

Extracting Dependency Relation from Digital Learning Content

Giovanni Adorni, Felice Dell'Orletta, Frosina Koceva,
Ilaria Torre and Giulia Venturi

Frosina Koceva - frosina.koceva@edu.unige.it

Department of Informatics, Bioengineering, Robotics, and System Engineering

University of Genova

Summary

- Objectives
- Research Problem
- State of Art
- Solution
- Conclusions

Introduction & Objectives

- Digital Libraries evolution
- Digital Libraries & Education
- Digital Content → Structured Content

Research Question

- Automatic learning of the knowledge structure of the content
 - Relevant Concept Extraction
 - Relation Identification → Dependency (prerequisite) relation

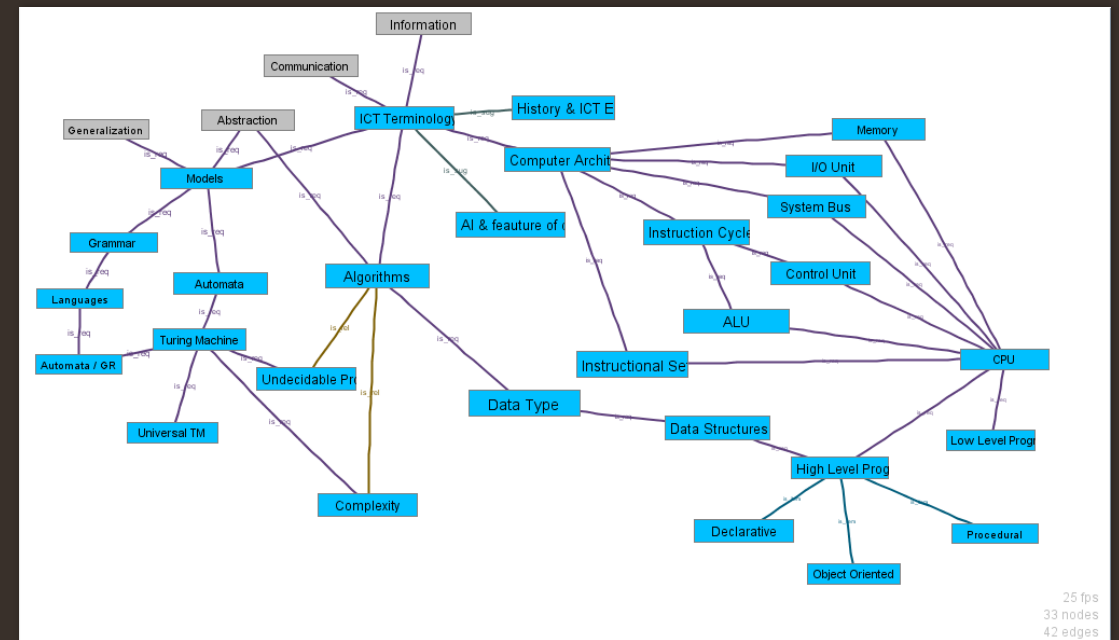
Prerequisite Relation & ECM

Prerequisite Relation

The prerequisite relation between two concepts A and B is a dependency relation which represents what a learner must know/study (concept A), before approaching concept B.



Knowledge Graph



State of the Art

- Exploit Knowledge Source:
 - Textbook (digital content): Table of Content; Analytical Index
 - Wikipedia, DBpedia, WordNet, MOOC course structure, ...
- NLP-based
- Statistical and Machine Learning Based:
 - Cross-Entropy; Jaccard Similarity
 - Unsupervised learning: agglomerative clustering methods for identifying hierarchical relations, etc.
 - Supervised learning: Naïve Bayes, SVM, etc.

Assumptions

- Co-occurrence of two concepts is necessary but not sufficient condition to identify the prerequisite relation.
- Temporal occurrence of terms and/or sections are taken into account to identify the direction of the prerequisite relation.

Candidate Prerequisite Relations

Method 1

Based on temporal order and co-occurrence of concept

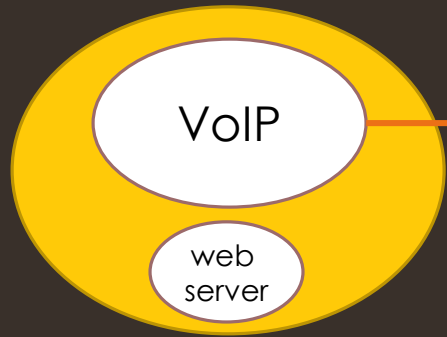
1. Order the concepts according to their first significant appearance in the text
2. Identify co-occurrence relation between concepts, and create a undirected graph
3. Assign direction of the arcs based on the precedence order in step 1

Method 2

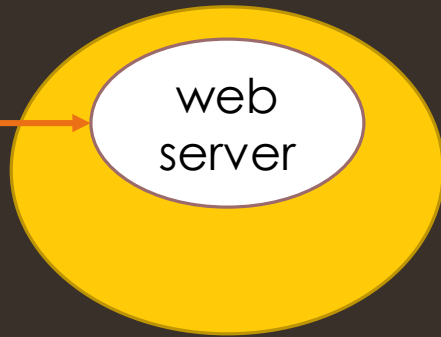
Based on Table of Content and relevance of a concept.

