

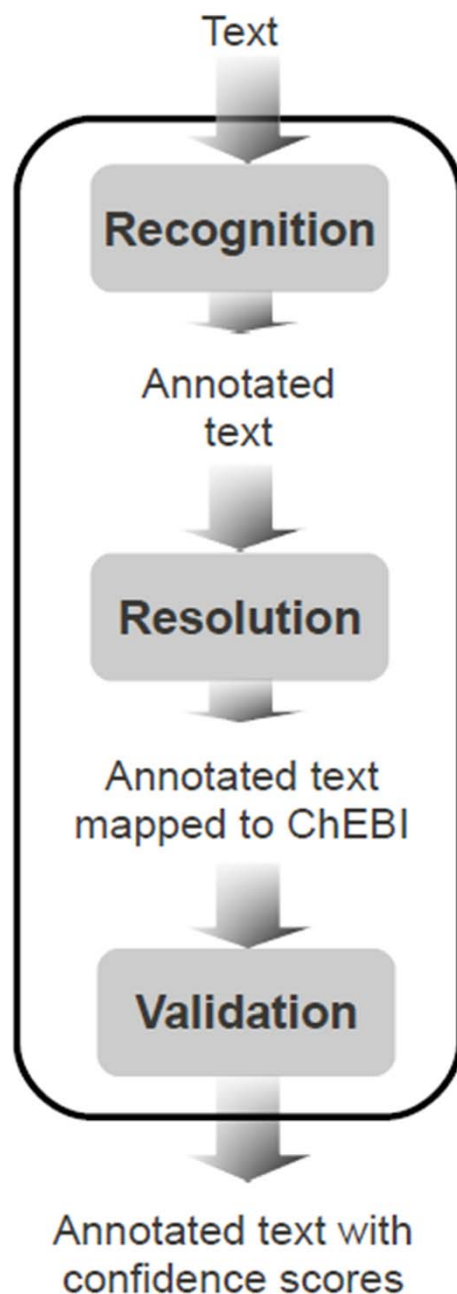


# Identifying Chemical Entities based on ChEBI

Tiago Grego, Francisco Pinto, Francisco Couto

Universidade de Lisboa

July, 2012 at ICBO Graz Austria



- Conditional Random Fields (CRF)
  - ChEBI patents gold standard
- Lexical similarity method
  - based on evidence content
  - Mapping to ChEBI
- Semantic similarity measures

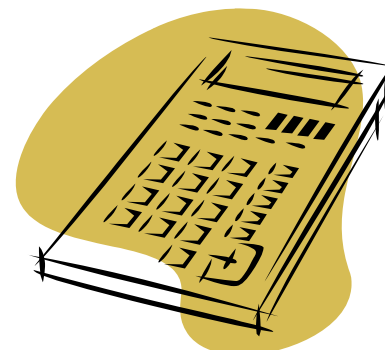
*Grego, T., Pesquita, C., Bastos, H., and Couto, F. (2012). Chemical entity recognition and resolution to ChEBI. ISRN Bioinformatics, 2012.*

# Validation

- Assumption
  - Chemical entities mentioned in a fragment of text are at some extent semantically related
- At least they share
  - the subject of the fragment of text
  - their interest by the authors

