Computational Argumentation Seminar

Basics of Scientific Presentation

Henning Wachsmuth

November 1, 2017

www.webis.de

henning.wachsmuth@uni-weimar.de

Bauhaus-Universität Weimar

partly based on slides of
Prof. Dr. Gregor Engels and
Prof. Dr. Steffen Becker

Outline

- **Literature research**
  Types, quality, reading, acquisition, and organization

- **Oral presentations**
  Content, structure, style, talking, and timing

- **Written presentations**
  Content, structure, style, citations, and plagiarism

Literature research

Doing literature research

- **Literature research**
  - Fundamental task in science
  - Time-intensive and hard — but necessary
  - You are rarely the first to work on a problem
    If you are, what does that tell you?
  - Don't reinvent the weel

- **Why literature research**
  Not all relevant for the seminar talk
  - Find out if your approach to a problem is new
  - Find alternative approaches or perspectives
  - Widen the scope of the problem
  - Obtain background information
  - Obtain evidence for your or other's claims
    ... and many other reasons

Selecting literature

- **Types of literature (and similar)**
  1. Books. Theory, basics, approved techniques
  2. Scientific journal papers. Completed research lines
  3. Conference papers. State-of-the-art research
     In our field, major publication type
  4. Workshop papers. New ideas, ongoing research
  5. Conference/Online tutorials. Easy access to basics and techniques
  6. Popular science magazines. Easy access to research lines
  7. Other websites. Anything

- **What type to prefer**
  - Generally, literature should be peer-reviewed
    Most literature of types 1–4 is peer-reviewed, but not all
  - Rule of thumb: books > journals > conferences > workshops
    > tutorials > magazines > websites > other
  - But, for example, top conferences > average journals

Assessing quality of literature

- **Conference and journal rankings**

  - Top tier ranked A^+/A* or A; B still good

  - Unranked conferences/journals may be doubtful
    No ranking achieves complete coverage, though

  - One of the most reputable rankings is CORE
    core.edu.au/conference-portal

- **Number of citations**

  - Roughly indicate importance

  - Rather for *relative* comparisons within a topic

  - Notice: Newer papers naturally tend to have less citations

  - Good resource for citation numbers is Google Scholar scholar.google.de
    Journals also have so called impact factors derived from citation numbers

- **Disclaimer**

  - Good and bad research appears at all places

  - Often, only reading helps
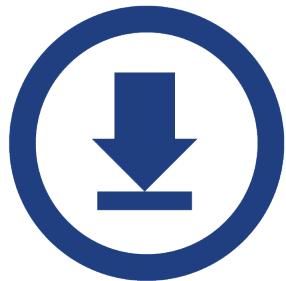
Reading and finding literature

- **Reading papers efficiently**

  1. Read abstract, introduction, and conclusion
  2. Look at figures and tables
  3. Decide whether worth reading everything
  4. Read goal-driven
     Specify questions to be answered during reading

- **Finding the next paper**

  - Follow promising references at the end of a paper
  - Find promising papers citing a paper
  - Learn to identify the best search terms
    Rule of thumb: As specific as possible, but as abstract as needed

- **Getting started in the seminar**

  1. First read the literature that we provide
  2. Then find further literature
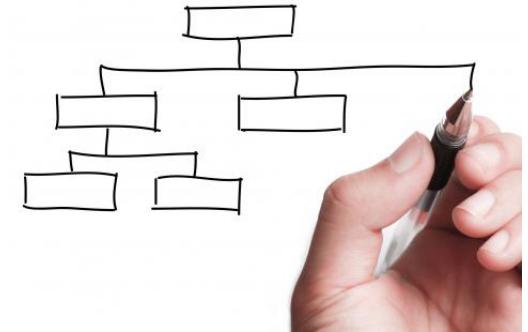
Acquiring literature
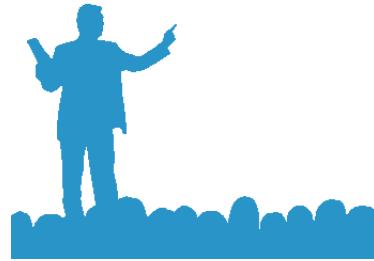
- **Obtaining papers**
  - Many papers freely available online
  - Others might be free from a university network
    ... although our university not the best in this regard ☹
  - If neither, maybe your advisors can help

