

Writing Perfect Papers

Baochun Li

Department of Electrical and Computer Engineering

University of Toronto

A long talk today on writing perfect
papers

“Perfect papers” don’t exist

But let's try our best

And try to write good papers quickly

1

Writing perfect papers

2

Writing good papers fast!

Our goal today: improve both the skills 1
and the process 2

1

Writing perfect papers

A **top-down** approach: from the big picture to tiny details

1

Writing perfect papers

I am going to talk about **four** things

1.1

The story

Writing perfect papers

1.2

The work

Writing perfect papers

1.3

The art

Writing perfect papers

1.4

The details

Writing perfect papers

1.1

The story

Writing perfect papers

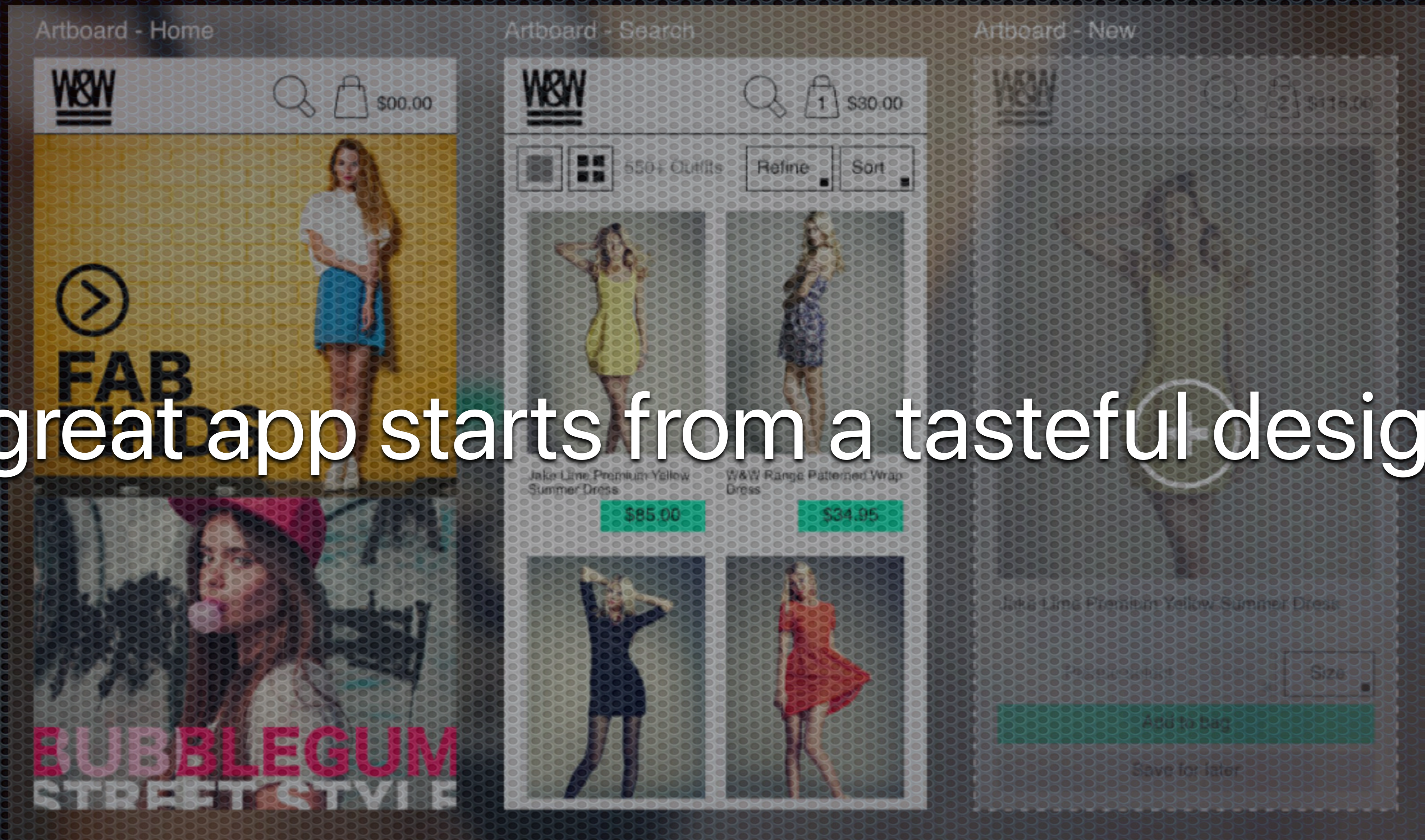


How do I make a great movie?



A great movie starts from a powerful and
convincing story

A great app starts from a tasteful design



Three essential elements of a great story



My
solution

My math
proofs

My
simulation
results



I am very
smart!

My math
proofs

My
simulation
results

The image features three overlapping circles on a dark blue background with a fine, repeating dot pattern. Each circle has a light blue outline and contains white text. The circles are arranged in a triangular pattern, with one at the top and two below it.

I am very
smart!

I am
smarter
than you
think

My
simulation
results

I am very
smart!

I am
smarter
than you
think

The thing
actually
works!



I am very
smart!

I am
smarter
than you
think

The thing
actually
works!

Your paper is not really the best place to
show your theoretical prowess

It's about advancing the state-of-the-art

The image features three overlapping circles on a dark blue background with a fine, repeating dot pattern. Each circle has a light blue outline and contains white text. The circles are arranged in a triangular pattern, with one at the top and two below it.

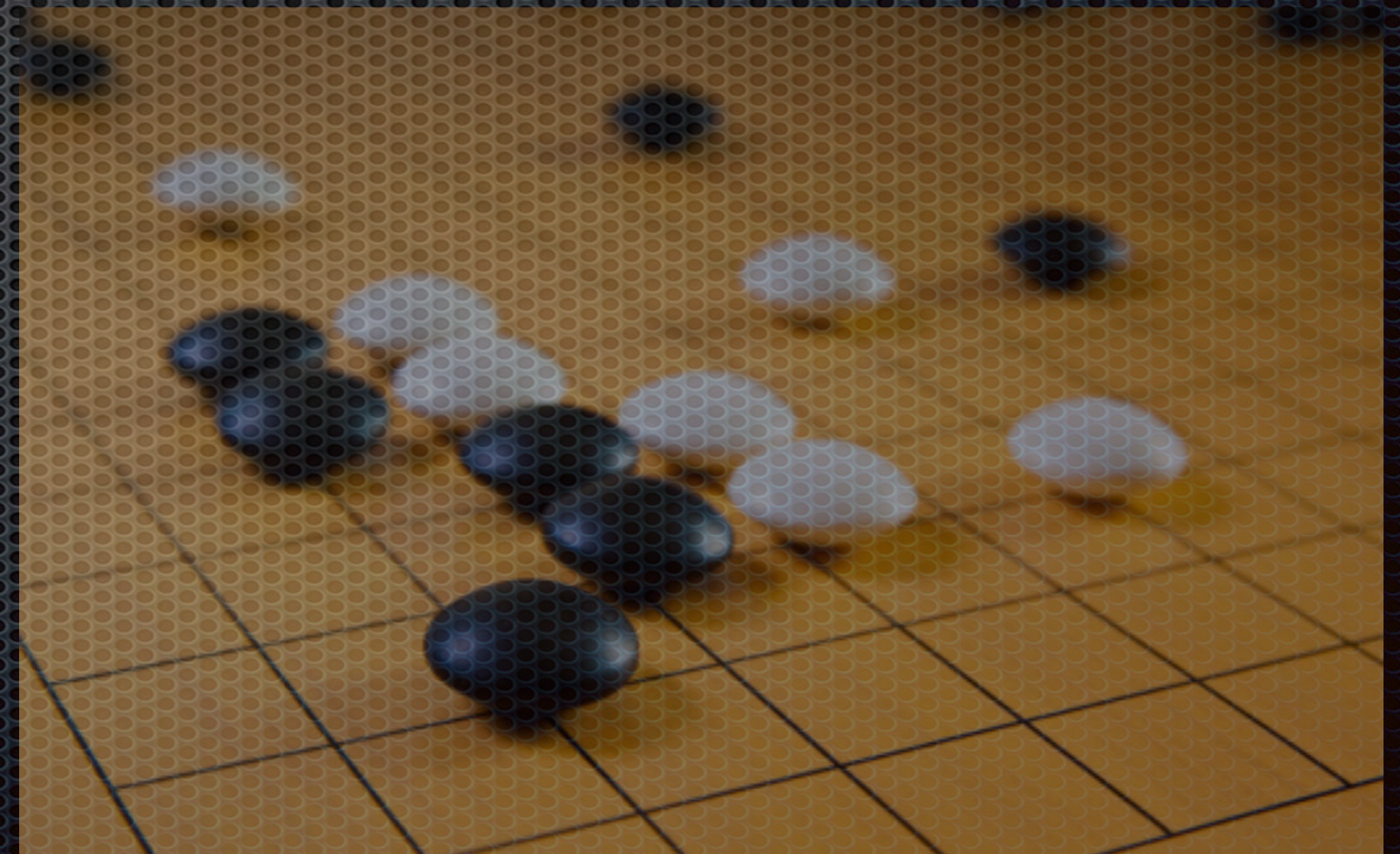
The
problem

What's
new?

How is it
better?

The
problem

Doesn't have to be
something "trendy"



