CSE227 – Graduate Computer Security

Side Channels

UC San Diego

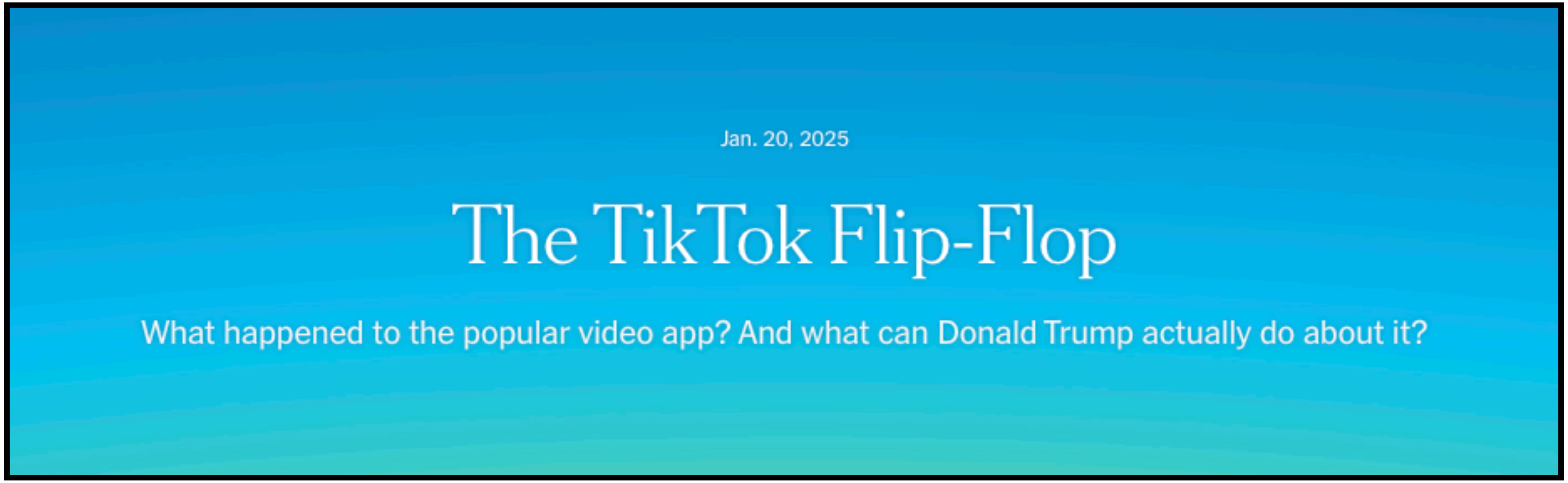
Housekeeping

General course things to know

- Course projects
  - I will provide some initial thoughts and feedback on each of your project ideas by **this Friday** via email
  - Start meeting with your teams, ideating, and reaching out to me if you have things you want to chat about
- All of you have notecards now (courtesy of me)

News

TikTok went and came back again in the US



Jan. 20, 2025

The TikTok Flip-Flop

What happened to the popular video app? And what can Donald Trump actually do about it?

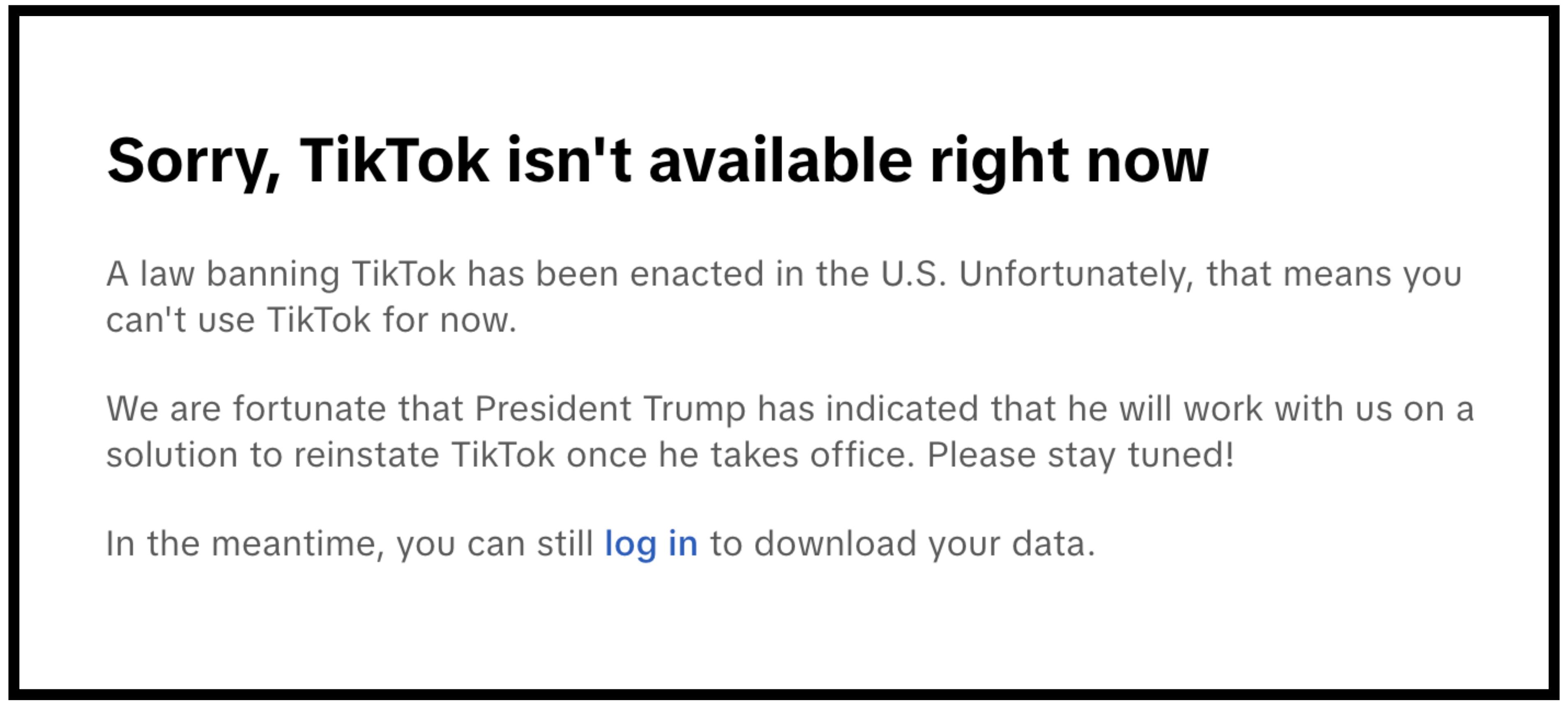
What happened?

- In March / April of 2024, U.S. lawmakers decided to focus on the most pressing issue facing Americans: TikTok
    - A bill was passed that would either ban TikTok or force a sale of TikTok to an American entity
    - Key issue is that TikTok is owned by ByteDance, a Chinese company, and there were concerns of data privacy leakage and potential influence + persuasion control from a foreign adversary (i.e., China)
    - Cite "national security" as the biggest reason for this
- Other countries already do this: Pakistan, Afghanistan, India, etc.

What happened?

- The sale did not happen, so TikTok was slated to be shut down on Sunday, January 19th – service was halted

- 12 hours later, service was restored: notably **before Trump was sworn into office** (i.e., he wasn't president)

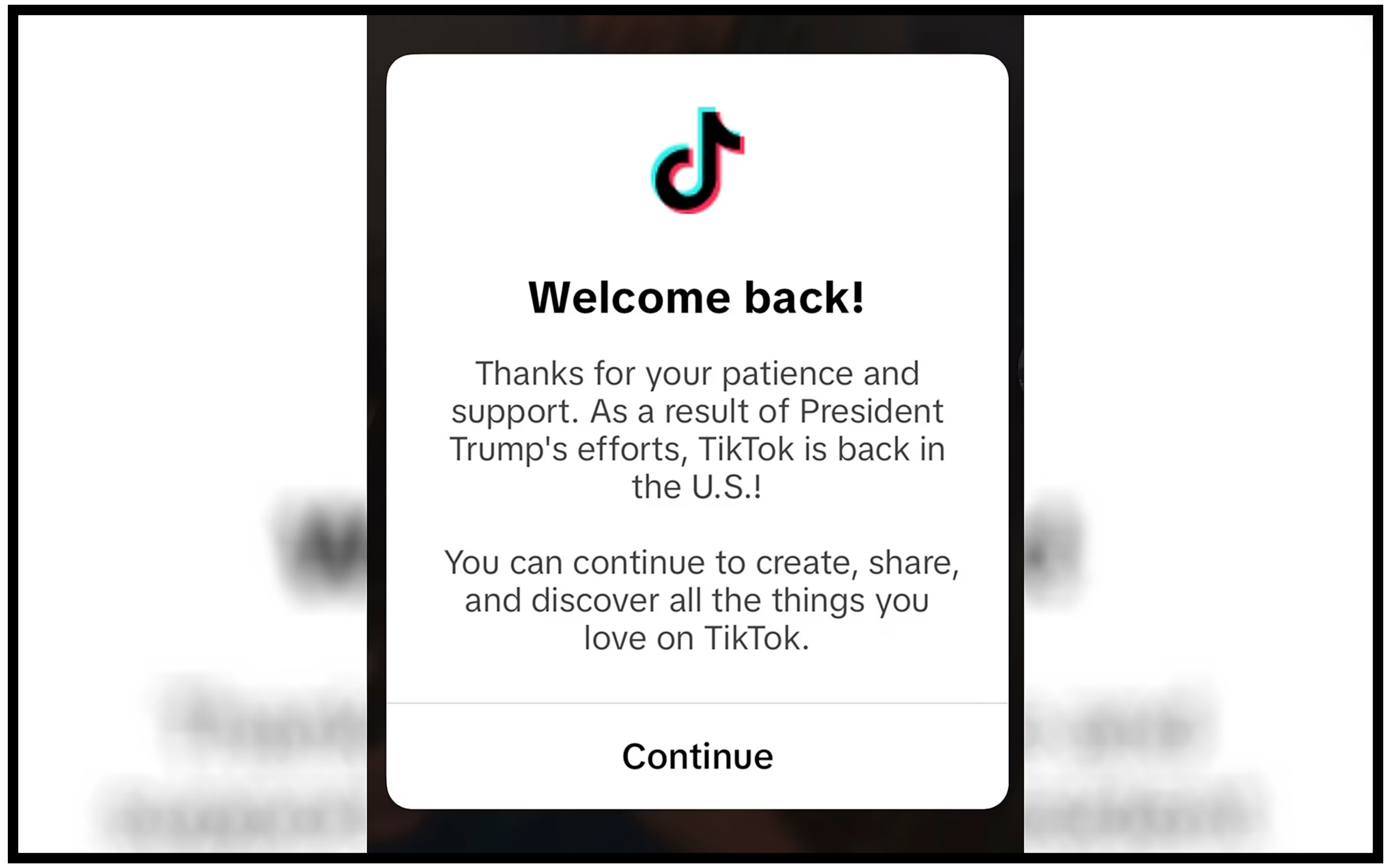


Sorry, TikTok isn't available right now

A law banning TikTok has been enacted in the U.S. Unfortunately, that means you can't use TikTok for now.

We are fortunate that President Trump has indicated that he will work with us on a solution to reinstate TikTok once he takes office. Please stay tuned!

In the meantime, you can still log in to download your data.



Welcome back!

Thanks for your patience and support. As a result of President Trump's efforts, TikTok is back in the U.S.!

You can continue to create, share, and discover all the things you love on TikTok.

What are these national security issues?

- TikTok collects data on its users
  - So do lots of American companies…. aka every social media app
- TikTok is owned by China, whose gov't will get all this data
  - Chinese gov't already has your data, because American third-party data brokers sell it to them routinely :)
- TikTok has enormous influence + persuasion control over Americans, in particular youth
  - See: Instagram, Snapchat, Facebook, etc.

Will banning TikTok solve our problems?

My take on the matter

- Banning TikTok before this weekend felt mostly like "security theater" – today is feels like both security *and* political theater
- The federal government **does not** have nationwide *privacy legislation* or meaningful *social media* legislation – both of which would be much more effective here
- "Data brokers" remain a huge problem to user privacy, but no one is doing anything about it because the US govt also buys data from data brokers :)

Today

Today's lecture – Side Channels

Learning Objectives

- Learn what a side channel is, why side channels exist, and how side channels gets implemented or exploited in practice
- Walk through some modern versions of side channels and understand how to find them in the wild
- Discuss the "keyboard emanations" attack
- Discuss the "Cold boot" attack on DRAM

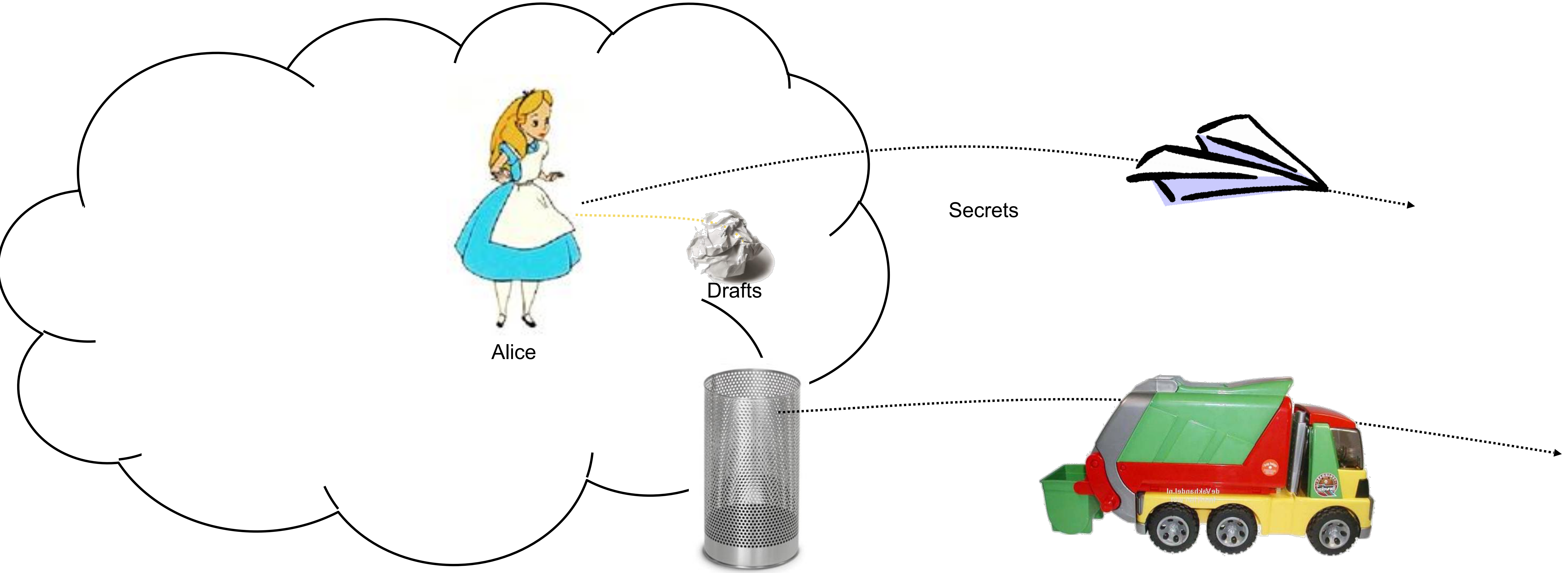
Preliminaries

Our typical model of security

Safe havens of goodness
The Badlands
Secrets
Alice the Good
Bob the Good
Zurg the Evil

Things are more complicated in practice…

Side channel: Secrets may leak *outside* the protocol because of how its implemented in practice



Side Channel Scenario

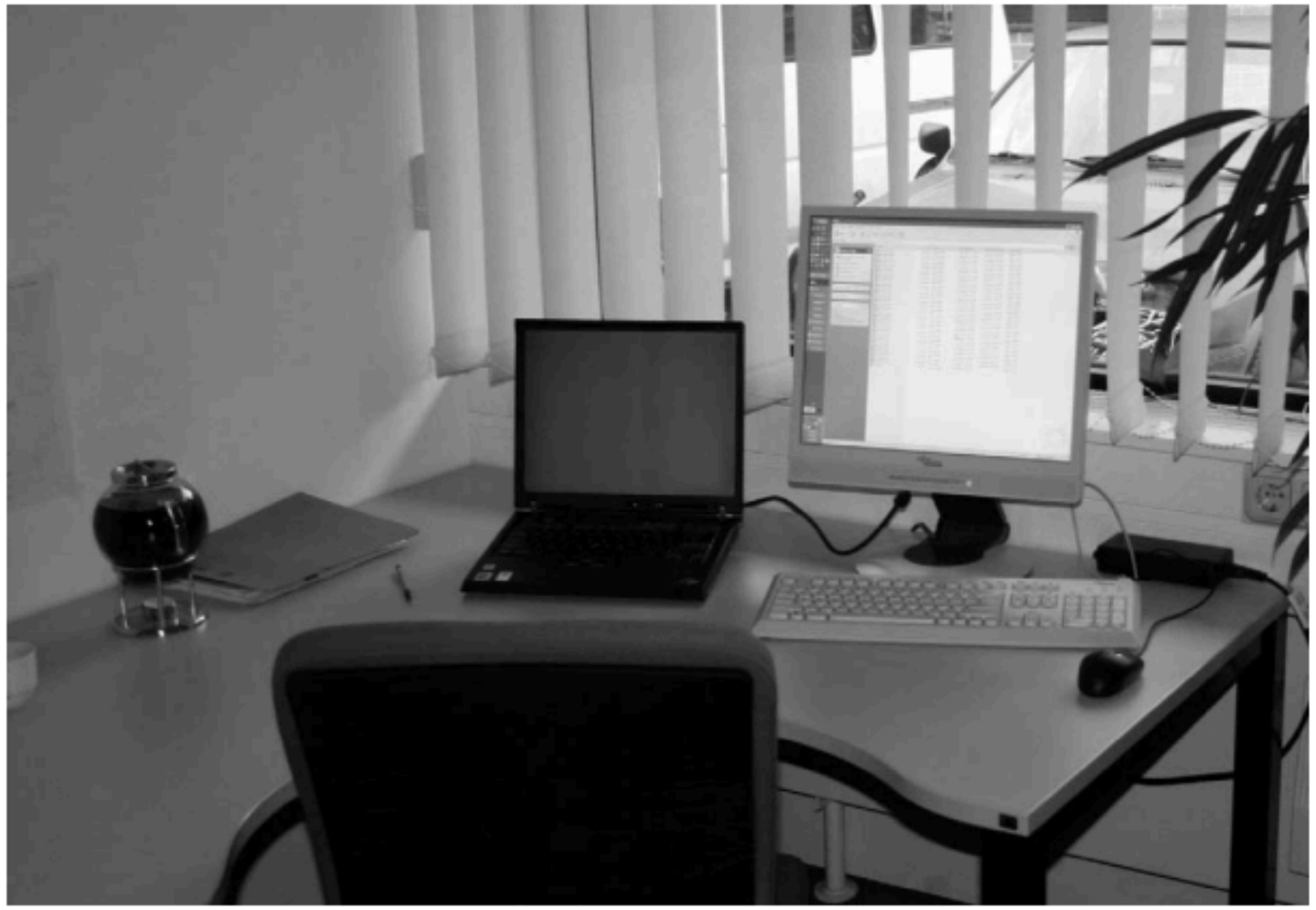


Figure 2. The basic setting: The monitor faces away from the window in an attempt to hide the screen's content.

From: Compromising Reflections – or – How to Read LCD Monitors Around the Corner

How might you read the computer screen from the window?