- **Important sources**
  - ACL Anthology for computational linguistics papers aclanthology.coli.uni-saarland.de
  - ACM Digital Library for many important computer science papers dl.acm.org
  - dblp for any literature related to computer science dblp.dagstuhl.de
  - Google Scholar for any scientific literature scholar.google.de
    ... and general web search, of course

- **Accessing books**
  - Check if available in the library
  - Some accessable online, for example, on Google Books books.google.de
    Purchasing books can make sense when of continuous importance for you

Organizing literature

- **Literature organization**
  - Maintain overview, start from the beginning
  - "Extra" effort will pay off

- **Create logical folder structure**
  - Build your own view of the field
  - Logically subdivide topics, but don't over-engineer
    For instance ./literature/computational-argumentation/argument-mining/ — but maybe not deeper

- **Rename all literature consistently**
  - Simplifies browsing and finding
  - We use <1stauthor><2digityear>-<full-title-lower-case-no-special-chars>.pdf
    For example: ajjour17-unit-segmentation-of-argumentative-texts.pdf

- **Organizing meta-information**
  - Bibliographical information needed when citing literature
  - Store bibtex of literature whenever available
    Learn more on en.wikipedia.org/wiki/BibTeX; many pages such as dblp provide bibtex's

Oral presentations

Content of your talk

- **Scientific presentation is storytelling**
  - Plan what points to make and how to get there
  - Tell a coherent story with a central theme
  - Make it exciting, show importance
  - Don't be complete, be selective
    Holds for talks; different in articles (see below)
  - Avoid surprise: Clarify why you tell something

"Sometimes **reality** is too complex.

Stories give it a form."

Jean-Luc Godard
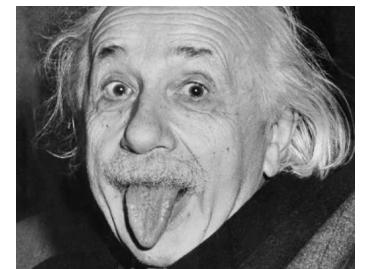
- **Science needs to be understood**
  - Adjust complexity to audience
  - Leave out formal things, unless needed
    Holds for talks; may be different in written articles (see below)
  - Be precise and clear
  - Introduce terms, use consistently
  - Figures and examples help

"Everything should be as **simple** as possible, but **not simpler.**"

Figures
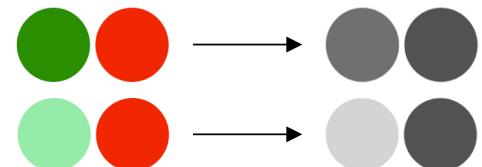
- **Figures**
  - Charts, diagrams, graphs, pictures, drawings, ...
  - Slides are visual
  - Rule of thumb: No slide without figure

- **What to use figures for**
  - Primary. Replace text; visually explain concepts, ...
  - Secondary. Support your message with pictures
    (as often done in this presentation)

- **Formats**
  - Vector graphics whenever possible
  - Others: Optimize sharpness, scale down smartly
   Never scale > 100%; 50% is better than 53% — why?

