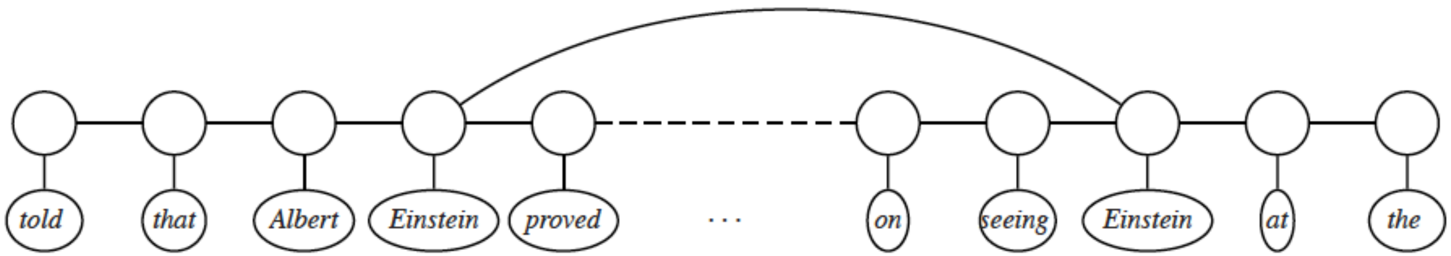


Figure 1: An example of the label consistency problem. Here we would like our model to encourage entities *Albert Einstein* and *Einstein* to get the same label, so as to improve the chance that both are labeled *PERSON*.

[Krishnan & Manning ACL 2006]

Constrained output of local classifiers

- Template filling:
  - Local extractors and templates are filled based on enforcing constraints in a late fusion approach [Minkov & Zettlemoyer ACL 2012]
  - Possibility of integer linear programming formulation

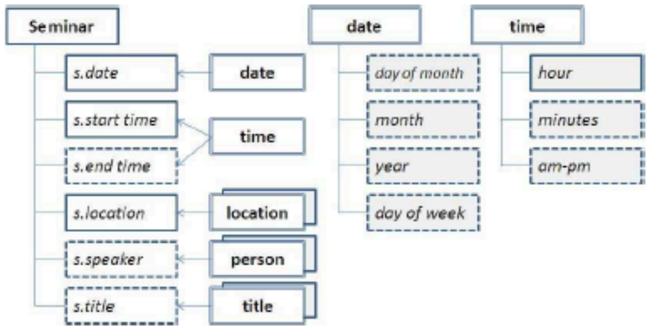


Figure 2: The relational schema for the seminars domain.

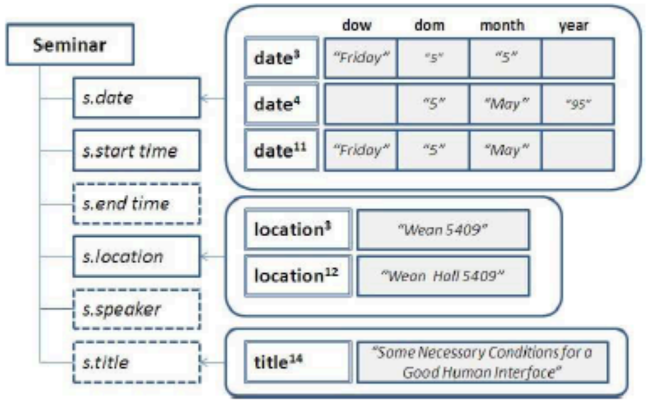


Figure 3: A record partially populated from text.