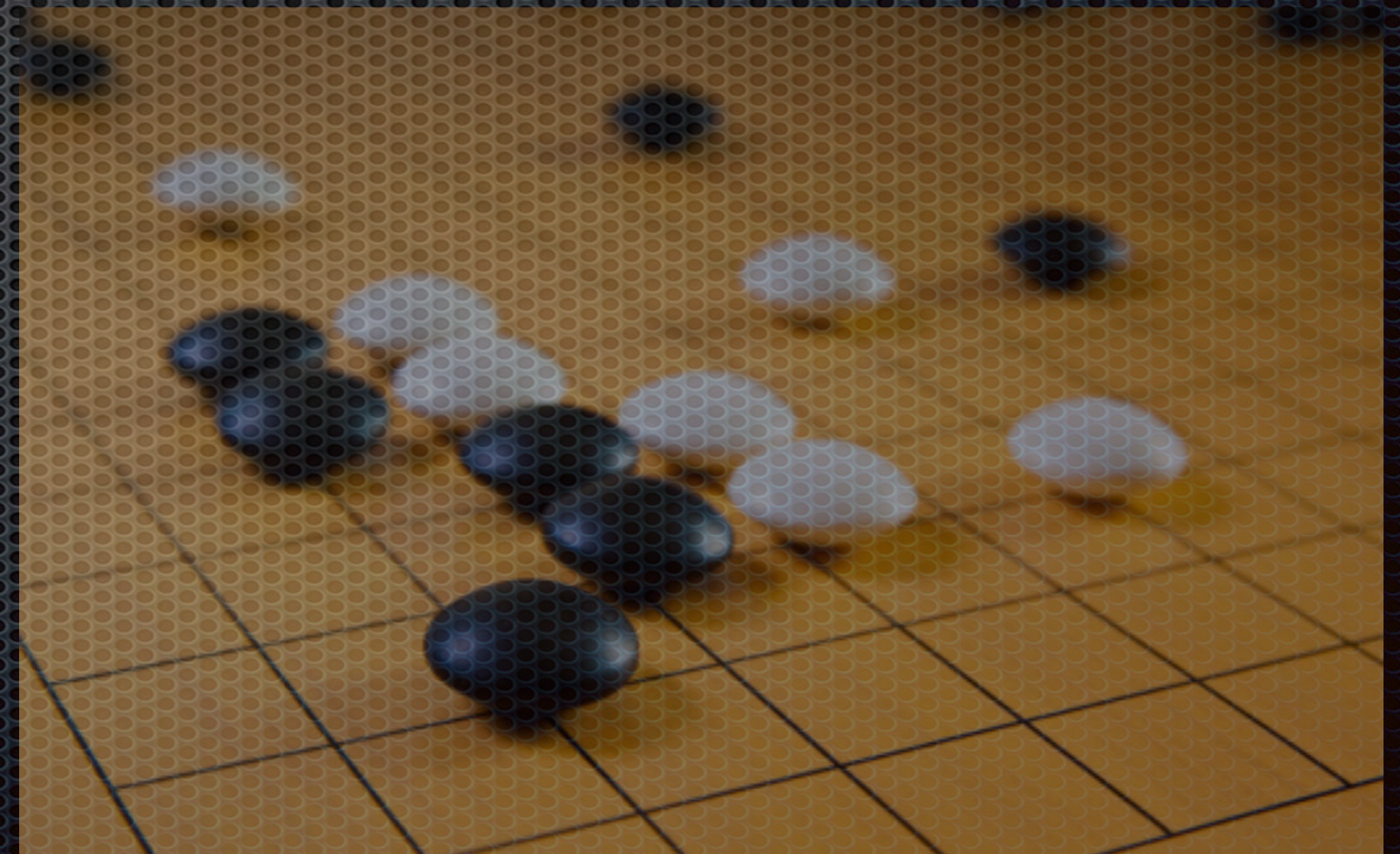


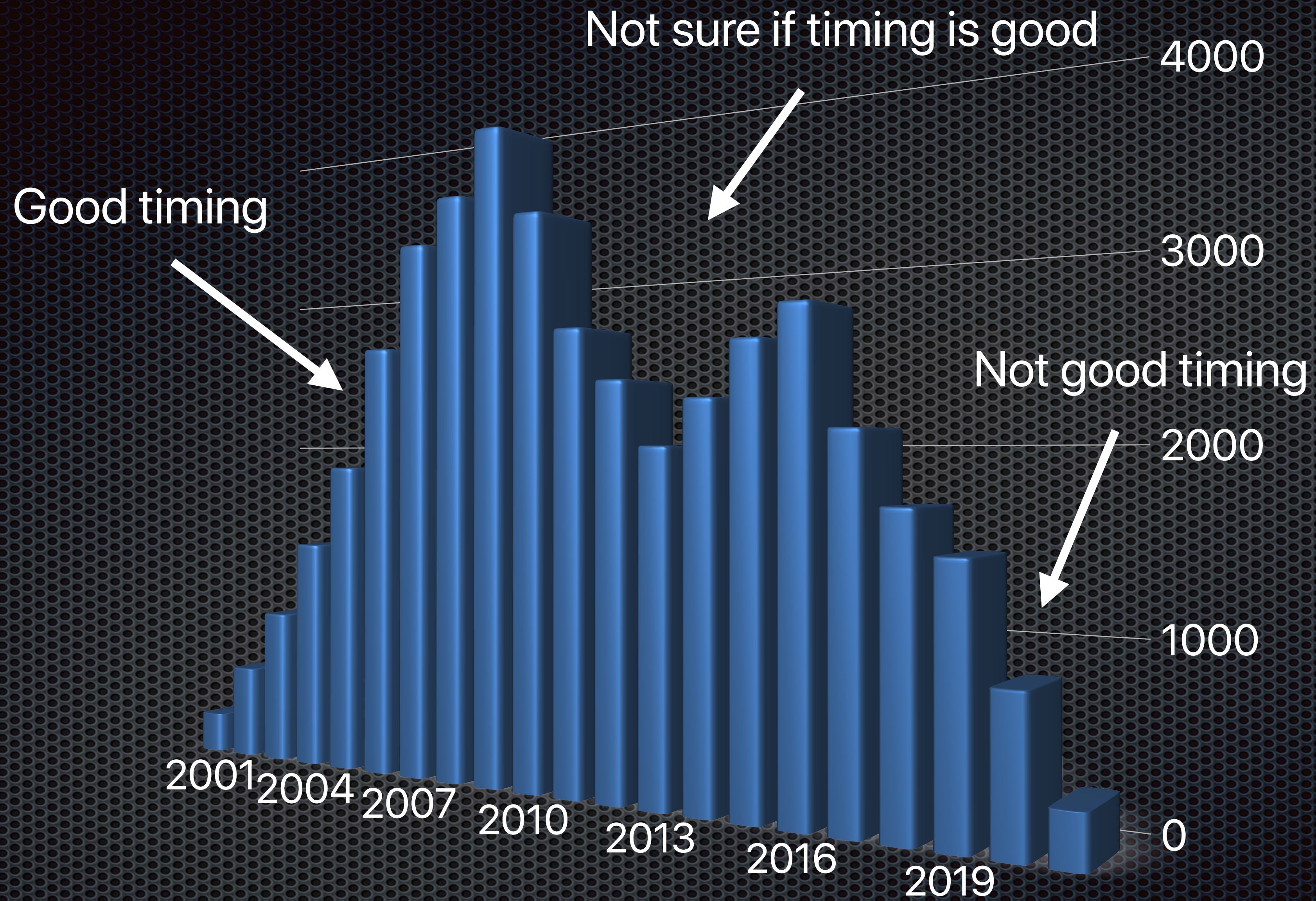
The
problem

But it has to be something you really enjoy
working on

The
problem

The “trend” may not
last forever





Google Scholar search: "P2P" in the paper title

Following the trend may lead to more
“incremental” results — less exciting and
less important

Be a **contrarian** and work against the trend

In fact, a recent 2017 paper, titled “BBR:
Congestion-based Congestion Control,”
brought congestion control in the general
case back to the spotlight

So, be a **contrarian** and working against
the trend may be a good idea



What's
new

In the context of what's already done in
the related work, **what's new** in your work?

First, briefly introduce the **state-of-the-art** in the literature, solving the same problem you wish to solve

Then write about what is **new** in your work that is different from previous work

If the problem is **different** from previous work — different assumptions, for example — provide justifications **why** this is **important** and **necessary**



How is it
better?

Now that your ideas are new, they should
also lead to **better** results than before

Show **how** your solution is **better** than previous work: such as stronger theoretical properties or better experimental results



The
problem

What's
new?

How is it
better?

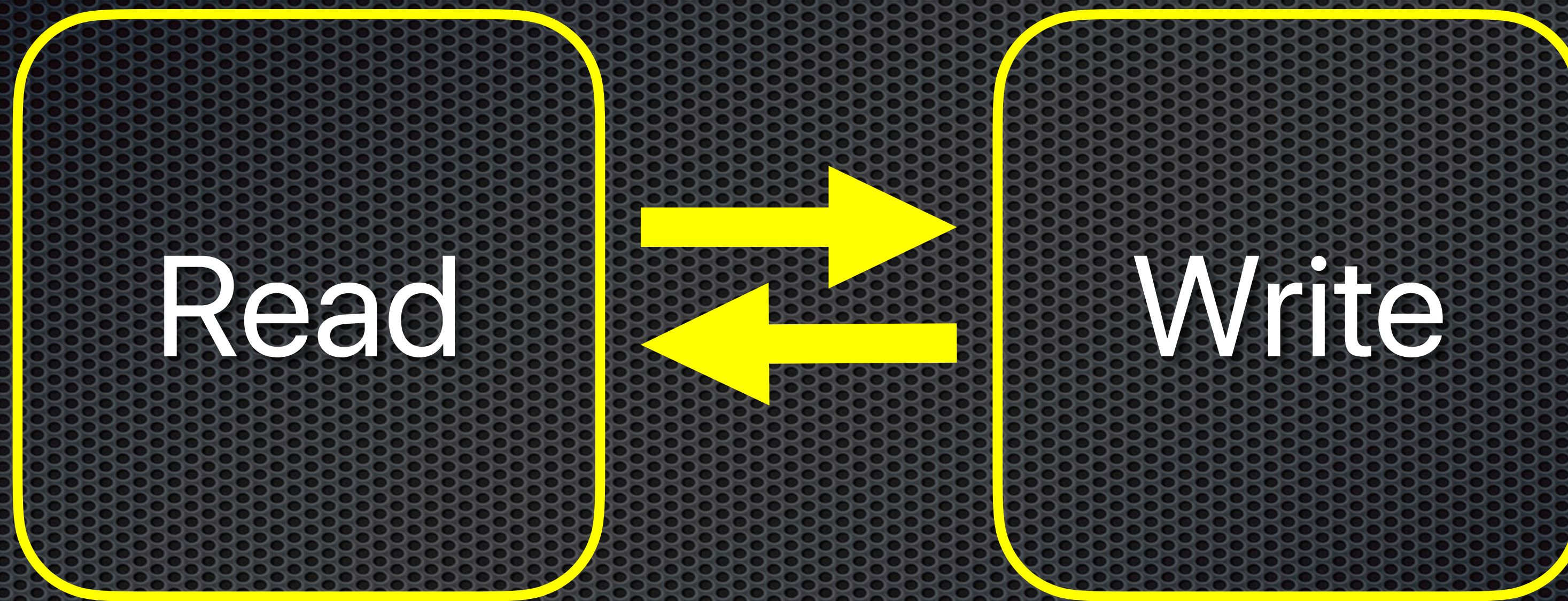
Three essential elements

1.2

The work

Writing perfect papers

Two things that you must do



Read

Goal: become an
expert on the problem

Read

But how?

Read

Start from one paper

Read

Perhaps a highly cited
seminal paper

Read

Then do an expanded-
ring search

Read

Papers that cited this
paper

Read

Papers that it cited

Read

Papers authored by the
same researchers

Read

Try not to depend only
on a search engine

Read

Read quickly first

Read

Read more carefully if
necessary later

Write about what you
understood

Write

Write every day

Write

idea

```
graph TD; A[idea] --> B[do research]; B --> C[write a paper];
```

A vertical flowchart with three rectangular boxes. The top box contains the word 'idea'. A white arrow points down from the bottom of this box to the top of the middle box. The middle box contains the words 'do research'. Another white arrow points down from the bottom of the middle box to the top of the bottom box. The bottom box contains the words 'write a paper'.

do research

write a paper

idea



do research



write a paper

idea



write a paper



do research

If I don't have an idea,
what do I write about?

Write

You write anyway

Write

Write about related work

Write

Write about what's
challenging

Write

Write about why the
problem is important

Write

Write about what needs
to be fixed

Write

Write a survey paper

Write

Submit your survey
paper for publication

Write

Write about **what's new** in
the context of related work

Write

The
problem

What's
new?

How is it
better?

