



FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Writing for Computer Science

Prof. Dr. Viktor Leis

Professur für Datenbanken und Informationssysteme

Low-Level Aspects

“The northern United States and Canada are places where herons live and breed. Spending the winter here has its advantages. Great Blue Herons live and breed in most of the northern United States. It’s an advantage for herons to avoid the dangers of migration. Herons head south when cold weather arrives. The earliest herons to arrive on the breeding ground have an advantage. The winters are relatively mild in Cap Cod.”



“Great Blue Herons live and breed just about anywhere in the northern United States and most of Canada. When the cold weather arrives, the herons head south. A few come to Cape Cod where winters usually aren’t too bad. Most of these herons are either inexperienced young birds or lost adult males too stubborn to ask for directions south. Spending the winter here has its advantages, and I’m not talking about the free off-season parking in Provincetown. Herons are able to avoid the dangers of migration, plus they can be one of earliest to arrive on the breeding grounds.

However, there is a risk with staying this far north. Yes, our winters are often mild and pleasant. Then there is this winter, the winter that never ends. Snow, ice and cold are not kind to birds and I’d bet many herons won’t be booking a visit to Cape Cod next year.

Herons have on thing their favor: they are excellent hunters and are total opportunists. When the fish are frozen out, they’ll eat other things, including crustaceans, mice, voles and small birds. One hungry heron was seen chowing down a litter of feral kittens. I know, I know, I too was upset to read about herons eating small birds.

Herons also have one odd behavior that is not in their favor. In the winter they seem to choose and defend a favorite fishing hole. When these areas become frozen solid, some herons don’t seem to catch on and often will stand over a frozen stream for days waiting for the fish to return. Boy, talk about stubborn.”

Low-Level Aspects

Connecting Sentences

Splicing Sentences

- . (two separate sentences):
“Herons have one thing in their favor. They are opportunistic hunters.”
- ; (sentences belong together, but no explicit coherence relation):
“Herons have one thing in their favor; they are opportunistic hunters.”
- : (elaboration, exemplification):
“Herons have one thing in their favor: they are opportunistic hunters.”
- , (with coordinator (“and”, “or”, “but”, “yet”, “so”, “not”) or preposition (“although”, “except”, “if”, “before”, “after”, “because”, “for”))¹:
“Herons live in the northern United States, but they don’t live in South America.”

¹“however”, “nonetheless”, “consequently”, “therefore” are separate sentences

Similarity and Contrast

- similarity
 - “and”, “similarly”, “likewise”, “too”, symmetry
 - “Herons live in the northern United States. Herons live in most of Canada.”
 - “Herons live in the northern United States, and they live in most of Canada.”
- contrast
 - “but”, “in contrast”, “on the other hand”, “alternatively”
 - “Herons are opportunistic hunters, but they sometimes defend a fishing hold even when it is frozen.”
 - “Herons have one thing in their favor: they are opportunistic hunters. Herons have one thing not in their favor: they defend a fishing hold even when it is frozen.”

Elaboration, Exemplification, Generalization

- elaboration
 - : (colon), “that is”, “in other words”, “which is to say”, “also”, “furthermore”, “in addition”, “notice”, “that”, “which”
 - “Herons have one thing in their favor: they are total opportunists.”
- exemplification
 - “for example”, “for instance”, “such as”, “including”
 - “Herons are total opportunists. When the fish are frozen out, they’ll eat other things, including crustaceans, mice, voles, and small birds.”
- generalization
 - “in general”, “more generally”
 - “When the fish are frozen out, herons will eat other things, including crustaceans, mice, voles, and small birds. They are total opportunists.”

Exception and Sequence

- exception (generalization first)
 - “however”, “on the other hand”, “then there is”
 - “Cape Cod winters are often mild and pleasant. Then there is this winter, the winter that never ends.”
- exception (exception first)
 - “nonetheless”, “nevertheless”, “still”
 - “This winter seems like it will never end. Nonetheless, Cap Cod winters are often mild and pleasant.”
- sequence: before-and-after
 - “and”, “before”, “then” (usually easier to understand)
 - “The cold weather arrives and then the herons head south.”
- sequence: after-and-before
 - “after”, “once”, “while”, “when”
 - “The herons head south when the cold weather arrives.”

