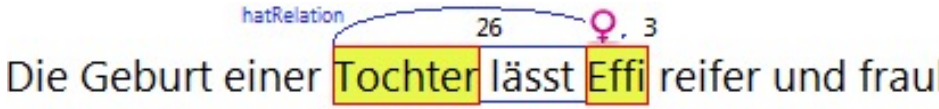
Information Extraction

Relation Extraction

Relation Detection and Classification

- What is “Relation Detection”
  - This task determines, whether (usually) two entities are part of a relation (e.g. Family relation)



The birth of a daughter makes Effi more mature and womanly;

  - Or not:

the letters that Crampas wrote to Effi in Kessin.

➔ Binary classification

Relation Detection and Classification

- What is “Relation classification”
  - This task determines the label of a previously detected relation
  - Not used consistently in the literature

  - Multi-class classification

Some clarification

- In the literature you might find:

- Relation classification
- Relation detection

+

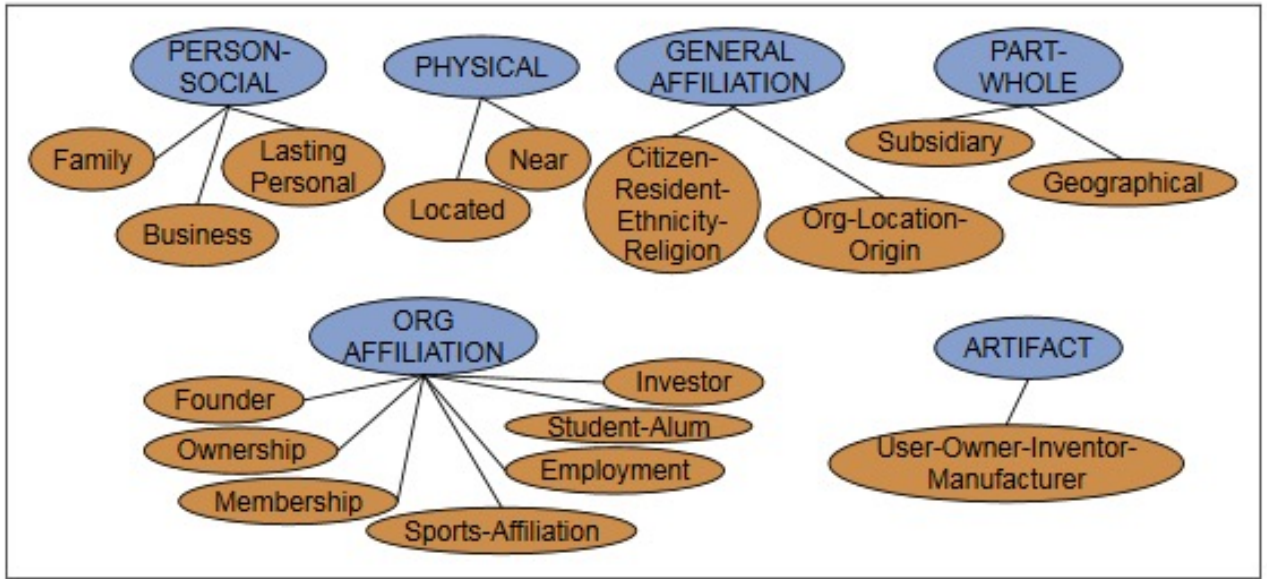
- Relation extraction

=

- In this lecture, the term „extraction“ refers to the sum of the other two, that is to detect, whether entities are in a relation and if so also the type of relation!

Relation Classification –Label sets

- The most well known label set originate the ACE conferences and comprises 17 relations:



- Annotated on an English data set with about 9600 annotated relations

Relation Classification –Label sets

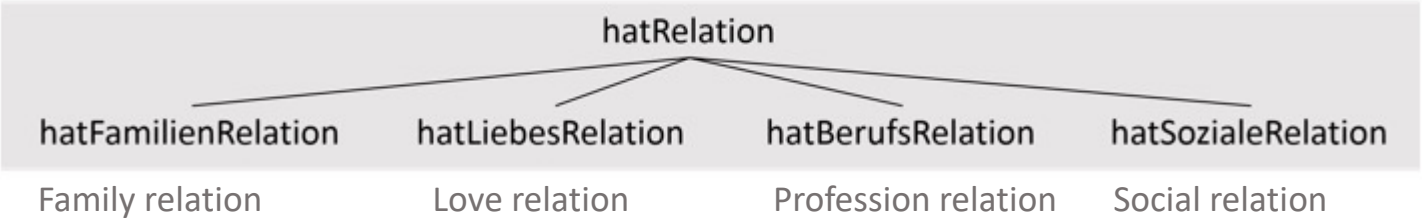
- Medical Relations, as defined by the UMLS

isa
associated_with
  physically_related_to
    part_of
    consists_of
    contains
    connected_to
    interconnects
    branch_of
    tributary_of
    ingredient_of
  spatially_related_to
    location_of
    adjacent_to
    surrounds
    traverses
  functionally_related_to
    affects
      manages
      treats
      disrupts
      complicates
      interacts_with
      prevents
    brings_about
      produces
      causes