Creating ideas interacts
with writing closely

Write

Writing is the best way to
force yourself to think
clearly and be focused

Write

and to **crystallize** what
you don't quite
understand yet

Write

Writing also opens
a dialogue

Write

For others to read what
you wrote

Write

To stimulate
discussions with others

Write

That's why **writing** is a
slow and **painful** process

Write

Write in a crystal clear
way

Write

Write about **one** problem
and **one** (potential)
solution

Write

Write with a flow of ideas
that's easy to follow

Write

Keep your readers
engaged throughout the
paper

Write

Write your paper so that
it's as easy to understand
as **absolutely possible**

Write

Your readers don't have
to work hard

Write

Write your paper **slowly**,
so that your readers can
read **quickly**

Write

1.3

The art

Writing perfect papers

Pay attention to the title, abstract,
introduction, and flow of ideas

Title: 2 lines, 10,000 readers

Abstract: 10 lines, 1,000 readers

Introduction: 100 lines, 100 readers

The rest of the paper:
1,000 lines, 10 readers



The title

Writing perfect papers: the art

What's a good title?

It attracts a reader to read the abstract

It reflects the essence of the new idea

It is as simple as possible

does not have to be a precise summary of the
paper

does not need to include all the keywords

The best title is one that is the easiest to
understand at a glance

Let's say we have a new idea that uses a new coding technique to reduce the latency when delivering data over the Internet

Which title is the best?

LLRC: A Low-Latency Rate-Controlled System for Fast Data Delivery over the Internet

Coded Information Distribution: New Content-Delivery Protocols for the Internet

Coding Reduces Latency over the Internet

Which title is the best?

LLRC: A Low-Latency Rate-Controlled System for Fast Data Delivery over the Internet

Coded Information Distribution: New Content-Delivery Protocols for the Internet

Coding Reduces Latency over the Internet

S. U. Stich. "Local SGD Converges Fast and Communicates Little," Int'l Conference on Learning Representations (ICLR) 2019.

Here's another new paper on the wireless spectrum market, where buyers and sellers are matched in a nicely optimized fashion

Which title is the best?

Extended Deferred Acceptance: Interference-Free
Matching between Spectrum Buyers and Sellers

Stable Matching for Wireless Spectrum Markets

Spectrum Matching

Which title is the best?

Extended Deferred Acceptance: Interference-Free
Matching between Spectrum Buyers and Sellers

Stable Matching for Wireless Spectrum Markets

Spectrum Matching

I have a co-authored paper in 2016 titled
“Spectrum Matching” — so though it’s a
bit risky, it can work out



B

The abstract

Writing perfect papers: the art

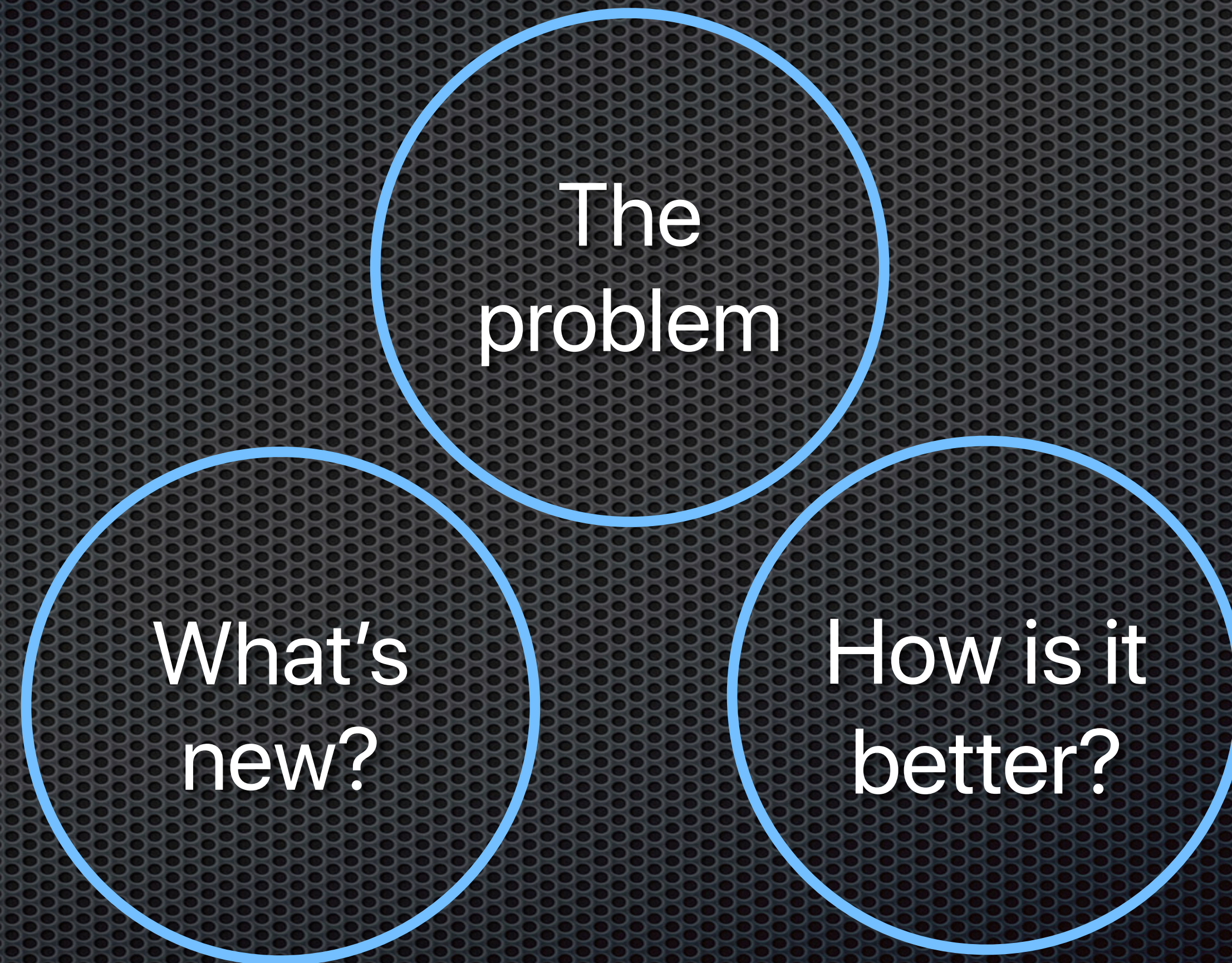
The abstract

The abstract is a very important tool to attract readers to read the introduction

Again, it conveys essential information about the paper

It should be concise, and does not need to be long

Remember the three essential elements



A typical structure

One sentence to state the background

One sentence to state **what the problem is**

Two to four sentences to state the original contributions in the paper — **what's new?**

One sentence to state **how the solution is better,**
validated using analyses, simulations, or experiments

Here is an abstract from one of my co-
authored papers

We will use this as our running example
throughout this talk

It's title: "Oruta: Privacy-Preserving Public Auditing for Shared Data in the Cloud"

It currently has the highest number of citations in all the papers I co-authored

So let's take a look at how it's written

Background

With cloud storage services, it is commonplace for data to be not only stored in the cloud, but also shared across multiple users. However, public auditing for such shared data — while preserving both data and identity privacy — remains to be an open challenge.

Problem

With cloud storage services, it is commonplace for data to be not only stored in the cloud, but also shared across multiple users. However, public auditing for such shared data — while preserving both data and identity privacy — remains to be an open challenge.

Contribution: Overview

In this paper, we propose the first privacy-preserving mechanism that allows public auditing on shared data stored in the cloud. In particular, we exploit ring signatures to compute the metadata needed to verify the integrity of shared data. With our mechanism, the identity of the signer of each data block is kept private from a third party auditor (TPA), who is still able to publicly verify the integrity of shared data without retrieving the entire file. In addition, we utilize homomorphic MACs and homomorphic hash functions to mitigate the overhead of storing our signatures.

Contribution: Highlights

In this paper, we propose the first privacy-preserving mechanism that allows public auditing on shared data stored in the cloud. In particular, we exploit ring signatures to compute the metadata needed to verify the integrity of shared data. With our mechanism, the identity of the signer of each data block is kept private from a third party auditor (TPA), who is still able to publicly verify the integrity of shared data without retrieving the entire file. In addition, we utilize homomorphic MACs and homomorphic hash functions to mitigate the overhead of storing our signatures.

Contribution: Highlights

In this paper, we propose the first privacy-preserving mechanism that allows public auditing on shared data stored in the cloud. In particular, we exploit ring signatures to compute the metadata needed to verify the integrity of shared data. With our mechanism, the identity of the signer of each data block is kept private from a third party auditor (TPA), who is still able to publicly verify the integrity of shared data without retrieving the entire file. In addition, we utilize homomorphic MACs and homomorphic hash functions to mitigate the overhead of storing our signatures.

Contribution: Additional features

In this paper, we propose the first privacy-preserving mechanism that allows public auditing on shared data stored in the cloud. In particular, we exploit ring signatures to compute the metadata needed to verify the integrity of shared data. With our mechanism, the identity of the signer of each data block is kept private from a third party auditor (TPA), who is still able to publicly verify the integrity of shared data without retrieving the entire file. In addition, we utilize homomorphic MACs and homomorphic hash functions to mitigate the overhead of storing our signatures.

Validation by experiments

Our experimental results demonstrate the effectiveness and efficiency of our proposed mechanism when auditing shared data.