- **Readability**
  - Think of color blind people — contrast helps
  - Check readability of included text

"a **picture** is worth a **1000 words**"

"**unsharpness** is the mistake that even **lay persons** see"
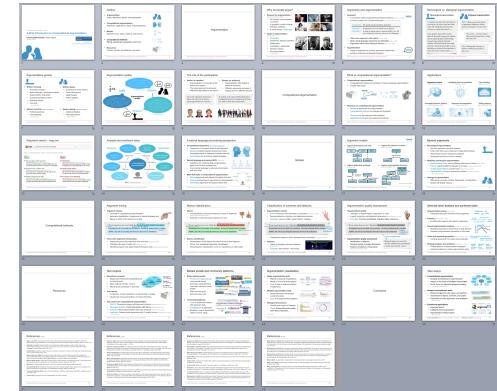
Herbert Kania

Structure of your slides

- **Overall stucture of presentation**
  - Title slide. Title, authors, maybe date
  - Outline slides. Only for longer talks (as of ~20 slides)
  - Content slides. Your story
  - Conclusion slide. Take aways, future work
  - References. Only shown when asked for

- **Structure of content slides**
  - Header. Clear unique title
    Notice: Titles often not read by the audience
  - Body. Bullet points, figures, tables, etc.
  - Footer. Page number, title, presenter, progress?, no date!
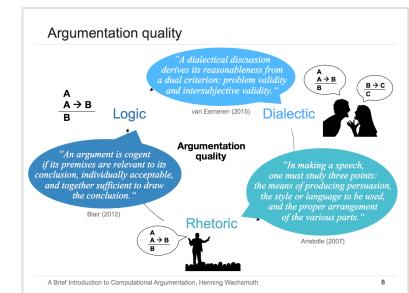
- **Space for separation**
  - Leave space between different slide parts
  - Leave some space to slide borders
    Harder to read there + border sometimes clipped
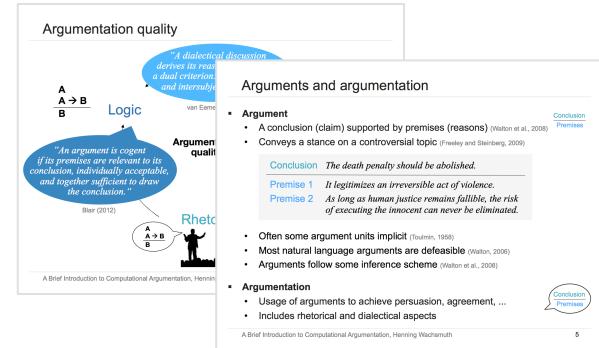
Style of your slides

- **General slide style**
  - Vary slides to maintain attention
  - Decide what to put on slide and what to say
  - Animations only when useful, use consistently
    Avoid playful ones, unless they match your message
  - Clarify what is from you and what from others
    Also see notes on citations below

- **Text style**
  - No full sentences, rather key phrases
  - Avoid grammar and spelling errors
  - AIA & AUA
    Always introduce acronyms & Avoid unnecessary acronyms

- **Amount of text**
  - Some say 7x7 — maximum 7 bullet points per slide, 7 words per point
  - I'd rather say 3x3 — 3 top-level points with 3 sub-points

Fonts

- **Fonts**
  - Sans-serif fonts (Arial, Helvetica, ...) much more readable on slides
  - I use *serif* fonts (Times, Garamond, ...) only for example texts

- **Font size**
  - This text is written in 26 pt — for titles and stressing
  - This text is written in 24 pt
  - This text is written in 21 pt
  - This text is written in 18 pt — minimum for text that should be read
  - This text is written in 16 pt
  - This text is written in 14 pt
  - This text is written in 12 pt — minimum for extra information that may be skipped
  - This text is written in 10 pt
  - This text is written in 8 pt
  - This text is written in 6 pt — maybe for texts that should on purpose not be readable

- **Font colors**
  - Use colors consistently
  - Not too colorful

Talking and timing

- **Giving a talk**
  - Look at audience, speek to everybody
  - No pre-phrased sentences
  - Match words on slides, but complement them
  - Don't be too formal, but be serious, avoid slang
    Jokes may be nice if you know how to use them

