Circuit Complexity: Autumn 2024

Course Summary & Review

Instructor: William Hoza

The University of Chicago

Circuits vs. Turing machines

- Let $f: \{0,1\}^* \rightarrow \{0,1\}$
- **Theorem:** The following are equivalent:
  - f can be computed by poly-size circuits ($f \in \mathrm{PSIZE}$)
  - f can be computed by a poly-time Turing machine with a poly-length advice string ($f \in \mathrm{P/poly}$)
- **Adleman's Theorem:** $\mathrm{BPP} \subseteq \mathrm{P/poly}$

Circuit complexity and P vs. NP

- **Shannon's Counting Argument:** For most functions $f: \{0,1\}^n \to \{0,1\}$, the circuit complexity of f is $\Omega(2^n/n)$
- If you can show $\exists f \in \mathrm{NP}$ with circuit complexity $n^{\omega(1)}$, then it follows that $\mathrm{P} \neq \mathrm{NP}$
- So far, the best circuit complexity lower bound for a function in NP is approximately $3.1 \cdot n$ [Li, Yang 2022]

Shallow circuits

- We have better tools for reasoning about shallow circuits
- Constant-depth circuits represent ultra-fast parallel algorithms
- Depth ≈ Time
- Size ≈ Work

Shallow circuits can do interesting stuff

- Examples of problems in NC^0:
  - Three-to-two addition
- Examples of problems in AC^0:
  - Integer addition
  - Promise majority (Exercise 4)

- *"Local functions"*
- *Each output bit depends on $O(1)$ input bits*

Shallow circuits can do interesting stuff

- Examples of problems in $\mathrm{AC}^0[\oplus]$:
  - Nisan-Wigderson PRG
- Examples of problems in TC^0:
  - All symmetric functions ($SYM \subseteq \mathrm{TC}^0$)
  - Iterated integer addition (Exercise 5)
  - Candidate cryptographic PRFs

$\mathrm{TC}^0 \approx$ *Neural Networks*

Shallow circuits can do interesting stuff

- Examples of problems in NC^1:
  - Majority ($\mathrm{TC}^0 \subseteq \mathrm{NC}^1$)
- Examples of problems in AC^1:
  - s-t connectivity ($\mathrm{NL} \subseteq \mathrm{AC}^1$)

The complexity class $\mathbf{AC^0}$

- AC^0 is one of my favorite complexity classes!
- The theory of AC^0 is a "mini complexity theory"
- Maybe someday, your great-grandchildren will understand $\mathrm{P/poly}$ as thoroughly as we understand AC^0 today…
- Studying AC^0 gives us a taste of that glorious future 🥹

The Razborov-Smolensky method

- Let $C: \{0,1\}^n \to \{0,1\}$ be an AC_d^0 circuit of size $S \geq n$
- Let $\mathbb{F}$ be any field and let $\epsilon \in (0,1)$
- **Theorem:** There exists a probabilistic polynomial P over $\mathbb{F}$ that computes C with error ϵ and degree $O(\log S \cdot \log(S/\epsilon))^d$
- In contrast, the parity function cannot be approximated by low-degree polynomials over $\mathbb{F}_3$, hence $\mathrm{PARITY} \notin \mathrm{AC}^0$

Weak polynomial representations

- Let $C: \{0,1\}^n \to \{0,1\}$ be a $\mathrm{MAJ} \circ \mathrm{AC}_d^0$ circuit of size $S \geq n$
- Let $f: \{0,1\}^n \to \{0,1\}$ be a function that agrees with C on $1/2 + \epsilon$ fraction of inputs
- **Theorem:** The function f has a weak polynomial representation of degree $n - \Omega(\epsilon \cdot \sqrt{n}) + (\log S)^{O(d)}$
- In contrast, the parity function has no nontrivial weak polynomial representation, hence $\mathrm{PARITY} \notin \mathrm{MAJ} \circ \mathrm{AC}^0$