The introduction

Writing perfect papers: the art

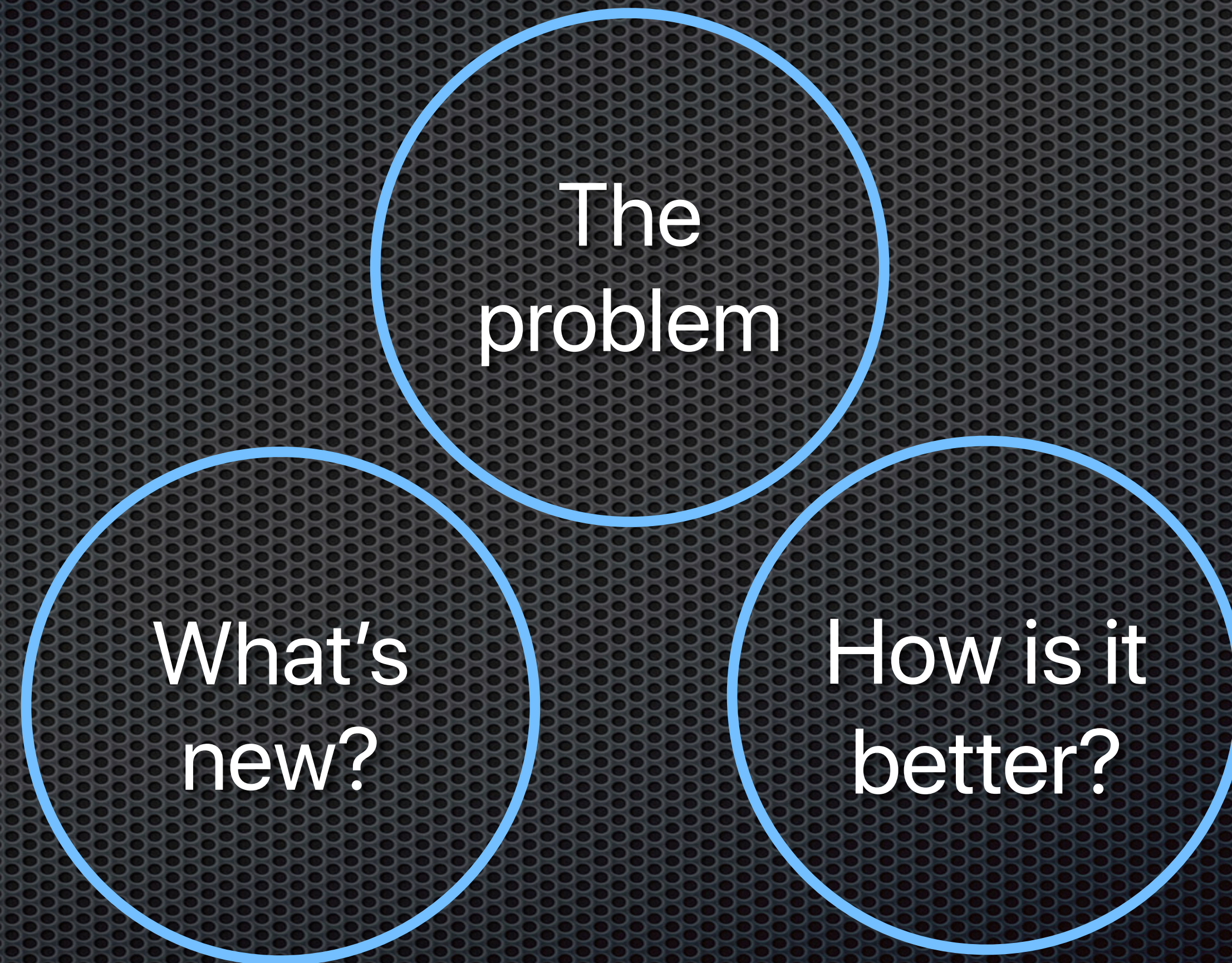
If your title and abstract can get a reader to
read the introduction, you are half way there

The remaining job: impress the reader
with an exciting and clear introduction

The introduction is so important that I
rewrote the introduction in many of my
papers

It is important because we wish our readers to understand our ideas, even if they only read the introduction and nothing else

Remember the three essential elements



Writing a good introduction

It should be self-contained, so that a reader short on time doesn't need to read the rest of the paper

It should be clear **what the problem is**

It should be easy to understand **why our solution is new, and how it's better** than previous work

Writing a good introduction

Don't make it too long

Don't spend too much space on the background and related work

There can be a separate "Related Work" section

Don't make your work sound more original than it really is

Introduction: a typical structure

First (opening) paragraph

A **general overview** of the research field — basic facts needed to “warm up” the reader and to prepare for the problem statement

Not too long — 2-3 sentences are good enough

For our auditing paper, here is the first paragraph that the student author at the time wrote —

Overview: First draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. For instance, a group of user can share and edit the same document simultaneously using Google Apps. Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Overview: First draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. For instance, a group of user can share and edit the same document simultaneously using Google Apps. Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Overview: First draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. For instance, a group of user can share and edit the same document simultaneously using Google Apps. Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Overview: First draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. For instance, a group of user can share and edit the same document simultaneously using Google Apps. Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Overview: First draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. For instance, a group of user can share and edit the same document simultaneously using Google Apps. Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Now let's find out all the problems in this paragraph together

This paragraph is simply too long — the author is asking the reader to “decipher” it

The first two sentences tried to warm-up
the reader with an overview

And the last three sentences actually tried to describe the problem — the **problem statement**

In this case, they should be separated
into **two** paragraphs

The second paragraph: **problem statement**

A clear description of the challenges to be addressed and the **problems** to be solved

Appeal to the intuition of the reader

Describe at a high level

It is fine (and even a good idea) to include an **intuitive solution** to the problem

Getting back to our example

Overview: Second draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. For instance, a group of user can share and edit the same document simultaneously using Google Apps.

Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

The two paragraphs are not balanced in length

And the second paragraph is pretty hard
to understand

Let's first take a look at the first paragraph
to see if we can improve it further

Overview: Second draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. For instance, a group of user can share and edit the same document simultaneously using Google Apps.

Overview: Second draft

Nowadays, data storage **and** sharing **is** one of the most popular **service** in cloud computing. For instance, a group of **user** can share and edit the same document simultaneously using Google Apps.

Grammatical mistakes

Overview: Second draft

Nowadays, data storage and sharing is one of the most popular service in **cloud computing**. For instance, a group of user can share and edit the same document simultaneously using Google Apps.

What's "cloud computing?" Not perfectly self-contained

Overview: Second draft

Nowadays, data storage and sharing is one of the most popular service in cloud computing. **For instance**, a group of user can share and edit the same document simultaneously using Google Apps.

Poor transition: "for instance" with respect to what?

Now let me show you the first paragraph
I wrote

Overview: Second draft

Cloud service providers manage an enterprise-class infrastructure that offers a scalable, secure, and reliable environment for users, at a much lower marginal cost due to the sharing nature of resources.

It is routine for users to use cloud storage services to share data with others in a team, as data sharing becomes a standard feature in most cloud storage offerings, including Dropbox and Google Docs.

First, "warm-up" to cloud computing

Overview: Second draft

Cloud service providers manage an enterprise-class infrastructure that offers a scalable, secure, and reliable environment for users, at a much lower marginal cost due to the **sharing nature** of resources. It is routine for users to use cloud storage services to share data with others in a team, as data sharing becomes a standard feature in most cloud storage offerings, including Dropbox and Google Docs.

At the same time, "warm-up" to the idea of sharing

Overview: Second draft

Cloud service providers manage an enterprise-class infrastructure that offers a scalable, secure, and reliable environment for users, at a much lower marginal cost due to the sharing nature of resources. It is routine for users to use cloud storage services to **share data with others** in a team, as data sharing becomes a standard feature in most cloud storage offerings, including Dropbox and Google Docs.

Formally introduce the idea of "sharing data"

Now let's take a look at the **next two** paragraphs that the student wrote (after detaching the original first paragraph)

Problem statement: First draft

Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Second paragraph

Problem statement: First draft

To protect the integrity of the cloud data, an efficient method is to enable public auditing by introducing a third party auditor (TPA). This public auditor should have more powerful computation and communication abilities than regular users and is able to offer security services to cloud users [1].

First part of the third paragraph

Observation: both are somewhat related to the challenges of storing data in the cloud

The **intuitive** idea is that data stored in the cloud may be lost or corrupted, and the users don't know

So a third-party auditor can be used to check the integrity of data on behalf of the users

But what the student wrote is asking a reader to do the hard work to “decipher” the writing and get the idea

The problem with his writing was that he
tried to mix the basic idea with
unnecessary details

Problem statement: First draft

Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Redundant sentence

Problem statement: First draft

Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Three citations to the same paper in the same paragraph

Problem statement: First draft

Although the cloud can provide more reliable services than personal devices, the integrity of the cloud data still faces challenging security threats [1]. Due to human errors, device failures or internal/external attacks, the outsourced data on the cloud may easily lost or corrupted [1]. Another threat for the integrity of the cloud data is that the cloud service provider may delete rarely used data without informing users to save the storage space and reduce the commercial cost on those data [1].

Unnecessary details

Why can't we just describe the **basic idea** of the problem and **its existing solution** intuitively — and remove all the details that make it harder to read?

Unnecessary details inhibit the **natural**
and **brisk** flow of the main ideas

Now let me show you the second
paragraph I wrote

Problem statement: Second draft

The integrity of data in cloud storage, however, is subject to skepticism and scrutiny, as data stored in an untrusted cloud can easily be lost or corrupted, due to hardware failures and human error [1]. To protect the integrity of cloud data, it is best to perform public auditing by introducing a third party auditor (TPA), who offers its auditing service with more powerful computation and communication abilities than regular users.

Highlighting the problem: "the integrity of data"

Problem statement: Second draft

The integrity of data in cloud storage, however, is subject to skepticism and scrutiny, as data stored in an untrusted cloud can easily be lost or corrupted, due to hardware failures and human error [1]. To protect the integrity of cloud data, it is best to perform public auditing by introducing a third party auditor (TPA), who offers its auditing service with more powerful computation and communication abilities than regular users.

Mentioning two of the several reasons of lost data

Problem statement: Second draft

The integrity of data in cloud storage, however, is subject to skepticism and scrutiny, as data stored in an untrusted cloud can easily be lost or corrupted, due to hardware failures and human error [1]. To protect the integrity of cloud data, it is best to perform public auditing by introducing a third party auditor (TPA), who offers its auditing service with more powerful computation and communication abilities than regular users.

Smooth transition to a typical solution: "TPA"

Now let's see the new 1st and 2nd
paragraphs in the second draft

Now, the two paragraphs are about the same length

Cloud service providers manage an enterprise-class infrastructure that offers a scalable, secure, and reliable environment for users, at a much lower marginal cost due to the sharing nature of resources. It is routine for users to use cloud storage services to share data with others in a team, as data sharing becomes a standard feature in most cloud storage offerings, including Dropbox and Google Docs.

The integrity of data in cloud storage, however, is subject to skepticism and scrutiny, as data stored in an untrusted cloud can easily be lost or corrupted, due to hardware failures and human error [1]. To protect the integrity of cloud data, it is best to perform public auditing by introducing a third party auditor (TPA), who offers its auditing service with more powerful computation and communication abilities than regular users.

You may say, oh this is hard!

But remember, we only went through the
first two paragraphs in the introduction

The gap between mediocre and excellent
writing: only a few words here and there

Returning to the typical structure of the
introduction

General overview to warm-up the reader (the opening paragraph)

State the problem (challenge) and existing solutions (1-2 paragraphs)

Limitations of existing solutions that motivate this paper (2-3 paragraphs)

Proposed solution: main idea (1-2 paragraphs)

Now it's time to write about what's new

What are the **original** highlights of the proposed solution? (1-2 paragraphs)

Why is the proposed solution different from and better than existing solutions?

State 1-2 most impressive highlights, not all of them

Make the originality of the paper crystal clear and stand out

You may use the sentence: "Highlights of our original contributions in this paper are as follows. First, ... Second, ... Finally, ..."

The list of original contributions drives
the **entire paper** — the rest of the paper
substantiates the claims you have made

Your readers may think: "Hey, if they can really deliver this, that'll be very exciting! I'd better read the rest of the paper."