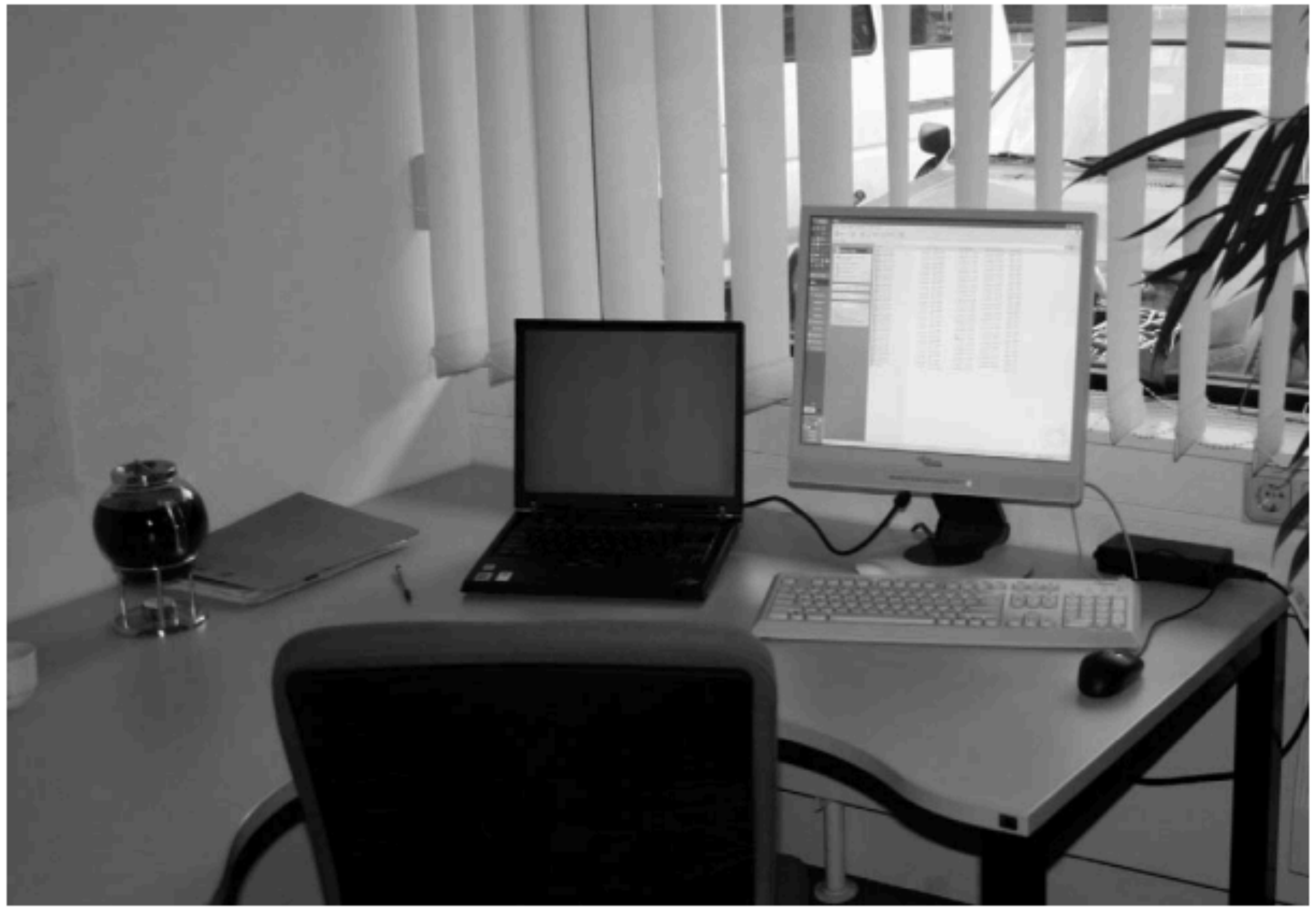


Figure 2. The basic setting: The monitor faces away from the window in an attempt to hide the screen's content.

From: Compromising Reflections – or – How to Read LCD Monitors Around the Corner

Reflections are your friend!

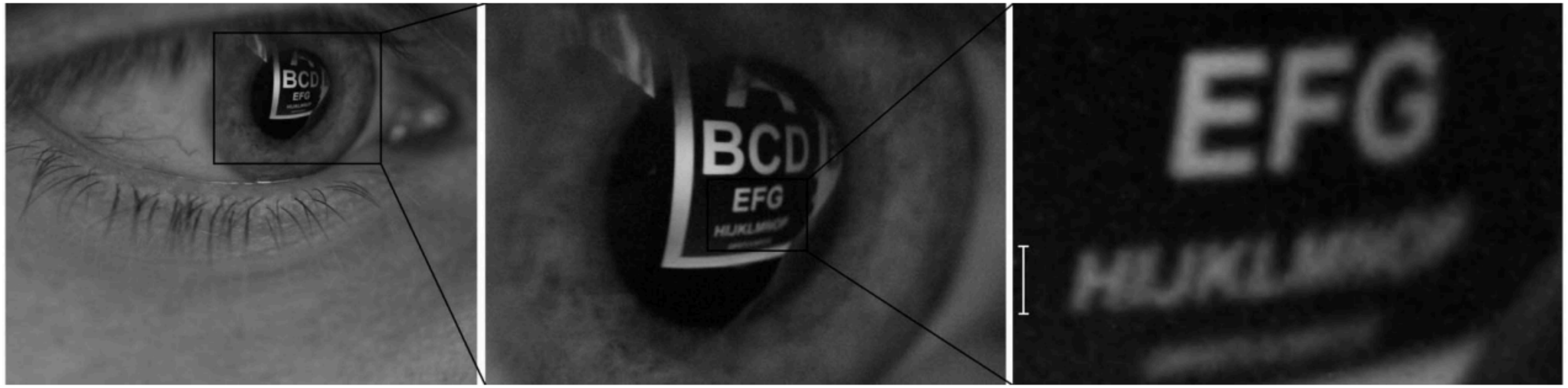


Figure 1. Image taken with a macro lens from short distance; the distance between the eye and the monitor was reduced for demonstration. Readability is essentially limited by the camera resolution.

From: Compromising Reflections – or – How to Read LCD Monitors Around the Corner

Reflections are your friend!

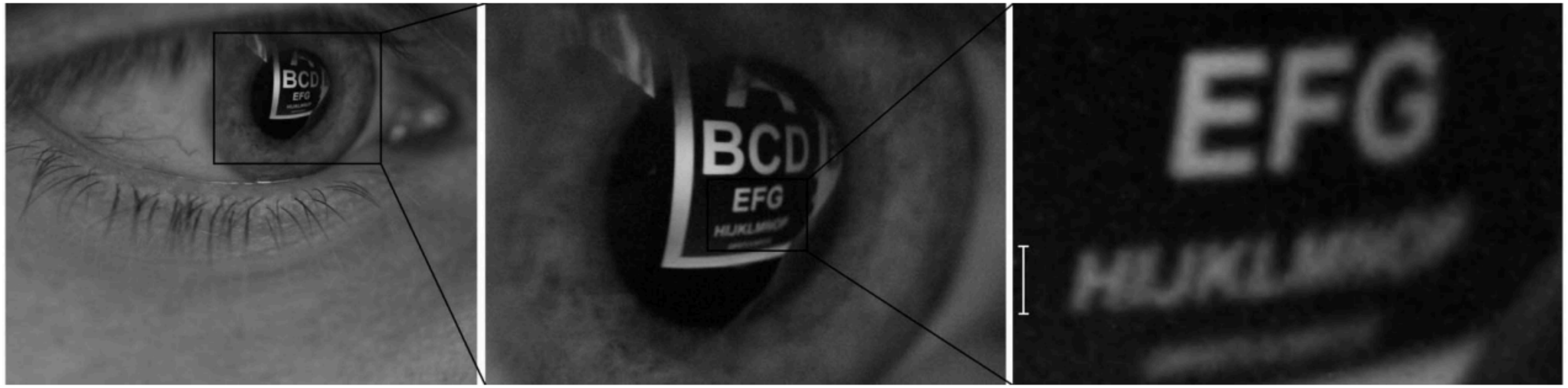


Figure 1. Image taken with a macro lens from short distance; the distance between the eye and the monitor was reduced for demonstration. Readability is essentially limited by the camera resolution.

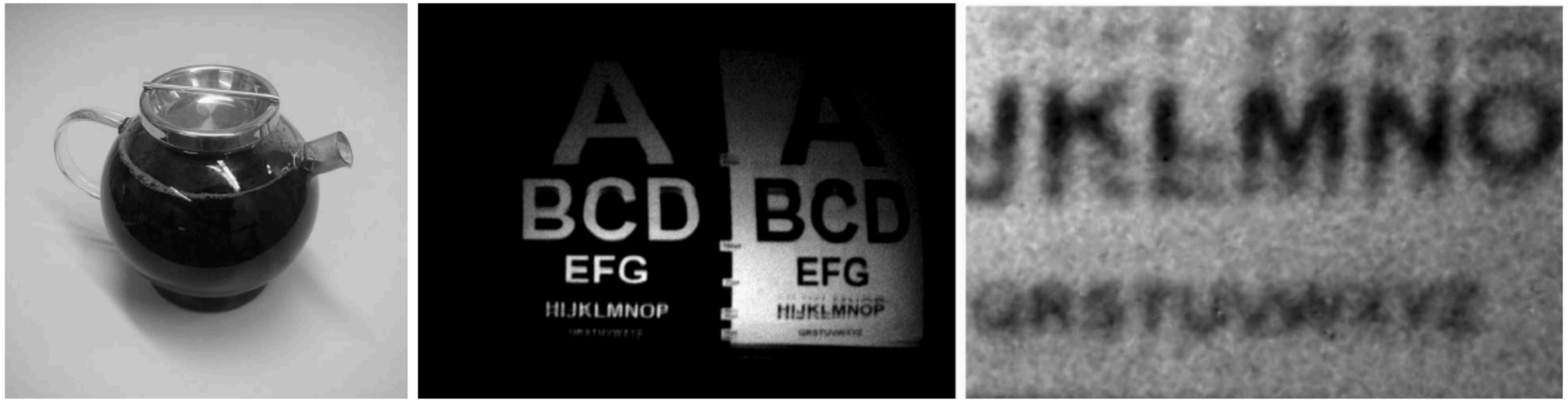


Figure 5. Reflections in a tea pot, taken from a distance of 10m. The 18pt font is readable from the reflection.

The craziest reflection of them all…

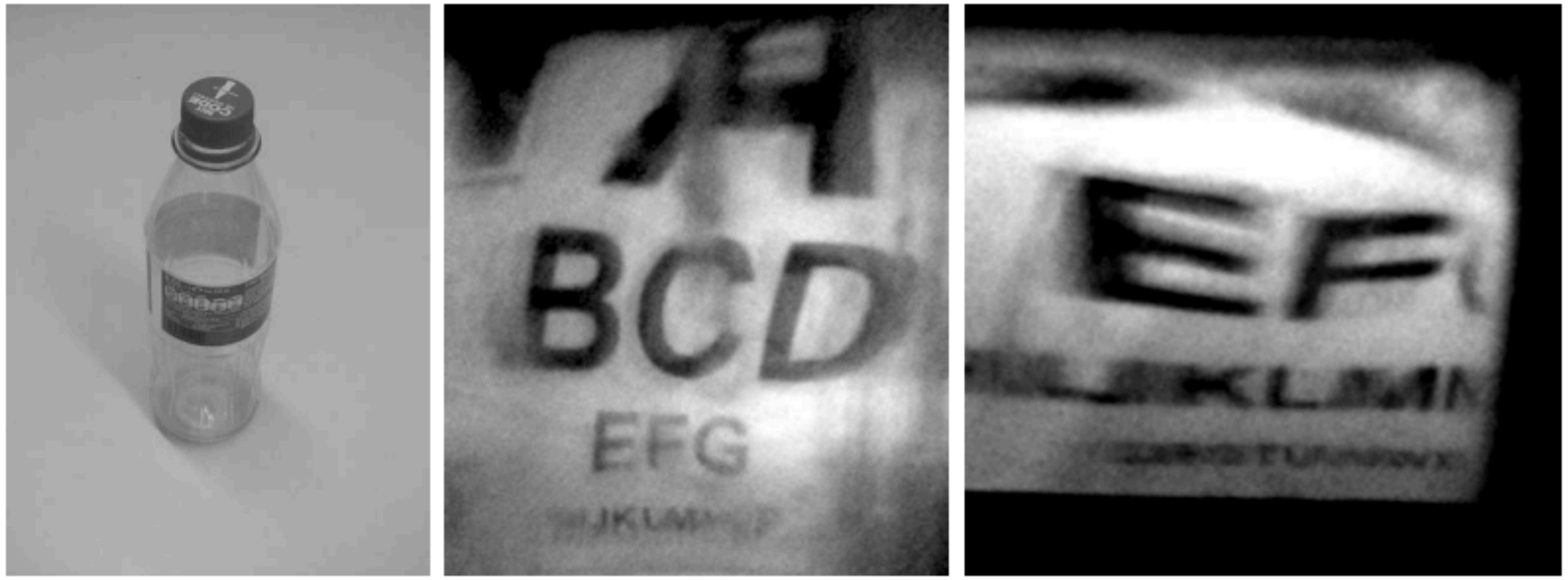


Figure 12. Reflections in a 0.5l plastic Coca-Cola bottle, taken from a distance of 5m. Because of the irregular surface, only parts of the text are readable.

From: Compromising Reflections – or – How to Read LCD Monitors Around the Corner

Keyboard Acoustic Emanations

What is an acoustic emanation?

What is an acoustic emanation?

Acoustic emanation: A sound made by an object in the course of normal use of that object

Keyboard acoustic emanations

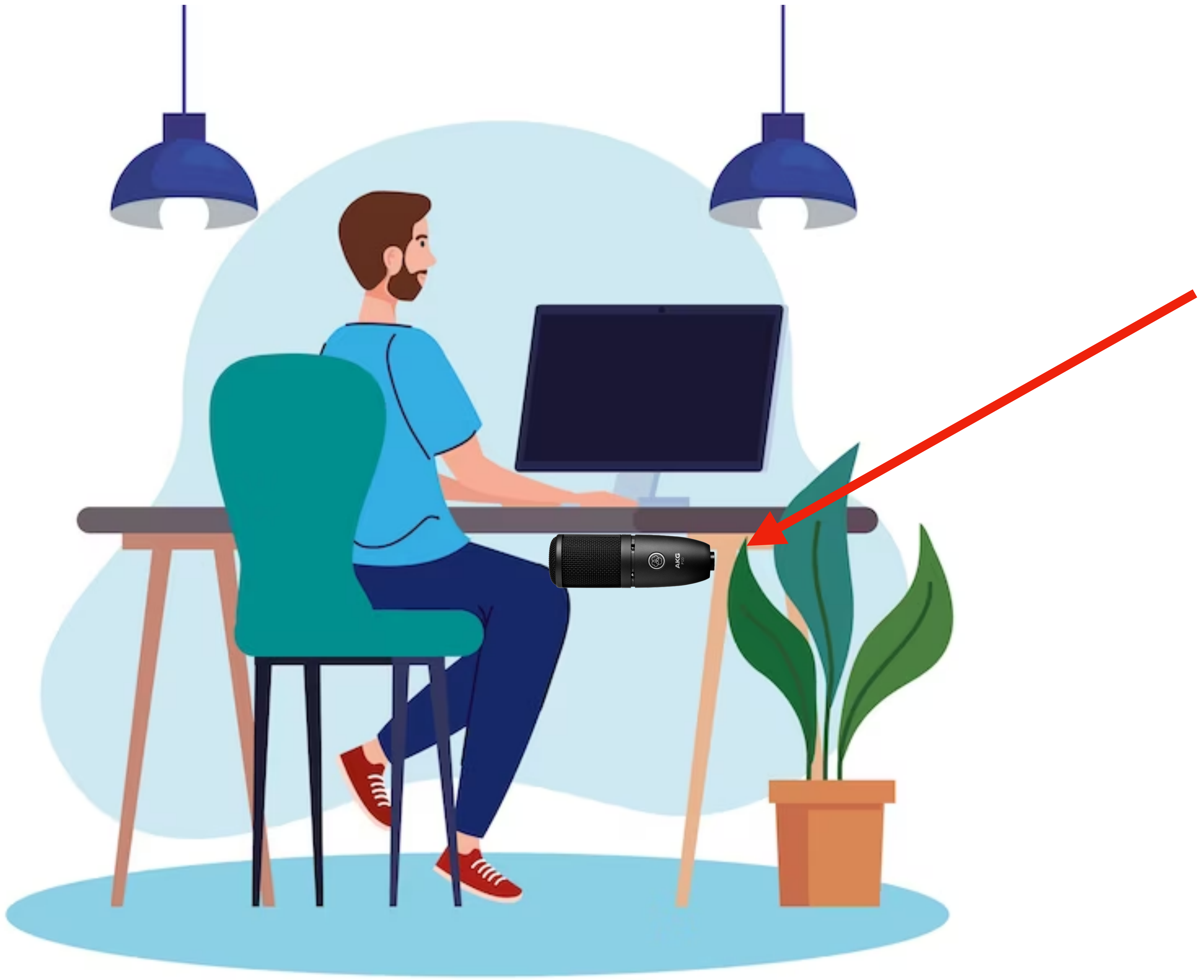
What is the attack the authors want to conduct?

Keyboard acoustic emanations

What is the attack the authors want to conduct?

Keyboard acoustic emanations

What is the attack the authors want to conduct?



Covertly or overtly record keystrokes

Keyboard acoustic emanations

What is the attack the authors want to conduct?

Attacker can use keystrokes to recover passwords or other secrets

Keyboard acoustic emanations

What is the attack the authors want to conduct?

What makes the attack possible? What’s the side channel here?

Keyboard acoustic emanations

What is the attack the authors want to conduct?

What makes the attack possible? What's the side channel here?

Why do keystrokes make different sounds?

Keyboard acoustic emanations

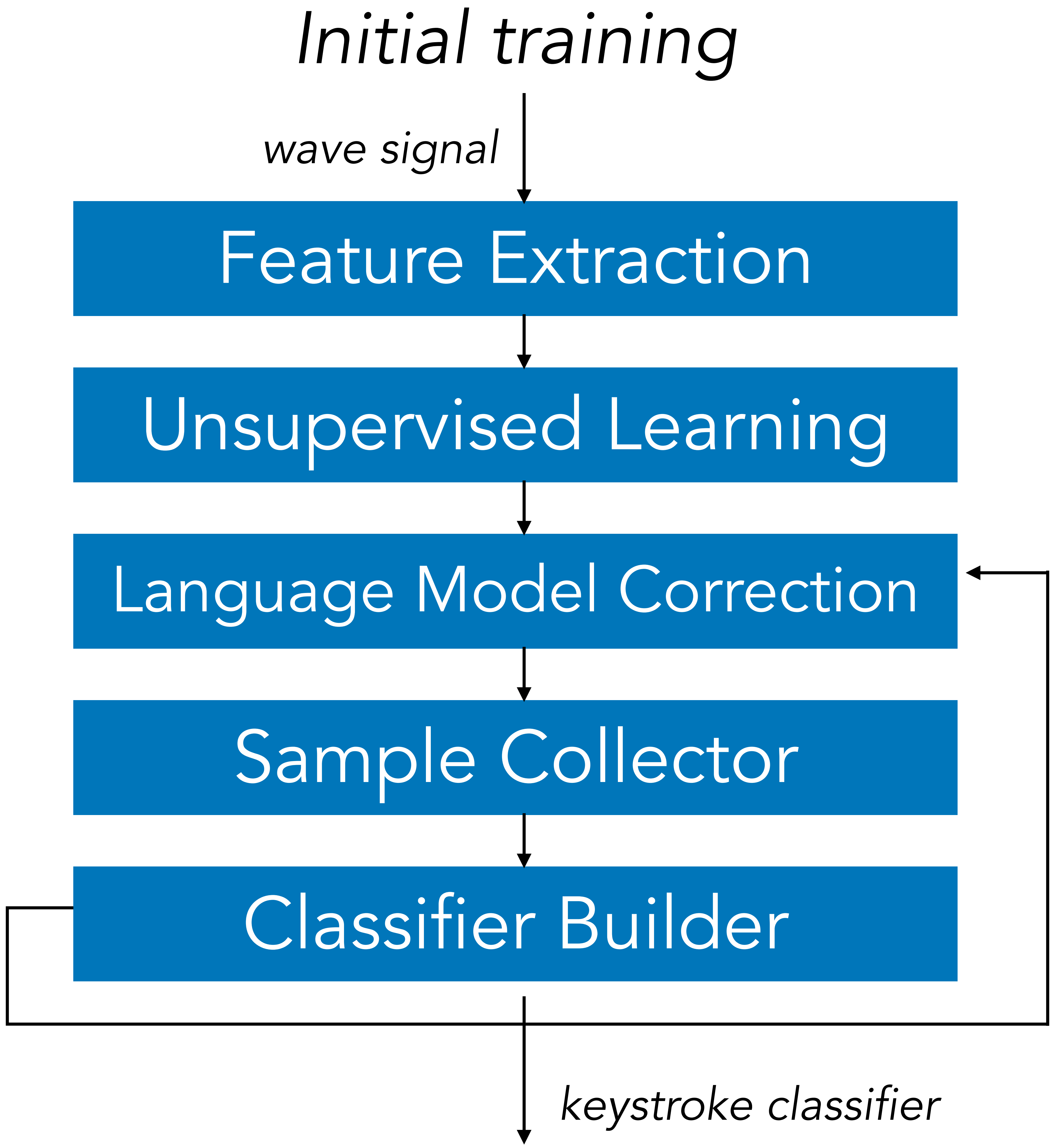
What is the attack the authors want to conduct?

