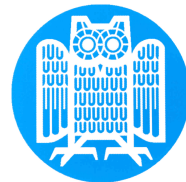# Seminar Textual Entailment

Introduction

Manfred Pinkal
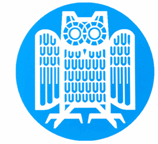Saarland University
Summer 2009

Inference

- Can I take Aspirin for headache?
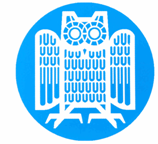- Aspirin is used for headache.

- Will there be classes on May 1?
- May 1 is a holiday.

- Should we meet at 11 a.m.?
- I have a class starting at 10.

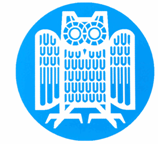# Logical Inference/ Entailment

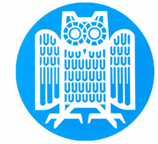| Text | Relation of
Inference/ Entailment? | Hypothesis |

Should I take Aspirin for headache?

Aspirin is used for headache	Relation of Inference/ Entailment?	One should take Aspirin for headache

Should I take Aspirin for headache?

SR_{Text}

$SR_{Hypothesis}$

Semantic
Interpretation

Aspirin is used
for headache

Relation of
Inference/ Entailment?

One should take
Aspirin for headache

Should I take Aspirin for headache?

Used(aspirin, headache)		Take(aspirin, headache)

Semantic
Interpretation ↑

Aspirin is used for headache	Relation of Inference/ Entailment?	One should take Aspirin for headache

Should I take Aspirin for headache?

Logical Entailment?

Used(aspirin, headache) $\models$ Take(aspirin, headache)

Semantic
Interpretation

Aspirin is used
for headache

Relation of
Inference/ Entailment?

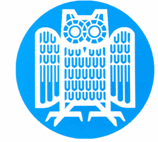One should take
Aspirin for headache

Should I take Aspirin for headache?
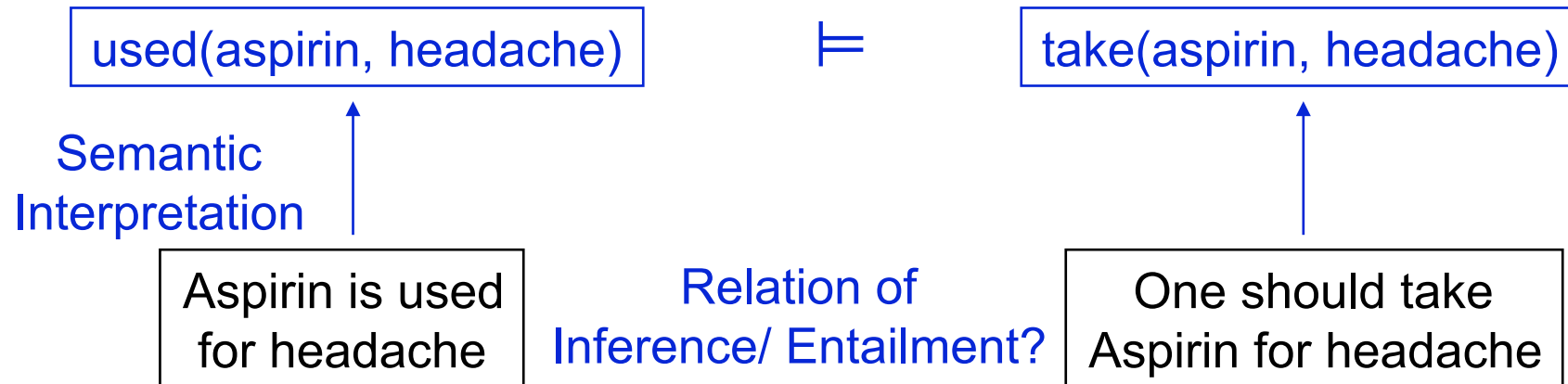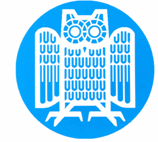
Logical Entailment?
Computed through deduction systems/ theorem provers

$\forall x \, \forall y(\text{used}(x,y) \rightarrow \text{take}(x,y))$
used(aspirin, headache) $\rightarrow$ take(aspirin, headache), used(aspirin, headache)
take(aspirin, headache)

used(aspirin, headache) $\models$ take(aspirin, headache)

Semantic
Interpretation

Aspirin is used
for headache

Relation of
Inference/ Entailment?