It is a good idea to include a **table** to compare important properties of the proposed solution with its “direct competitors” in the existing literature, highlighting your **advantages**

Show a table

It is a good idea to include a **table** to compare important properties of the proposed solution with its “direct competitors” in the existing literature, highlighting your **advantages**

Use an example

It's an even better idea to show an
intuitive **example**

Use an example

Your **example** shows how your **main idea**
works in a **special case**

Use more examples

Continue with **more** examples
throughout the paper

Use more examples

After each theorem is proved — or each algorithm shown — explain their **implications** or **intuition** with an example

Use more examples

Your examples need to be really simple

Use more examples

Your examples need to be really simple

Use more examples

They help readers to understand your solution well

Use well-designed figures

It'll be a great idea to include a well-designed **figure** to illustrate your example

Use well-designed figures

Include as many figures as you can
throughout the paper

D

The flow of ideas

Writing perfect papers: the art

Just like the storyboard when making a movie,
the flow of ideas needs to be carefully designed

Make sure that each paragraph contains one **complete** idea — **split** a paragraph that contains more than one idea, and **merge** short paragraphs that cover the same idea

Flow of ideas

Where is the section on related work?

Flow of ideas

Some prefer to place it **after the introduction**

Rationale: the section can be used to “warm up” the reader

Some others prefer to place it before the conclusion

Rationale: After the introduction, the readers don't understand your main ideas yet, there's no point in talking about their differences from related work

It depends on the design of your flow of
ideas

Flow of ideas

Where do you write about experimental results?

Most papers collect all the experimental results and put them at the end

But you don't have to

Some negative or preliminary
experimental results may be important
for motivating the main idea in the paper

Your flow of ideas can then be: “initial results — idea to improve — more results to show better performance”

You can even **interleave** experimental results with descriptions of your idea, if this provides the best flow

Flow of ideas

How do you allocate space across sections?

The short answer is: no one knows it better than you, because it depends on the flow of ideas

General rules of thumb

Don't write a very long abstract (200 words),
introduction (one page), related work (half a page),
or concluding remarks (1-2 paragraphs)

Throughout the paper, make it self-contained, but
don't use a lot of space for unnecessary background

Keep the motivation short and concise

1.3

The art

Writing perfect papers

1.4

The details

Writing perfect papers



Use a **git repository** to
manage your workflow

Writing perfect papers: the details

Git repositories help you to track all histories and collaborate with others

Place everything — source code,
references — into your git repository

Use any of the public cloud services
such as [GitHub](#)

Use GitHub Desktop

Current Repository
wenting-ieeeenetwork18

Current Branch
master

Fetch origin
Last fetched 7 days ago

Changes

History

Select Branch to Compare...

Wenting Wei • Aug 21, 2020

Merge branch 'master' of https://github.com/baoc...
Baochun Li • Aug 21, 2020

Revised the conclusion.
Baochun Li • Aug 21, 2020

Uploaded Figure 4 again.
Wenting Wei • Aug 21, 2020

Added Orca in Figure 4.
Wenting Wei • Aug 21, 2020

Finished Section 3.
Baochun Li • Aug 21, 2020

Revised To Do list in README.md.
Baochun Li • Aug 21, 2020

Finished the discussions of Orca.
Baochun Li • Aug 21, 2020

Added details about 3 new references in Table 2.
Wenting Wei • Aug 21, 2020

Added citation information about 3 new references.
Wenting Wei • Aug 20, 2020

Added new references.
Wenting Wei • Aug 20, 2020

Added Orca to Table 2 and refreshed To Dos.
Baochun Li • Aug 20, 2020

Finished discussing MVFST-RL.
Baochun Li • Aug 20, 2020

Started to work on MVFST-RL.
Baochun Li • Aug 20, 2020

Finished the discussions of R3Net.

Finished Section 3.

Baochun Li 2c7f75f 3 changed files

datacenter.tex

drl.tex

intro.tex

65

Performance-wise, after only 6 hours of training using emulated network environments, Orca is able to achieve the best performance to date in a variety of typical network environments, compared with both traditional hand-tuned heuristics (such as BBRv2) and recent DRL-based congestion control protocols (such as Aurora). It also incurs very little overhead, on par with hand-tuned heuristics such as TCP CUBIC and BBR. With respect to fairness, it is friendly to competing TCP CUBIC flows, most likely because it uses TCP CUBIC itself as the underlying congestion control protocol, and therefore does not show aggressive behavior when trying to saturate the available bottleneck bandwidth.

66

67

- {\bf Remarks.} While there exists several alternative congestion control protocols based on deep reinforcement learning in the literature \cite{eagle-infocom20,mptcp-jsac19}, the protocols we have sampled in this paper are able to represent the state-of-the-art fairly well. For a summary comparison between existing DRL-based congestion control protocols, refer to Table \ref{tab:4} for key differences between Aurora, R3Net, MVFSF-RL, Eagle \cite{eagle-infocom20}, and DRL-CC \cite{mptcp-jsac19}, and Orca. In addition, Figure \ref{fig:drl-cc} provides an illustrative comparison with respect to the performance metrics that they included in their state spaces, as well as the actions they are designed to take.

65

Performance-wise, after only 6 hours of training using emulated network environments, Orca is able to achieve the best performance to date in a variety of typical network environments, compared with both traditional hand-tuned heuristics (such as BBRv2) and recent DRL-based congestion control protocols (such as Aurora). It also incurs very little overhead, on par with hand-tuned heuristics such as TCP CUBIC and BBR. With respect to fairness, it is friendly to competing TCP CUBIC flows, most likely because it uses TCP CUBIC itself as the underlying congestion control protocol, and therefore does not show aggressive behavior when trying to saturate the available bottleneck bandwidth.

66

67

+ {\bf Open challenges.} While there exists several alternative congestion control protocols based on deep reinforcement learning in the literature (such as Eagle \cite{eagle-infocom20} and DRL-CC \cite{mptcp-jsac19}, the protocols we have sampled in this paper are able to represent the state-of-the-art fairly well. For a summary comparison between existing DRL-based congestion control protocols, refer to Table \ref{tab:4} for key differences between Aurora, R3Net, MVFSF-RL, Eagle \cite{eagle-infocom20}, DRL-CC \cite{mptcp-jsac19}, and Orca. In addition, Figure \ref{fig:drl-cc} provides an illustrative comparison with respect to the performance metrics that they included in their state spaces, as well as the actions they are designed to take.

68

69



B

Eliminate typographical
and grammatical mistakes

Writing perfect papers: the details

First, eliminate all spelling mistakes by running your paper through a spell checker

Use a UNIX **command-line tool** such as spell
to check spelling — not Microsoft Word

Then fix grammar and usage problems
by proofreading

Proofreading also helps you to fix the remaining spelling mistakes that a spell checker cannot catch

Example: instead of "must," you wrote "mist"

Proofreading also helps you to fix the remaining spelling mistakes that a spell checker cannot catch

Example: instead of "must," you wrote "mist"

You cannot rely on someone else —
certainly not the reviewers — to
proofread for you, it's your paper!



Use **transitional** words,
phrases, and sentences

Writing perfect papers: the details

Use **transitions** across the boundary of sentences, paragraphs and sections

Important for readers to follow your flow
of ideas

If you don't add transitions, readers will need to
add them **subconsciously** in their mind, anyway

You are asking readers to do the hard
work!

Examples of transitions

Connecting two halves of a sentence —

as, since, or else

Examples of transitions

Connecting sentences —

However, In addition, Further, Nevertheless,
Fortunately, Unfortunately, Surprisingly,

To make matters worse, to further exacerbate the
problem, The bad news is,

The implications are two-fold, It turns out that, As an
example, To take ... a step further,

Examples of transitions

Connecting paragraphs and sections —

It only remains to see...

The simple answer to this question is,

The only challenge that remains now is,

To address this challenge,

We first present...

Next, we evaluate...

We are now ready to...



D

Use correct English

Writing perfect papers: the details

Keep punctuation marks inside the
closing quotation mark

...making it a “shared secret key”.

...making it a “shared secret **key**”.

Keep the punctuation mark inside
the closing quotation mark

...making it a "shared secret **key**".

Keep the punctuation mark inside
the closing quotation mark

...making it a “shared secret key.”

Don't use long sentences

Don't use long sentences with more than one comma in the middle of the sentence — and abuse "where," "in which," "whose," "so that," "such that"

Don't be too colloquial

Don't use words that are too informal and colloquial

"a lot of" is more colloquial than "a large number of"

Instead of "big," use "substantial" or "large"

Don't be too formal either

Overly formal words and phrases sometimes feel awkward

"We endeavour to ascertain that..." — "We show that..."

"It can be ascribed to..." — "It is due to..."

"The overwhelming quantity of..." — "The exceedingly large number of..."

Don't use words that are too emotional

An "easy" solution to solve the problem above is to share the private key of the original user with other group members as a group private key, however, it is like "suicide"...

A sentence from the first draft of our example paper

Similar emotional words: "kill," "crazy," "happy,"
"fantastic," "marvellous," or "breathtaking"

An "easy" solution to solve the problem above is to share the private key of the original user with other group members as a group private key, however, it is like "suicide"...

A sentence from the first draft of our example paper

Countable vs. uncountable nouns

If a word is countable, when using its plural form, remember to use "a few," "a number of," "fewer"

Instead of "less bits," use "fewer bits"

If not, do not invent its plural form (such as "performances," "advices," "equipments," "informations"), and use "less" or "lower" rather than "fewer"

Agreement of the verb with the subject

"The figures above **shows**" vs. "The figures above **show**"

"These problems that lead to lower efficiency **shows** that" vs.

"... **show** that"

"Its efficiency in energy savings **are** remarkable" vs. "... **is** remarkable"

Slows down reading — your paper can be rejected just because of these problems!

Watch out on common words and phrases

It almost takes a lifetime to master the usage of common words, such as "with", "of", "for", "at", "against", "in", "on"

Or to master the large number of phrases in English
"in other words" or "in another word"?

Oxford Dictionary of English (macOS)

Dictionary (26 found)

<>A A

perfect

AllEnglishEnglish ThesaurusSimplified ChineseSimplified Chinese - EnglishAppleWikipedia

perfect

perfect binding

perfect cadence

perfect competition

perfect crime

perfect fifth

perfect fifths

perfect fourth

perfect fourths

perfect gas

perfect market

perfect pitch

perfect square

perfect storm

perfecta

perfectibility

perfection

perfectionism

perfectionist

perfective

perfectly

perfecto

perfecter

perfectible

perfectionistic

perfectness

perfect

adjective | 'pə:fɪkt |

1

having all the required or desirable elements, qualities, or characteristics; as good as it is possible to be: *life certainly isn't perfect at the moment | a perfect summer's day.*

- free from any flaw or defect in condition or quality; faultless: *the equipment was in perfect condition.*
- precisely accurate; exact: *a perfect circle.*
- highly suitable for someone or something; exactly right: *with a little help you can create a room that is perfect for you | the perfect present for golfers everywhere.*
- dated* thoroughly trained in or conversant with: *she was **perfect in** French.*

2

[*attributive*] absolute; complete (used for emphasis): *a perfect stranger | all that Joseph said made perfect sense to me.*

3

Mathematics (of a number) equal to the sum of its positive divisors, e.g. the number 6, whose divisors (1, 2, 3) also add up to 6.

4

Grammar (of a tense) denoting a completed action or a state or habitual action which began in the past. The perfect tense is formed in English with *have* or *has* and the past participle, as in *they have eaten* and *they have been eating* (**present perfect**), *they had eaten* (**past perfect**), and *they will have eaten* (**future perfect**).