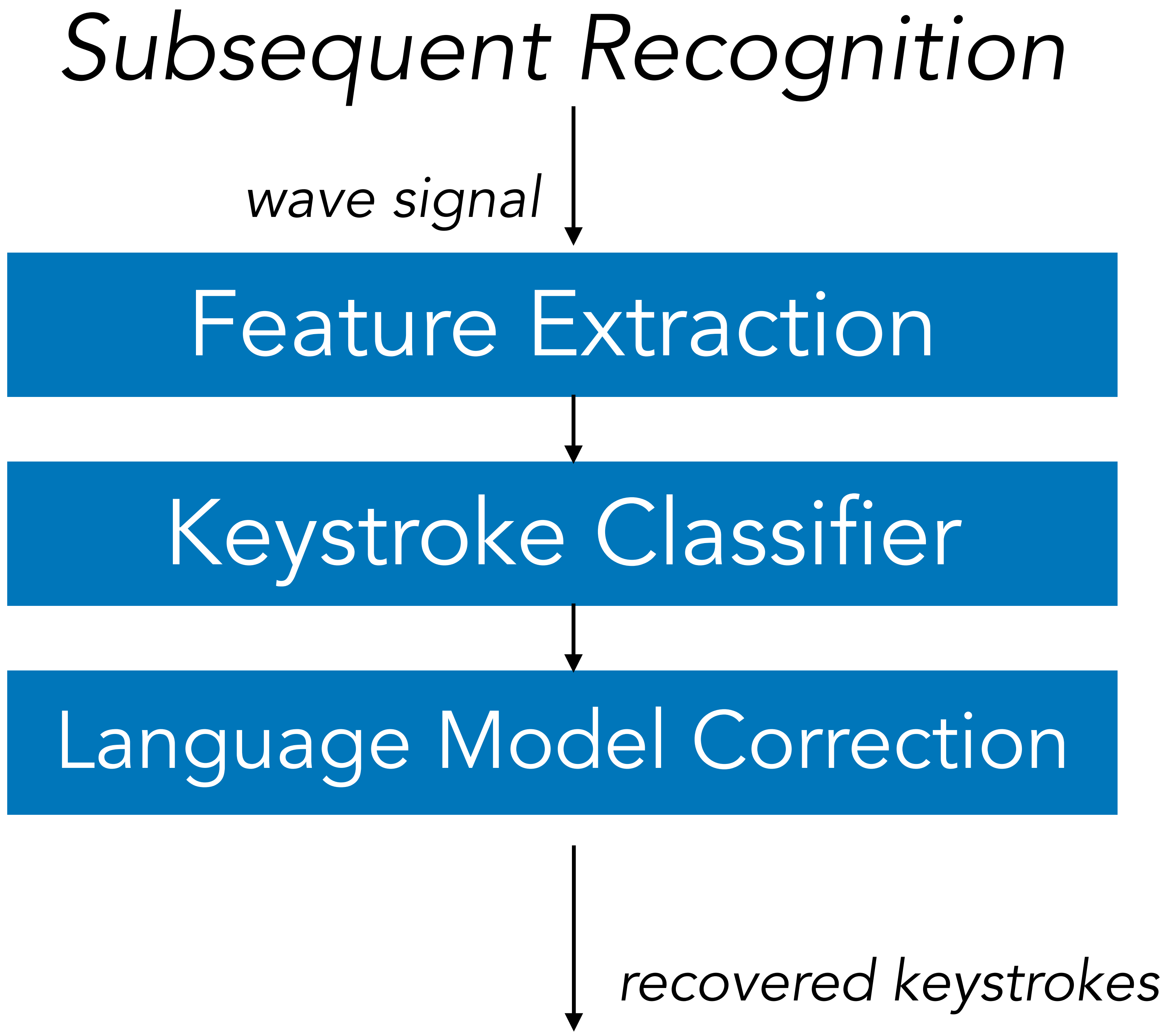
What makes the attack possible? What's the side channel here?

Why do keystrokes make different sounds?

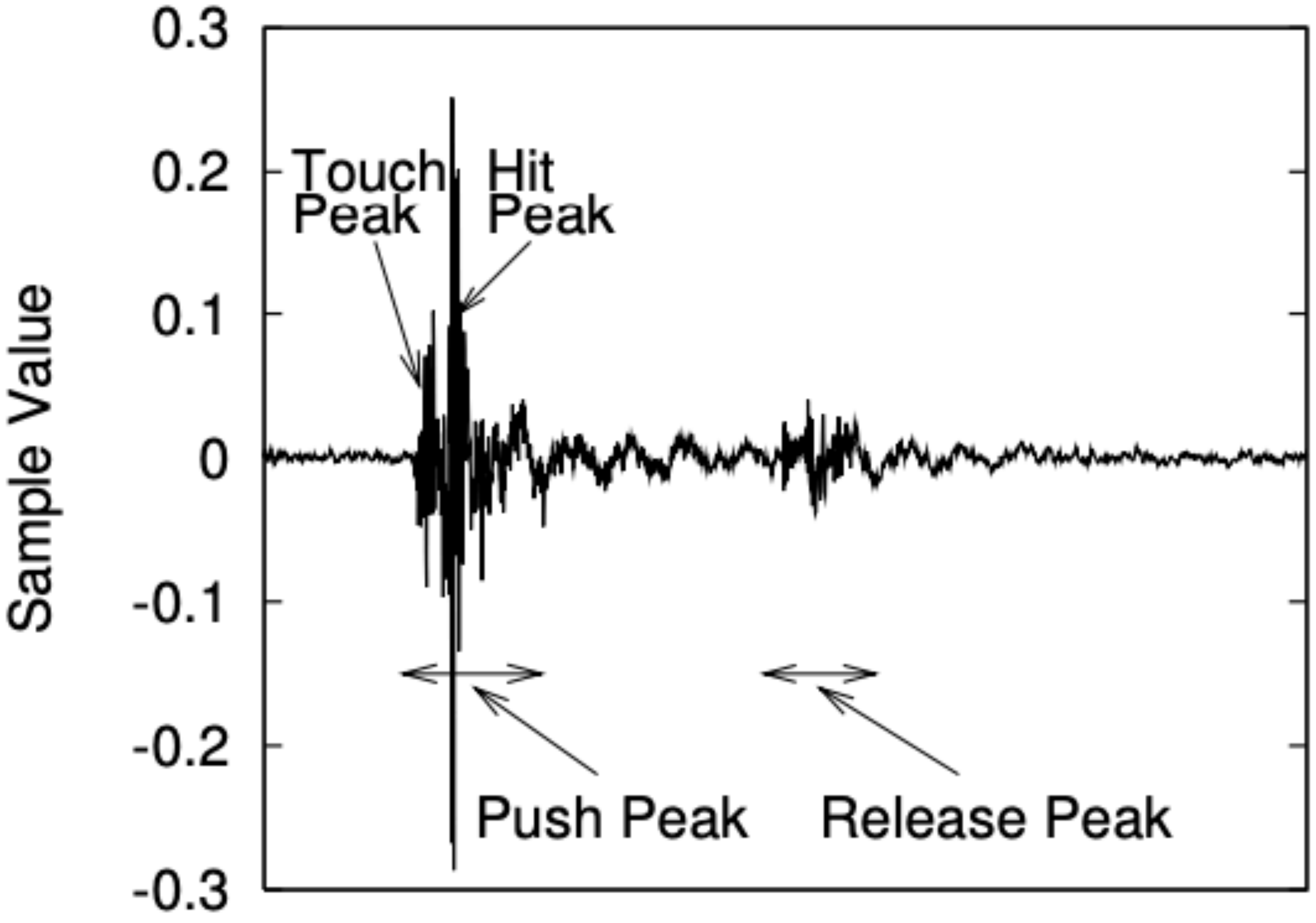
What capabilities (think threat model) are required of the attacker?

The attack flow



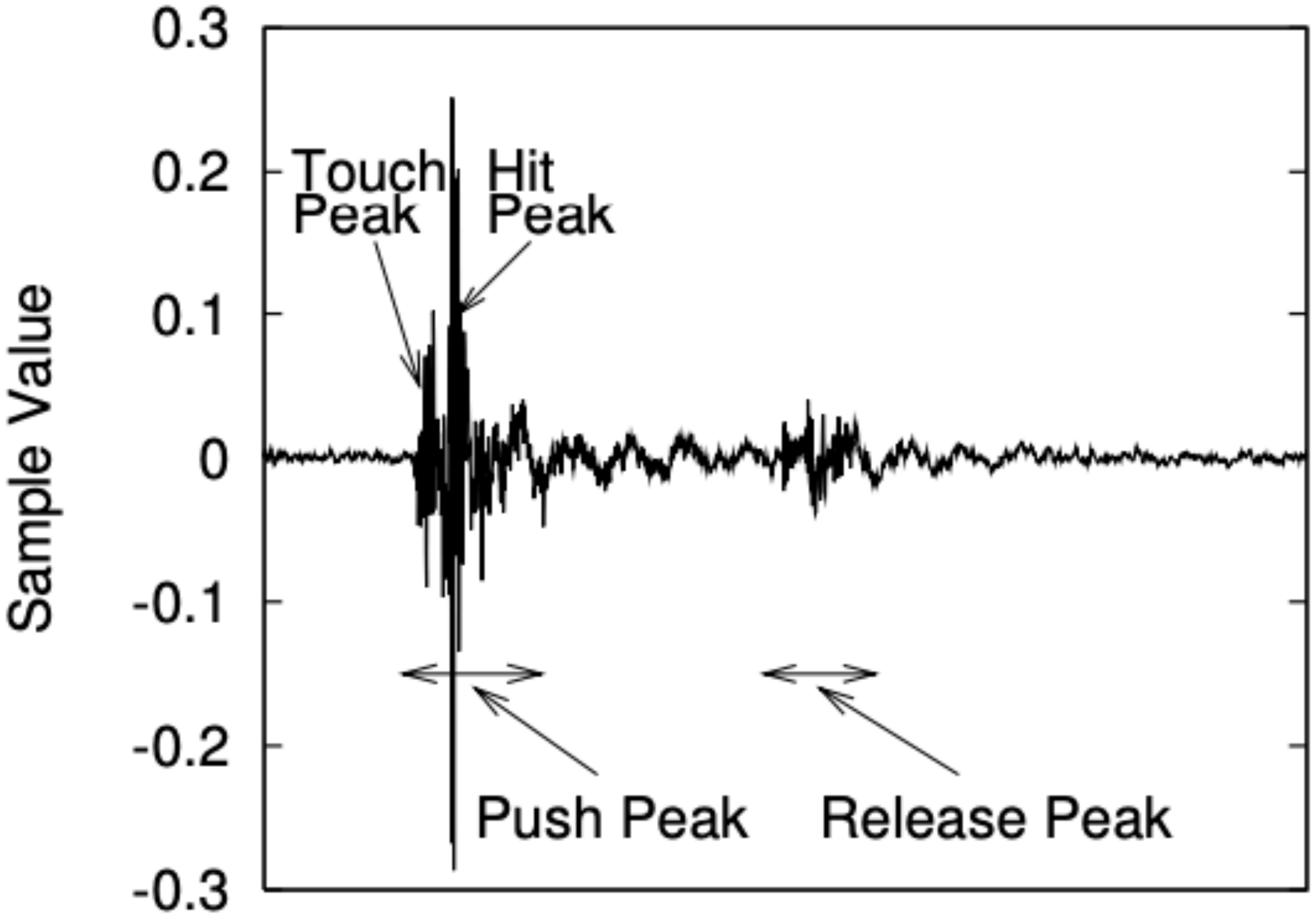


Extracting Keystrokes



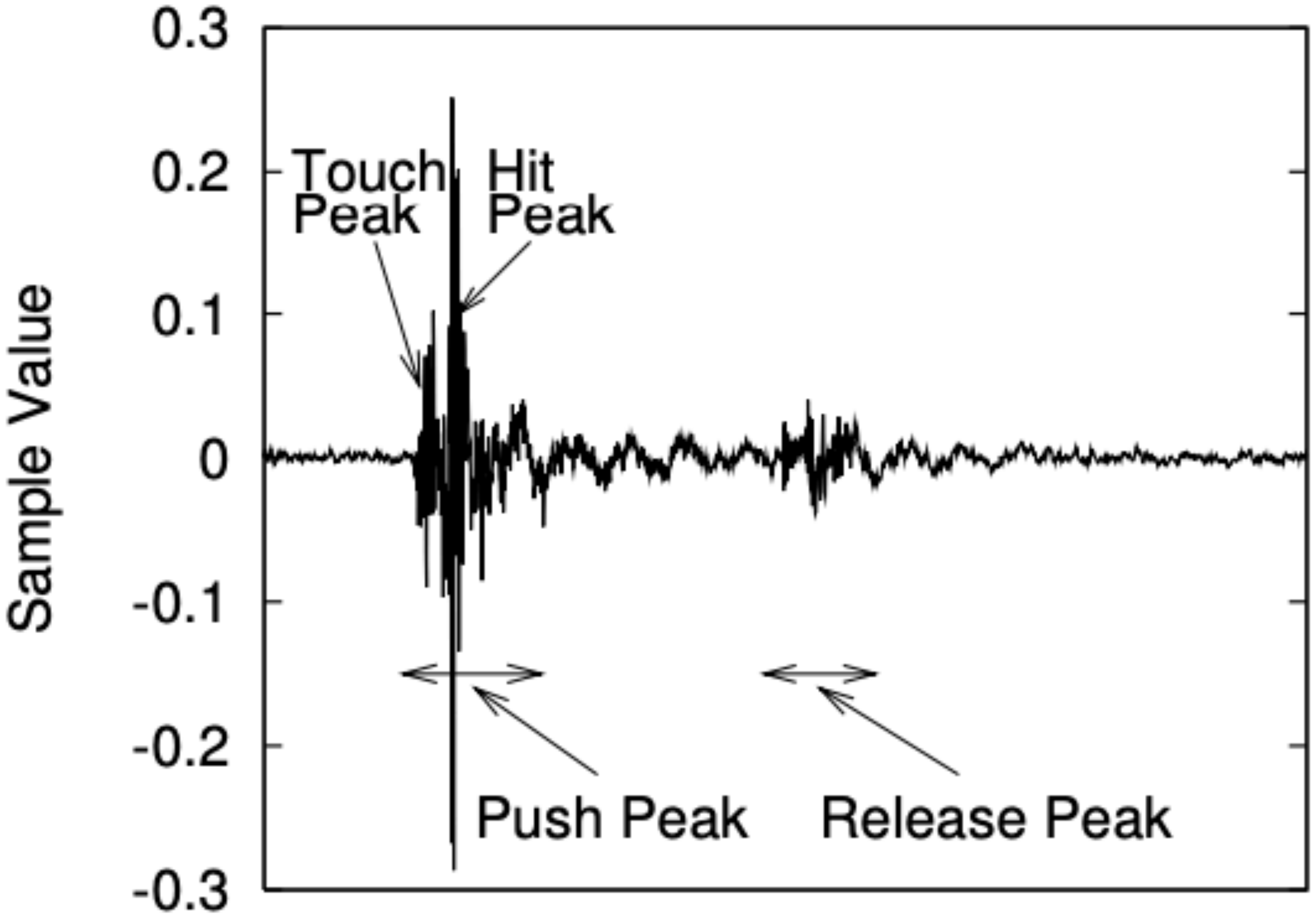
Extracting Keystrokes

- What is the touch peak?



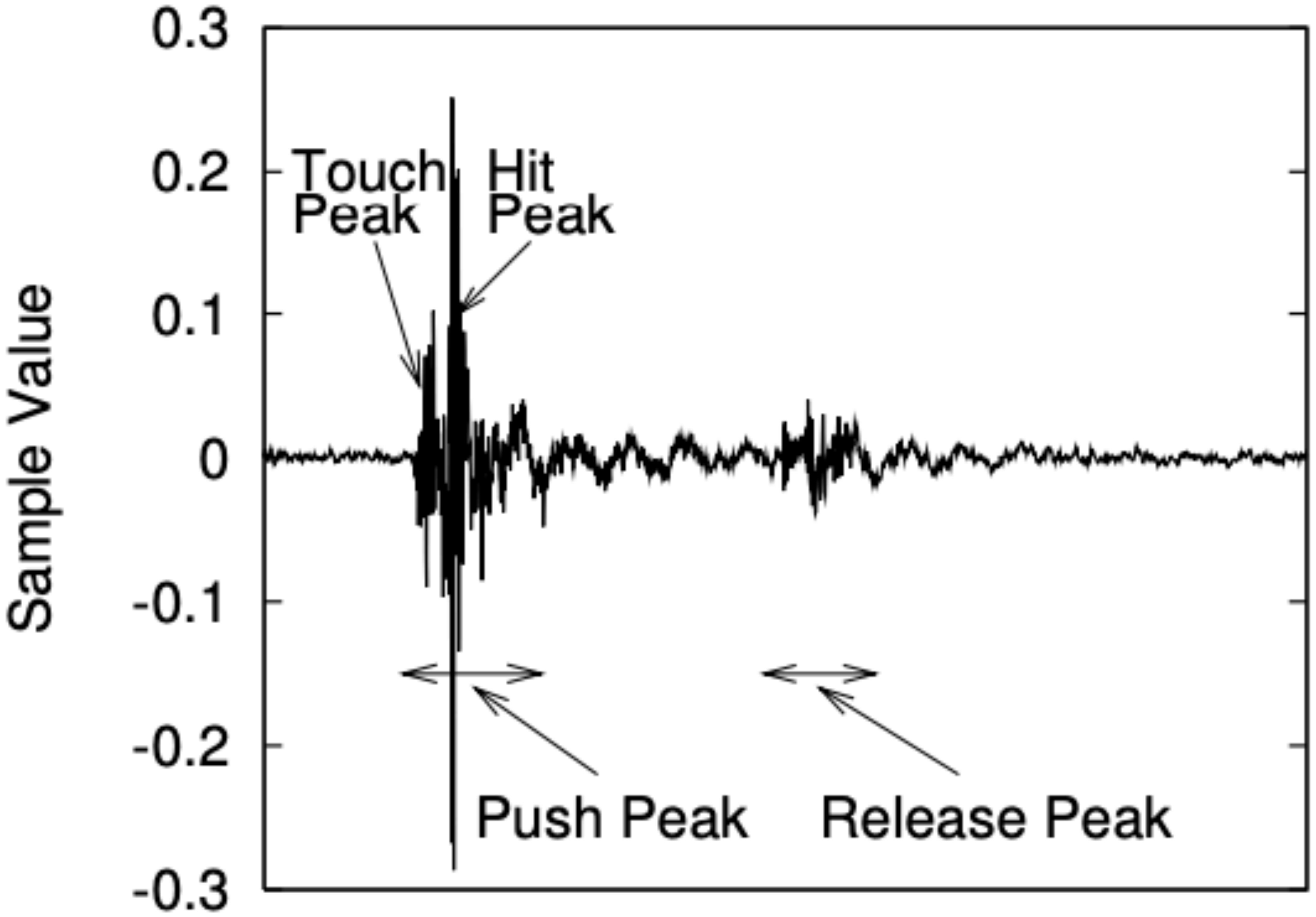
Extracting Keystrokes

- What is the touch peak?
- What is the hit peak?