Prediction of a structured output

- Potential of
  - Extension of dependency framework – graph based approach now used in TERENCE (PhD of Oleksandr Kolomiyets)
  - Semi-supervised dependency parsing [Mirroshandel et al. ACL 2012]
  - Tensor form frameworks [Cohen et al. ACL 2012]
  - Decomposition models [Samdani & Roth ICML 2012]
  - Communicative optimization [current PhD research of Parisa Kordjamshidi]

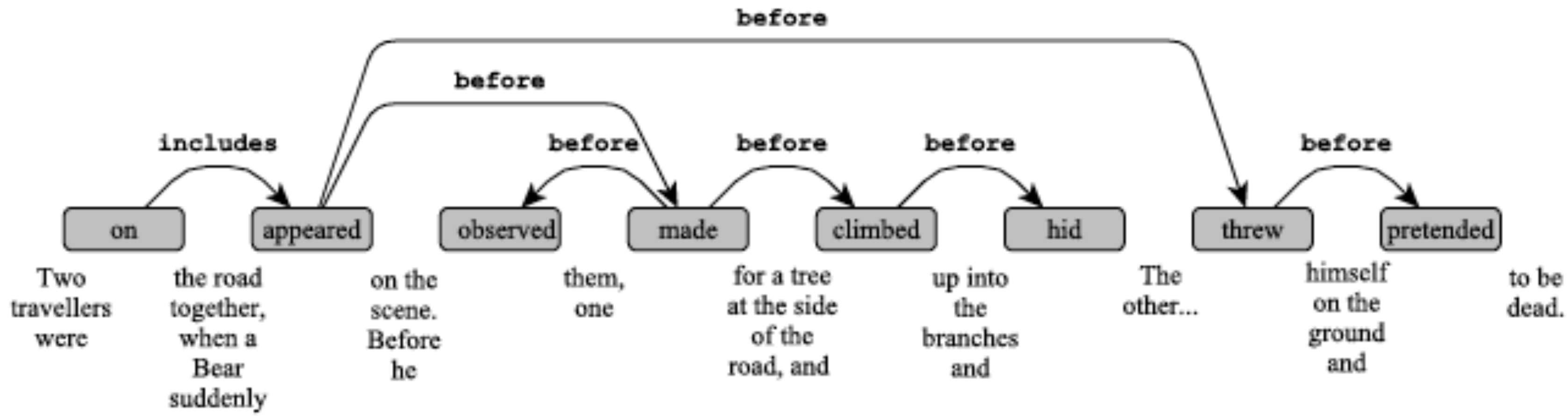


Figure 1: Event timeline for the story of the Travellers and the Bear. Nodes are events and edges are temporal relations. Edges denote temporal relations signaled by linguistic cues in the text. Temporal relations that can be inferred via transitivity are not shown.

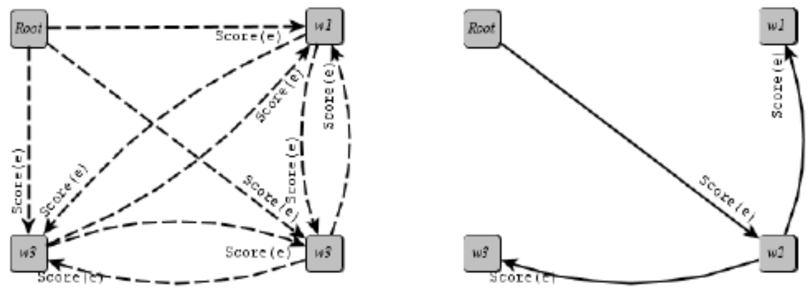


Figure 2: A setting for the graph-based parsing model: an initial dense graph G (left) with edge scores $\textsc{Score}(e)$. The resulting dependency tree as a spanning tree with the highest score over the edges (right).

Learning of a semantic dependency structure: graph based mining: [Kolomiyets et al. ACL 2012]