5

Botany (of a flower) having both stamens and carpels present and functional.

- Botany* denoting the stage or state of a fungus in which the sexually produced spores are formed.
- Entomology* (of an insect) fully adult and (typically) winged.

verb | pə'fɛkt | [*with object*]

make (something) completely free from faults or defects; make as good as possible: *he's busy perfecting his bowling technique.*

- archaic* bring to completion; finish: *then urg'd, she perfects her illustrious toils.*
- complete (a printed sheet of paper) by printing the second side: *the heap was normally printed as white paper in the morning, turned at the midday break, and perfected in the afternoon.*
- Law* satisfy the necessary conditions or requirements for the transfer of (a gift, title, etc.): *equity will not perfect an imperfect gift.*

noun | 'pə:fɪkt | (the perfect) Grammar

the perfect tense.

DERIVATIVES

perfecter | 'pə:fɛktə | noun

perfectible | pə'fɛkɪb(ə)l | adjective

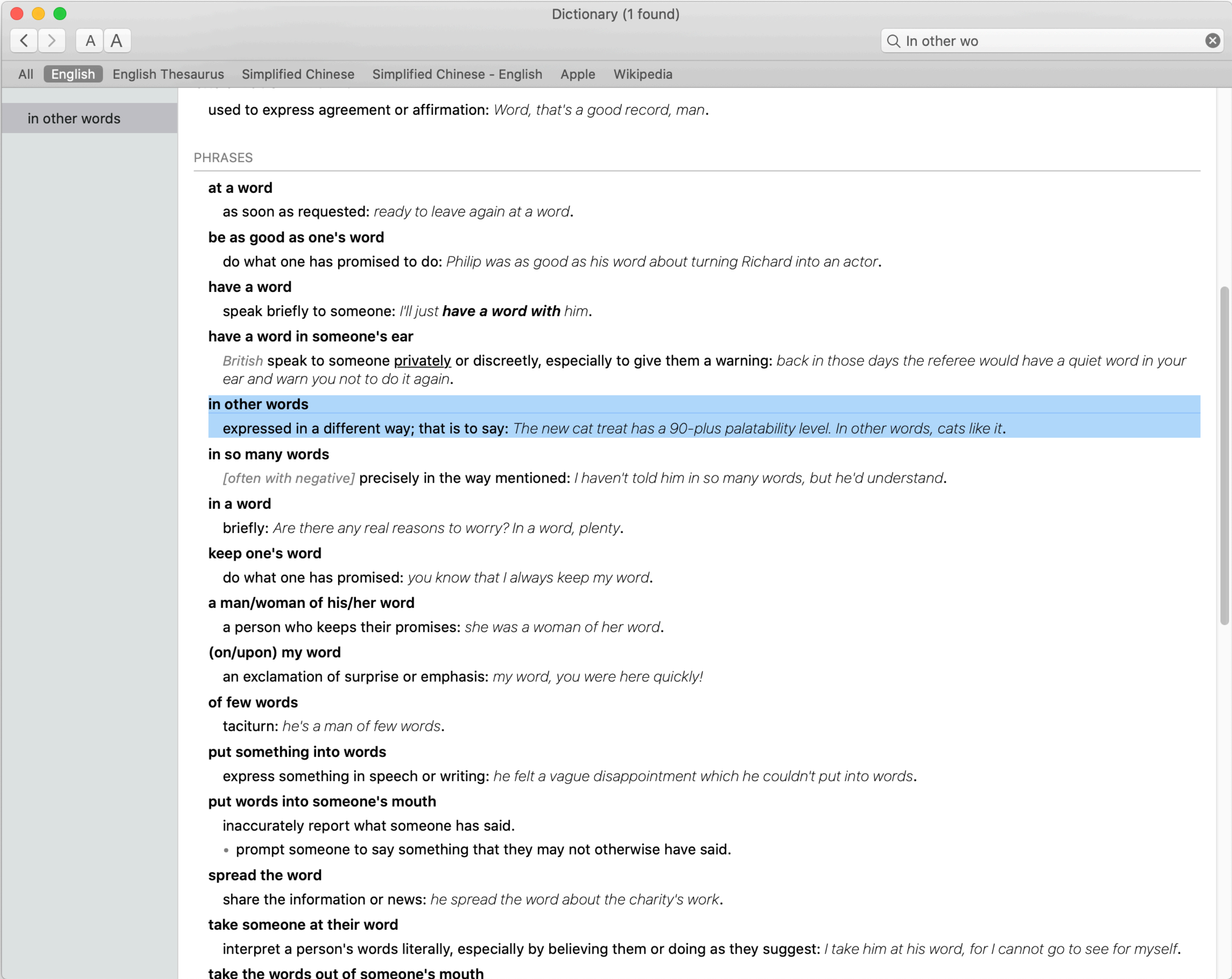
perfectness | 'pə:fɪktnəs | noun

ORIGIN

Middle English: from Old French *perfet*, from Latin *perfectus* 'completed', from the verb *perficere*, from *per-* 'through, completely' + *facere* 'do'.

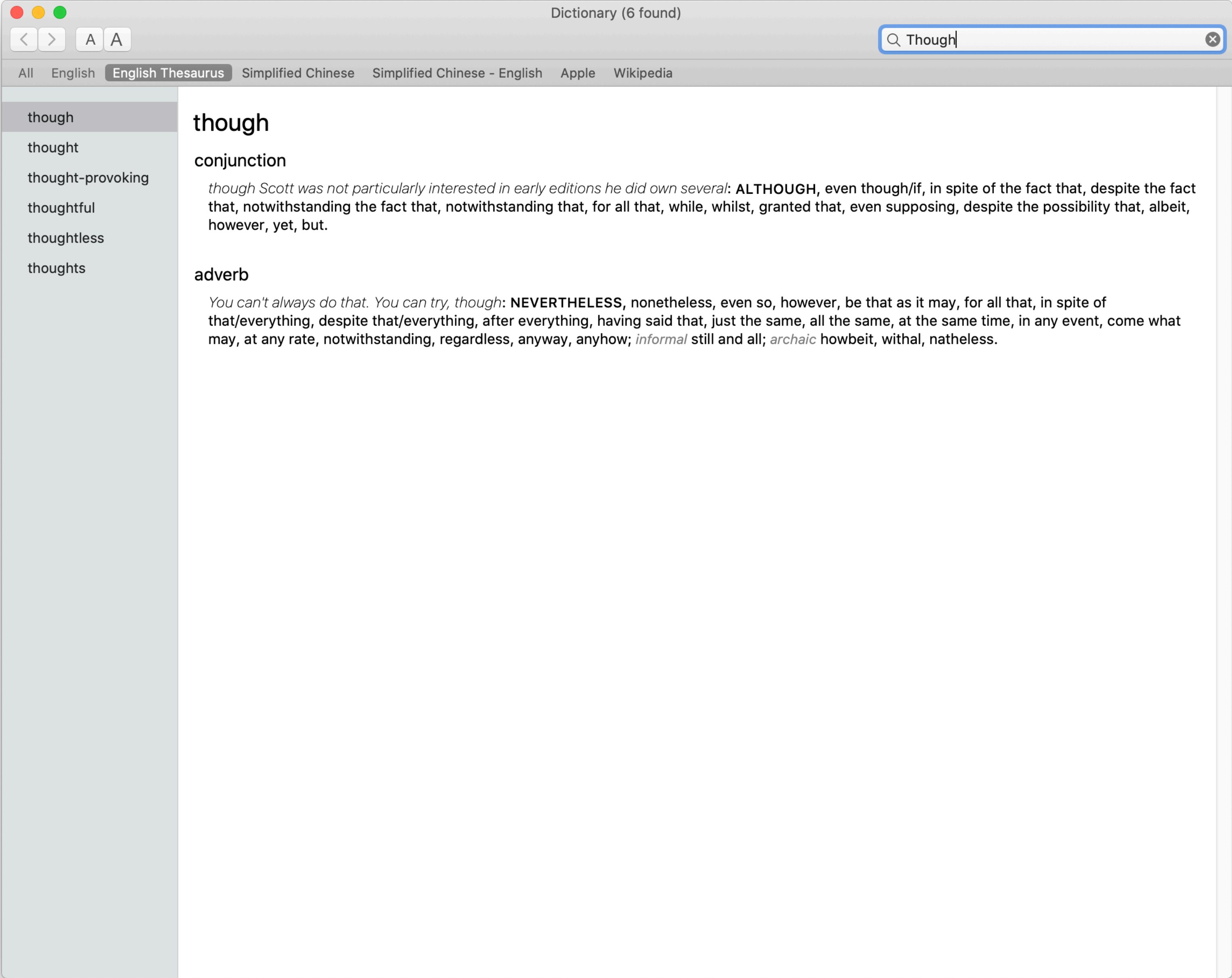
Example sentences

Oxford Dictionary of English (macOS)



Phrases

Oxford Dictionary of English (macOS)



Thesaurus

Articles as determiners

The articles "a" / "an" (the indefinite article) and "the" (the definite article) are frequently used incorrectly by new students

Plural nouns are typically used without an article:

"The source node receives acknowledgments" (not "the acknowledgments")

The indefinite article is weaker than the definite article:

"a large portion of" (not "the large portion of")

But there's no need for "the" in section titles — instead of "The System Model," just say "System Model."

Articles can be tricky to use, but there are too many of them in a paper — **pay attention!**

Let's take a look at one example
sentence

The extent to which the users can effectively communicate with the service providers depend on the size of community.

The extent to which **the users** can effectively communicate with the service providers depend on the size of community.

Plural noun: "users" —
the definite article is not needed

The extent to which the users can effectively communicate with the **service providers** depend on the size of community.

Plural phrase: "service providers"

The extent to which the users can effectively communicate with the **service providers** depend on the size of community.

Plural phrase: "service providers" —
the definite article is not needed, but it feels
strange without it, too

The extent to which the users can effectively communicate with **their service providers** depend on the size of community.

It would only be correct to use "their"
as the determiner

The extent to which users can effectively communicate with their service providers depend on the size of community.

Good enough?

The extent to which users can effectively communicate with their service providers depend on the size of community.

The subject needs to agree with the verb

The extent to which users can effectively communicate with their service providers depends on the size of community.

The subject needs to agree with the verb

The extent to which users can effectively communicate with their service providers depends on the size of community.

Perfect?

The extent to which users can effectively communicate with their service providers depends on the size of community.

The noun "community" needs a determiner in front of it

The extent to which users can effectively communicate with their service providers depends on the size of a community.

The extent to which users can effectively communicate with their service providers depends on the size of a community.

The community in question is a specific community, not an arbitrary one

The extent to which users can effectively communicate with their service providers depends on the size of the community.

The extent to which users can effectively communicate with their service providers depends on the size of the community.

This is good enough, but it's even better to reinforce the idea of which community is being discussed

The extent to which users can effectively communicate with their service providers depends on the size of their community.

The extent to which users can effectively communicate with their service providers depends on the size of their community.

Perfect!

Good news: most English problems are not hard to fix — just proofread every sentence with plenty of time!



