Combining methods to learn feature-norm-like concept descriptions

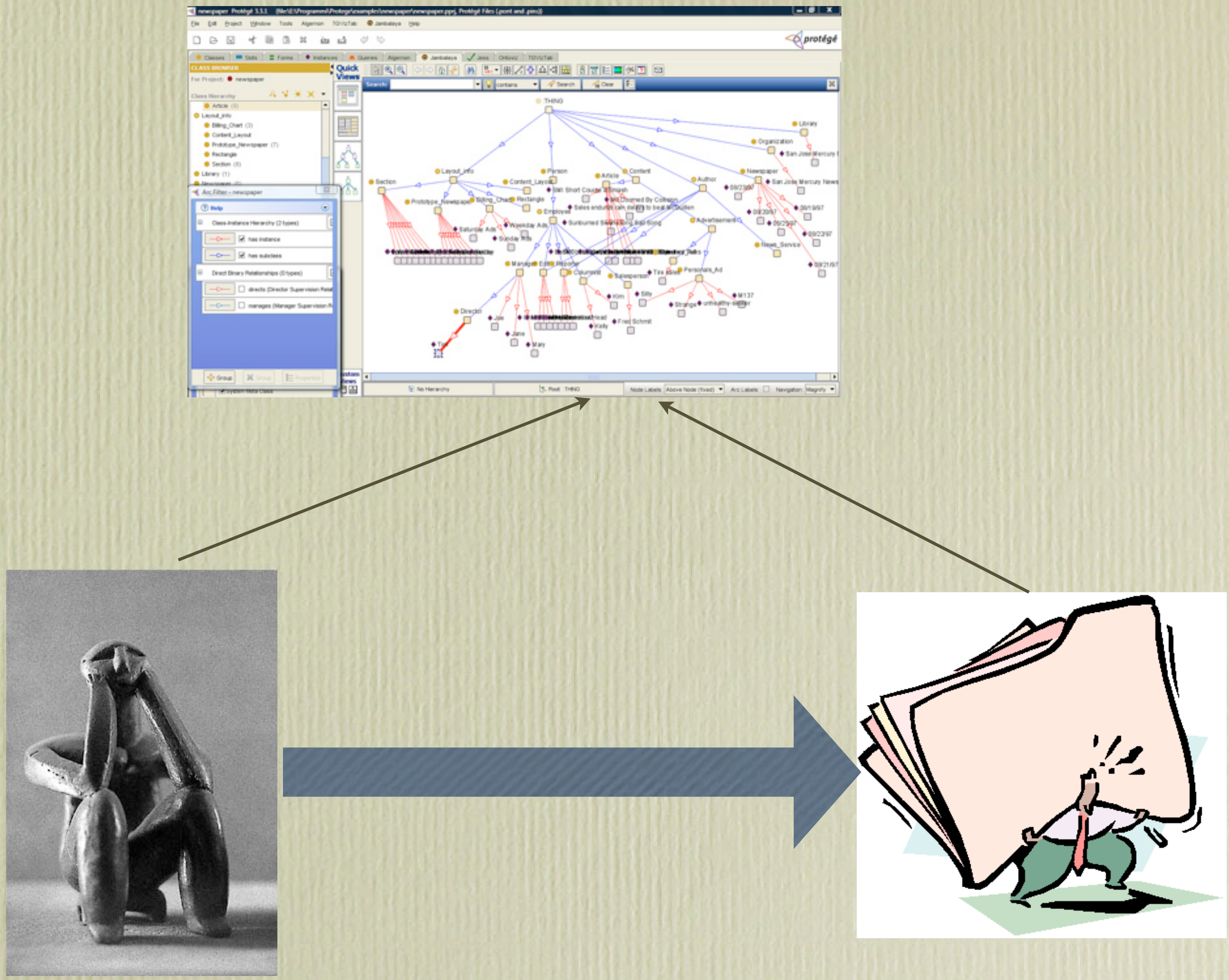
Eduard Barbu
Center for Mind/Brain Sciences
Rovereto, Italy