Connecting Sentences: Result, Explanation, Expectation

- result (cause-effect)
 - “and”, “as a result”, “therefore”, “so”
 - “Young herons are inexperienced, so some of them migrate to Cape Cod.”
- explanation (effect-cause)
 - “because”, “since”, “owing to”
 - “Some herons migrate to Cape Cod, because they are young and inexperienced.”
- violated expectation (preventer-effect)
 - “but”, “while”, “however”, “nonetheless”, “yet”
 - “Herons have tough time when the ponds freeze over. However, they will hunt and eat many other things.”
- failed prevention (effect-preventer)
 - “despite”, “even though”
 - “Herons will hunt and eat many things in winter, even though the ponds are frozen over.”

Low-Level Aspects

Commas

Which vs. That

- the which/that distinction does not exist in German
- “that” is restrictive:
“Herons that live in North America have been studied extensively.” (only herons in NA have been studied)
- “which” is nonrestrictive:
“Herons, which live in North America, were studied extensively.” (all herons have been studied and they live in NA)
- parenthetical comments:
“Herons, a common bird in North America, have been studied extensively.” =
“Herons (a common bird in North America) have been studied extensively.”

Commas

- indicate a pause in pronunciation
- required commas:
 - enumeration:

“I like to eat fish, meat, and vegetables.”
 - introductory phrase:

“On April 1, the world pretends to be funny.”

“Furthermore, there is winter.”

“Living everywhere in North America, herons have been studied extensively.”
 - parenthetical comments/nonrestrictive clauses:

“Herons, which live in North America, have been studied extensively.”

“Herons, living in North America, have been studied extensively.”
 - two full sentence connected through coordinator:

“Spending the winter here has its advantages, and I’m not talking about the free off-season parking in Provincetown.”
- when in doubt remove commas

Low-Level Aspects

Common Problems

- “X is faster than Y. This is not well known.” ✗
- “X is faster than Y. This fact is not well known.” ✓
- “X is faster than Y. This data structure is not well known.” ✓
- “X is faster than Y. Y is not well known.” ✓
- also: “these”, “they”, “it”, “its”, etc.

- “don’t” = “do not”
- “it’s” = “it is” or “it has”
- contractions are common in colloquial text and speech, but should be avoided in formal writing

- used differently in German and English
- German: connect nouns
 - Duden: “Der Bindestrich *kann* zur Hervorhebung einzelner Bestandteile in Zusammensetzungen und Ableitungen verwendet werden, die normalerweise in einem Wort geschrieben werden.”
 - “Soll-Stärke” vs. “Sollstärke”
- English: create compound adjective
 - “The Nobel Prize–winning author will be reading from her book at the library tonight.”
 - “X is a state-of-the-art query processing technique.” vs. “X is the state of the art for query processing.”
 - “data set” vs. “data-set” vs. “dataset”
- also: “There are 2-5 items per node.”

- interjection that intentionally interrupts flow
- used for emphasis
- “Most ideas are worthless – except for those that change the world.”

Mistakes (By German Speakers)

- “This approach allows to reduce overhead.” ✗
“This approach allows us to reduce overhead.” ✓
“This approach allows reducing overhead.” ✓
- “There are less than five options.” ✗
“There are fewer than five options.” ✓
- “We setup the server ...” ✗
“We set up the server ...” ✓
“The setup was ...” ✓(also “lookup” vs. “look up”)

- most examples from: Steven Pinker, *Sense of Style: The Thinking Person's Guide to Writing in the 21st Century*, Penguin, 2014

Technical Writing

How to Read a Paper Efficiently (By Prof. Pete Carr)

<https://www.youtube.com/watch?v=IeaD0ZaUJ3Y>

Technical Writing

Paper Structure

Standard Structure

1. Title
2. Abstract
3. Introduction
4. Related Work
5. The Meat
6. Evaluation
7. Conclusions and Future Work
8. References