Typeset your paper
correctly and beautifully

Writing perfect papers: the details



E.1

Use **LaTeX**, no matter what

You do need to spend some time
learning how to use **LaTeX**

But the typesetting results are
dramatically better

No ligature

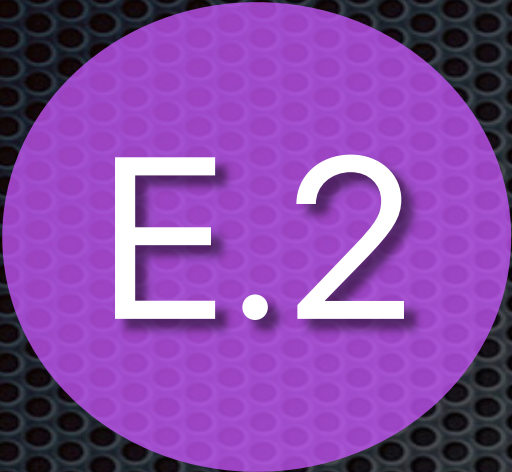
Efficient

efficient

Correct ligature

In fact, my mind doesn't work efficiently
when reading a paper typeset in Word

There is a huge amount of useful information
about LaTeX on the web — no excuse!



E.2

Draw figures using a **vector-based** application

Image-based drawing applications
produce images that become fuzzy
when zoomed in on a reader's iPad

Recommendation: draw.io

<https://diagrams.net>

Search Shapes

▼ Scratchpad

Drag elements here

▼ General

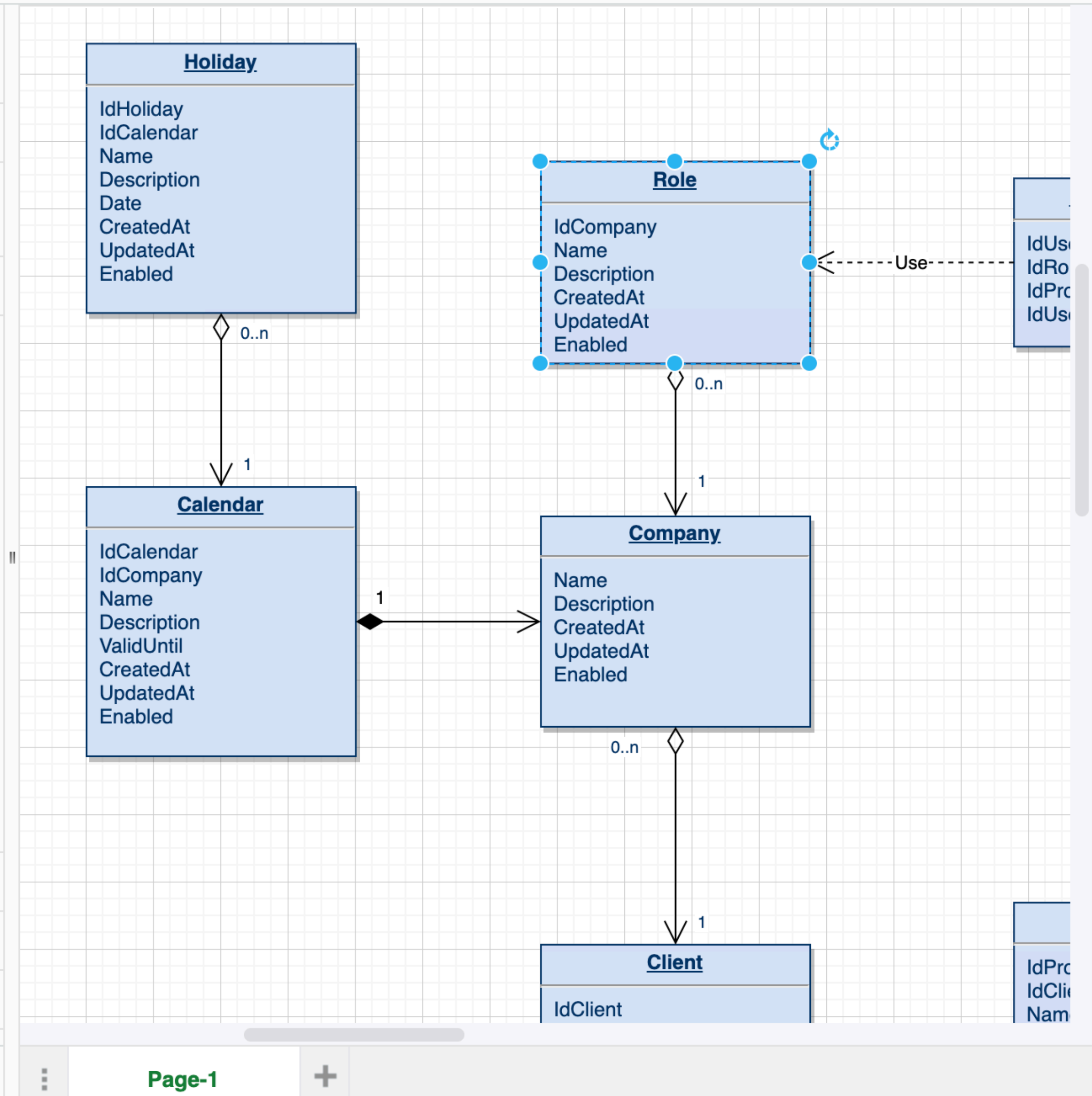
[illegible]

► Misc

► Advanced

► UML

+ More Shapes...



Style	Text	Arrange
<div> <div> <div></div> <div></div> <div></div> <div></div> </div> <div> <div></div> <div></div> <div></div> <div></div> </div> </div> <div> <div></div> <div></div> <div></div> <div></div> <div></div> <div></div> </div>		
<div> <input checked="" type="checkbox"/> Fill <div></div> </div> <div> <input type="checkbox"/> Gradient </div>		
<div> <input checked="" type="checkbox"/> Line <div></div> </div> <div> <div> <div></div> <div></div> </div> <div> <div>1 pt</div> <div></div> </div> </div> <div> <div>Perimeter</div> <div> <div>0 pt</div> <div></div> </div> </div>		
<div> <div>Opacity</div> <div> <div>100 %</div> <div></div> </div> </div>		
<div> <input type="checkbox"/> Rounded <input checked="" type="checkbox"/> Shadow </div> <div> <input type="checkbox"/> Glass <input type="checkbox"/> Sketch </div>		
<div> <div>Edit Style</div> <div>Edit Image</div> </div> <div> <div>Copy Style</div> <div>Paste Style</div> </div> <div> <div>Set as Default Style</div> </div>		
<div> <div>Property</div> <div>Value</div> </div>		



E.3

Proofread your bibliography and
make it consistent

Use BibTeX to typeset the bibliography in
your paper

After downloading BibTeX entries from the web, proofread to get them consistent

Keep a **consistent** style of abbreviating journal
and conference titles throughout the bibliography

Try your best to find out the venue where the paper was published in, rather than citing arxiv.org

1

Writing perfect papers

1.1

The story

Writing perfect papers

1.2

The work

Writing perfect papers

1.3

The art

Writing perfect papers

1.4

The details

Writing perfect papers

2

Writing good papers fast!

Writing **good** papers is hard

That's why you are advised to write
slowly

In the good old days, academics can take
years to write **one** research paper

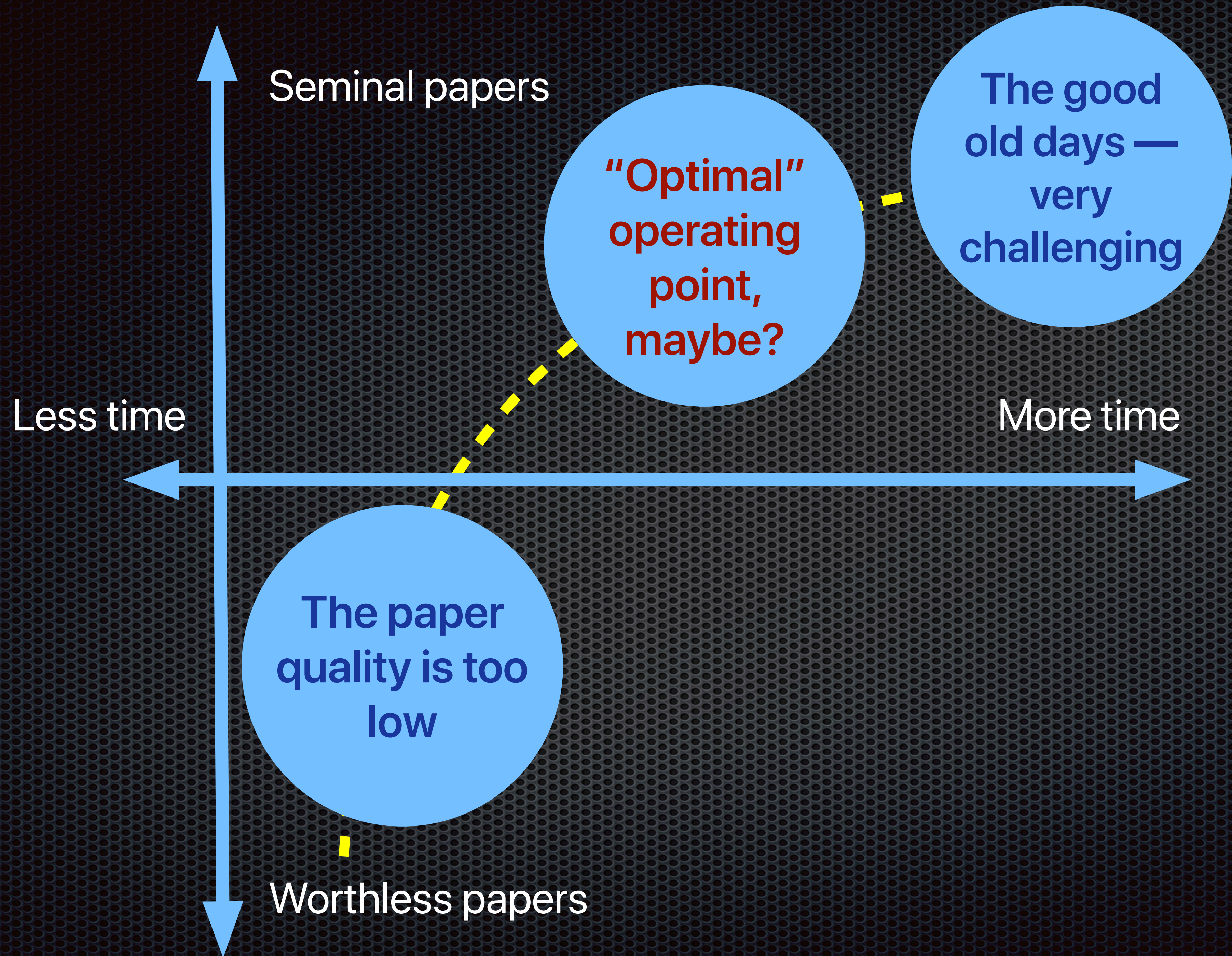


But we have arxiv.org today, with a
“timestamp” on papers

We also have the “publish or perish” culture, where the number of papers is an indication of quality