Talk Outline

- **Introduction**
- **Property Classification**
- **Property Learning**
- **Results and conclusion**

Introduction

- To build an ontology one needs to identify the main concepts and the relations holding for the domain of interest.
- It is easier to formalize a narrow domain like physics or marketing than the domain of common-sense knowledge
- Attempts to formalize the common-sense knowledge domain (Cyc, Open Mind)



protégé
Quick Views
THING
Library
Organization
Section
Layout_info
Person
Article
Content
Author
Newspaper

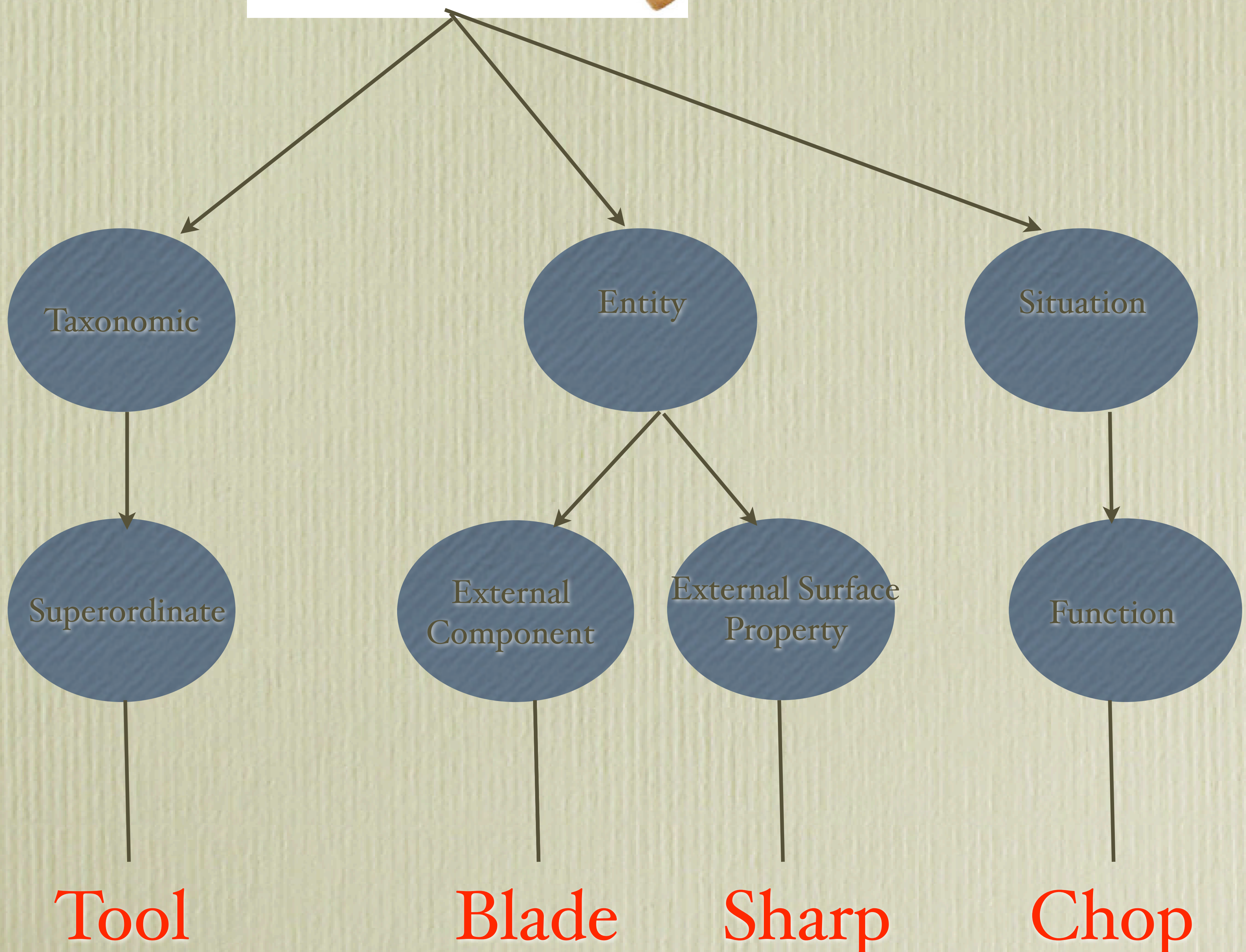
Introduction

- Feature norms can be regarded as repositories of common sense knowledge
  - An **apple** (concept) is a **fruit** (property)
  - An **airplane** (concept) is used for people **transportation** (property)
- Our aim is to classify the features (properties) in the feature norms and then devise shallow methods to learn them.

Property classification

- We choose a feature norm by McRae and colleagues
    - 725 subjects listed features for 541 living and not living basic level concepts
- The norm is annotated with Wu and Barsalou taxonomy. The taxonomy is inspired by principle governing human perception.

Part of the conceptual description for the concept AXE
The features are annotated using WB taxonomy



Property classification

- We cannot directly use this taxonomy in the learning process because some of the distinctions it makes are too fine grained.
- We map the original WB taxonomy on a better set of relations (better from the point of view of learning task)
- As we want to learn the features produced in the feature generation task does not matter which classification schema we use.

Property classification

- We classify the properties at two levels: a semantic level and a morphological level

Semantic	Morphological
Superordinate	Noun
Part	Noun
Stuff	Noun
Location	Noun
Action	Verb
Quality	Adjective

Property classification

- **Superordinate properties**. Classification from a taxonomic point of view
    - Example: The **dog** is an **animal**
- **Part.** This category collapses the internal and external components in WB taxonomy
    - Example: **Axe** has as part **blade**
- **Stuff.** The substance an object is made from.