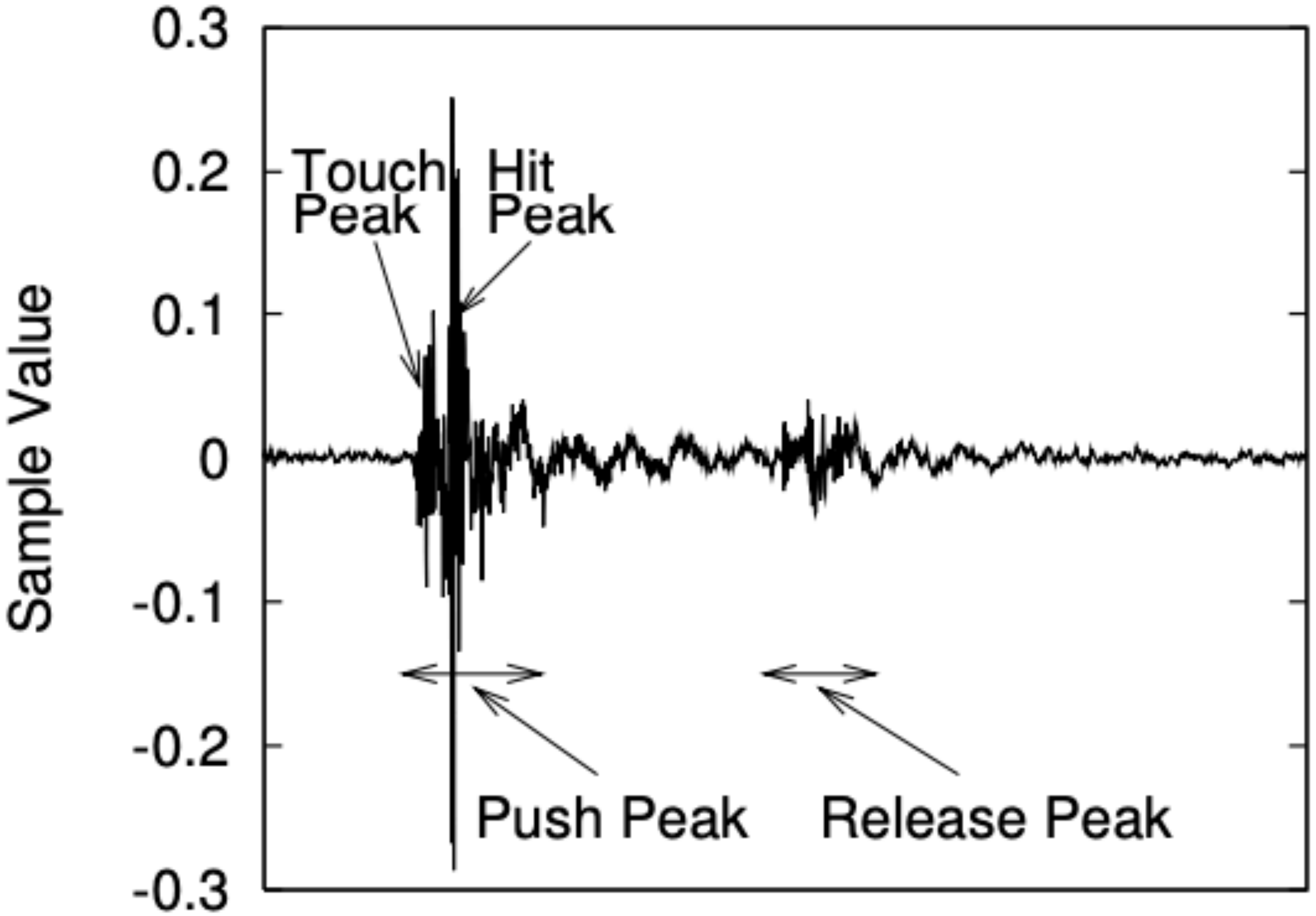
Extracting Keystrokes

- What is the touch peak?
- What is the hit peak?
- What is the release peak?



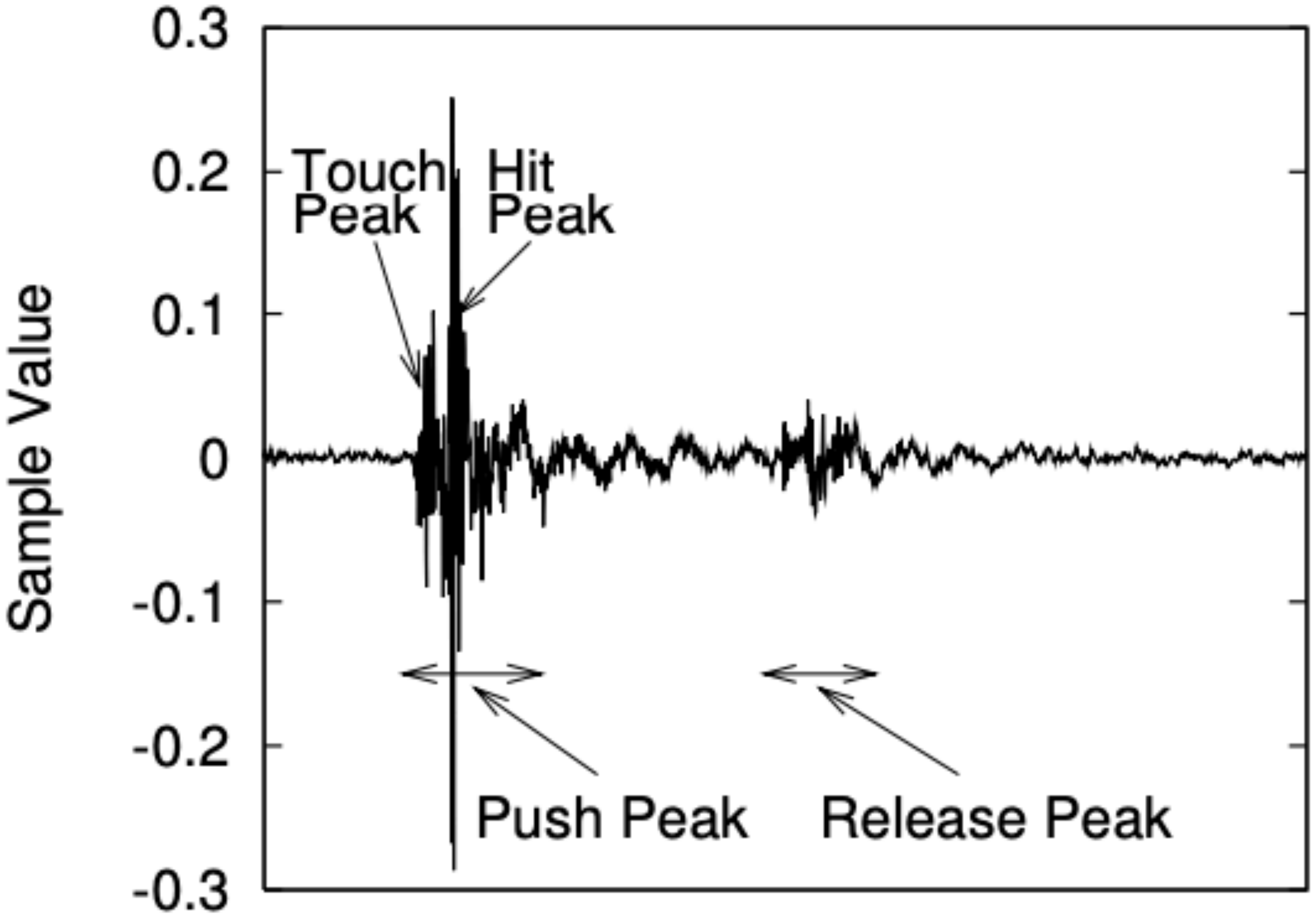
Extracting Keystrokes

- What is the touch peak?
- What is the hit peak?
- What is the release peak?
- How much time from push to release, on average?



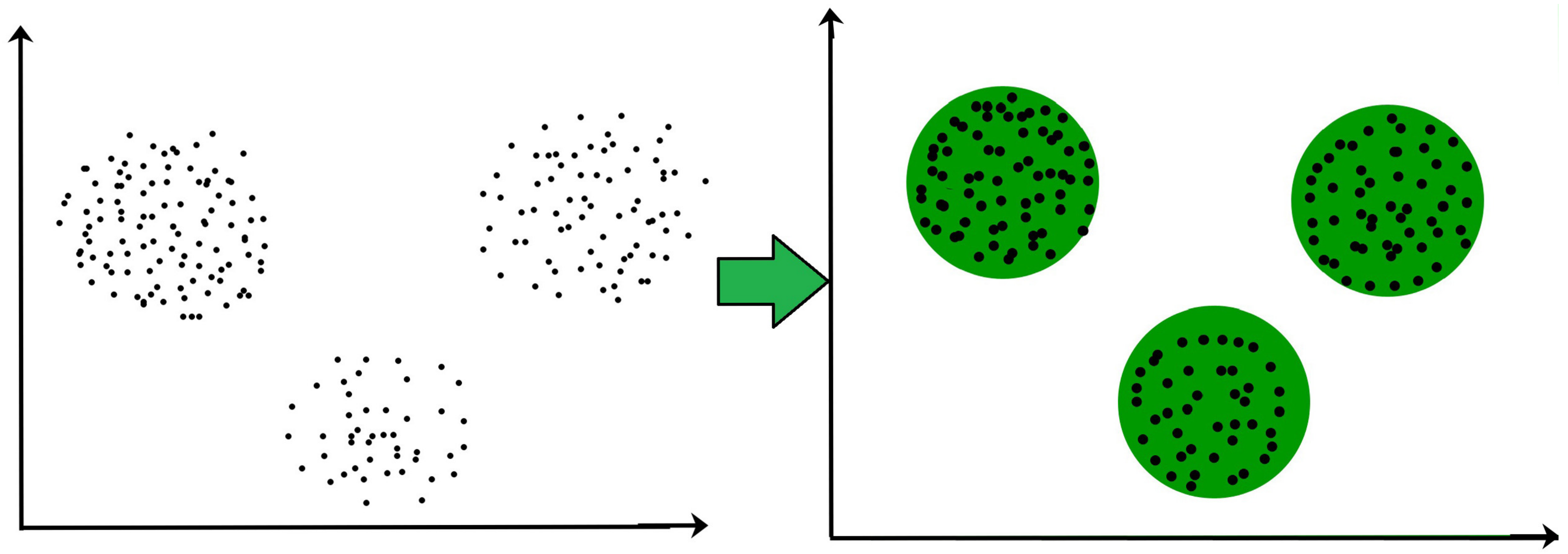
Extracting Keystrokes

- What is the touch peak?
- What is the hit peak?
- What is the release peak?
- How much time from push to release, on average?
  - **100 milliseconds… enough for a computer to discern!**



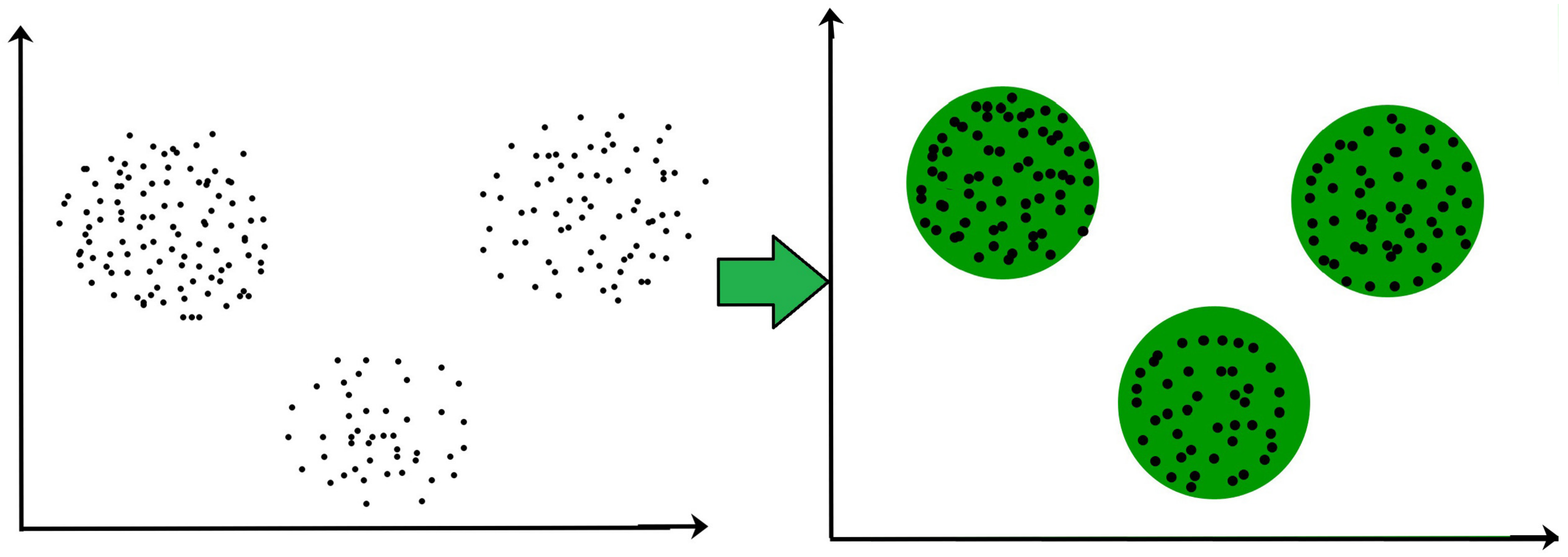
Unsupervised Learning

- What is unsupervised learning?



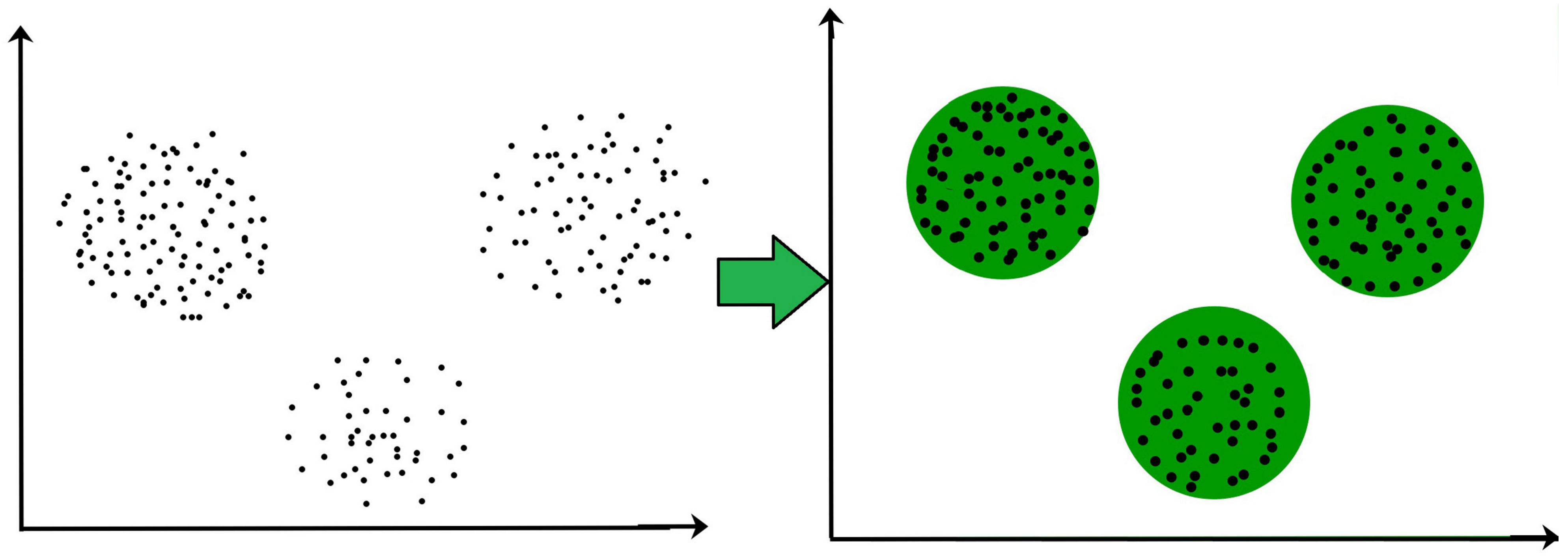
Unsupervised Learning

- What is unsupervised learning?
- What is clustering?



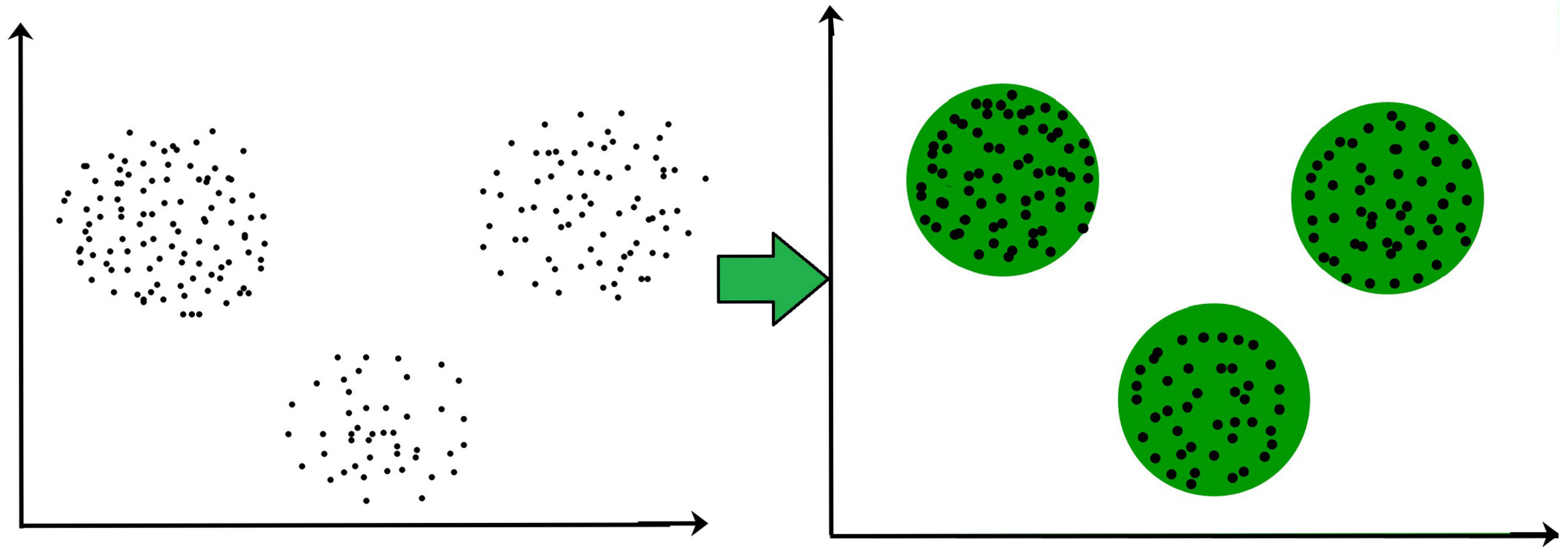
Unsupervised Learning

- What is unsupervised learning?
- What is clustering?
- How do the authors use unsupervised learning in their paper?