- **Timing**
  - Use your time, but stick with time limit
  - Expect ≥ 2 minutes per content slide
  - Rule of thumb: Audience can read slide twice
  - Leave time for questions and discussion

- **Practice your complete talk**
  - How much time do you need?
  - Does your story work?
  - Can you explain everything well?
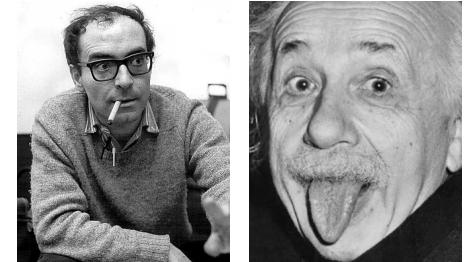
Written presentations

Content of articles

- **Most hints on talks also hold for articles**
  See above!

  - Science is storytelling
    Seminar: No scientific break-through expected

  - Science needs to be understood

- **Articles more complete**

  - Tell the whole story, avoid gaps in argumentation

  - But: Include only relevant content

  - Don't expect too much prior knowledge

  - But: No details on knowledge that can be presupposed

"**Don't** make me **think.**"

Steve Krug

- **Articles should be sound**

  - Need to be precise more than in talks

  - Use logical arguments, from broad context to deep details

  - Formalize concepts if needed/helpful

Structure of articles

- **High-level stucture**

  - Title and author information

  - Abstract

  - Usually 4–7 sections

  - References

  - (Appendices)

- **Section structure**

  - Often numbered subsections (2.1, 2.2, ...)

  - If any, subsubsections unnumbered

  - Always have text before sub[+]sections

- **Section headings**

  - Conventional: First is "Introduction", last is "Conclusion"

  - Other sections topic-specific

    Some semi-convential content sections exist, but not really suitable for seminar articles

  - Short misleading headings worse than long clear ones

Abstract

- **Abstract**
  - A concise high-level summary of the paper
  - Usually 5–10 sentences

- **My view of a good abstract**
  - Motivation and context (1 sentence)
  - Problem and why not solved (1–2 sentences)
  - Question addressed in the paper (1 sentence)
  - Approach in general, some details (2–3 sentences)
  - Evaluation, results, conclusion (1–3 sentences)
    For seminar articles, may differ a bit though

- **My former professor's view**
  - What is the problem? Why is it a problem?
  - What is the solution? Why is it a solution for the problem?
    Notice that this view is NOT in conflict with mine
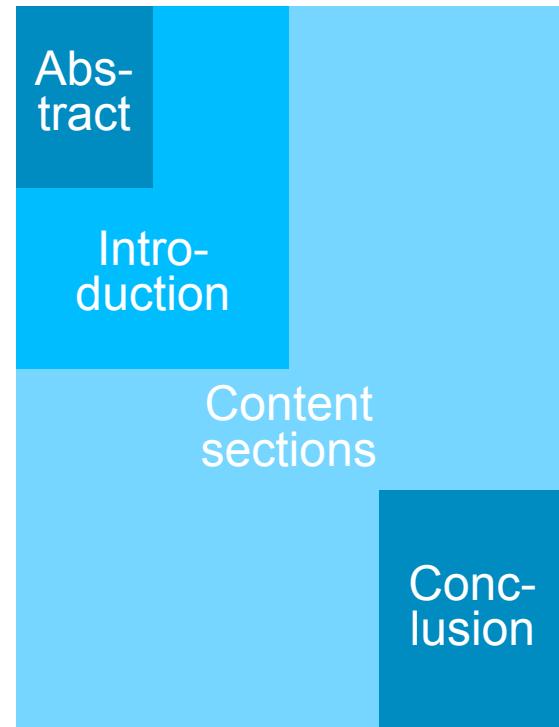
Abstract

Several approaches have been proposed to model either the explicit sequential structure of an argumentative text or its implicit hierarchical structure. So far, the adequacy of these models of overall argumentation remains unclear. This paper asks what type of structure is actually important to tackle downstream tasks in computational argumentation. We analyze patterns in the overall argumentation of texts from three corpora. Then, we adapt the idea of positional tree kernels in order to capture sequential and hierarchical argumentative structure together for the first time. In systematic experiments for three text classification tasks, we find strong evidence for the impact of both types of structure. Our results suggest that either of them is necessary while their combination may be beneficial.