    - Example: **Bottle** is made of **glass**

Property classification

- **Location properties**. Typical places where the instances of the concept are found
  - Example: The **airplane** is found in **airport**
- **Action.** The characteristic actions defining the behavior of an entity or the function instances of the concepts typically fulfill.
  - Example: **The cat meow**
  - Example: **The heart pumps blood**
- **Quality.** This class of properties denote the qualities (color, taste) of the objects instances of concepts.
  - Example: **Apple** is **red**
  - Example: **Apple** is **sweet**

Property Learning

- For Learning the a Feature Norm like structure we employ two different strategies.
    - Superordinate, Part, Stuff and Location properties are learnt using a pattern based approach.
    - Quality and Action properties are learnt using a novel method that quantifies the strength of association between the nouns representing the focal concepts and the adjectives and verbs co-occurring with them in a corpus.
- The best patterns are identified using a semi-automatic procedure. For the automatic part of our procedure we automatize the framework introduced by Hearst

Pattern induction. For each relation of interest we collect a set of instances, extract the sentences containing the instances from a corpus and identify their commonalities.

The upper portion of a typical reactor vessel is shown in the form of the bottle made of glass, for example.

The upper portion of a typical reactor vessel is shown in the form of the N made of N, for example.

The bottom part of the ship made of steel sank immediately

The bottom part of the N made of N sank immediately

of the N made of N	100
...	

Property Learning

- **Pattern Ranking and selection.** We pursue the hypothesis that the most salient patterns are those highly associated with the instances. We use (and benchmark) 4 association measures at this task.
  - *Frequency* .
  - *(Pointwise) mutual information*
  - *Chi-squared (with Yates continuity correction)*
  - *Log-Likelihood*

	p	~p	Row sum
i	O_{11}	O_{12}	$R_1 = O_{11}+O_{12}$
~i	O_{21}	O_{22}	$R_2 = O_{21}+O_{22}$
Column sum	$C_1 = O_{11}+O_{21}$	$C_2 = O_{12}+O_{22}$	$N = R_1+R_2$