Prediction of a structured output

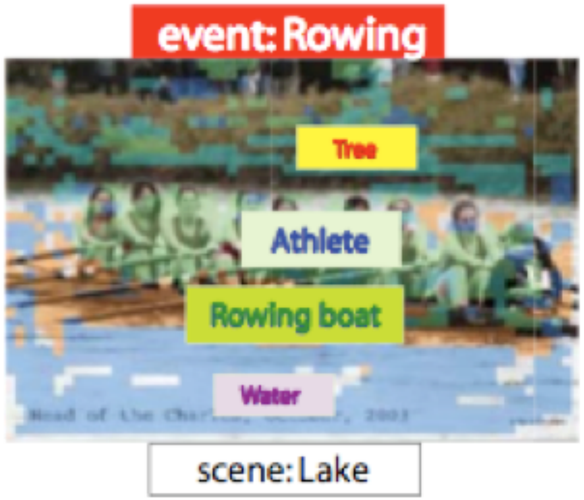
56 .... To be fully published soon

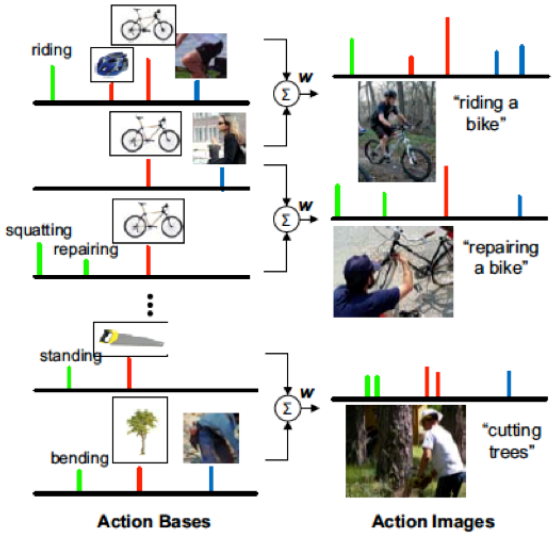
Prediction of a structured output

[PhD of Parisa Kordjamshidi] [Kordjamshidi & Moens ROKS 2013]

Supporting and combining evidences

- Guessing labels in contexts:
  - Current work in computer vision could be inspiring





Work of Fei-Fei Li and her group at Stanford

MUSE: Learning of complex related structures

- Recognition of the **narrative structure** [cf. Mani Morgan&Claypool 2013]
  - Connected events
  - Protagonists
  - Temporal and causal relations/preconditions - postconditions
  - Character traits
  - …