Impagliazzo's Hard-Core Lemma

- Let $\mathcal{C}$ be a circuit class and let $h: \{0,1\}^n \to \{0,1\}$
- Assume that $\forall C \in \mathrm{MAJ}_t \circ \mathcal{C}$, we have $\Pr_x[C(x) = h(x)] \leq 0.9$
- **Impagliazzo's Hard-Core Lemma:** There exists a set $H \subseteq \{0,1\}^n$ of size $\Omega(2^n)$ such that $\forall C \in \mathcal{C}$, we have

$$\Pr_{x \in H}[C(x) = h(x)] \leq \frac{1}{2} + O\left(1/\sqrt{t}\right)$$

Ignoring some technicalities…

Yao's XOR Lemma

- Let $\mathcal{C}$ be a circuit class and let $h: \{0,1\}^n \to \{0,1\}$
- Assume that $\forall C \in \mathrm{MAJ}_t \circ \mathcal{C}$, we have $\Pr_x[C(x) = h(x)] \leq 0.9$
- **Yao's XOR Lemma:** $\forall C \in \mathcal{C}, \forall k \in \mathbb{N}$, we have

Ignoring some technicalities…

$$\Pr_x\left[C(x) = h^{\oplus k}(x)\right] \leq \frac{1}{2} + 2^{-\Omega(k)} + O\left(1/\sqrt{t}\right)$$

- Consequence: Correlation between PARITY and AC^0 is exponentially small

Nisan-Wigderson Pseudorandom Generator

- Let $n, S, d \in \mathbb{N}$ and $\epsilon \in (0, 1)$ where $S \geq n$
- **Theorem:** There exists a PRG $G: \{0,1\}^s \to \{0,1\}^n$ such that:
  - (Fooling) For every AC_d^0 circuit C of size at most S, we have

$$\left|\Pr_x[C(G(x)) = 1] - \Pr_y[C(y) = 1]\right| \leq \epsilon$$

  - (Efficiency) Given n, S, d, ϵ, x, the string $G(x)$ can be computed in $\mathrm{poly}(n)$ time
  - (Seed length) We have $s = (\log(S/\epsilon))^{O(d)}$

More sophisticated PRGs are known, with better seed lengths

AC^0 can be computed by probabilistic polynomials

$\Downarrow$

$\mathrm{MAJ} \circ \mathrm{AC}^0$ has weak polynomial representations

$\Downarrow$

Correlation between PARITY and $\mathrm{MAJ} \circ \mathrm{AC}^0$ is $o(1)$

Von Neumann's Minimax Theorem

$\Downarrow$

Impagliazzo's Hard-Core Lemma

$\Downarrow$

Yao's XOR Lemma

$\Searrow$ $\Swarrow$

Correlation between PARITY and AC^0 is $2^{-n^{\Omega(1)}}$

$\Downarrow$

Nisan-Wigderson PRG

The Switching Lemma

- Distribution R_p over $\{0,1,\star\}^n$: For each variable independently, keep it alive with probability p, otherwise assign a random value
- **The Switching Lemma:** If C is a width-w DNF/CNF, then

$$\Pr_{\rho \sim R_p}\left[\mathrm{DTDepth}\left(C|_\rho\right) \geq D\right] \leq O(pw)^D$$

- For example, when $D = 1$, we get $\Pr\left[C|_\rho \text{ is nonconstant}\right] \leq O(pw)$

The AC^0 Criticality Theorem

- Let $C: \{0,1\}^n \to \{0,1\}$ be an AC^0_d circuit of size S
- **$\mathbf{AC}^0$ Criticality Theorem:** $\Pr_{\rho \sim R_p}\left[\mathrm{DTDepth}\left(C|_\rho\right) \geq D\right] \leq \left(p \cdot O(\log S)^{d-1}\right)^D$
- In contrast, the parity function does not simplify under restrictions, hence

$$\Pr_x[C(x) = \mathrm{PARITY}_n(x)] \leq \frac{1}{2} + 2^{-n/O(\log S)^{d-1}}$$

Fourier analysis of Boolean functions

- **Fact:** Every function $C\colon \{\pm 1\}^n \to \{\pm 1\}$ can be uniquely written as a multilinear polynomial:

Correlation

$$C(x) = \sum_{S \subseteq [n]} \hat{C}(S) \cdot \chi_S(x)$$

Parity functions

- **Parseval's Theorem:** $\sum_S \hat{C}(S)^2 = \mathbb{E}_x[C(x)^2] = 1$

AC^0 Fourier tail bound

- Let $C: \{\pm 1\}^n \to \{\pm 1\}$ be an AC^0_d circuit of size S
- $\mathbf{AC^0}$ **Fourier Tail Bound**, aka **LMN Theorem**: For all $k \in \mathbb{N}$, we have

$$\sum_{S \subseteq [n], |S| \geq k} \hat{C}(S)^2 \leq 2 \cdot 2^{-k/O(\log S)^{d-1}}$$