- For each concept c_u identify the section s_i in which c_u is relevant
- For each concept c_v identify the section s_j in which c_v is relevant
 - Identify one of the four following cases (see next slide)

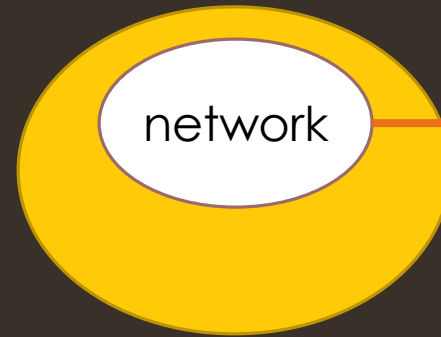
Subchapter 2. The Internet



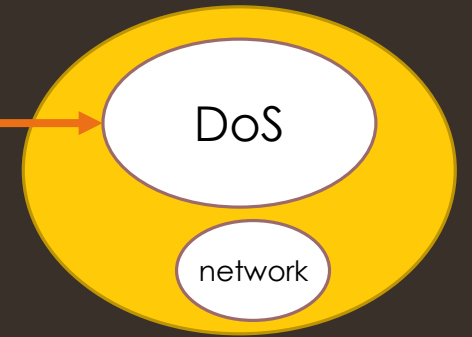
Subchapter 3. World Wide Web



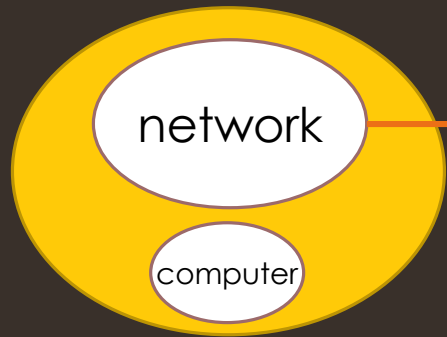
Subchapter 1. Network Fundamentals



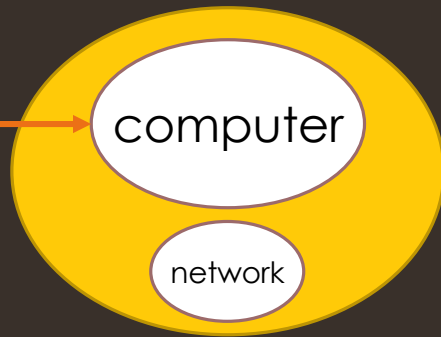
Subchapter 5. Security



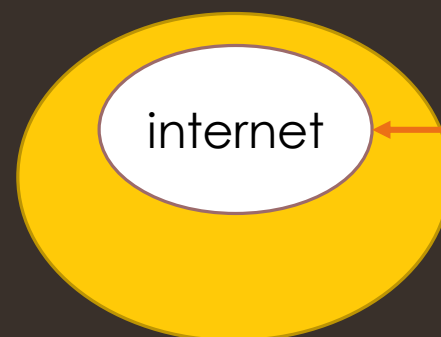
Subchapter 1. Network Fundamentals



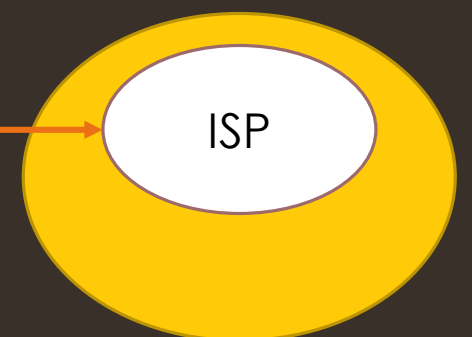
Subchapter 5. Security



Subchapter 2. The Internet



Subchapter 2. The Internet



Feature Improvements

- Evaluation
- Method 1: Burst Analysis instead of co-occurrence
- Reduce the biases due to author's subjective choice in structuring the book

Conclusions

- There is no canonical way of representations of knowledge structure
 - The prerequisite relation depends on the hole course structure
- Building Prerequisite Dependency structure is important for:
 - Processing knowledge → Creating a good sequence of teaching activities → Creating Learning Paths

Acknowledgments

The authors thank prof Carlo Tasso for making available Distiller System during the initial experiments of the described methodology

Thank you for your attention