Unsupervised Learning

- What is unsupervised learning?
- What is clustering?
- How do the authors use unsupervised learning in their paper?
- Why is clustering *hard* in this context?



Recovering text from clusters

- How do the authors map clusters to letters?

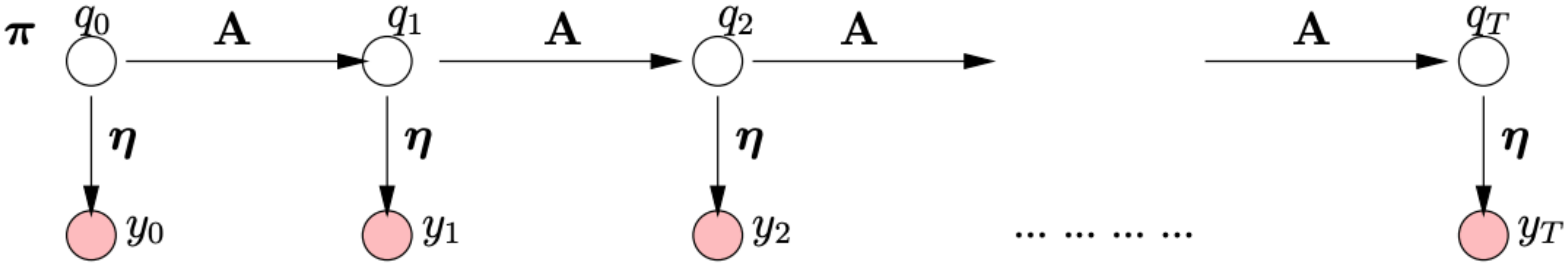
Recovering text from clusters

- How do the authors map clusters to letters?
- What is a Hidden Markov Model?

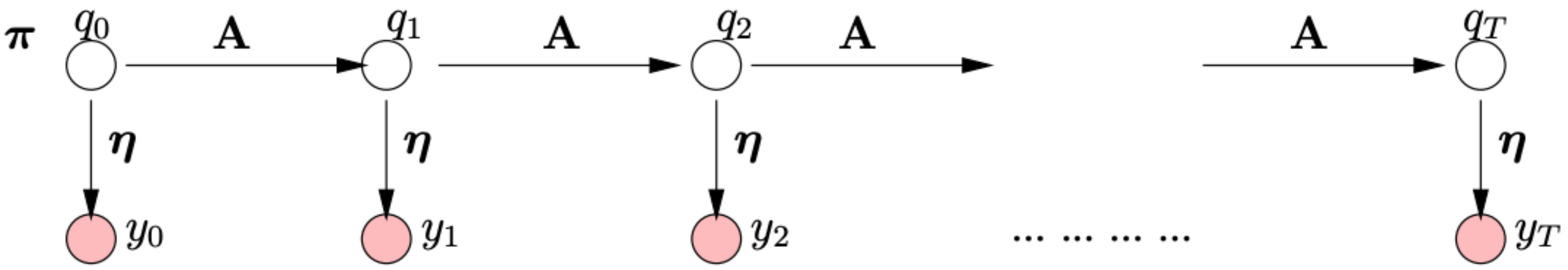
Recovering text from clusters

- How do the authors map clusters to letters?
- What is a Hidden Markov Model?
    - Authors “embed” English probabilities here – **th** is much more likely than **tj**

Bi-grams of characters

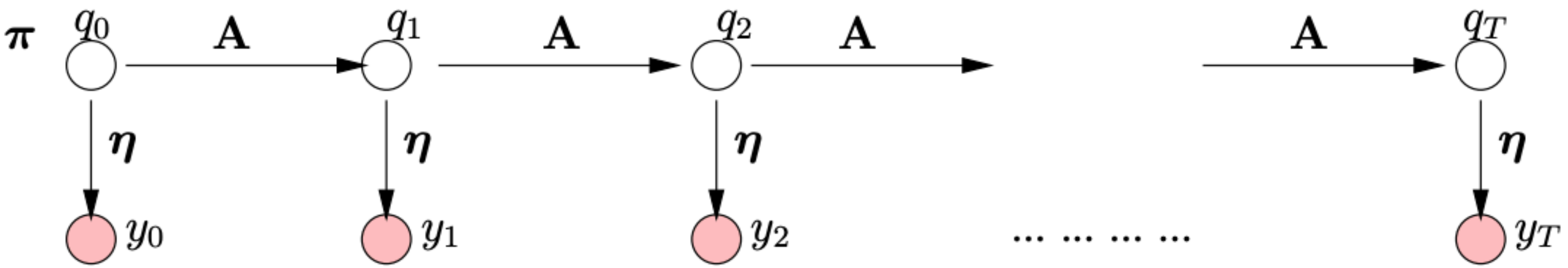


Bi-grams of characters



- What do the circles represent in the model? What are shaded and unshaded?

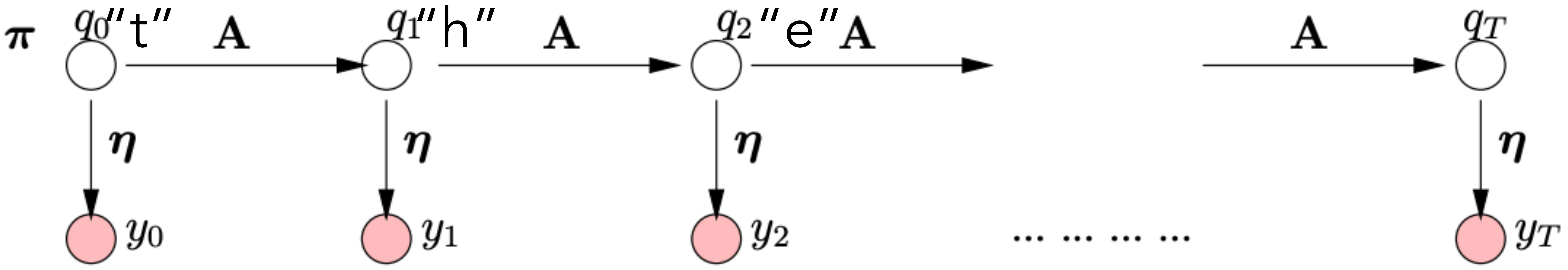
Bi-grams of characters



- What do the circles represent in the model? What are shaded and unshaded?

- What is *q*?

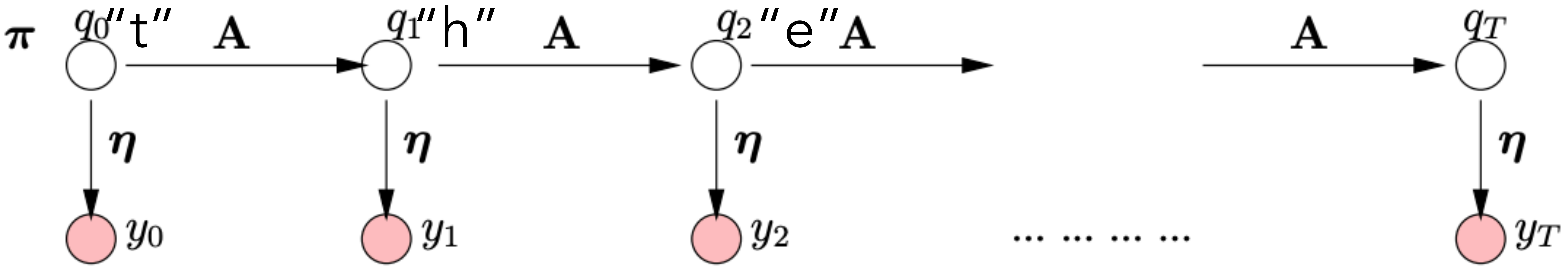
Bi-grams of characters



- What do the circles represent in the model? What are shaded and unshaded?

- What is q? ***Characters pressed***

Bi-grams of characters

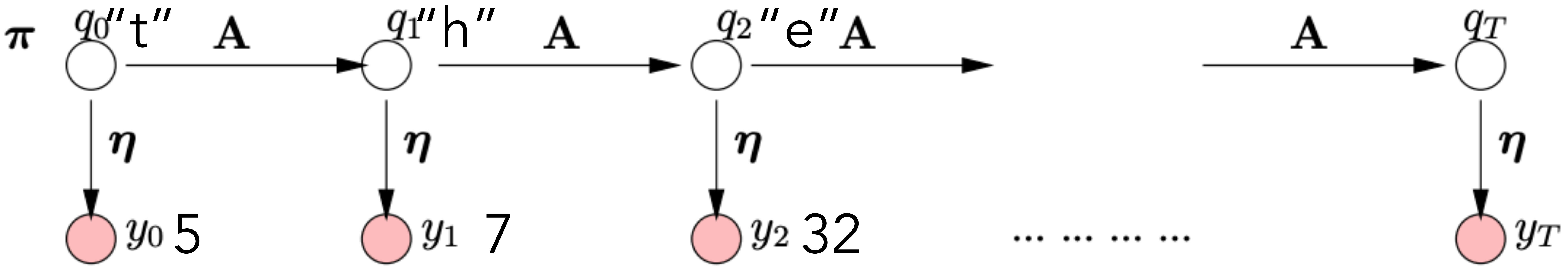


- What do the circles represent in the model? What are shaded and unshaded?

- What is *q*? ***Characters pressed***

- What is *y*?

Bi-grams of characters

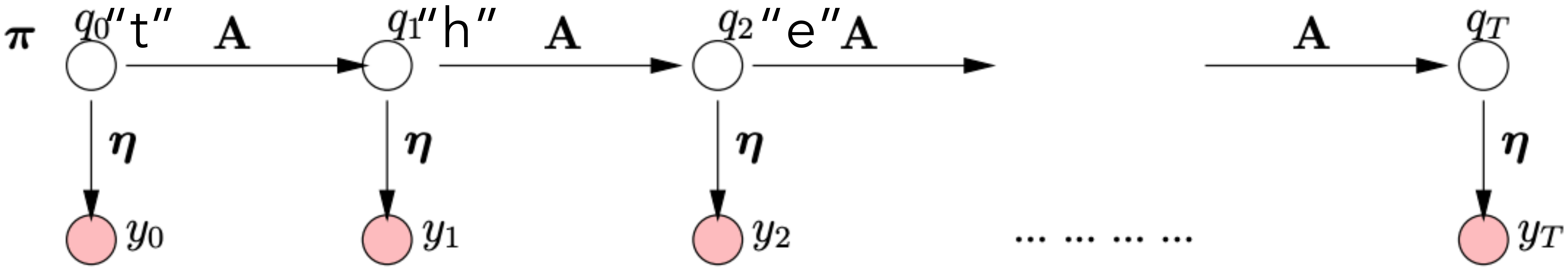


- What do the circles represent in the model? What are shaded and unshaded?

- What is *q*? ***Characters pressed***

- What is *y*? ***Cluster labels***

Bi-grams of characters



- What is **A**, otherwise known as the *transition matrix?*
- How do we populate **A?**
- What is *eta*?
- How do the authors populate *eta*?

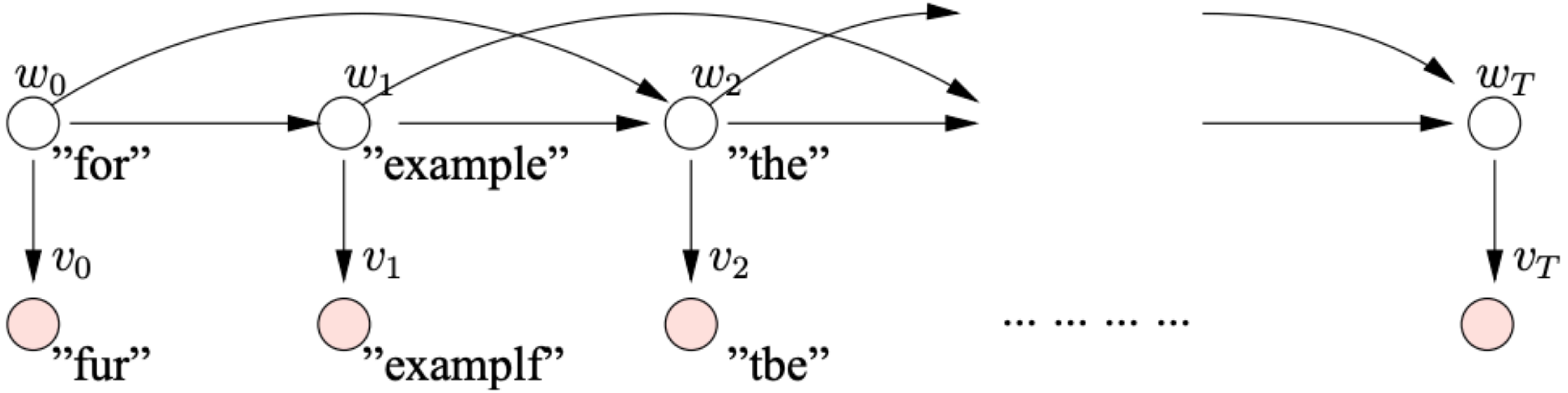
So now we have characters…. are they right?

- Authors used *spellcheck* to try and make the text more readable, but it still made mistakes
  - e.g., “fur example” vs. “for example”
- Can get more readable text using an n-gram language model
  - What’s an n-gram language model?

So now we have characters…. are they right?

- Authors used *spellcheck* to try and make the text more readable, but it still made mistakes
    - e.g., “fur example” vs. “for example”
- Can get more readable text using an n-gram language model
    - What’s an n-gram language model?
- n-gram: sequence of *n* adjacent items in text, speech, genomes, etc.

Tri-grams of words (HMMs to the rescue, again)



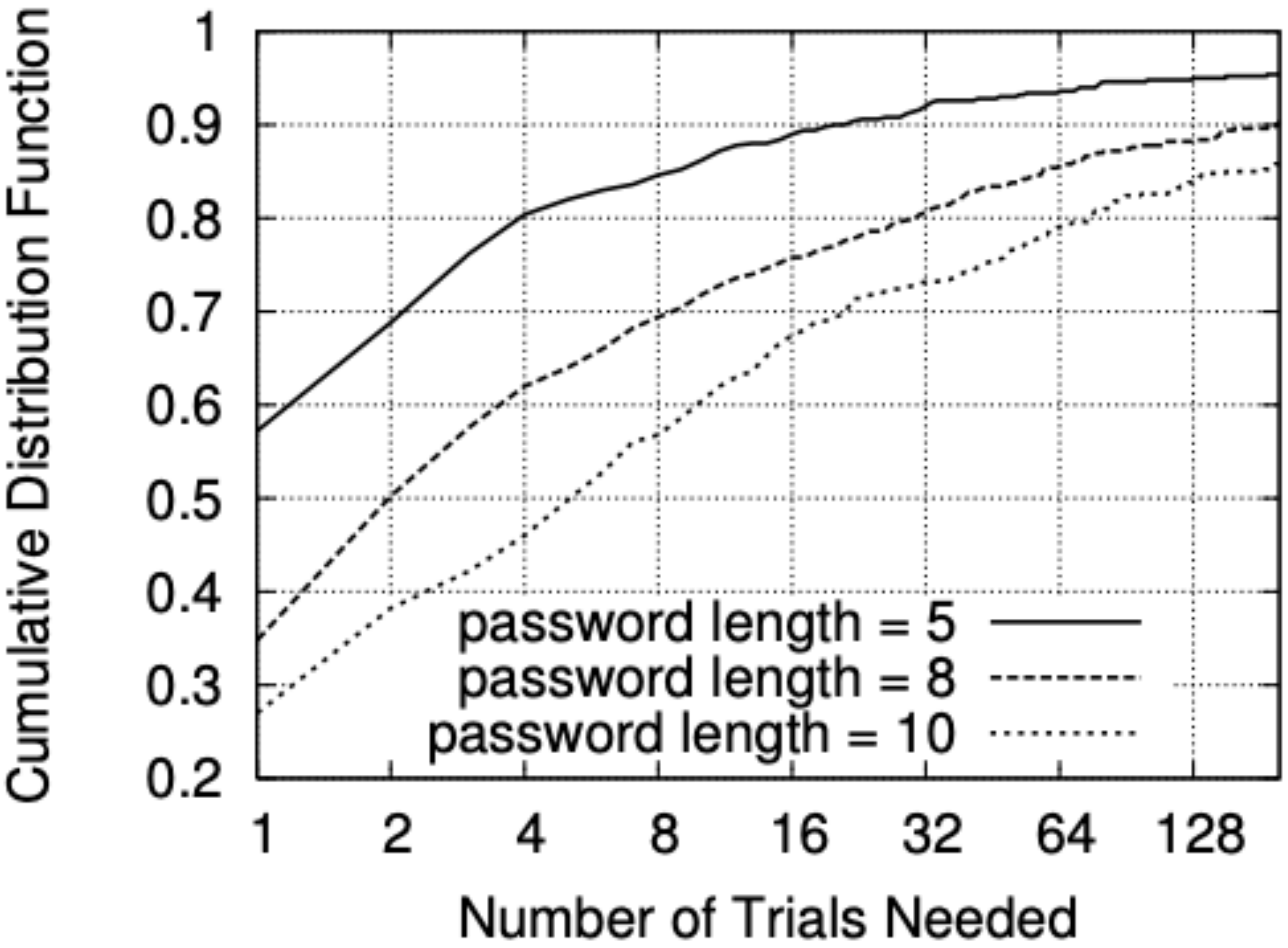
- Hidden variables are the original words
- This HMM depends on *two* layers (previous word **and** previous previous word)
- This is a great strategy if you have **unknowns** you want to predict from a known distribution!

Evaluation

- What was the evaluation setup for the attack?
- How many environments did the authors test their attack in?
- Are all keyboards vulnerable to this kind of attack? How did the authors evaluate this?

So it works on english, does it work on passwords?