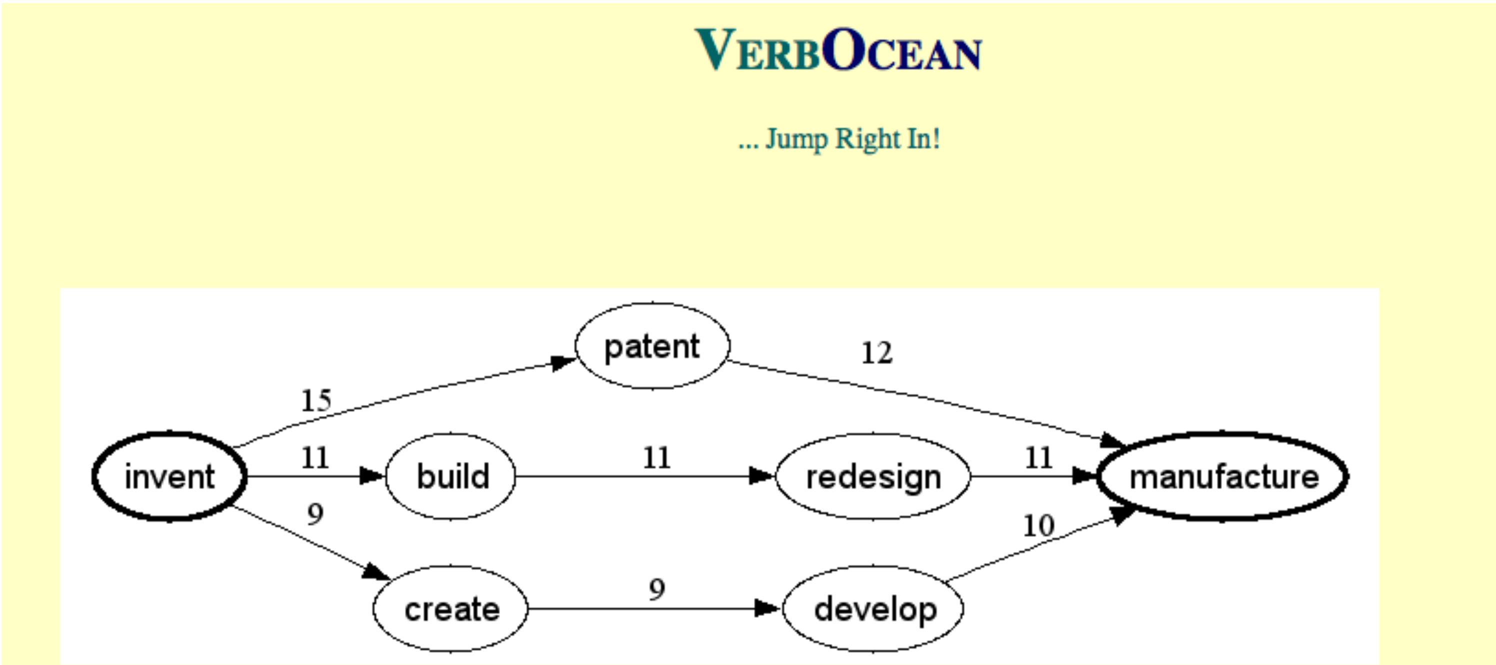
Challenge 3:
Lack of world knowledge

Examples of world knowledge

- Multiple examples of world knowledge:
  - It is much more likely that a person says something than a location
  - It is much more likely that a chair is next to the table than on top of the table
  - It is much more likely that you jump in the water and then swim than vice versa
  - …
  - Could be learned from text, video, databases (Freebase – Krishnamurthy and Mitchell 2012), etc.

Learning temporal relations between events

- VerbOcean



Example: VERBOCEAN's temporal precedence chains (the "happens-before" relation) between *invent* and *manufacture* shown with edge weights

[Chklovski & Pantel EMNLP 2004]

Learning of narrative scripts

row	s_1	s_2	s_3	s_4
1	∅	walk into restaurant	∅	enter restaurant
2	∅	∅	walk to the counter	go to counter
3	∅	find the end of the line	∅	∅
4	∅	stand in line	∅	∅
5	look at menu	look at menu board	∅	∅
6	decide what you want	decide on food and drink	∅	make selection
7	order at counter	tell cashier your order	place an order	place order
8	∅	listen to cashier repeat order	∅	∅
9	pay at counter	∅	pay the bill	pay for food
10	∅	listen for total price	∅	∅
11	∅	swipe credit card in scanner	∅	∅
12	∅	put up credit card	∅	∅
13	∅	take receipt	∅	∅
14	∅	look at order number	∅	∅
15	∅	take your cup	∅	∅
16	∅	stand off to the side	∅	∅
17	∅	wait for number to be called	wait for the ordered food	∅
18	receive food at counter	get your drink	get the food	pick up order
19	∅	∅	∅	pick up condiments
20	take food to table	∅	move to a table	go to table
21	eat food	∅	eat food	consume food
22	∅	∅	∅	clear tray
22	∅	∅	exit the place	∅

Figure 2: A MSA of four event sequence descriptions

[Chambers & Jurafsky ACL 2009; Regneri et al. ACL 2010; Jans et al. EACL 2012; Li et al. AAAI 2012]

1

World knowledge in MUSE

- World knowledge in MUSE:
  - Some very basic knowledge (e.g. physics) will always be manually encoded in the VR environment anyway
  - Automatically acquiring the world knowledge can reduce the need for hand-coding (but in MUSE we can still fall back on hand-coding world knowledge when necessary)
- Knowing when world knowledge is relevant

Challenge 4: Matching text semantics to the planning representation

Text semantics vs. planning representation

- A central problem is learning the correspondence between a rich world state and a stream of text referring to that state
- Information in the text can be missing or being communicated **at a different level of detail** than is needed in the planning language (e.g., [Branahan et al. ACL 2012])
- A key challenge for MUSE will be to match the semantics we extract from text to the semantics of the story planner

    [related to Challenges 2 and 3]

Text semantics vs. planning representation

- **Virtual camera control**: e.g., translation of an image caption of museum object to guide video exploration of the object [Reiterer et al. 2010]