One should take
Aspirin for headache

Inference

- Aspirin is used for headache.
- Should I take Asprin for headache?

- May 1 is a holiday.
- Will there be classes on May 1?

- Should we meet at 11 a.m.?
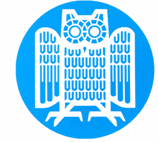- I have a class starting at 10.

Logical Inference

Problems with:

- Coverage
  - Linguistic knowledge/ world knowledge required

- Conceptual adequacy
  - Inferences often just plausible, not necessarily valid

- Ambiguity of natural language
  - In particular, lexical ambiguity

- An approach building on classical logic:

  → Interpretation by abduction (Hobbs et al.)

Textual Entailment

A novel framework to model linguistic inference, offering:

- An intuitive, pre-theoretic concept of entailment and inference

- Approximate, wide-coverage methods for checking entailment and inference

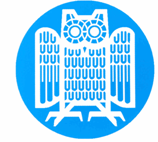- An evaluation method and shared tasks for recognizing textual entailment with objectively measurable results

Seminal paper on textual entailment:

→   Monz and de Rijke 2001

Recognizing Textual Entailment challenge (since 2005)
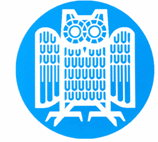
→   Dagan et al. 2006

Textual Entailment: Concept

- An intuitive, pre-theoretic concept of entailment and inference:

  „We say that *T entails H* if the meaning of *H* can be inferred from the meaning of *T*, as would typically be interpreted by people. This somewhat informal definition is based on (and assumes) *common human understanding of language* as well as *common background knowledge.*" (Monz and de Rijke 2001)

Textual Entailment: Concept

- An intuitive, pre-theoretic concept of entailment and inference:

    „We say that *T entails H* if the meaning of *H* can be inferred from the meaning of *T*, as would typically be interpreted by people. This somewhat informal definition is based on (and assumes) *common human understanding of language* as well as *common background knowledge.*" (Monz and de Rijke 2001)

- Discussion of the textual entailment concept in
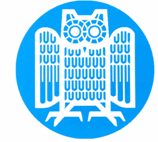   → *Zaenen et al. 2005, Manning 2006*

Textual Entailment

A rather novel framework to model inference, offering:

- An intuitive, pre-theoretic concept of entailment and inference

- Approximate, wide-coverage methods for checking entailment and inference

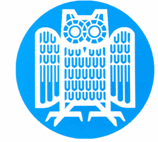- An evaluation method and shared tasks for recognizing textual entailment with objectively measurable results

Textual Entailment: Evaluation

Shared task – RTE challenge:

- Central task: Determine whether a pair of a text and a hypothesis stands in (textual) entailment relation.

- Training and test material taken from
  - Information Retrieval
  - Information Extraction
  - Summarisation
  - Question Answering

- Typically, hypothesis is manually constructed. Typically, "Text" is one (possibly long) sentence.

- Development set and evaluation set with 800 TH pairs each, balanced w.r.to entailed/ not entailed.

- Development set annotated with Yes/No

- Task: Achieve maximal accuracy on evaluation set
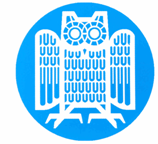
  … using arbitrary information.

Textual Entailment: Example (1)

Text: *The Arabic-language television network Al-Jazeera reports it has received a statement and a videotape from militants.*

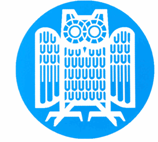Hypothesis: *Al-Jazeera is an Arabic-language television network.*

Textual Entailment: Example (2)

T: *With $549 million in cash as of June 30, Google can easily afford to make amends.*

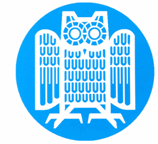H: *Some 30 million shares have been assigned to the company's workers.*

T: *Whooping cough, or pertussis, is a highly contagious bacterial infectioncharacterized by violent coughings, gasp for air that resemble `whoop' sounds, and vomiting*

