PloneOntology
Collaborative Keywordmanagement

Thomas Förster

t.foerster@biologie.hu-berlin.de

Institute for Theoretical Biology

Humboldt–University Berlin

1. Keyword Management Problems

2. The Approach of PloneOntology

3. How to Use it

Keyword Managing Problems

Thomas Förster, Vienna, September 2005

Classification is about relating things to concepts, not words!

"Zope" - one word, (at least) two concepts

- A highly recognized and reputable fish *(Abramis ballerus)*

- A web application server striving for world domination

A moderately large fish, the Zope has a life span of 18 years. It grows slowly, ...

http://earth.leeds.ac.uk/ jackson/shoal/z.html

A collection of keywords from scientific papers all containing the phrase **"visual cortex"**:

- Visual system

- V1

- Visual cortex

- Vision

- Primary visual cortex

The same applies for abbreviations, translations, ...

- Principal component analysis

- PCA

- Hauptkomponentenanalyse

- (add translation to your favourite language here)

Without a semantic understanding, there's no way to match identical or similar concepts.

Unsupervised learning algorithms cluster content into groups, but don't tell what the cluster is about.

For supervised learning you need a (preferably large) training set

In any case you'll probably end having natural language processing on your todo list.

The needed semantic understanding is provided in a so called **ontology**.

An ontology is a specification of a conceptualization.

Tom Gruber, Stanford

You can think of it as a graph with concepts as the nodes.
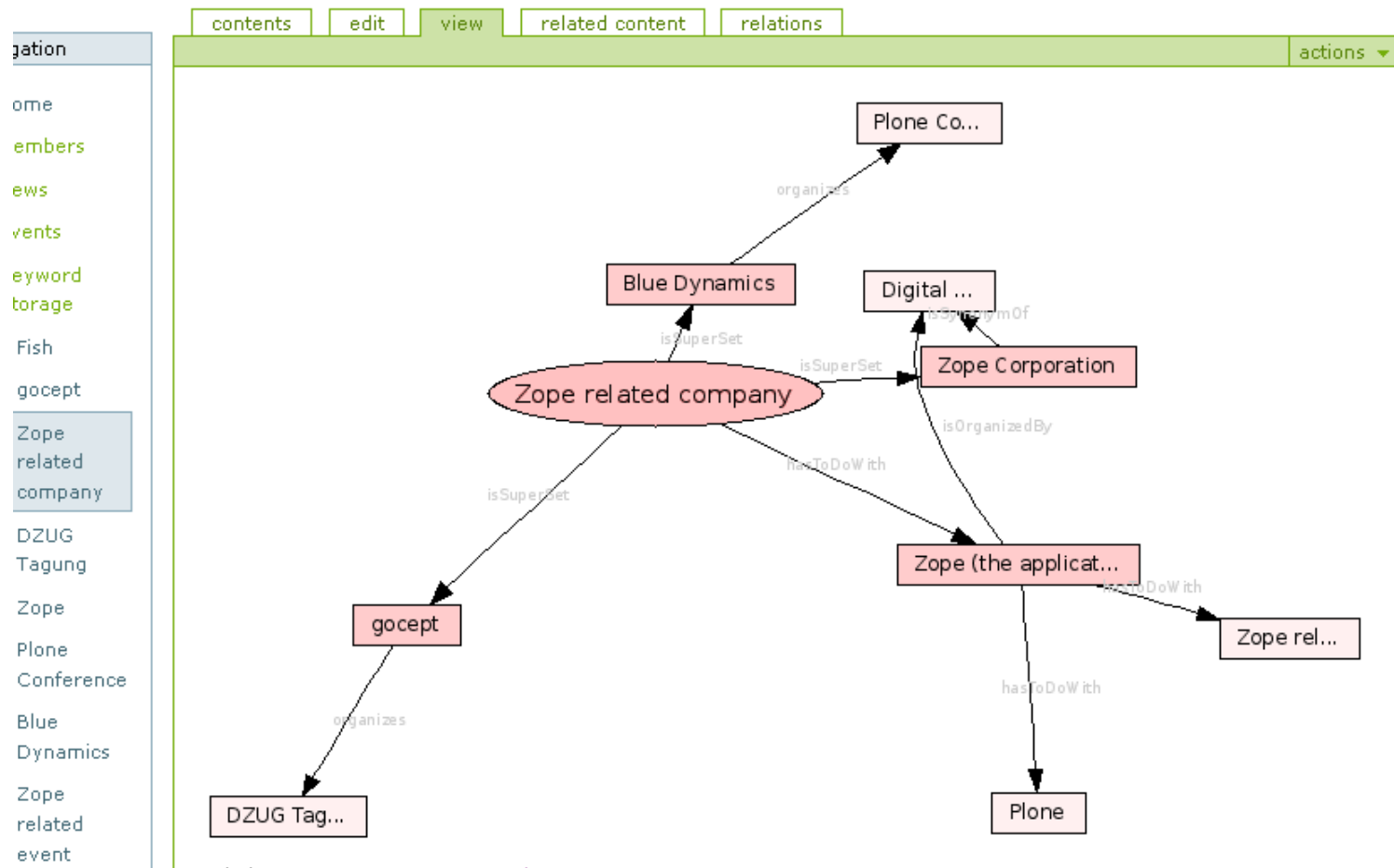
Ontologies are recognized as a crucial part of the *Semantic Web*.

Fortunately Tim Berners Lee is a big fan of Python.