I need to graduate — tell me how to write
papers quickly

If others travel at the speed of sound, you
wish to be **supersonic**



$$\max Q(p)$$

$$\text{subject to } t(p) \leq T$$

2

Writing good papers fast

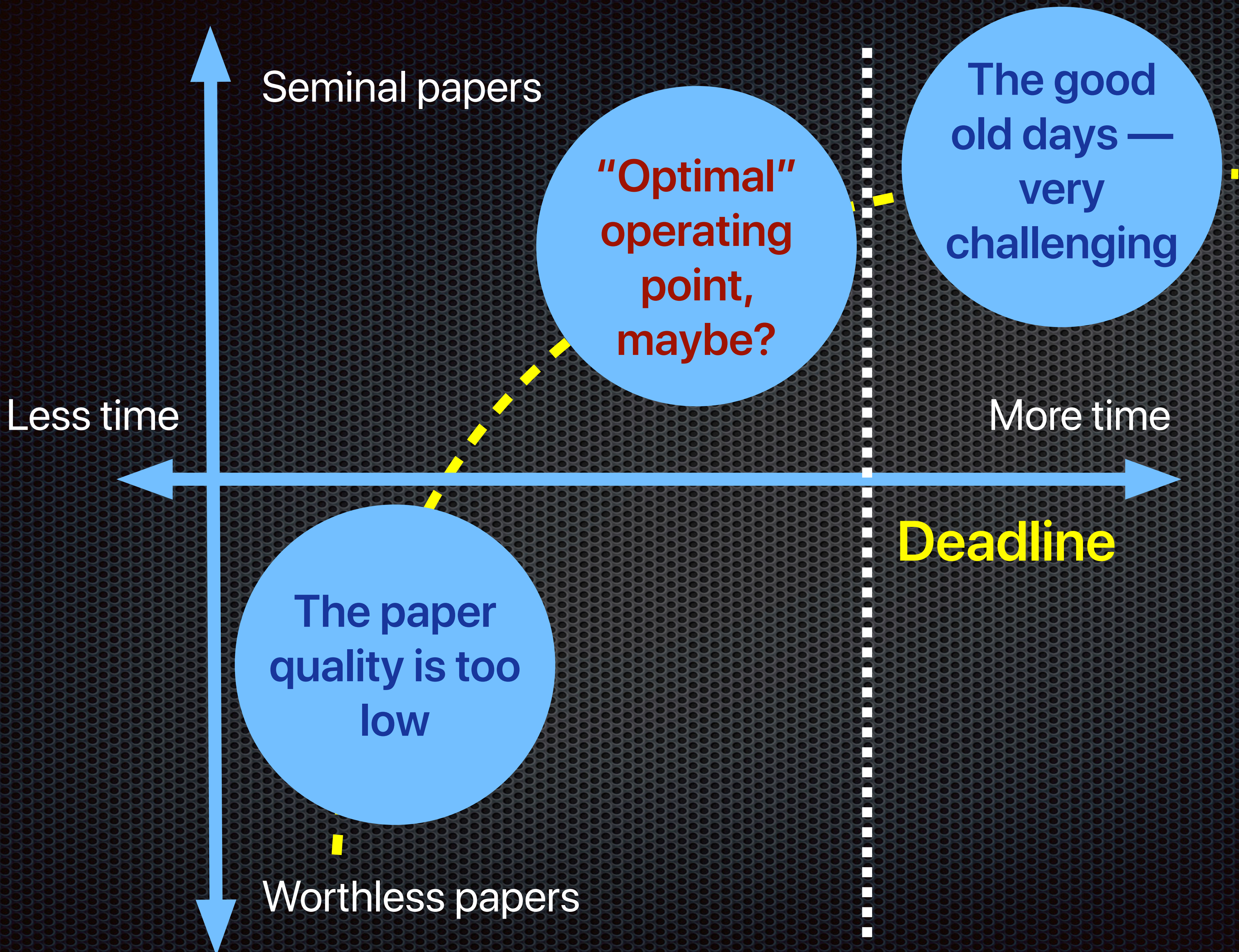
I am going to talk about **three** things

2.1

Know your **deadline**

Writing good papers fast

$$t(p) \leq T$$



At all times, always work towards a paper
deadline

Deadlines are all equal

It doesn't matter what kind of deadline it is —

a conference **deadline**

a **deadline** from a journal special issue

a **deadline** you set up yourself for submitting

a journal paper

dead · line — you will have to make it no
matter what happens

But how do I **guarantee** that I will make
my deadline?



A

Set your deadline
realistically

Writing good papers fast: know your deadline

Admit that you are not able to write a good
paper in a month or two

But a deadline **six months later** may not carry a lot of weight either — unless you are extremely disciplined



B

Make it public knowledge

Writing good papers fast: know your deadline

Tell as many people as you can when your
deadline will be

So that it feels like a **public disgrace** if
you fail to meet the deadline

As a researcher, you sure wish to be seen
as **responsible** and **dependable**

You work around the clock to make sure that others who depend on you will have the peace of mind that the work will be done

And done by the deadline



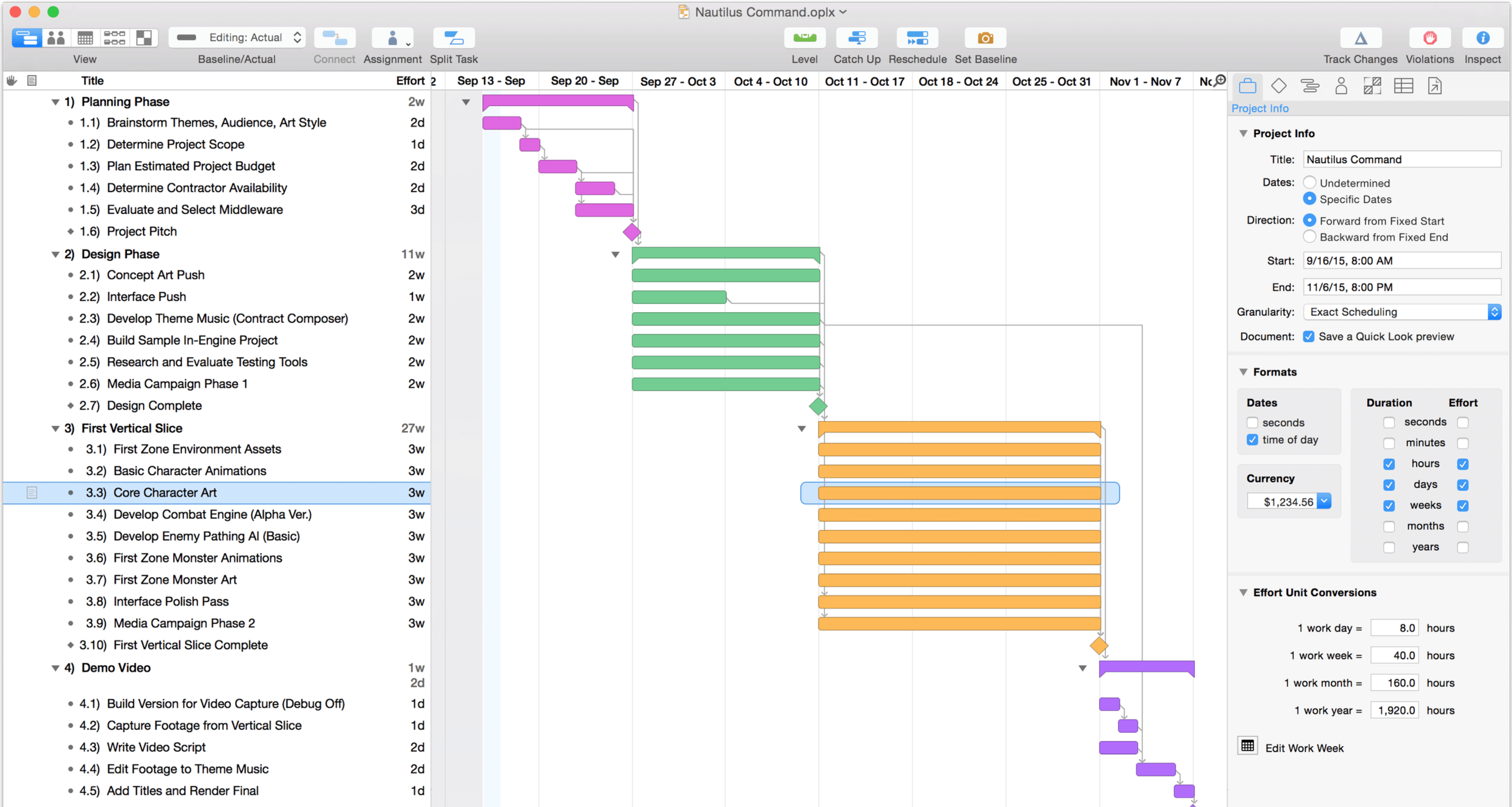
Allocate your time **well**

Writing good papers fast: know your deadline

Design a rough timeline at the very beginning

Assign deadlines to each milestone in
your timeline

Use a Gantt chart



If necessary, revise your timeline along
the way

But think seriously why you missed your
milestone deadlines

Don't procrastinate and delay all the
work to the days before the deadline

It may be feasible, but it will be **stressful**
and it **degrades** the quality of your work

2.2

Don't follow, **lead!**

Writing good papers fast

You are the first author of the new paper,
right?

Then you are the **leader**, the **boss**, and the person who cares most about this work

Start your team at the beginning of your project

Only invite **trustworthy** collaborators
when building your team

Your collaborators include your **advisors**
— get them to work for you

Get them to commit their time

Warn them of the **tough** work ahead
along the way

Promise them an **exciting** outcome

Assign your collaborators to milestones
— they are your “resources”

Remind them about **their** milestone
deadlines **often**

Persist if they forget about your requests

Cancel their “membership” in your team
if they consistently fail

Have **backup plans** if they fail to deliver
by the deadline

2.3

Choose a **problem** with the **right size**

Writing good papers fast

It must be **feasible** to be completed
before the **deadline**

Goal: make solid improvements over one paper, and one paper only

Choose a paper that you **enjoyed** reading

Or even better, a paper with **source code**

If its source code is not on **GitHub**, ask
the authors

If you don't have source code, estimate
the amount of time to **reproduce** the work

If it may take a long time, **give up** and
choose another paper

Read the paper along with its source
code

The source code contains secret ingredients in its “recipe”

It helps you to understand the paper
completely and **deeply**

And only with a complete and deep
understanding can you **create something new**

And create something new **quickly**

To do it, you will have to have a **deadline** in
your timeline for **creating something new**

A deadline forces you to think **long** and
hard

Revise the source code and do a
feasibility test for your idea

If its feasibility is promising, think about its theoretical properties — anything provably correct?

Your little problem must **fit into the scope**
of the conference you'll submit to

Goal: get the paper accepted

Many conferences have their **favoured**
research topics

2

Writing good papers fast!

2.1

Know your **deadline**

Writing good papers fast

2.2

Don't follow, **lead!**

Writing good papers fast

2.3

Choose a **problem** with the **right size**

Writing good papers fast

That's it!

Good luck with your papers!

Baochun Li

liqua.ece.toronto.edu