- “Analysis of Two Existing and One New Dynamic Programming Algorithm for the Generation of Optimal Bushy Join Trees without Cross Products” (2.6% are longer than 100 chars)
- “Only Aggressive Elephants are Fast Elephants”
- “Effective and Robust Pruning for Top-Down Join Enumeration Algorithms”
- “Managing Non-Volatile Memory in Database Systems”
- “Everything You Always Wanted to Know About Compiled and Vectorized Queries But Were Afraid to Ask”
- “The Adaptive Radix Tree: ARTful Indexing for Main-Memory Databases”
- “SuRF: Practical Range Query Filtering with Fast Succinct Tries” (32% contain :)
- “How Good Are Query Optimizers, Really?” (1.3% contain ?)
- “On the Stability of Plan Costs and the Costs of Plan Stability” (1.5% start with on)
- “Towards Scaling Blockchain Systems via Sharding” (0.9% start with towards)

- tell reader what the paper is about (title may not be enough to do that)
- after reading the title, readers may check the abstract whether the paper is indeed relevant for them
- I like it short (1/4 of a page)
- redundant: everything in it is also in the paper (introduction)
- abstract is copied to websites, booklets etc: do not use citations in it

Section: Introduction

- most important section of the paper, should be very well polished
- must motivate the reader to keep reading:
 - why is this an interesting and important problem?
- may be necessary to discuss some related work
- can be short (1 page of 12) or longer (2 pages of 12)
- reading only the introduction should give the reader about what the paper is about and what the paper achieves
- often hardest section to write, so start with other sections first

Typical Structure of Introduction Section

- usually no subsections, unless introduction becomes very long
 1. context, background
 2. problem description, research questions answered in the paper
 3. sometimes: crucial related work necessary to understand the paper or its contributions
 4. very brief description of key idea/technique (solution)
 5. key findings, experimental results
 6. optional: explicit list of contributions
 7. optional: explicit outline of the rest of the paper

Section: Related Work

- discuss prior work the paper is based on
- discuss alternative approaches
- mention papers that may appear similar, but are actually different (“Abgrenzung”)
- good to group similar approaches into subsections
- citations should be invisible:
 - “The undecidability of the halting problem was shown in [1].” ✗
 - “The halting problem is undecidable [1].” ✓
 - “Turing showed that the halting problem is undecidable [1].” ✓

Section: The Meat

- describe core contributions
- may be split into multiple sections
- each section is split into subsections
- top down: start with high-level design, then zoom in
- this is probably the section you want to start with

Section: Evaluation

- after reading the paper, the reader wants to know whether the approach is worth it
- start with a simple, easy-to-understand plot showing the benefits of the approach
- then look at specific aspects (e.g., varying parameters)
- if the paper has three technical contributions, evaluate them step-by-step
- good to state explicitly which questions will be answered in the section
- often easy to write, but also not as important as other introduction and meat

Section: Conclusions and Future Work

- briefly summarizes what the paper has achieved
- often a more appropriate name than conclusions: “Summary”
- future work describes extensions, modifications, follow-up research
- future work is optional

Typical Length for 12-Page Paper

1. Abstract ($\frac{1}{8}$ – $\frac{1}{2}$ page)
2. Introduction (1–2 page)
3. Related Work (1–2 page)
4. The Meat (2–6 pages)
5. Evaluation (1–4 pages)
6. Conclusions and Future Work ($\frac{1}{8}$ – $\frac{1}{2}$ page)

Alternative Structure

1. Abstract
2. Introduction
3. Background
4. main section(s)
5. Evaluation
6. Related Work
7. Conclusions and Future Work

Section: Background

- there may be prerequisites readers have to know to understand the paper
- related work can discuss these, but may become very long
- good solution: split Related Work into Background (at the beginning of the paper) and Related Work (at the end)
- Background should (only) contain crucial information necessary to understand the paper
- Related Work contains the rest

Technical Writing

Tips & Tricks

10 Tricks

1. Top-Down is Better Than Bottom-Up
2. Goal-Oriented Writing
3. Baseline Idea
4. Describe Challenges
5. Control Connotation
6. Avoid Wall of Text
7. Examples
8. Repeat Important Points
9. Naming
10. Figures