- **Robot control**: e.g., translating instructions to planning language

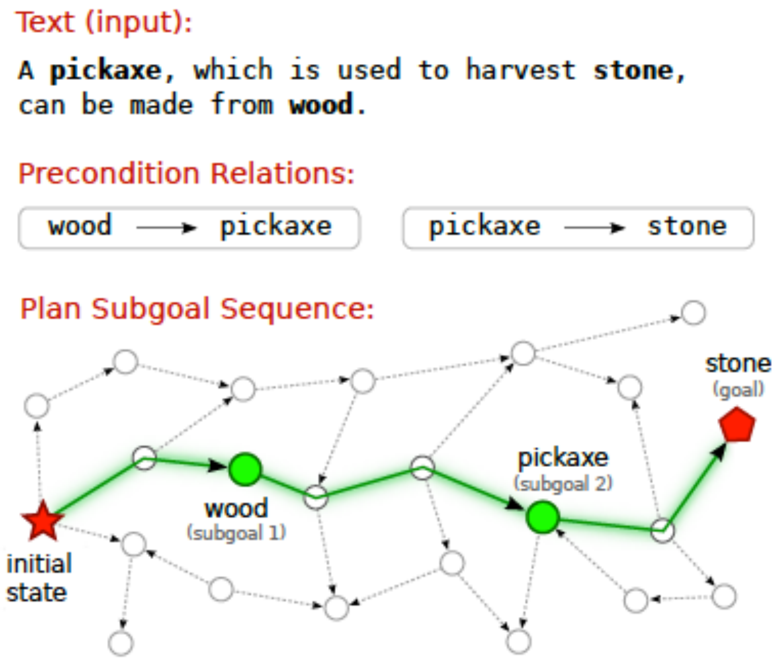


Figure 2: A high-level plan showing two subgoals in a precondition relation. The corresponding sentence is shown above.

A pickaxe, which is used to harvest stone, can be made from wood.

(a)

Low Level Actions for: wood → pickaxe → stone	
step 1:	move from (0,0) to (2,0)
step 2:	chop tree at: (2,0)
step 3:	get wood at: (2,0)
step 4:	craft plank from wood
step 5:	craft stick from plank
step 6:	craft pickaxe from plank and stick
. . .	
step N-1:	pickup tool: pickaxe
step N:	harvest stone with pickaxe at: (5,5)

(b)

[Branavan et al. ACL 2012]: (in reinforcement learning framework)

1) extract a set of precondition/effect relations implied by the text

2) use these induced relations to determine an action sequence for completing a given task in the environment

⇒ learning of general knowledge about the domain structures Cf. [Jans et al. EACl 2012], PhD thesis of Aparna Nurani

⇒ domain structure constraints the possible actions

Examples of the PDDL representation

- Text: ***Over several months, you will meet various professionals who are members of a multidisciplinary team (surgeon, nutrition doctor, dietician, psychiatrist or psychologist, anaesthetist, etc .) who will*** *provide you with information and* ***examine you.*** **They will also request various examinations (blood tests,** upper gastrointestinal endoscopy* and, if necessary, X-rays, assessments of respiratory and cardiac function, pregnancy test and an examination of the mouth and teeth**).**
- **Planning Domain Definition Language (PDDL) representation**:

```
(:action clinical-examination
        :parameters (?p - patient ?d - medicalprofessional ?l - medicalLoc)
        :precondition
                (and
                        (= ?l practitionersoffice)
                        (at ?d ?l)
                        (at ?p ?l)
                        (not (examined-by ?p ?d))
                )
        :effect
                (and
                        (examined-by ?p ?d)
                        (blood-test-required ?p)
                )
)
```

Conclusions

- Semantic parsing of sentences and discourses is still different from machine understanding of text
- **MUSE = test case for machine understanding**
- Promising directions:
  - Latent class models and other unsupervised techniques will help us handle the lack of training data
  - Structured learning allows different pieces to be recognized and integrated while reinforcing each other: will be useful in mapping to knowledge representation
  - Some world knowledge might be automatically acquired from large corpora or other resources

Our MUSE team

- Steve Bethard
- Wim De Mulder
- Oleksandr Kolomiyets
- Ngoc Quynh Do Thi

Thanks !

Questions?