# Semantic similarity measures

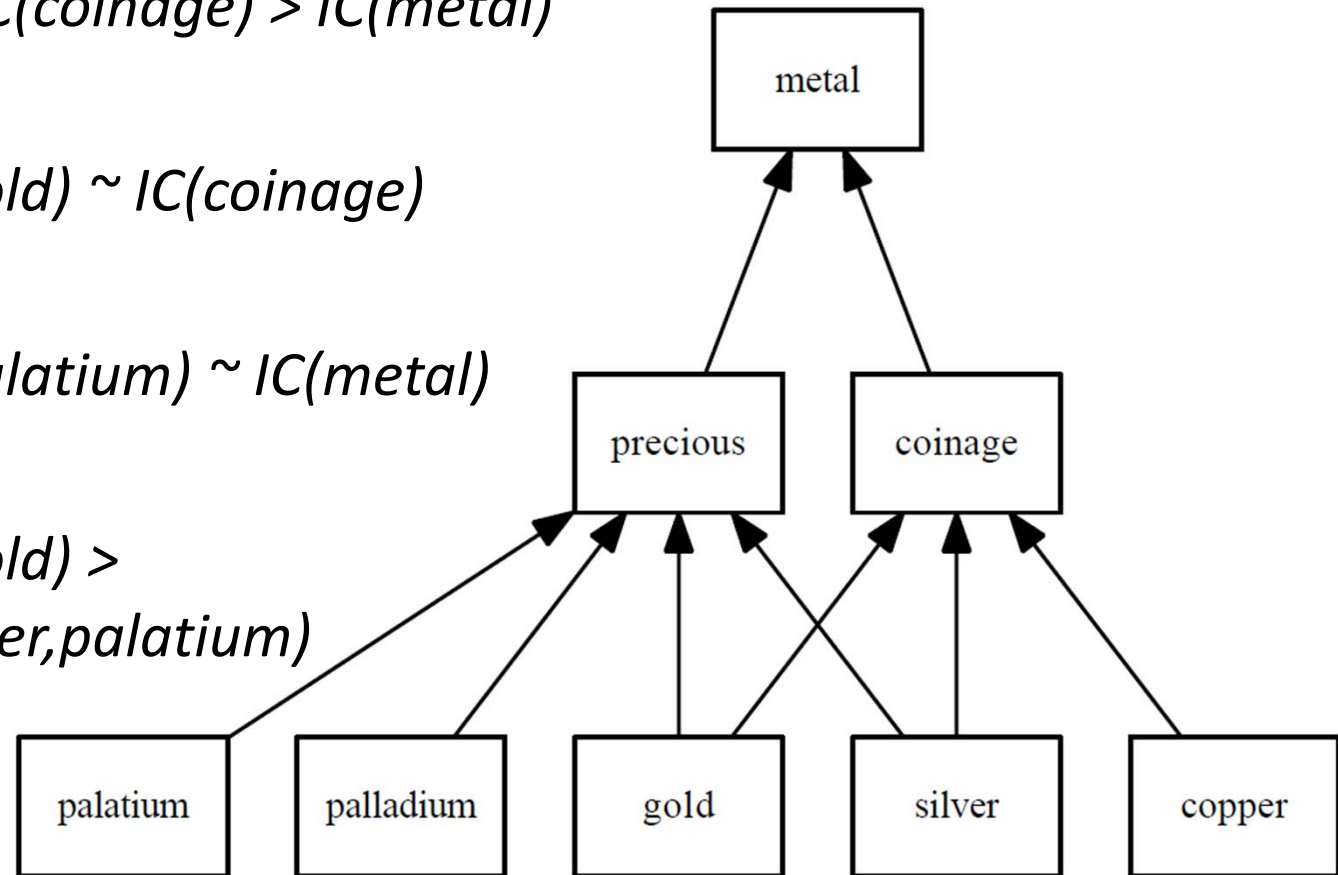
- Input:
  - two ontology concepts
  - or two sets of terms annotating two entities
- Output:
  - a numerical value reflecting the closeness in meaning between them



*J. Ferreira and F. Couto, Semantic similarity for automatic classification of chemical compounds, PLoS Computational Biology, vol. 6, no. 9, p. e1000937, 2010*

# Example

- $IC(\text{copper}) > IC(\text{coinage}) > IC(\text{metal})$
- $Sim(\text{copper}, \text{gold}) \sim IC(\text{coinage})$
- $Sim(\text{copper}, \text{palatium}) \sim IC(\text{metal})$
- $Sim(\text{copper}, \text{gold}) > Sim(\text{copper}, \text{palatium})$



# Demo

<http://lasige.di.fc.ul.pt/webtools/ice/>

<http://www.youtube.com/watch?v=hpxchOi-UwU>

# Identifying Chemical Entities

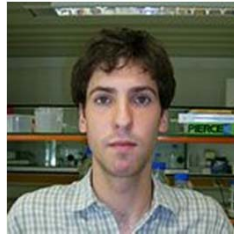
Recognition, Resolution, Validation

A mixture of ethanol, propanol and acetic acid with a small amount of sodium chloride.

Similarity: measure =  threshold =

**Submit**

# Thanks for your attention!



More information:

<http://lasige.di.fc.ul.pt/webtools/ice/>

Support: Portuguese Fundação para a Ciência e Tecnologia through the financial support of the SPNet project (PTDC/EIA-EIA/119119/2010), the SOMER project (PTDC/EIAEIA/119119/2010) and the PhD grant SFRH/BD/36015/2007 and through the LASIGE multi-annual support.