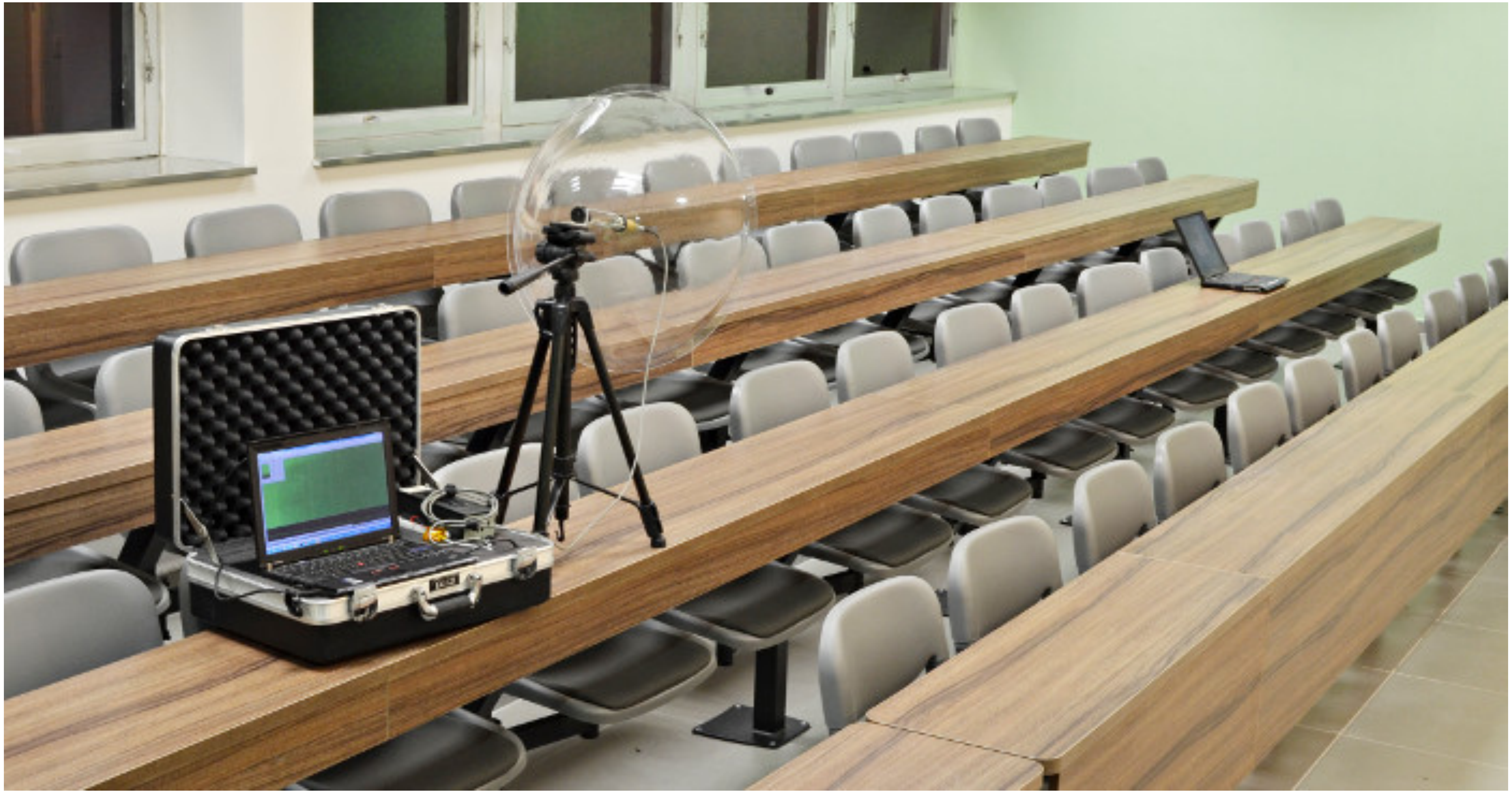
- Yes!
- Authors found that they could recover 90% of 5-character passwords, 77% of 8-character passwords, and 69% of 10-character passwords
- Probabilities form a “hit list” of potential passwords to try



5-minute discussion: Meta points

- How feasible is this attack?
- Do you believe this attack will work in practice? Why or why not?
- What do we think about side channel research?

Side channels can be even crazier….



RSA Key Extraction via Low-Bandwidth Acoustic Cryptanalysis

Any questions?

Break Time + Attendance

Codeword:
TikTok-Fail

https://tinyurl.com/cse227-attend

Lest We Remember: Cold-Boot Attacks on Encryption Keys

What is DRAM?

What is DRAM?

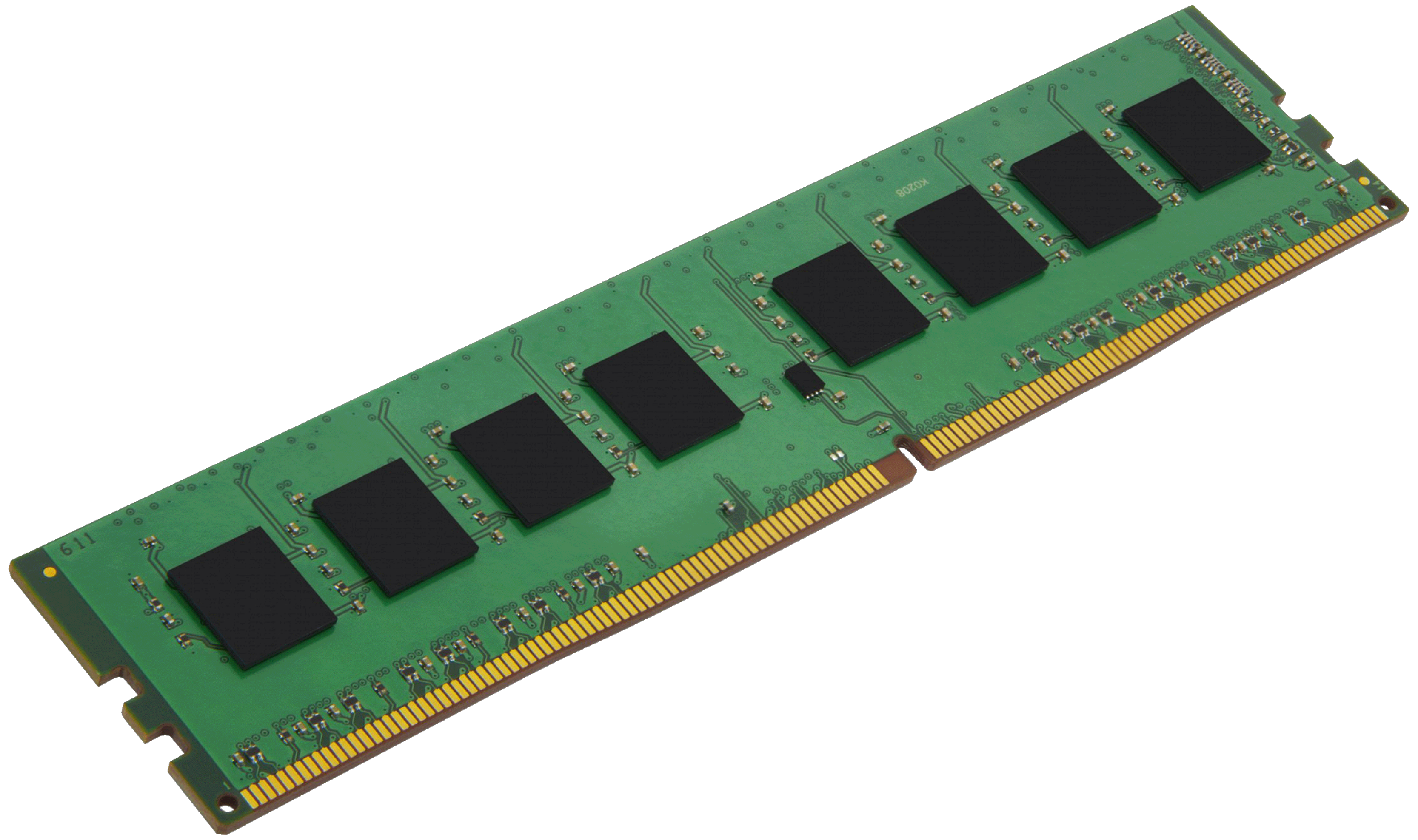
DRAM: Dynamic random-access memory – a type of computer memory (hardware)

Cold Boot Attack

What is the attack the authors want to conduct?

Cold Boot Attack

What is the attack the authors want to conduct?



Cold Boot Attack

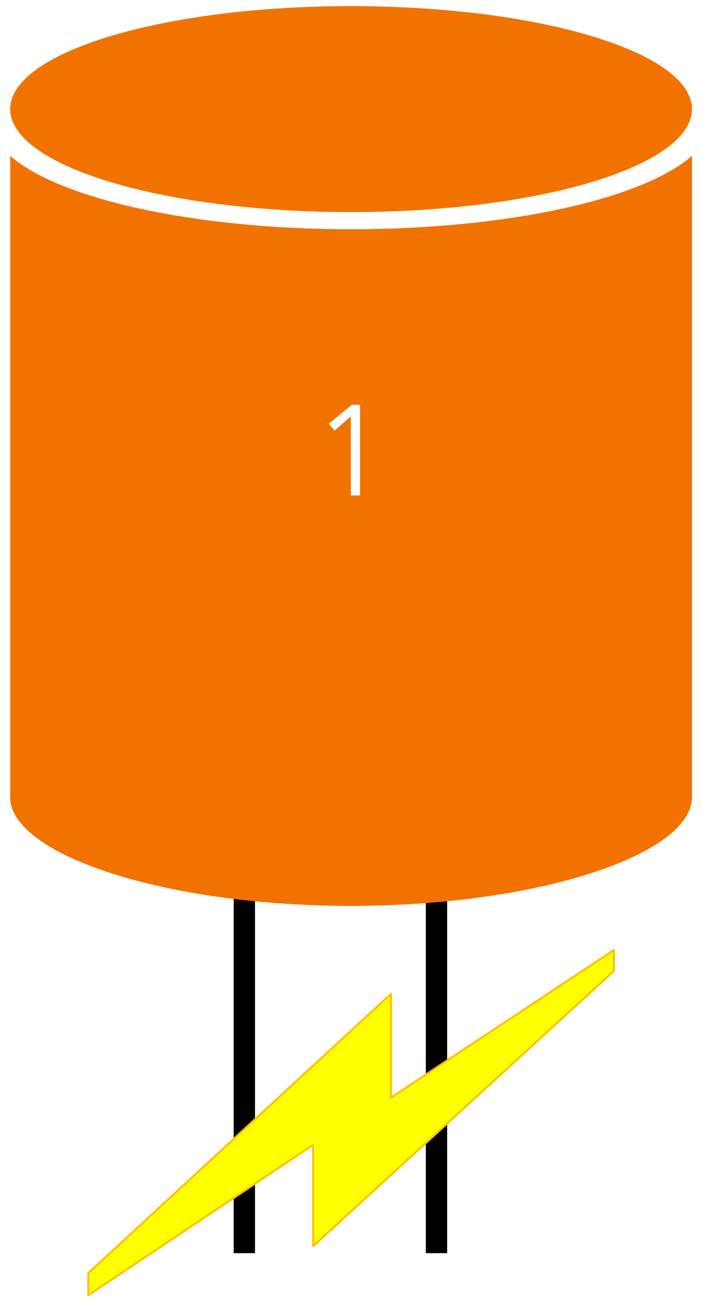
What is the attack the authors want to conduct?

What makes the attack possible?

Memory remanance: Most DRAM lose contents *gradually* over a period of seconds not all at once. This creates an opportunity to inspect what's in DRAM!

How does DRAM work?

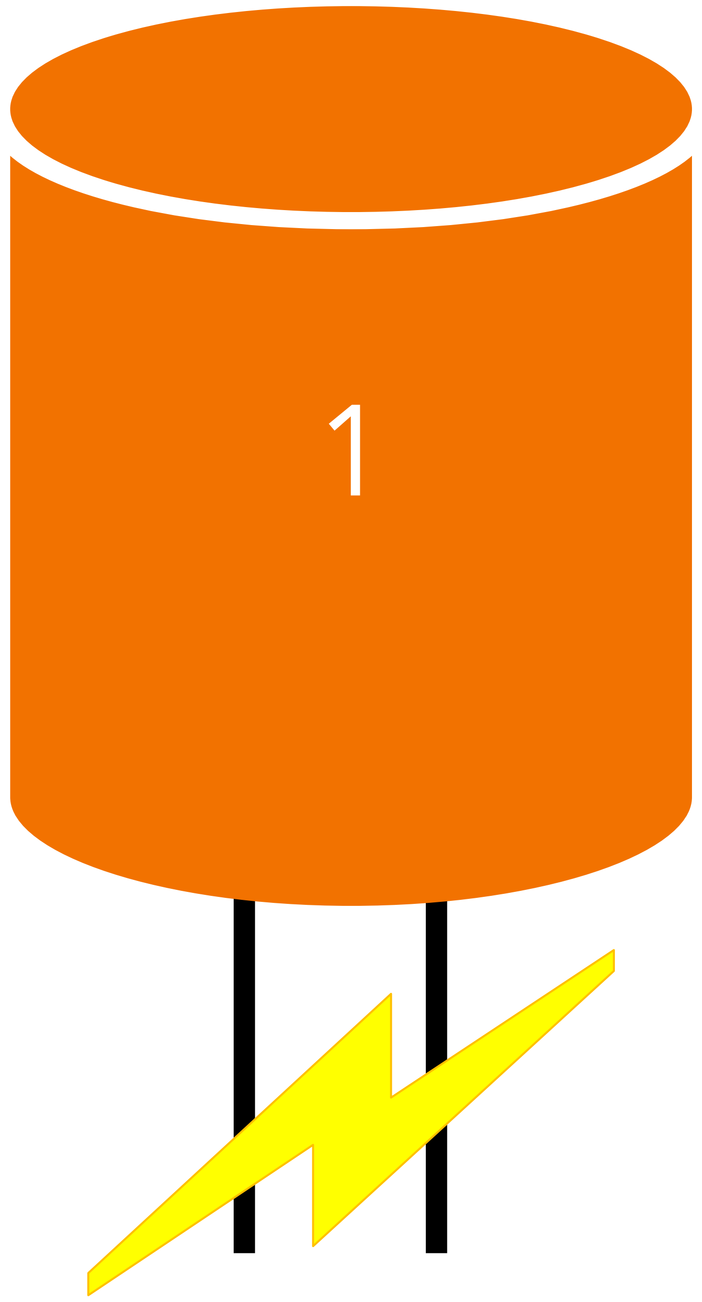
DRAM cells are essentially just capacitors. What is a capacitor?



Write “1”

How does DRAM work?

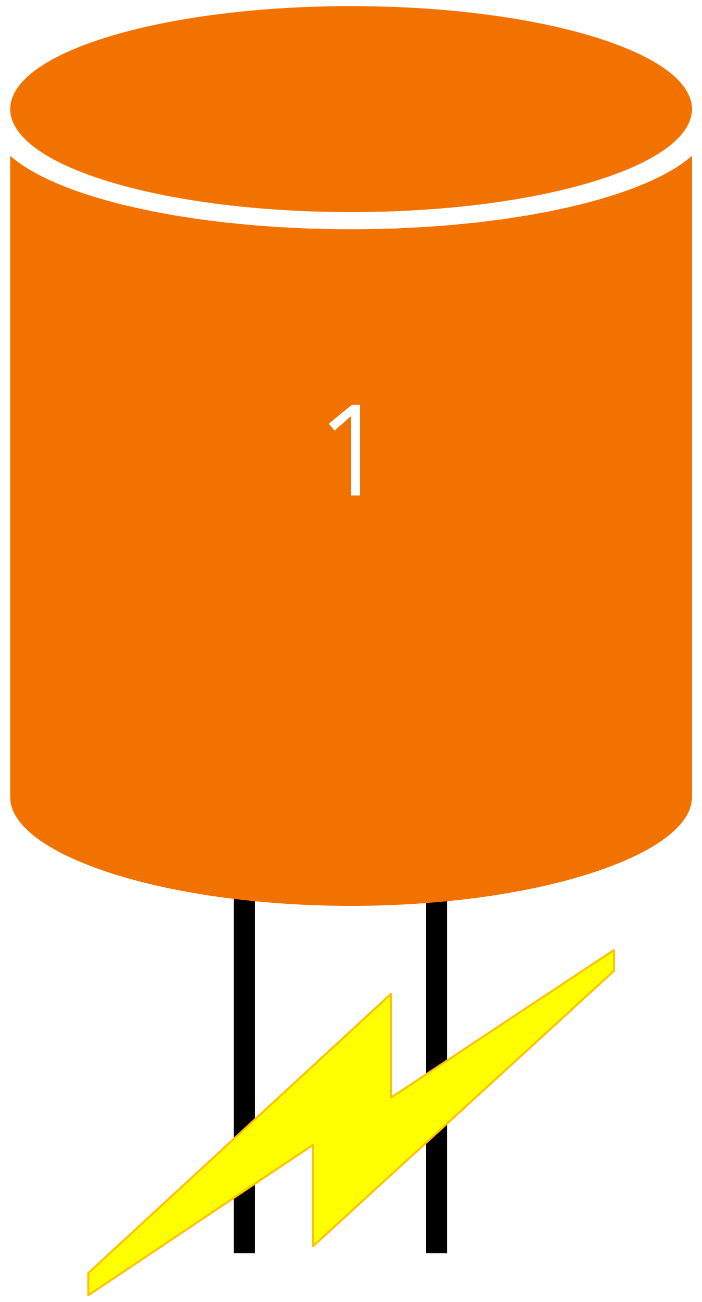
What is a *refresh* in DRAM, and how does it work?



Write “1”

How does DRAM work?

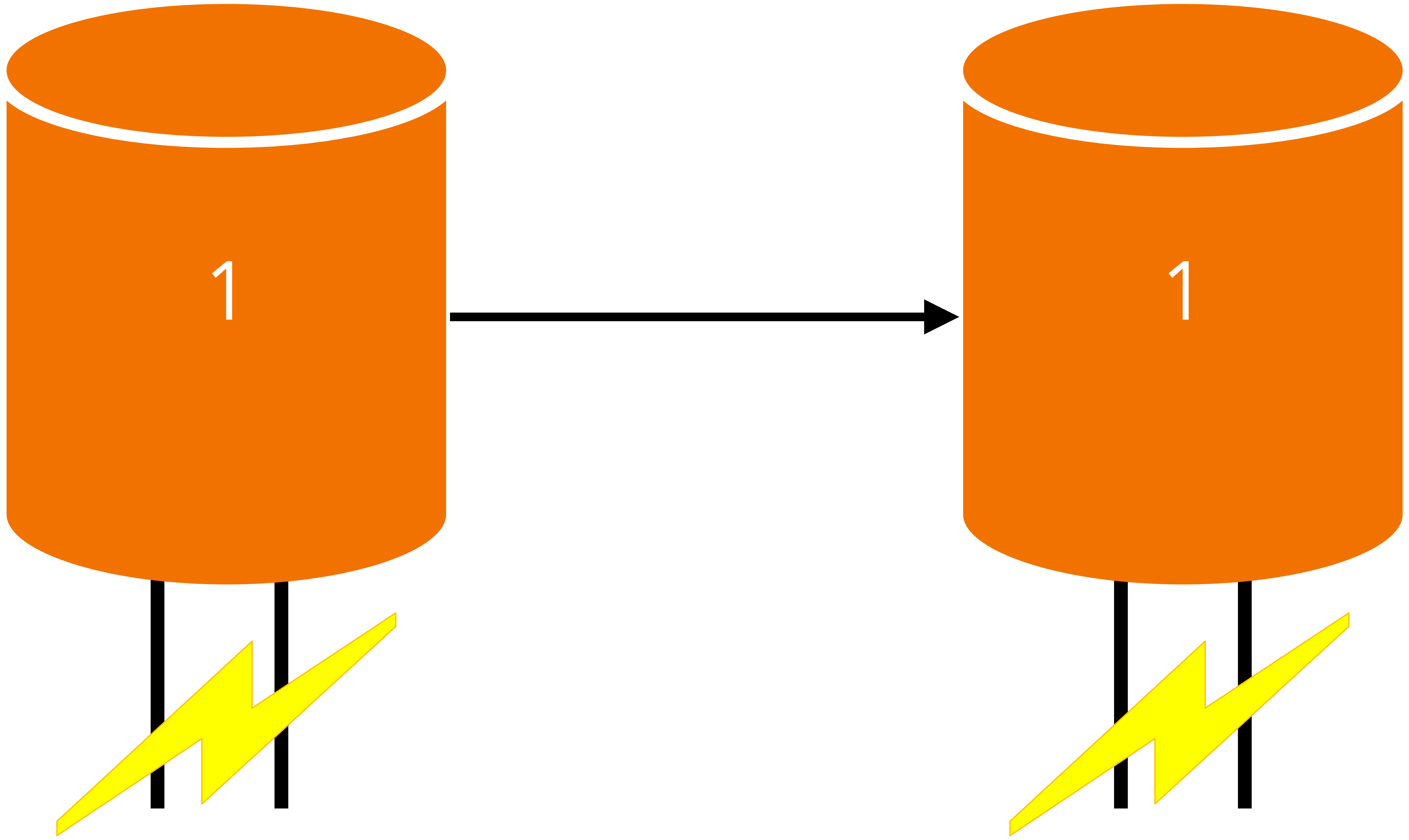
What is a refresh interval?



Write “1”

How does DRAM work?

What is a *refresh* in DRAM, and how does it work?