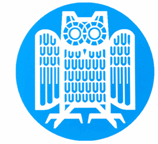H: *Pertussis is not very contagious*

Textual Entailment: Example (3)

T: *Scientists have discovered that drinking tea protects against heart disease by improving the functions of the artery walls*

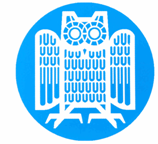H: *Tea protects from some diseases.*

Textual Entailment: Example (3)

T: *Scientists have discovered that drinking tea protects against some diseases by improving the functions of the artery walls*

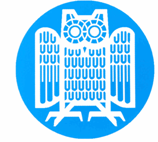H: *Tea protects against heart disease.*

Textual Entailment: Example (5)

T: *Aki Kaurismäki directed his first full-time feature*
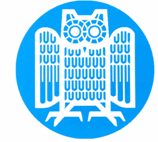
H: *Aki Kaurismäki directed a film*

Textual Entailment: Example (6)

T: *His wife Strida won a seat in parliament after forging an alliance with the main anti-Syrian coalition in the recent election.*
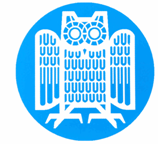
H: *Strida elected to parliament.*

Textual Entailment: Example (7)

T: *Wyniemko, now 54 and living in Rochester Hills, was arrested and tried in 1994 for a rape in Clinton Township.*
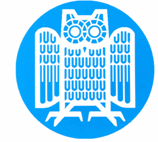
H: *Wyniemko was accused of rape.*

Textual Entailment: Example (8)

T: In 1954, in a gesture of friendship to mark the 300th anniversary of Ukrainian union with Russia, Soviet Premier Nikita Khrushchev gave Crimea to Ukraine.
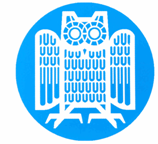
H: Crimea became part of Ukraine in 1954.

Textual Entailment: Example (9)

T: *Security authorities have declared a state of maximum emergency in Guatemala, which is located directly in the path of the hurricane.*

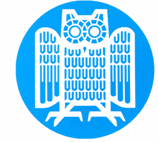H: *There is a state of maximum emergency in Guatemala because of the hurricane.*

T: *As a real native Detroiter, I want to remind everyone that Madonna is from Bay City, Mich., a nice place in the thumb of the state's lower peninsula.*

H: *Madonna was born in Bay City, Mich.*
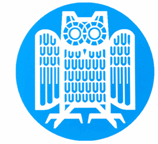
T: *For hundreds of family members holed up in a hotel not far from the Stockholm ferry terminal, where the boat Estonia was to have docked Wednesday morning after an overnight trip from the Estonian capital of Tallinn, the suspended rescue effort was just the latest in a day of heartbreaking reports.*