The contingency table

Simple Frequency

$$O_{11}$$

(Pointwise) Mutual Information

$$MI^2 = \log_2 \frac{O_{11}^2}{\frac{R_1 \cdot C_1}{N}}$$

Chi-Squared (with Yates continuity correction)

$$chi_{corr} = \frac{N(\left|O_{11} \cdot O_{22} - O_{12} \cdot O_{21}\right| - \frac{N}{2})^2}{R_1 \cdot R_2 \cdot C_1 \cdot C_2}$$

Log Likelihood

$$\log - likelihood = 2 \cdot \sum_{ij} \log \frac{O_{ij}}{\frac{R_i \cdot C_j}{N}}$$

Property Learning

- Once the strength of association between patterns and instances is quantified we compute the sum of association score with the instances in the training set. We then pick for evaluation the best 2 patterns and evaluate them.
- Evaluation. (For the First two patterns Selected By Each association measure and for the Manually selected Patterns). A set of 50 concept Feature (CF) pairs is extracted from a corpus using each pattern and pattern precision is evaluated.

Property Learning

- We noticed that the Quality Properties are in general expressed by the adjectives modifying the noun representing the focal concept
  - Example: She took the **red** **apple**
- The Action properties are expressed by verbs.
  - Example:The ugly **dog** is **barking**
- We hypothesize that the properties of type Quality and Action are the adjectives and verbs highly associated with the focal concepts in a corpus.
- To compute this association strength we use the same measures introduced before

Results and Conclusion

- Corpora used for the experiments are the following:
  - British National Corpus (BNC)
  - ukWaC a 2 billion words corpus constructed by crawling the web
- For evaluating the success of our method we used a set of 44 from McRae feature norm (the same set used for property evaluation task)
- Software: CWB and USC toolkits

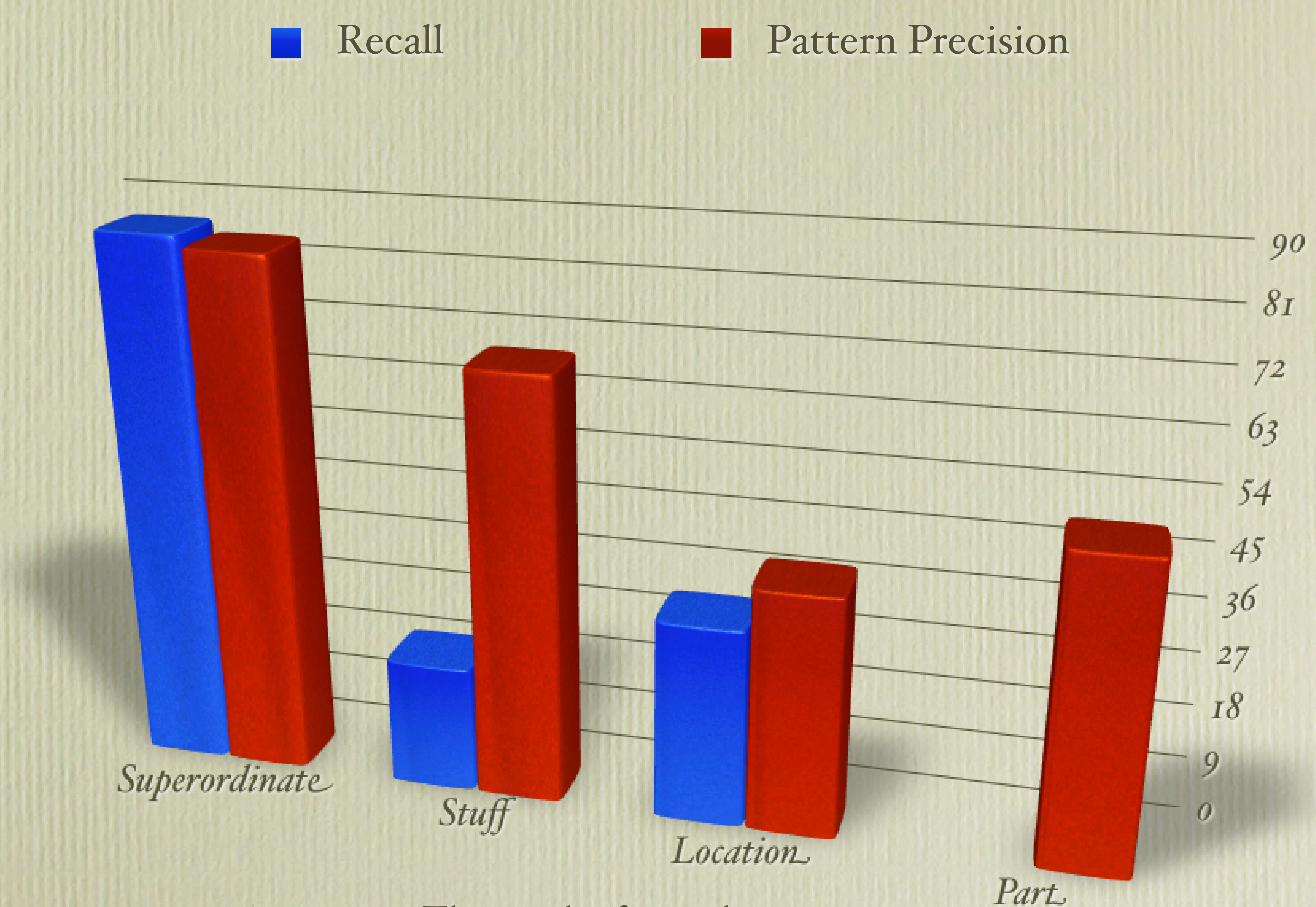
Learning Superordinate, Stuff, Location and Part Properties