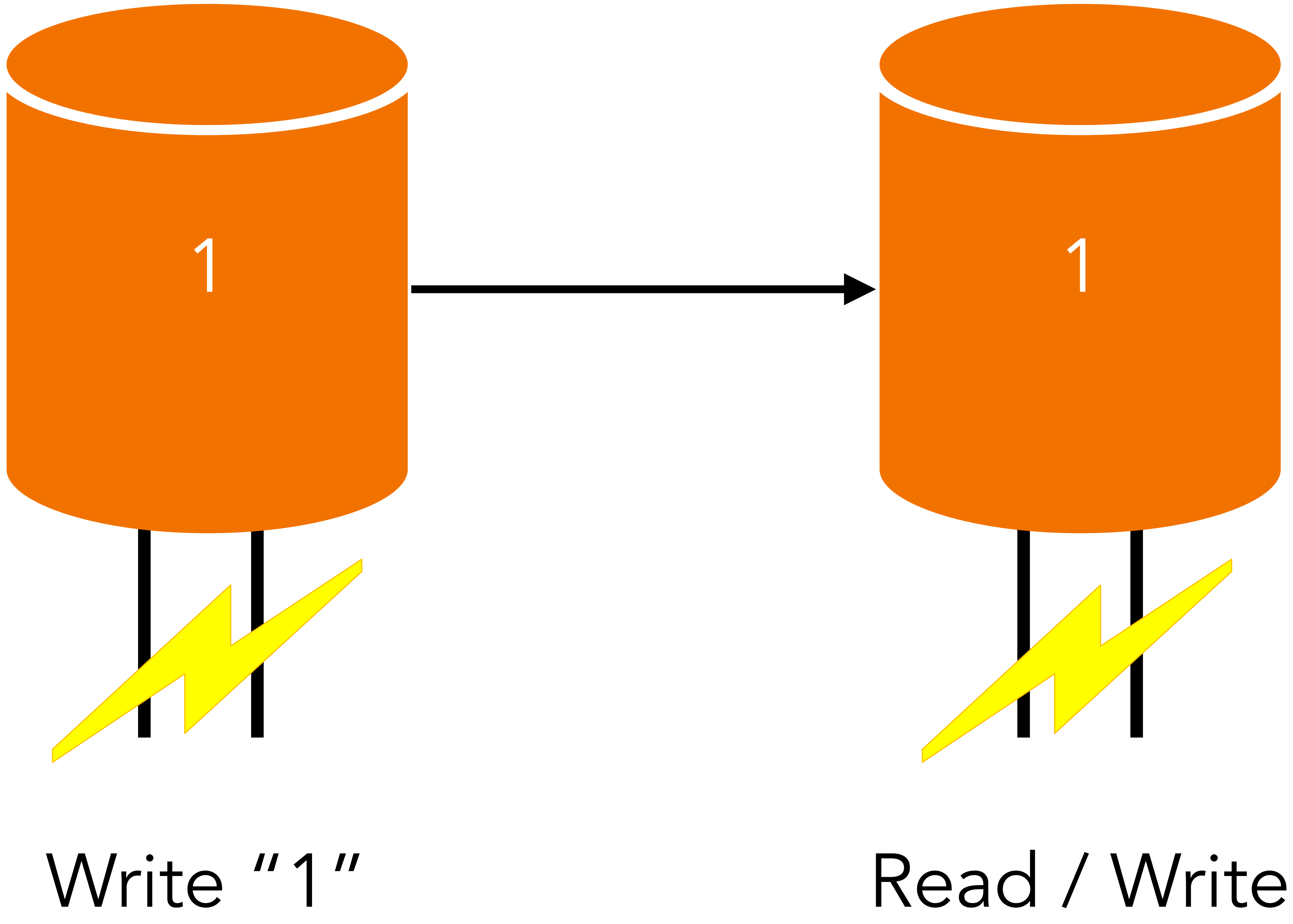


Write "1"

Read / Write

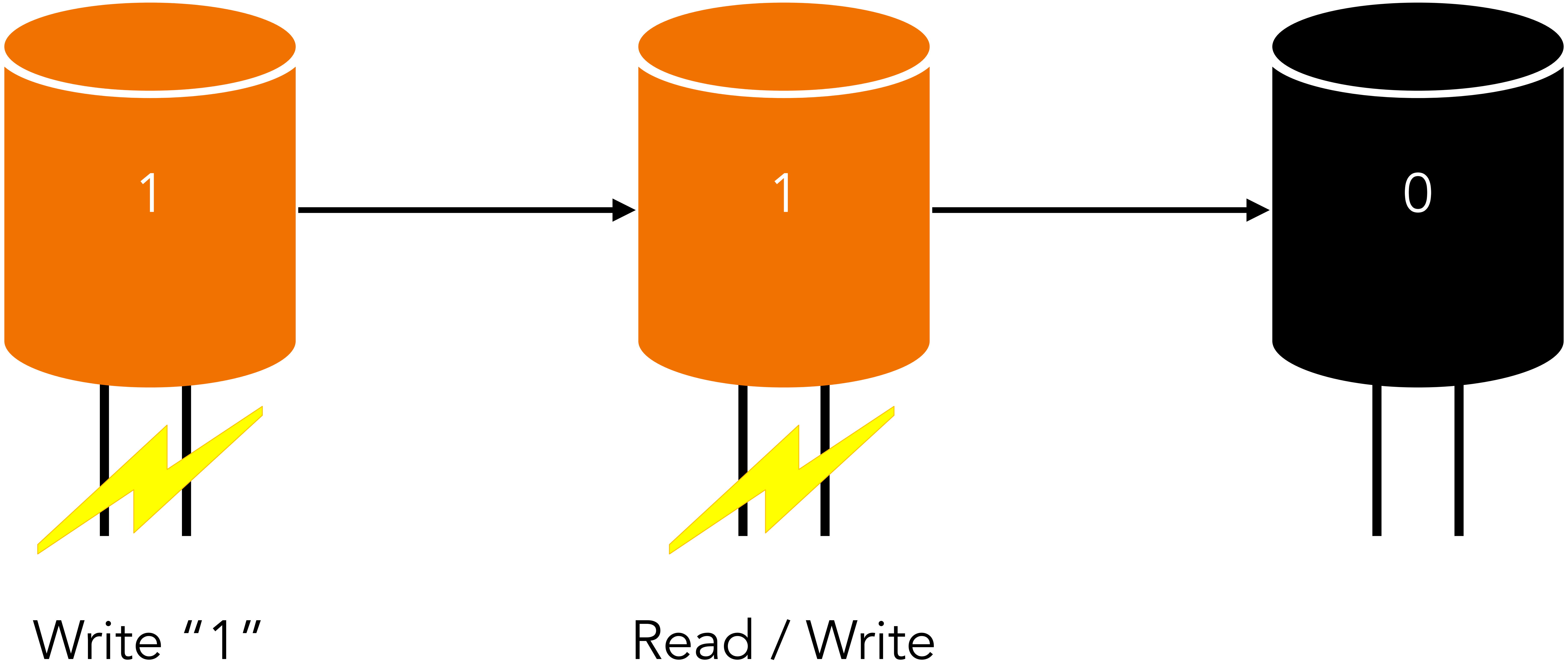
How does DRAM work?

What happens if we don’t refresh the cell?



How does DRAM work?

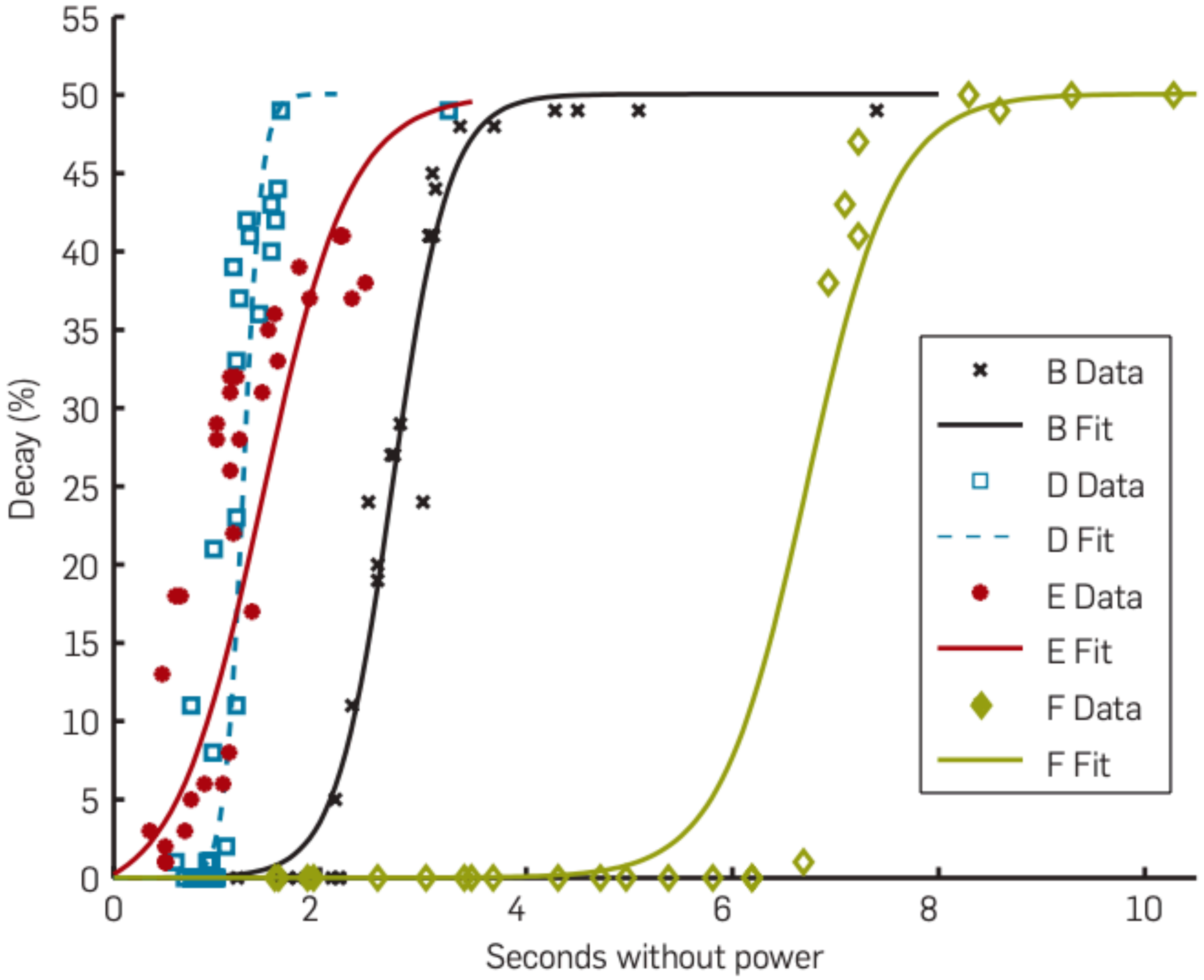
What happens if we don't refresh the cell?



DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?

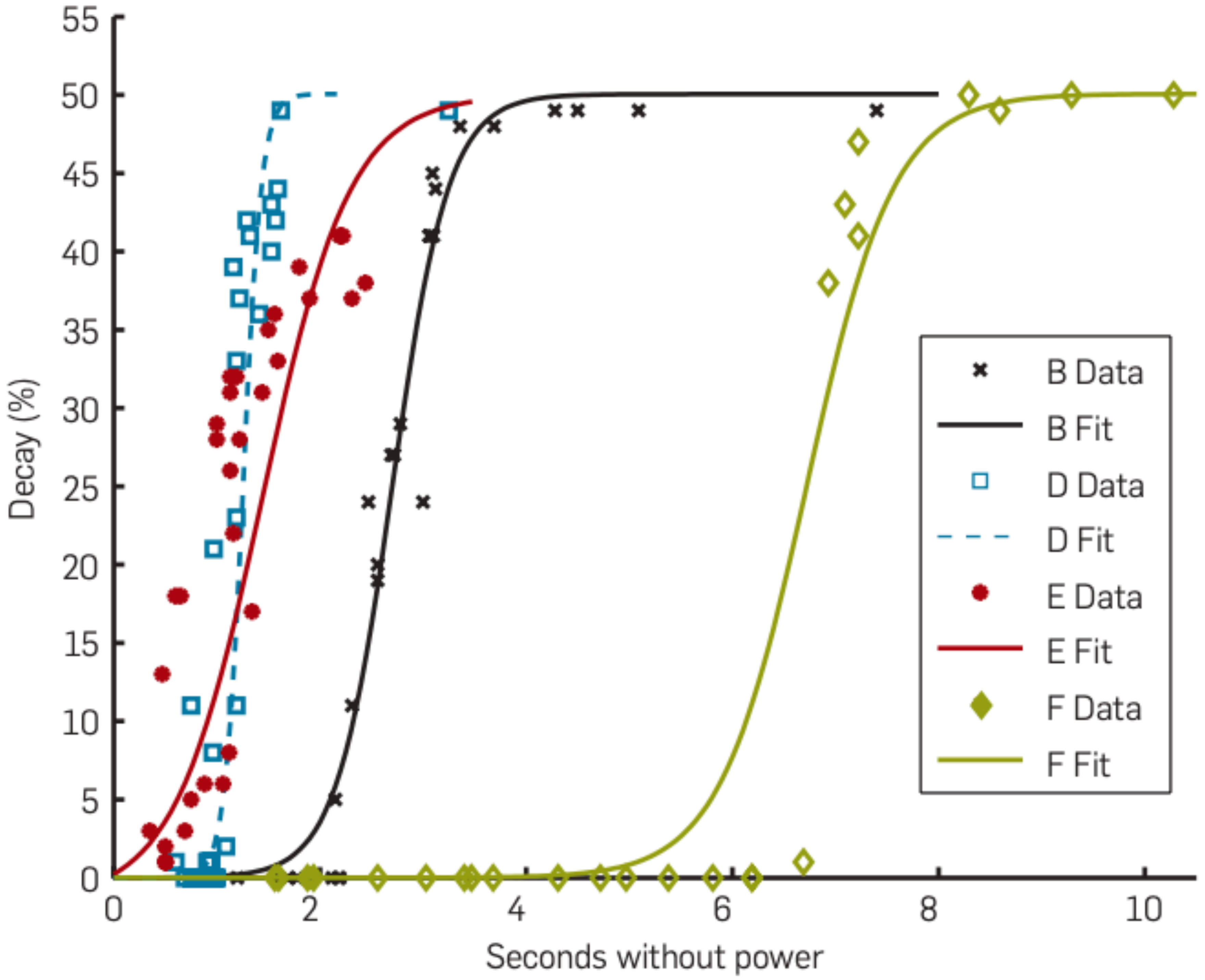
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?
- How did the authors measure errors?

Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.

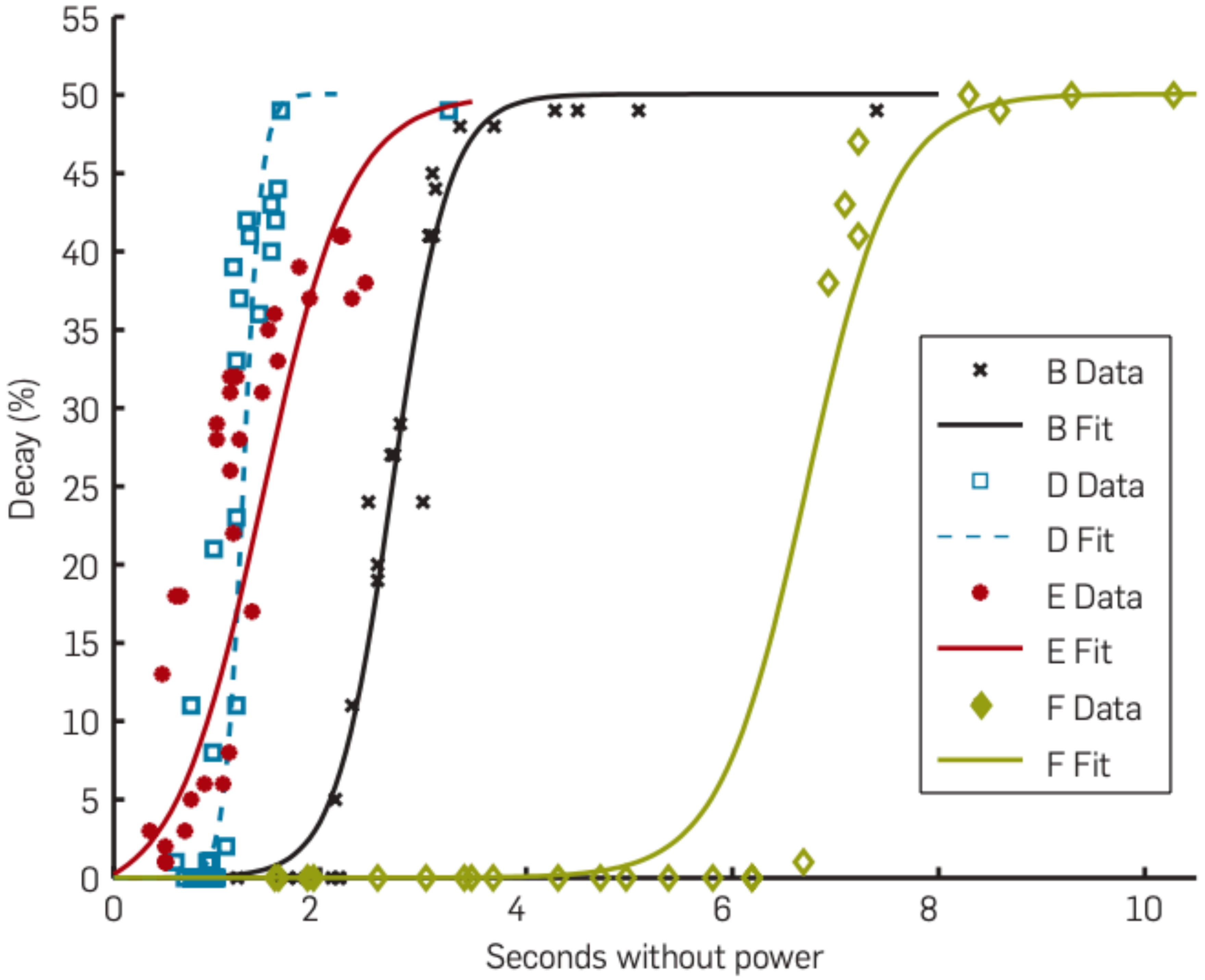


DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?