Top-Down is Better Than Bottom-Up

- it is natural to structure the paper bottom-up: first describe building blocks
- but this makes it hard to understand
- onion principle: write in layers (recursive structure), adding more details in each layer
- there may be cases where writing bottom-up in some section is more understandable
- but this should be done locally and after showing the big picture

Goal-Oriented Writing

- the world is complex and the world is large
- any text is limited in what it can explain/solve
- it is important to ask: what is this work trying to achieve?
- take a position and argue for it
- cut text that does not help
- stay on message (“kill your darlings”)
- move peripheral material to footnotes
- later: consider deleting footnotes

- when presenting a new approach it can be helpful to start with a simple baseline
- either a well-known approach from the literature or solution an intelligent reader would come up with in five minutes
- highlighting problems of this baseline hopefully naturally leads to the solution
- describe one's own proposal by comparing it with the baseline

Describe Challenges

- usually, the reader did not work on the problem for a long time
- just stating the solution may not convince the reader whether it is in fact a good solution
- one good approach is to explicitly discuss why the problem is difficult and what the challenges are (start with the problem, not the solution)
- this can also be done in steps
- helps understanding why the solution looks the way it does

Control Connotation

- same facts/ideas can be communicated very differently (denotation vs. connotation)
- use connotation that makes sense given your goals
- “The glass is half full.” vs. “The glass is half empty.”
- for many concepts there are positively- as well as negatively-associated words:
 - “simple”, “simplistic”, “trivial”, “naive”
 - “shrewd”, “egghead”, “bright”, “clever”, “brilliant”, “cunning”, “smart”, “intelligent”, “brainy”
 - “aggressive”, “assertive”, “domineering”, “dynamic”, “pushy”, “forceful”
 - “interested”, “questioning”, “nosy”
 - “unique”, “different”, “peculiar”
 - “thrifty”, “saving”, “stingy”

- readers get bored when confronted with lots of text: break up into subsections
- avoid long paragraphs
- add figures
- introduce bullet points for important points

- examples are crucial
- help readers to ensure that they understood the concepts
- running examples that are used throughout the paper can be great too
- may be used as a motivation

Repeat Important Points

- many readers only read superficially or only parts of the text
- it is therefore important to make sure important points do not get lost
- one way to do this is to repeat important points at several places in the paper (e.g., introduction, main chapter, conclusions)

- a new idea, concept, system, data structure should be given a name
- this makes it easier to talk about them and makes them a character of the story
- examples: RAID, cache-oblivious algorithm, ART, LeanStore
- pronounceable acronyms work well
- when first introducing the term write it in *italics* (and spell out the acronym)
- avoid introducing too many new names/jargon

Figures

- “a picture is worth a thousand words”
- many ideas can be much easier communicated using a picture than by text
- usually: describe idea in text *and* figure (reader can double-check understanding)
- each figure should be referenced and described in the text
- because some readers skip text it is a good idea to make figures+captions self-explanatory
- add textual annotations to the figure, colors too
- types of figures: example execution, abstract diagram, flow chart, code fragment, experimental results
- for difficult topics it often helps to start with the figure and then write the text about it

(Superficial) Details Matter

- minor issues like typos, errors, orphans and widows, and layout matter as they subconsciously signal lack of care
- sign of respect towards the reader
- it is fine for an early draft, but these issues must be addressed before submission
- always spell-check at the end

- after working on the paper for a while, it becomes to perceive it as a new reader would
 - dead-wood method: print on paper, read on couch (new perspective), mark issues with a pen
 - random method: jump to random (sub-)section and revise it
 - read-aloud method: helps catching many issues that are hard-to-spot in print

“I believed, in my early days, that you should spend at least as much time in the polish and presentation as you did in the original research. Now at least 50% of the time must go for the presentation. It’s a big, big number.” Richard Hamming, You and Your Research,
<https://www.cs.virginia.edu/~robins/YouAndYourResearch.html>