- We evaluate
  - the success of each association measure in finding good patterns
  - the success of the manually selected patterns in extracting good properties
- The input of the learning algorithm 200 instances (per relation)
- The pattern learning algorithm is run on BNC
- The voted patterns are then compared with the best patterns.

Relation	Pattern
Superordinate	N {JJ}-such {IN} of N N {CC}-and {JJ}-other N N {CC}-or {JJ}-other N
Stuff	N {VVN}-make {IN}-of N
Location	N {IN}-from {DT}-the N
Part	N {VVP}-comprise N N {VVP}-consists {IN}-of N

The manually selected patterns

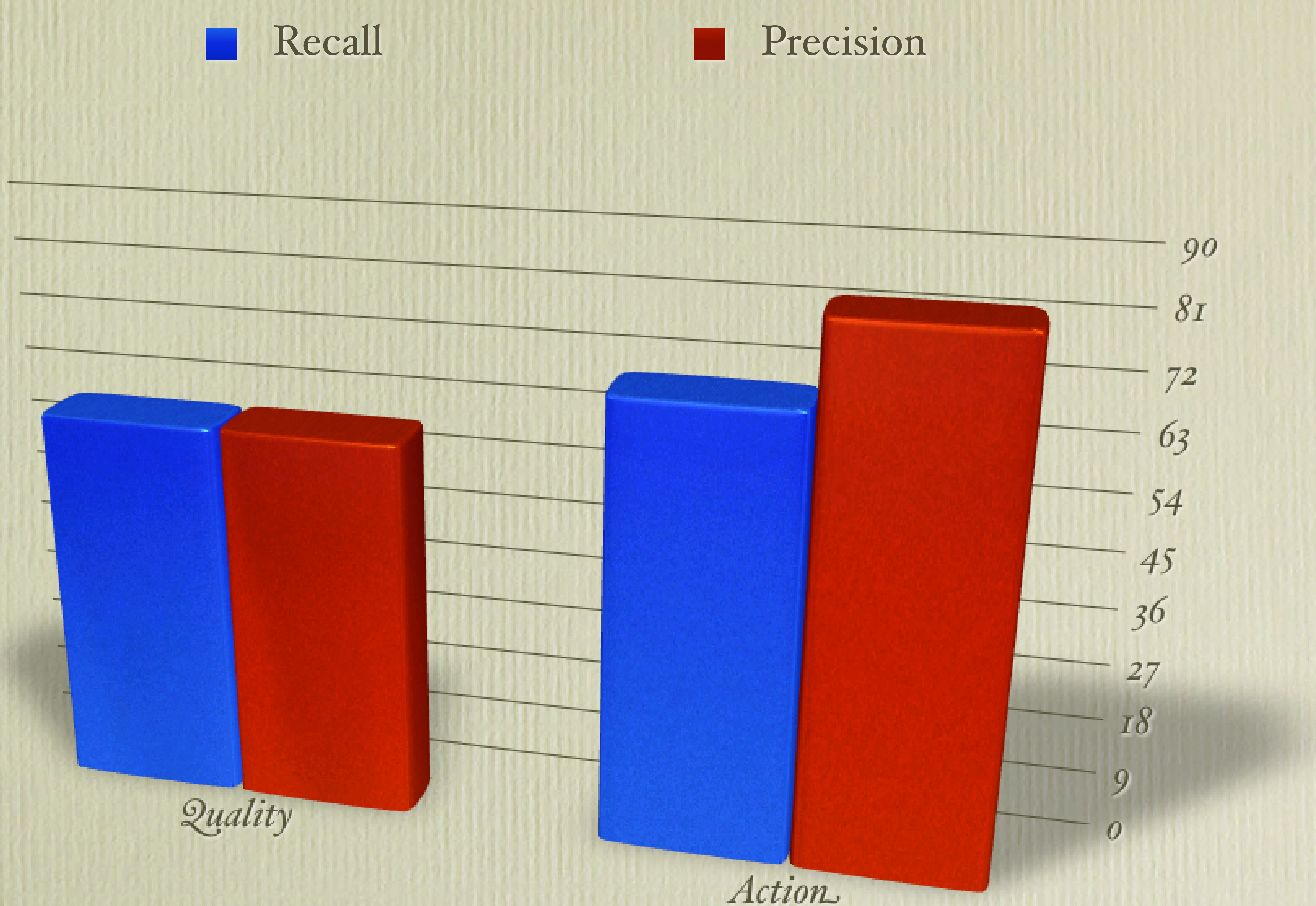
- **Superordinate patterns**
  - The manually selected patterns are voted by any association measure
- **Stuff patterns**
  - Simple frequency does not rank higher the pattern manually selected
- **Location patterns**
  - Chi-squared fails to vote the manually selected patterns
- **Part patterns**
  - None of the measures vote the patterns selected to represent the part relation



The results for each property class

- The patterns for **Superordinate Relation** blur the type -role distinction
  - Example: banana (fruit) vs (product)
- The pattern for **Stuff Relation** is sometimes used metaphorically
  - Example: car made of cheese
- A better pattern for **Location Relation** is N is found in N, however this pattern has 0 recall for the test set.

- The algorithm is run on ukWaC
- Computation of the association strength between the concepts in the test set and the co-occurring verbs and adjectives.
- The best recall for the test set was obtained by the log-likelihood measure
-



The results for Quality and Action Properties

- Many of the adjectives modifying nouns denoting concrete objects express the object's qualities, whereas the verbs usually denote actions different actors perform or to which various objects are subject.

- There are cases when the found properties are excellent candidates for the semantic properties of concepts.

  - Example: Quality. turtle {green, hard, small} vs {**marine**, green, giant}

  - Example: Action. {lays eggs, swims, walks slowly} vs {**dive**, nest, hatch}

Conclusions

- Superordinate, Quality and Action property can be learn with minimum effort
- To learn all other property classes a supervised method should be devised
- As in the case of ontology learning or qualia structure acquisition it seems that the best way to acquire feature norm like structure it is a semi-automatic one.