Introduction to Scientific Working

Cezary Kaliszyk

Summary of last PS

1. **Meta-cognition**
   Reflect on your own learning process
2. **Syntactic-semantic analysis**
   Clarify/Understand the used terms
3. **Reduction**
   Reduce the text to its main statements
4. **Reconstruction**
   Reconstruct the main features of the text using non-verbal methods
5. **Elaboration**
   Confront the text in a critical way

Lecture Content

Research and Understanding

Understanding and summarizing of scientific text, Literature research, Internet search, Citing, Practical scientific work

Structuring Scientific Works

Kinds: Seminar, Bachelor and Master theses, Topic analysis and structuring

LaTeX

Interaction, Typesetting of text, Images/Diagrams, Mathematical formulae, Lists, Tables, Fonts, Special cases

Evaluation, Checking and Presentation

Evaluation of work of others, Review system in computer science, Introduction to presentation

Meta cognition

Reflecting about your own mental process

Prior knowledge

- What do I know about the topic covered in the text?
- What kind of text am I analyzing?

Prerequisites / Assumed reader level

- What difficulties should I account for?
- Is the text created using a specialized scientific language?

Reading strategy

- In what way, in what order should I go about the actual reading?

Example

The article "Computation Takes Time, But How Much?" is a specialized article, that gives an overview of the whole domain

Homework

- Larger assignment: Quick look at the text.

1. Read "How to Read a Scientific Article" by Mary Purugganan and Jan Hewitt, Rice University
   `http://www.owlnet.rice.edu/~cainproj/courses/HowToReadSciArticle.pdf`

2. Use the method described there to analyze: "Computation Takes Time, But How Much?"

3. For each of the 5 steps proposed in the slides, find a text excerpt and analyze it accordingly. Pay special attention to reconstruction.

4. Find and prepare to recommend other texts on reading scientific works.

Syntactic and semantic analysis

Clarification of the used terms

Example

In "Computation Takes Time, But How Much?", the following terms need clarification:

- . . .
- . . .

Definitions?

Summary / Reduction

Definition (Margin notes)

- content division

  Subdivide the content; preferably with keywords that allow navigating the content

- logical division

  "How is a particular passage argumented?"
  1. Do the authors make a claim?
  2. Do they justify the claim?
  3. Do they draw conclusions?
  4. Do they results generalize?

- Could you justify / generalize it differently?

Summary / Reduction (cont.)

Definition (Excerpts)

selective repetition of a text, either copying or paraphrasing. It serves two goals:

1. a self made copy of a text (when original readable, but cannot be copied)
2. better understanding of the text

Definition (ways to make excerpts)

- general questions (little prior knowledge)
- specific questions (with prior knowledge)

Summary / Reduction (cont.)

Excerpts generally answer

1.
   - What is the **topic** of the paragraph?
   - What is the **message** of the paragraph?

   Both can be considered independently

2. Summarizing an excerpt

   On the levels of a paragraph, subsection, section

Example (Excerpts)

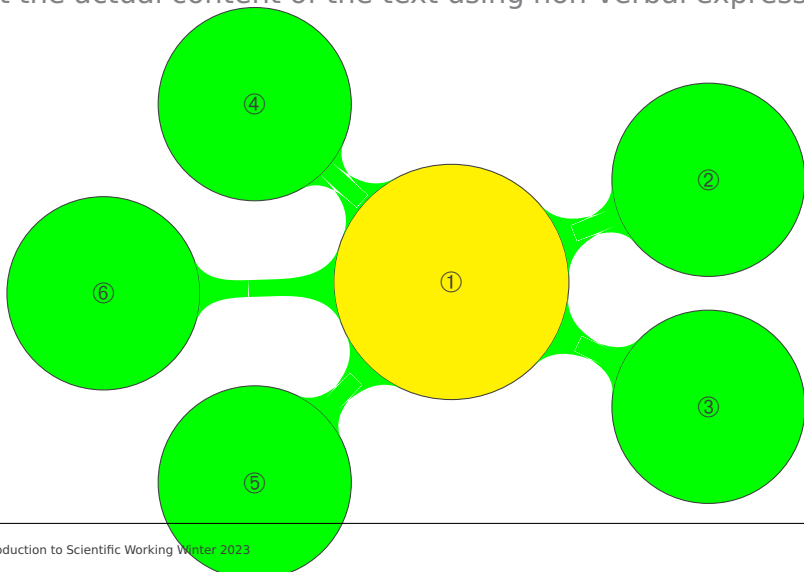①   ②

Example (Summary)

③

Reconstruction

Reconstruct the actual content of the text using non-verbal expressions

Elaboration

Go over the text in a critical way

Example

One needs to check:

- Is the text scientific, comprehensible?
- Is the author neutral?

Example

Consider the following questions about "Computation Takes Time, But How Much?":

- does the presented objective match the achieved goals?
- what is the class of analyzed programs?
- what are the challenges?
- automation level of the "automated" methods

Literature Research

Tools that simplify literature research:

1. General search engines

2. The DBLP Computer Science Bibliography

3. Online Library Catalogs

4. Electronic Libraries

5. Digital Archives

6. Publish or Perish

Google

Questions

- How do search engines it work?
- How does `google` order the results?

Answer

PageRank (eigenvector centrality of a graph) is a method that assigns weights to the nodes in an arbitrary directed graph based on the weights of the neighboring nodes ("incoming links"). In more detail, the weights are computed as the dominant eigenvector of the following set of equations:

$$PR_1(i) = \frac{1-f}{N} + f \sum_{i \in d(j)} \frac{PR_1(j)}{|d(j)|}$$

where N is the total number of nodes and f is a damping factor, typically set to 0.85.

The DBLP Computer Science Bibliography

DBLP

- DBLP gathers bibliography data for the imporant conferences and joirnals in computer science
- More than 2 mln articles indexed
- Originally a library service for DataBase) and Logic Programming
- DBLP offers:
  1. Electronic editions
  2. Citation statistics
  3. Bibliography
- Until 2011 administrated by Universität Trier, now combined project with Schloss Dagstuhl - Leibniz-Zentrum für Informatik GmbH

Homework / Work here

1 Read "Responsible and Efficient Literature Search" by R. Lewis and S. Sarli
https://becker.wustl.edu/sites/default/files/RespLitSearch.pdf

2 List sources (at least 5) of scientific literature useful in computer science

3 Can you give examples of misquotations / misrepresentations from recent international politics? What about science?

4 What are the legal consequences of plagiarism for UIBK students?