[associated_with] (continued)
  [functionally_related_to] (continued)
    performs
      carries_out
      exhibits
      practices
    occurs_in
      process_of
    uses
    manifestation_of
    indicates
    result_of
  temporally_related_to
    co-occurs_with
    precedes
  conceptually_related_to
    evaluation_of
    degree_of
    analyzes
      assesses_effect_of
    measurement_of
    measures
    diagnoses
    property_of
    derivative_of
    developmental_form_of
    method_of
    conceptual_part_of
    issue_in

Relation Classification –Label sets

- Relations for characters of the “Kallimachos” project



- 4 main categories, with in total 57 sub categories
- Annotated on 213 summaries of novels
- And about 2500 sentences from novels

Relation Classification -Techniques

- The techniques for relation extraction can be grouped into four categories:

1. Hand-written Patterns (rule-based approach)
2. Supervised Learning (classical Machine Learning)
   - also Kernel-Based Learning
3. Semi-Supervised Learning
4. Unsupervised Learning

➔ This lecture will mainly deal with the classical supervised setting

Relation Classification – Rule-Based

- Makes use of. “Lexico-Syntactic Expressions”
- Idea of Hearst in 1992, also called “Hearst Patterns”
- Example:

Agar is a substance prepared from a mixture of red algae, such as Gelidium, for laboratory or industrial use.

$$NP_0 \textit{ such as } NP_1\{,NP_2\ldots,(and|or)NP_i\}, i \geq 1$$

implies the following semantics

$$\forall NP_i, i \geq 1, \text{hyponym}(NP_i, NP_0)$$

allowing us to infer

$$\text{hyponym}(\text{Gelidium}, \text{red algae})$$

Relation Classification – Rule-Based

- If NER labels are available, rules can be formulated more precisely:

PER (named|appointed|chose|etc.) PER Prep? POSITION
Truman appointed Marshall Secretary of State

- This approach is a lot of work and usually results in:
  - Many rules with high Precision, but
  - Low overall Recall

Relation Classification –Supervised Learning

- If a data set with annotated relations is available, we can learn a classifier as follows:

1. Mark all relevant entities in the text fragment (usually a sentence)

Im Hause des **[Ritterschaftsrats von Briest]** auf Hohen-Cremmen hält ein **[Jugendfreund]** **[Frau von Briests]**, **[Baron von Innstetten]**, um die Hand der **[Tochter des Hauses]** an.

In the house of the [Ritterschaftsrats von Briest] at Hohen-Cremmen, a [childhood friend] of [Mrs. von Briest], [baron von Innstetten], asks for the hand of the [daughter of the house].

Relation Classification –Supervised Learning

2. Extract all pair of entities:

Ritterschaftsrat von Briest → **Jugendfreund** — **Label: O**

Jugendfreund → **Frau von Briest** — **Label: hasFriend**

- When we create the instances, we can either:
  - Ignore the order of the entities (would result in 10 instances)
  - Create pairs with entities in both positions (would result in 20 entities)
    - This could capture inverse relations (“hasDaughter” vs “hasMother”)

- The label is just the label of the relation

Relation Classification –Supervised Learning

3. Extract appropriate features for each pairing

- The head of an entity
- All words of the instance
- Words prior and after the instance
- Words in between „heads“
- N-Grams in between „heads“
- Lemmas, POS-Tags
- Entity-types
- Amount entities in between
- Phrases in between
- Path in the constituency tree
- Path in the dependency tree

Tochter des Hauses

Relation Classification –Supervised Learning

4. Application is no different than training

Relation Classification - Extensions

- Calculate more distinctive features using „Kernel machines“:
  - Easy to integrate in the Perceptron or the Support Vector Machine
- Enhance your train data using „unlabelled“ or „weakly-labelled“ data
  - This leads to „Distant Supervision“ or „Data Programming“
- Make use of the hierarchical structure of the label set!
  - We have more data to detect a general family relation, that just the relation „hasDaughter“

➔ This leads us into Non-structural hierarchical classification