PloneOntology's Approach

Provide a vizualization for the local ontology

Keywords are content (2)
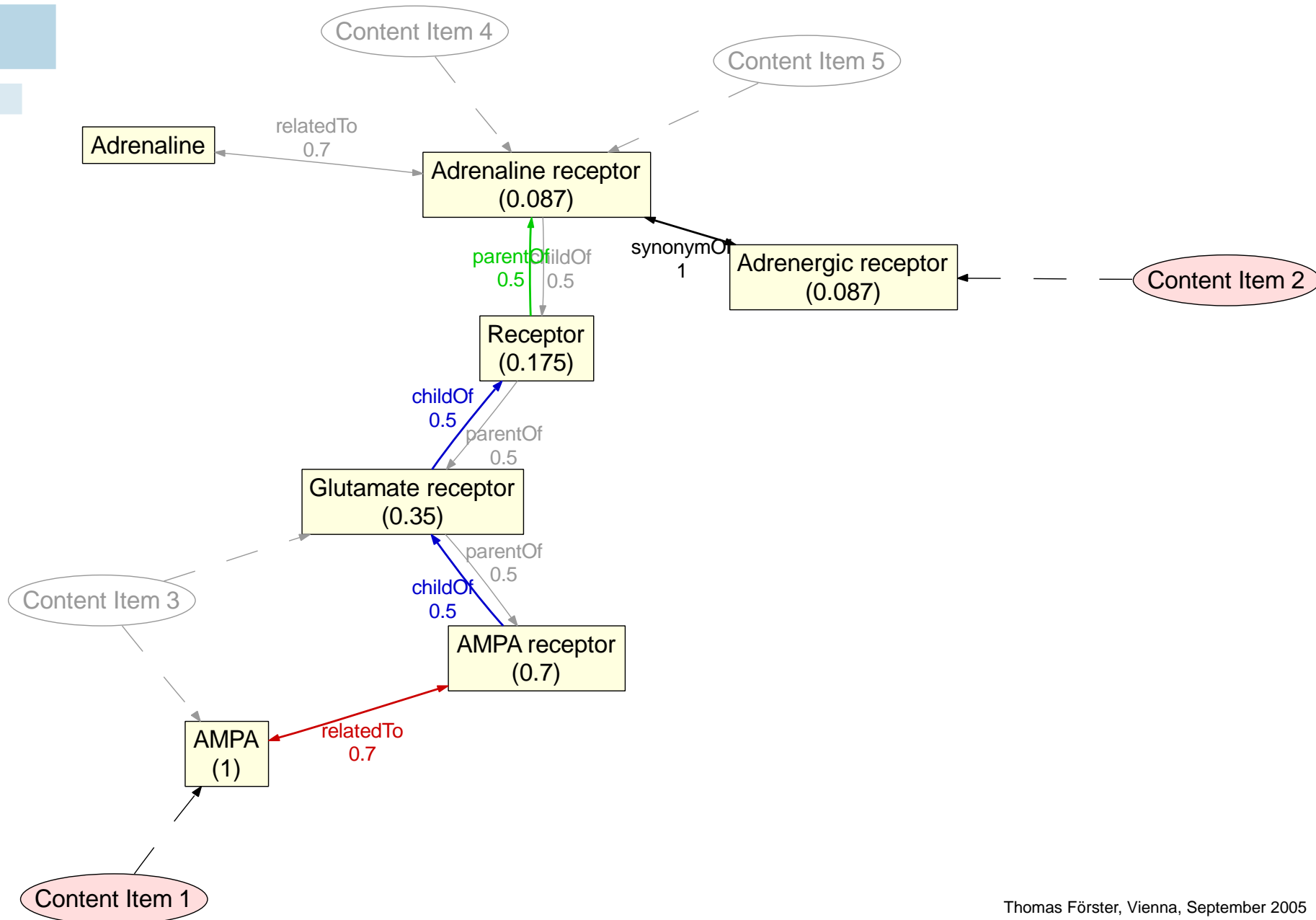
Provide links to content related to the keyword

Using 'Relations' for the references

The classification is just keywords referenced by content objects.

- 'Relations' allows putting constraints and attributes to the graph edges

- Inverse implicators are great for e.g. parent/child relationships

A simple relevancy ranking search

Usage

Classification (user side)

- seperate step during content creation/edit

- optional

- Sarissa provides search comfort

Classification (developer side)

- works only with Archetypes

- need to include some fields into content type

- already provided as ready to use Schema to include

```
from Products.PloneOntology.ClassificationSchema
                    import ClassificationSchema


myschema = BaseSchema + ClassificationSchema + Schema((...
```

Ontology Management via Proposals

- New terms and connections start out as a proposal, pushed through a reviewing workflow.

- On acception proposals are converted to Keywords resp. references between them.

- This allows to model draft ontologies before integration into main ontology.

PloneOntology has OWL (Web Ontology Language) import/export support.

Links

Zope fish http://earth.leeds.ac.uk/ jackson/shoal/z.html

Ontology Intro http://www-ksl.stanford.edu/kst/what-is-an-ontology.html

Semantic Web http://www.w3.org/2001/sw/

Semantic Web Intro http://infomesh.net/2001/swintro/

OWL Specification http://www.w3.org/2004/OWL/

.owl Ontology collection http://www.schemaweb.info/default.aspx

Relations http://svn.plone.org/view/archetypes/Relations/

Stefan Kröger → Proposals, Sarissa Integration and GraphViz
   support

David Baehrens → OWL Import/Export and Relations backend

And of course you for listening...