Sections

- **Introduction**
  - The abstract in more detail
  - Tell the whole story, from context to conclusion
  - High-level, understandable for computer scientists

- **Content sections**
  - The introduction in more detail
  - Elaborate on related work, concepts, models, data, approaches, experiments, and results
  - More technical, for researchers from the area

- **Conclusion**
  - The introduction in less detail
  - Summarize story in retrospective, give outlook
  - Semi-technical

Style of articles

- **Scientific writing style**
  - Write clearly, unambiguously, and concise
  - Don't make things complex
    (common misunderstanding)
  - Use impersonal or "we" form

- **Some guidelines**
  - Avoid pronouns with unclear references
  - Use explicit discourse markers, such as "because"
  - Blurring is non-scientific, such as "It could be..."
  - English sentences are short, one statement per sentence
  - Again: Avoid grammar and spelling errors
    Seminar: Too many of them will negatively affect your grade

- **Article format in the seminar**
  - Provided template predefines layout and its usage
  - 8 two-column pages of content, 1–2 pages of references

Tables, figures, terms, and footnotes

- **Tables and figures**
  - Explain in text and in caption
  - In papers, just number increasingly
    Figure 1, 2, ... Table 1, 2, ... (NOT: Figure 2.1, 2.2, ...)
  - Tables: Horizontal lines suffice
  - No included font larger than article font

- **Technical terms**
  - Introduce where needed, don't overformalize
  - Use well-defined terms, AIA & AUA
  - Don't use synonyms for terms
    Reader is misled to check whether intentional differences exist

- **Footnotes**
  - Only for secondary information
  - Reduce readability, should be an exception
  - Don't cite literature using footnotes

(a) monological argumentative text

[1] The death penalty is a legal means that as such is not practicable in Germany. [2] For one thing, inviolable human dignity is anchored in our constitution, [3] and furthermore no one may have the right to adjudicate upon the death of another human being. [4] Even if many people think that a murderer has already decided on the life or death of another person, [5] this is precisely the crime that we should not repay with the same.

(b)

Figure 1: (a) Example text with five argument units, taken from the *Arg-Microtexts* corpus introduced in Section 3. (b) Graph visualization of the sequential and hierarchical overall argumentation of the text.

alternatives by modeling the stance of each unit towards its parent in the associated tree. This stance can be derived in all corpora.[3] All other unit and relation types from the specific models are ignored, since there is no clear mapping between them.

[3] Alternatively, the stance towards the main claim could be modeled. We decided against this alternative to avoid possibly wrong reinterpretations, e.g., it is unclear whether a unit that attacks its parent always supports a unit attacked by the parent.

Citation

- **Citation**
  - In-text reference to a bibliographic source
  - We use ACL-style: Author names + year
    Other communities use numbers ([1], [2], ...) or acronyms ([ACW17], ...)

- **What to cite**
  - Any reuse, paraphrase, summary, and translation of text from another source
    Also have to cite yourself if you use your own sources
  - Rule of thumb: Always clarify what is from you and what from others
  - Better one citation too much than one too less
    Can also ask your advisors in cases of doubt

- **How to cite**
  - Direct reuse. Put in quotes (possibly shorten with [...]), give source
    Example: Unit segmentation is "[...] the splitting of a text into argumentative segments" (Ajjour et al., 2017).
  - Other citations. Give source close-by
    Example: As Ajjour et al. (2017) points out, segmentation is the first task of an argument mining pipeline.
  - Large text portions. Give source once in the beginning
    Example: In the following paragraph, we summarize the segmentation approach of Ajjour et al. (2017).

