

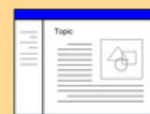
E-Course Authoring

Method for Automated Metadata Generation

Marián Šimko supervised by Professor Mária Bielíková



complicated boring difficult manual knowledge discovery lack of flexibility tiring non-trivial enormous load latent implicit connections complex exhausting frustrating time-consuming arduous



Learning objects
preprocessing

1 Text analysis

lexical analysis
stop words removal
lemmatization

2

Vector representation composition

3

Vector adjustment

index keyword processing
formatting processing

4

Relevant keyword selection

Learning objects - topics weights computation

5

tf-idf metric

6

Inter-topic relatedness computation

7

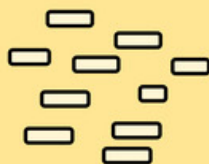
Top-K neighbors selection

Appropriate relations construction

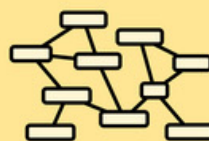
8



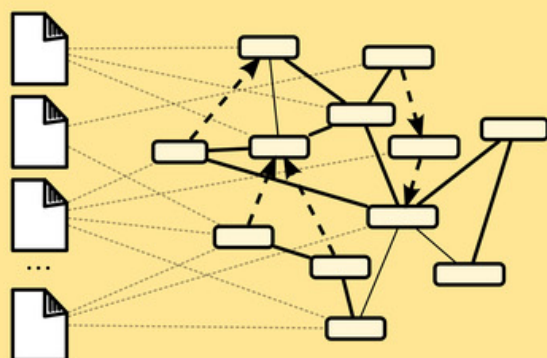
Topics
generation



Relations
generation



Semi-structured domain model



Learning Objects

Topics

Vector approach

- topic vector - weighted sum of learning object's term vector representation

$$\vec{t}_i = \sum_j w_{ij} \vec{o}_j$$

- relatedness computed using cosine similarity metric

A

Proposed variants

Spreading activation

- topics, learning objects
~ contextual network
- relatedness derived from each topic's accumulated energy ratio

$$asim_{ij} = \frac{E_j}{\sum_k E_k} \log(d_{ij})$$

B

PageRank-based analysis

- topic space similar to web space
- related neighbors search viewed as a Markov process
- relative relatedness determined implicitly as a prestige of node

C

Combining variants

- strengthening PageRank* starting point
- involving vector approach and spreading activation to build temporary metagraph

D

E

Evaluation

- programming learning domain: functional programming course
- 71 learning objects, 84 topics, 80,3% lemmatization success
- expected relations count: 234

Variant	A	B	C	D	E
Correctly generated relations*	119	98	134	141	131
Ratio [%]:	50,9	41,9	57,3	60,3	56,0

* At the time of the publication of the poster it has not been possible to conduct a complete evaluation of the proposed method. A more accurate evaluation matching the specific needs of an e-learning environment domain will be performed after the completion of the research.