Tips on Writing Papers with Mathematical Content

John N. Tsitsiklis

May 2019

http://www.mit.edu/~jnt/write.html

Writing is a serious affair

- Why?

- Efficient use of your time
- Efficient transmission of your message

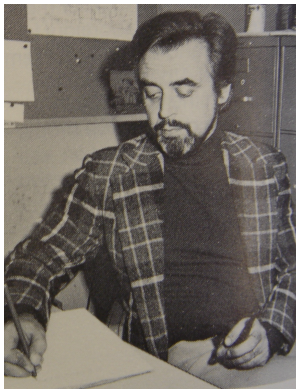
- All scales matter (micro/macro, details/ideas)

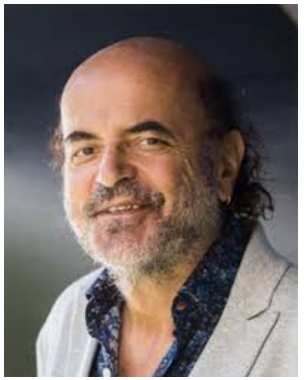
Overview

- Highest-level advice
- Before you start
- Document structure
- Modularity and guidance
  - Abstract, introduction, sections, appendices
- Good English language and style
- Mathematical style
- Typesetting

Highest level advice

- Do not overestimate the reader's ability
  - They should enjoy reading
- Be insecure
- Learn from "good examples"





- Spend time thinking before you start

Before you start

- Who is your audience?
- Why does this paper exist?
  - Main takeaways?
- Collect precise statements of key results (on paper)
- Make a table with your notation

  random variable X, takes values x

  x_t, $x(t)$, $x(n)$, $x[n]$

  $a_{i_k(t),j}$ ✗
- Settle on terminology, and stay consistent

  links, arcs, edges
  non-negative, nonnegative
  agent, node, sensor
  queueing, queuing
  multi-agent, multiagent

Document structure

1. Abstract
2. Introduction
3. The Model
4. Preliminaries (optional)
5. Results (usually 1-4 sections)
6. Conclusions
7. Appendices

- Modularity: subsections, subsubsections, examples, etc.
  - Titles (in bold) serve as sign-posts
  - Modules: 1-3 pages
    - with clear purpose (“In this subsection, we will …”)

Abstract

- **Declarative. Short and to the point; no background info**
  - NO: “In recent years, there has been an increased interest on … But the problem of … remained open…”
  - YES: “We consider a collection of agents who … We establish (i) …; (ii) …; (iii) … As a corollary, we settle an open problem posed by Fermat in 1637.”

Abstract: Reinforcement learning (RL) offers great promise in dealing with previously intractable control problems involving nonlinear dynamical systems. Modern RL methods, based on policy-space optimization, rely on a guarantee that stochastic gradient descent converges to local minima. Unfortunately, this guarantee fails to apply in settings involving open-loop unstable systems. The behavior of RL algorithms in such a context is poorly understood, and this is an important issue if RL-based controllers are to be deployed. In this paper, we address this issue. More specifically, we show that (i)…, (ii) …, and (iii) …

Introduction

- This is what most people will read…
- Each paragraph should have a clear purpose
  - Framing the paper (“In this paper, we …”)
  - Motivation
  - Background and history; literature review
  - Preview of main results
  - List of key contributions
  - Outline: “The rest of the paper is organized as follows”

Modularity within sections

- Section = a collection of items
    - Intro to the section; how it ties to the rest
    - Initial discussion, to set the reader's mind
    - Theorem
    - Interpretation of the theorem
    - Idea of the proof
    - Limitations of the theorem; counterexamples
    - Examples
    - Illustration through figures (long captions are fine)

Proofs

- We discover proofs by going backwards
  - To get to D, I need to show C, which I can establish through Lemmas A and B
- We write proofs by going forward, linearly
  - Prove Lemmas A and B
  - Use them to establish C
  - Prove D
- Outline this structure before starting the proof

(c) Tigatelu | Dreamstime.com

- No rabbits out of a hat: 5 rambling pages, followed by: “We just managed to establish the following amazing result”
- Long, technical arguments -> Appendices
  - Main text should be self-contained (no references to lemmas or notation that are local to an appendix)
- Alert the reader when skipping steps!

Language

Maman died today, but I do not know for sure, as it could also have been yesterday, based on the fact that I am only relying on a telegram from the Home saying that “mother deceased.”

Maman died today. Or yesterday maybe, I don't know. I got a telegram from the Home: "Mother deceased…”

- Break up sentences!

Language

- Active voice: “We show” vs. “It is shown”

- Pronouns must be **unambiguous pointers**
  - “When a message from a server arrives to the dispatcher, it ? stores the header…”

  - Remove redundant words
    - “If we define x=2y, we have that 2x=4y.”
      “If x=2y, then 2x=4y.”
    - “The proof rests on the idea of employing the triangle inequality.”
      “The proof employs the triangle inequality.”
    - “Using ~~the result in~~ Lemma 3, Lemma 4 follows.”
    - But: “Assume that… ”

Math language

- Aim for linear structure at the micro level too
  - Lemma 1: If *n* is even, then n is composite.
  - By Lemma 1, 2k is composite, because 2k is even.
  - Note that 2k is even. By Lemma 1, 2k is composite.

- Ideal: "If …, then …"
  "Define … Then, Lemma 2 implies that…"

- Short and crisp lemmas, theorems
  - Do not define terms or add discussion inside the statement
  - Introduce terms and assumptions outside/earlier

- Aim for parallel constructions

(a) For all even integers n, property P_n holds.
(b) However, property Q_n holds if n is odd.

(a) For all even integers n, property P_n holds.
(b) For all odd integers n, property Q_n holds.

- Math should read like English

~~"For every 1<k<10"~~

Quantifier ambiguities are common

for every n, we have $n < c$, for some c

for every n, there exists some c such that $n < c$

there exists some c such that for every n, we have $n < c$

$T = O(n^d)$

$T \leq cn^d$

There exists some c such that
for all large enough n and d,
we have $T \leq cn^d$

For any d, there exists some c such that
for all n large enough,
we have $T \leq cn^d$

Typesetting

- Beauty

  Avoid inline fractions such as $\frac{x+2}{x+3}$, which result in small fonts and interfere with proper line spacing, unless there is a compelling reason. Instead, write $(x+2)/(x+3)$.

- Make parsing easier

$$\mathbf{E}[X + 3 + k^2 | Y = 3 + \log k + n^2]$$

$$\mathbf{E}[X + 3 + k^2 \mid Y = 3 + \log k + n^2]$$

- And many more suggestions in the references

Sources

The essay "How to write Mathematics," by Paul Halmos, available at http://www.math.washington.edu/~lind/Resources/Halmos.pdf is a gem.

"Mathematical Writing," by Knuth et al., available at http://tex.loria.fr/typographie/mathwriting.pdf is very thorough. For the impatient, the 27 rules offered in the first 6 pages are very valuable.

Dimitri Bertsekas, "Ten Simple Rules for Mathematical Writing,"
available at http://www.mit.edu/~dimitrib/Ten_Rules.pdf.