References

- **List of references**
  - Bibliographical information at end of paper
  - Exactly those references cited in the text
  - References should be complete and consistent

- **Needed meta-information**
  - All literature. Author, year, title
  - Conferences/Workshops. Proceedings, pages
  - Journals. Journal name, issue, number, pages
  - Books. Edition if any, publisher
  - Only online. Give URL with access date
  - Other meta-information optional

- **Bibtex**
  - LaTeX handles references automatically using bibtex
    See part on organizing literature above

Aristotle. 2007. *On Rhetoric: A Theory of Civic Discourse* (George A. Kennedy, translator). Clarendon Aristotle series. Oxford University Press.

Beata Beigman Klebanov, Christian Stab, Jill Burstein, Yi Song, Binod Gyawali, and Iryna Gurevych. 2016. Argumentation: Content, structure, and relationship with essay quality. In *Proceedings of the Third Workshop on Argument Mining (ArgMining2016)*, pages 70–75. Association for Computational Linguistics.

Stefanie Brüninghaus and Kevin D. Ashley. 2003. Predicting outcomes of case based legal arguments. In *Proceedings of the 9th International Conference on Artificial Intelligence and Law*, pages 233–242.

Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3):27:1–27:27.

Plagiarism

- **Plagiarism**

  - To sell another's ideas or expressions as one's own
    See en.wikipedia.org/wiki/Plagiarism

  - On purpose or due to lack of giving sources

  - Plagiarism *not* a trivial offense
    In some countries considered as crime

  - Proper citing avoids all plagiarism issues

- **Consequences in the seminar**

  - Major cases lead to the denial of being graded

  - Minor cases can still negatively affect your grade

- **Webis and plagiarism**

  - Webis does research on plagiarism detection

  - See our tool picapica www.picapica.org

Sum up

Take aways

- **Literature research**
  - Fundamental part of scientific work
  - Literature varies in quality and suitability
  - Find, read, and organize literature efficiently

- **Oral and written presentation**
  - Science is storytelling, needs to be understood
  - Several best practices for content, structure, and style
  - Proper citation is a must
  - Practice oral and written presentation early

- **For the seminar**
  - Consider hints in this presentation
  - Notice that some are subjective, much is missing
  - Develop your own way of presenting

References

- **Several slides reuse content from:**

  - **Engels (2010).** Gregor Engels. Einführung in wissenschaftliches Schreiben und Präsentationstechniken. Presentation within the Seminar "Information-Driven Software Engineering". Paderborn, 2010. https://cs.uni-paderborn.de/fileadmin/informatik/fg/dbis/Lehre/ws10_11/PG_IDSE/Dokumente/ 2010-04-15_Schreiben_Praesentieren.pdf

  - **Becker (2012).** Steffen Becker. Scientific Working. Presentation within the Seminar "Model Driven Software Engineering with Eclipse. Paderborn, 2010. www.hni.uni-paderborn.de/fileadmin/Fachgruppen/Softwaretechnik/Lehre/ Proseminar_Model_Driven_Software_Engineering/ProSem_MDSD_Guidelines.pdf

- **Examples are taken from:**

  - **Ajjour et al. (2017)**. Yamen Ajjour, Wei-Fan Chen, Johannes Kiesel, Henning Wachsmuth, and Benno Stein. Unit Segmentation of Argumentative Texts. In Proceedings of the Fourth Workshop on Argument Mining, pages 118–128, 2017. http://aclweb.org/anthology/W17-5115

  - **Wachsmuth et al. (2017f).** Henning Wachsmuth, Giovanni Da San Martino, Dora Kiesel, and Benno Stein. The Impact of Modeling Overall Argumentation with Tree Kernels. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2369–2379, 2017. http://aclweb.org/anthology/D17-1252