H: *The Estonia set sail from Tallinn.*

Textual Entailment: Method
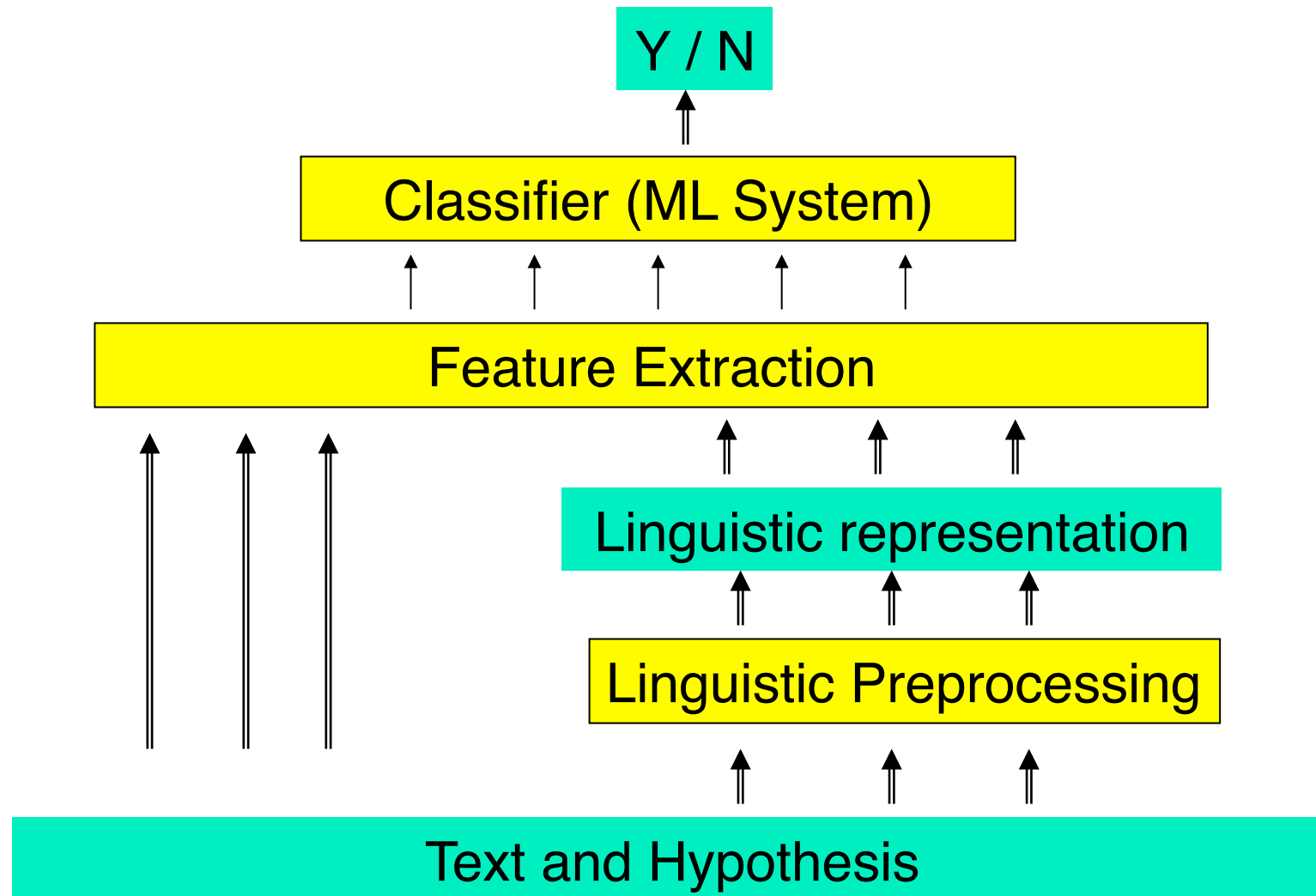
Textual Entailment: Results

- Using completely shallow methods only (e.g., word overlap) gives accuracy on the 60% level.

- To get significantly beyond 60%, you have to add knowledge of some kind.

Options:

- Extraction of lexical-semantic information (paraphrases, inference relations) from corpora (and contextualising this information) → SIM

- Use and enhancement of large existing lexical-semantic knowledge bases (WordNet, FrameNet) → LEX

- Using aspects of logical meaning structure. → LOG

Inference and Semantic Similarity

- Automatic Acquisition of Paraphrases: DIRT and TEASE (Lin&Pantel 2001, Szpektor et al. 2004)

- Context and Word-Sense specific Inference (Mitchell&Lapata 2008, Erk&Pado 2008,Dinu&Wang 2009, Wang&Neumann 2007)

- Inference as Directional Similarity (Bhagat et al. 2007, Geffet&Dagan 2005, Weeds&Weir 2003, Pantel et al. 2007)

Using lexical–semantic resources

- Lexical knowledge for inference (Moldovan et al. 2003, Tatu et al. 2006, Clark et al. 2008)

- Frame-semantic information for textual inference (Burchardt et al. 2005, <u>Burchardt&Frank 2006</u>, Burchardt et al. 2008)

Adapting Logical Approaches

- Computing and exploiting polarity information (Nairn et al. 2006)

- Natural Logic (<u>MacCartney&Manning 2007</u>, MacCartney&Manning 2009)

- Abductive Reasoning (<u>Hobbs et al. 1988</u>, Hobbs et al. 1993, Raina et al. 2005)

- DRT-Based Reasoning (Bos 2001, <u>Bos&Markert 2006</u>)