- How did the authors measure errors?
  - **Hamming distance:** number of bit errors divided by the total number of bits

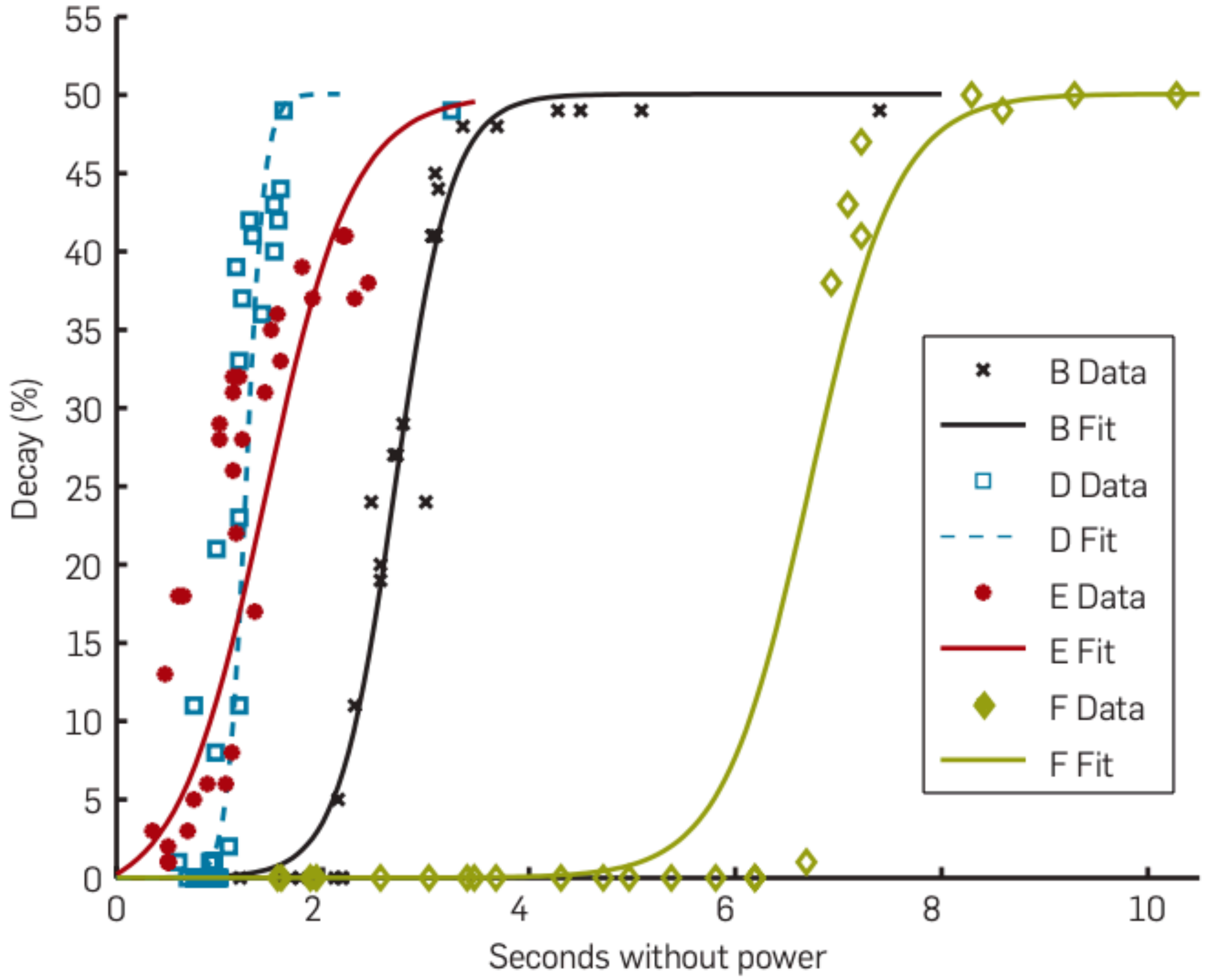
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



DRAM Decay Curves

- How did the authors test the time for DRAM cells to decay?
- How did the authors measure errors?
  - **Hamming distance:** number of bit errors divided by the total number of bits
- What would the error rate be if memory had fully decayed?

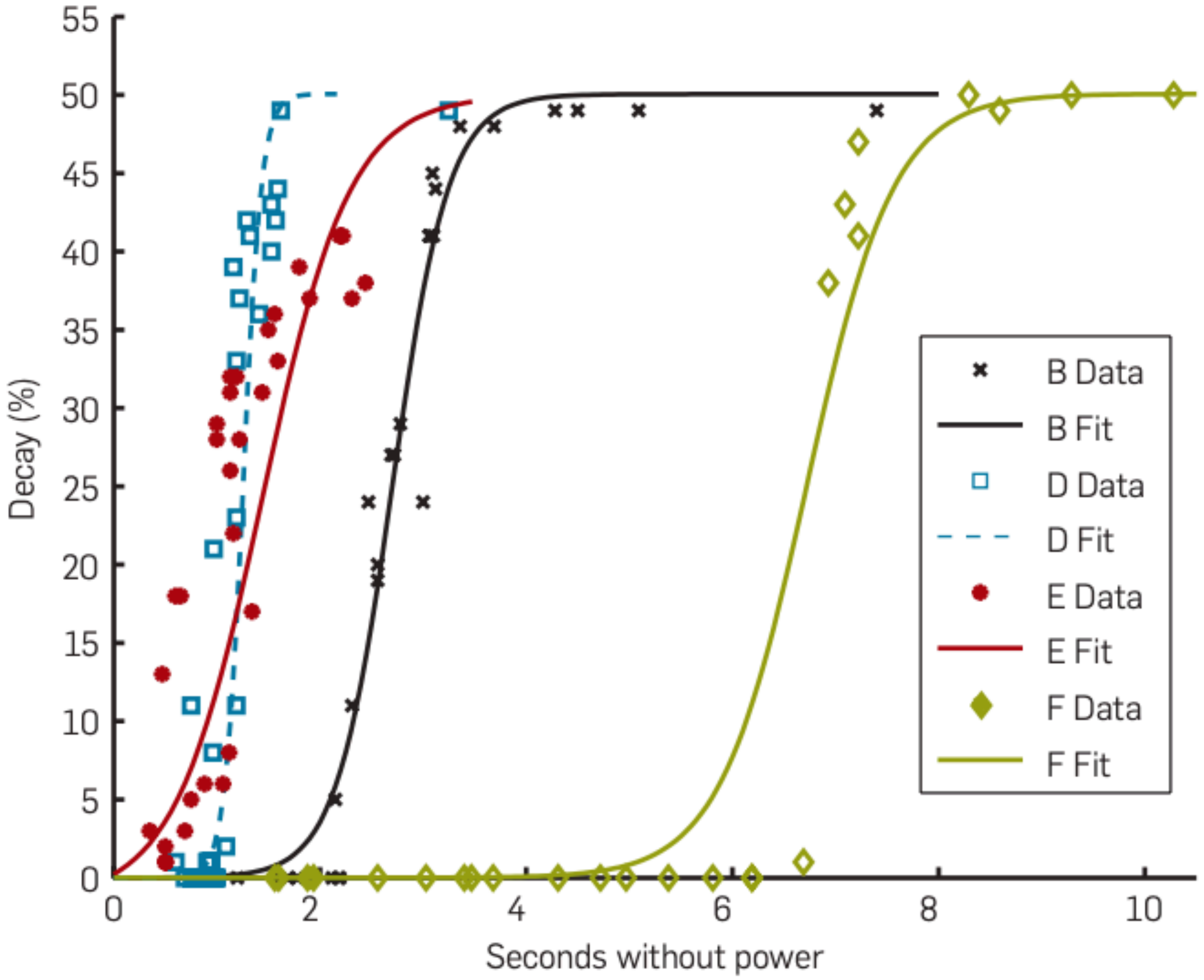
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



Reduced Temperature Experiments

- How did changing the temperature affect decay times?

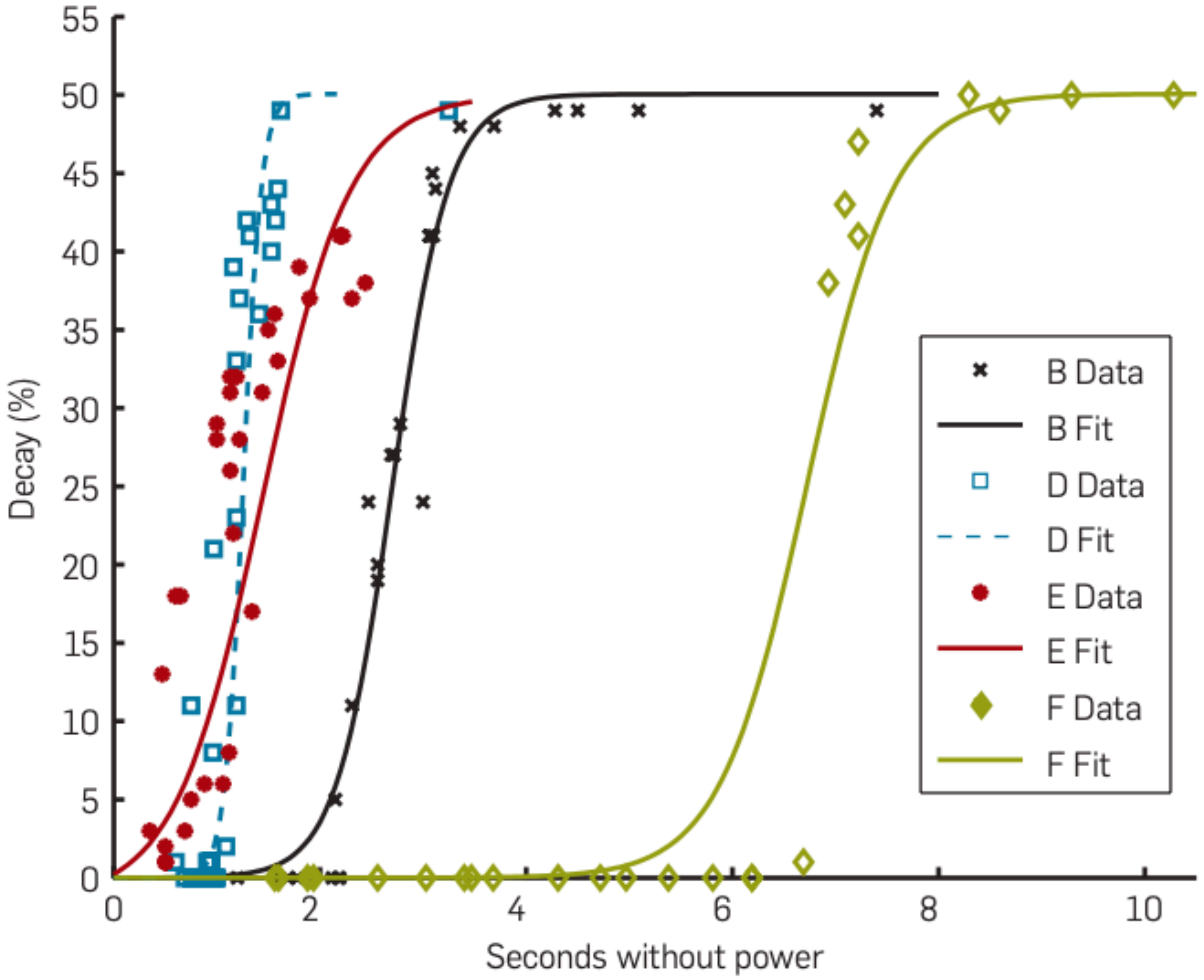
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.



Reduced Temperature Experiments

- How did changing the temperature affect decay times?
  - Reduced temperature to -50 degrees celsius (-58 degrees F) – attacker could cut power for 1 minute and recover at least 99.9% of bits correctly

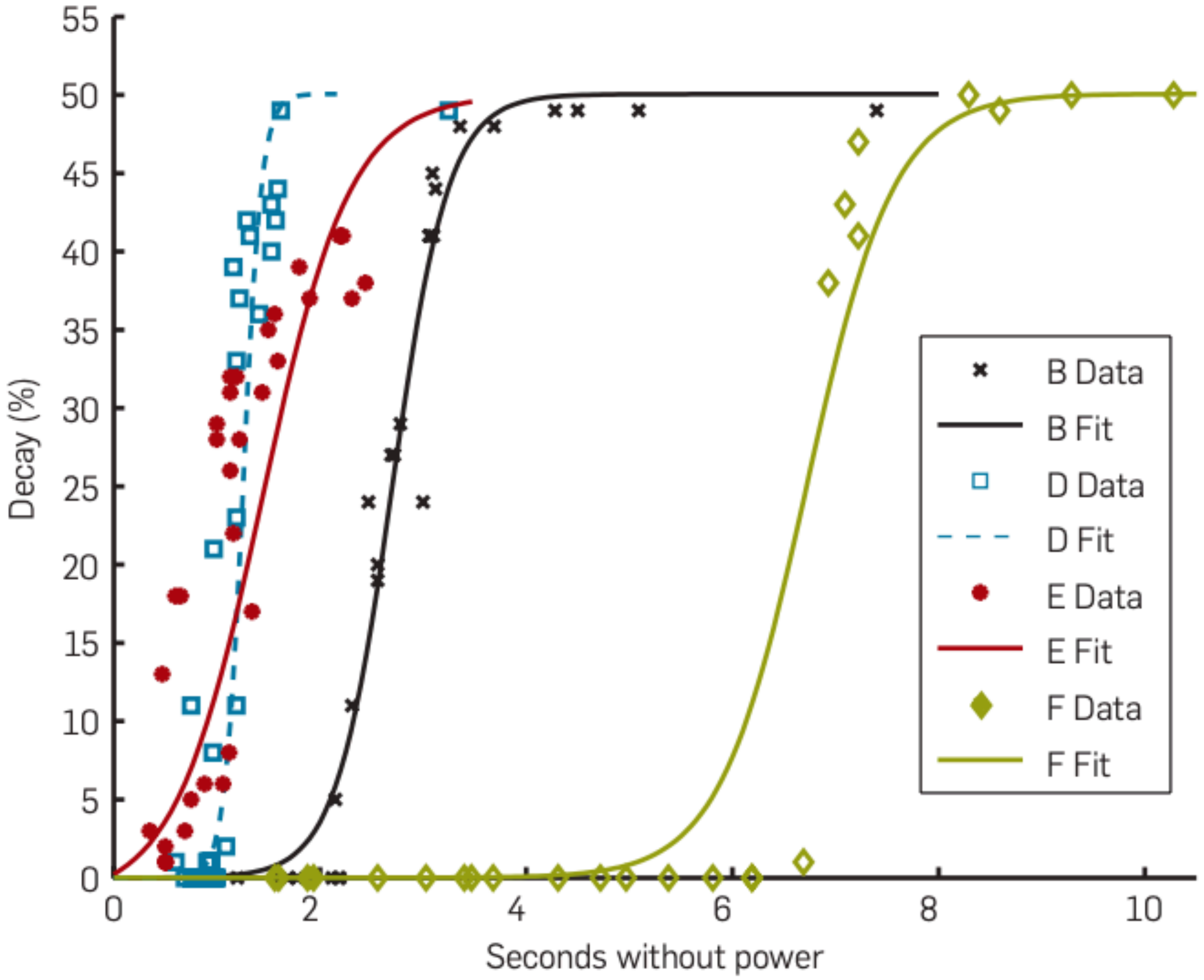
Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.

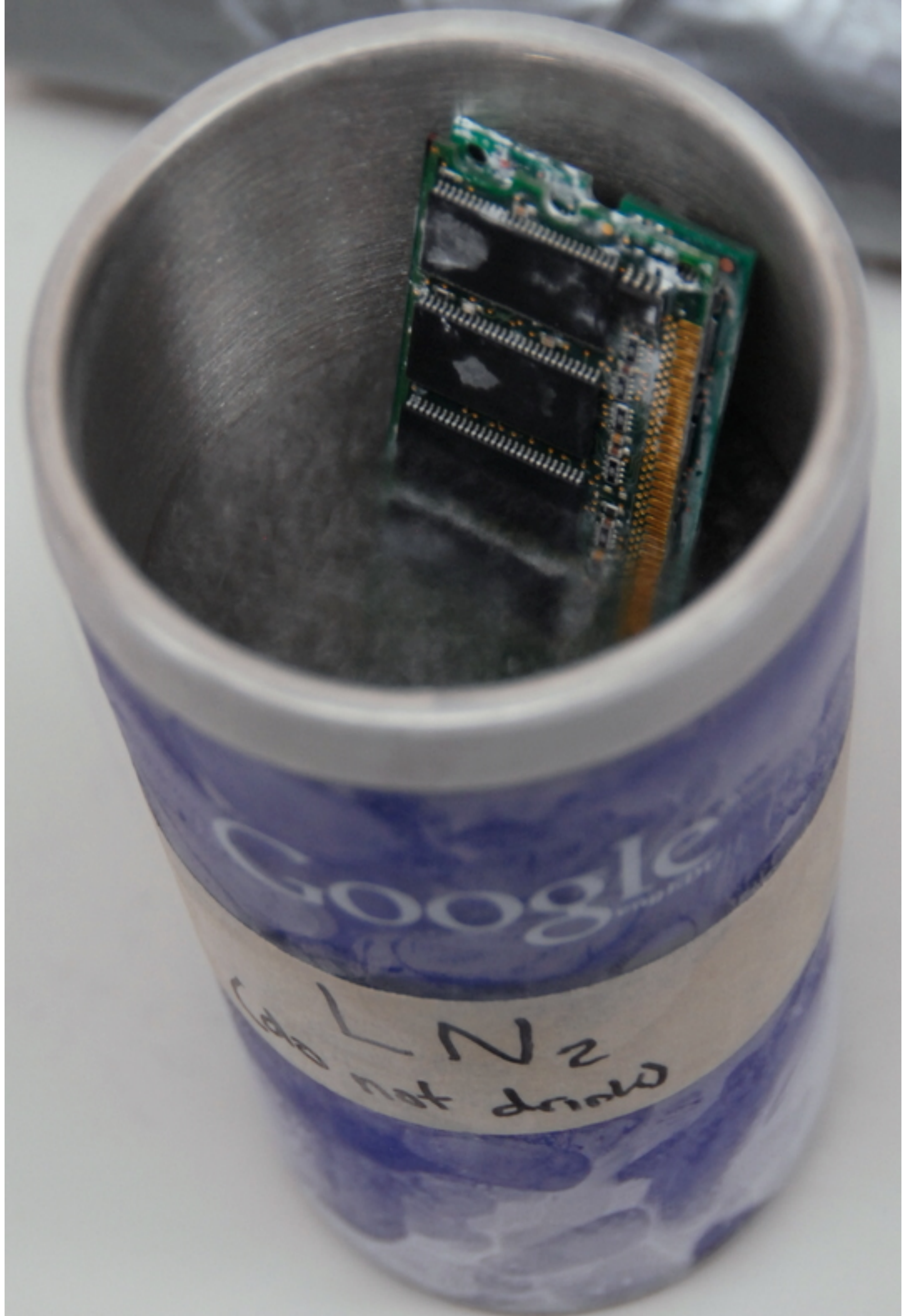


Reduced Temperature Experiments

- How did changing the temperature affect decay times?
    - Reduced temperature to -50 degrees celsius (-58 degrees F) – attacker could cut power for 1 minute and recover at least 99.9% of bits correctly
- How did the authors make it even colder?

Figure 2: Measuring decay. We measured memory decay after various intervals without power. The memories were running at normal operating temperature, without any special cooling. Curves for machines A and C would be off the scale to the right, with rapid decay at around 30 and 15 s, respectively.




Google
LN2
do not drink

Getting data out of DRAM

- Turns out, this is nontrivial
    - Rebooting immediately starts to refresh DRAM cells, which could **erase** memory that was previously written in those cells
- You could write a small program that can copy memory to another medium
    - Authors essentially created a few extremely tiny programs to do this to show proof of concept – we won't discuss this but it's pretty cool

What can we do with this?

- Paper goes into lots of fun detail about reconstructing cryptographic keys
  - DES, AES, RSA —> all of these can be reconstructed
- What is the fundamental reason why keys can be reconstructed?

What can we do with this?

- Paper goes into lots of fun detail about reconstructing cryptographic keys
  - DES, AES, RSA —> all of these can be reconstructed
- What is the fundamental reason why keys can be reconstructed?
  - Keys have unique signatures: AES has a **key schedule** with a repeatable pattern that you can exploit as a side channel to find an AES key :)

Attacking Encrypted Disks

- What is an encrypted disk?

Attacking Encrypted Disks

- What is an encrypted disk?
- What is on-the-fly encryption?

Attacking Encrypted Disks

- What is an encrypted disk?
- What is on-the-fly encryption?
- How does BitLocker encrypt data on the disk?

Attacking Encrypted Disks

- What is an encrypted disk?
- What is on-the-fly encryption?
- How does BitLocker encrypt data on the disk?
  - **Authors defeated BitLocker, FileVault, TrueCrypt, dm-crypt, Loop-AES, and could probably have done a lot more!**

Feasibility

- How feasible is this attack?
- Do you believe this attack will work in practice? Why or why not?

Discussion

What about these attacks *surprised* you?

What do these attacks teach us about *trust*?

What can we do about side channels?

For next time…

- Read two IoT papers (spoiler alert: one of them is mine)
- Keep meeting with your teams!