- Consequence: AC^0 circuits are learnable in quasipolynomial time under the uniform distribution, given random labeled examples

Limited independence fools $\mathbf{AC}^0$

- Let $d \in \mathbb{N}$ be a constant
- **Braverman's Theorem:** $\forall S \in \mathbb{N}, \forall \epsilon \in (0,1), \exists k = \mathrm{polylog}(S) \cdot \log(1/\epsilon)$ such that if $C: \{0,1\}^n \to \{0,1\}$ is an AC_d^0 circuit of size $S \geq n$ and X is k-wise uniform, then

$$|\Pr[C(X) = 1] - \Pr[C(U_n) = 1]| \leq \epsilon$$

- Follows from construction of low-degree sandwiching polynomials

(Multi) Switching Lemma

$\Downarrow$

AC^0 Criticality Theorem

$\Downarrow$

AC^0 Fourier tail bound

$\swarrow \quad \Downarrow \quad \searrow$

Optimal bound on the correlation between parity and AC^0

Learnability of AC^0

Limited independence fools AC^0

Beyond $\mathbf{AC}^0$: Sipser's program

- Strategy for proving $\mathrm{P} \neq \mathrm{NP}$: Prove $\mathrm{NP} \nsubseteq \mathcal{C}$ for stronger and stronger $\mathcal{C}$ until eventually we prove $\mathrm{NP} \nsubseteq \mathrm{P/poly}$
- $\mathrm{PARITY} \notin \mathrm{AC}^0$ ✔
- If p is a power of a prime, then $\mathrm{MAJORITY} \notin \mathrm{AC}^0[p]$ ✔
- Open problem: Prove $\mathrm{NP} \nsubseteq \mathrm{AC}^0[6]$…

The frontier of Sipser's program: ACC

$\mathrm{NQP} =$ *Nondeterministic Quasipoly Time*

$\mathrm{ACC} = \bigcup_m \mathrm{AC}^0[m]$

- **Theorem** [Murray, Williams 2018]: $\mathrm{NQP} \not\subseteq \mathrm{ACC}$
- Proof step 1: Every $C \in \mathrm{AC}^0[m]$ can be computed by a SYM of AND of literals, where the SYM has quasipoly fan-in and each AND has polylog fan-in
- Proof step 2: There is a nontrivial satisfiability algorithm for $\mathrm{AC}^0[m]$ circuits
- Proof step 3: Nontrivial satisfiability algorithms imply lower bounds
  - This last step is not specific to ACC

Natural properties

Can also define AC^0*-natural,* NC^1*-natural, etc.*

- Why has Sipser's program stalled? How can we make progress?
- We say that H is a P-natural property of Boolean functions if:
  - Density: If we pick $f: \{0,1\}^n \to \{0,1\}$ u.a.r., then $\Pr[f \text{ has property } H] \geq 2^{-O(n)}$
  - Constructivity: Can determine whether f has property H in time $2^{O(n)}$, given the 2^n-bit truth table of f
- We say H is useful against $\mathcal{C}$ if functions in $\mathcal{C}$ do not have property H

How powerful are natural proofs?

- **Theorem:** There exists an AC^0-natural property that is useful against AC^0

Random restrictions

- **Theorem:** There does not exist an AC^0-natural property that is useful against $\mathrm{AC}^0[\oplus]$

Nisan-Wigderson PRG

Naor-Reingold PRF

- **Theorem:** Under appropriate cryptographic assumptions, there does not exist a P-natural property that is useful against TC^0

Natural proofs: Interpretation

- Conventional interpretation:
  - We ought to study non-natural proof techniques
  - That way, someday, we can prove $\mathrm{NP} \not\subseteq \mathrm{TC}^0$, and eventually $\mathrm{NP} \not\subseteq \mathrm{P}/\mathrm{poly}$
- Another possibility: Candidate PRFs such as Naor-Reingold are insecure
- Yet another possibility: $\mathrm{NP} \subseteq \mathrm{TC}^0$

The complexity class $\mathbf{NC}^1$

- **Theorem:** For any $f: \{0,1\}^* \to \{0,1\}$, the following are equivalent:
  - $f \in \mathrm{NC}^1$ (log-depth poly-size circuits with bounded fan-in)
  - f can be computed by a De Morgan formula with poly leafsize
    - **"Formula Balancing Lemma"**
  - f can be computed by poly-length constant-width branching programs
    - **"Barrington's Theorem"**

Computing with $O(1)$ bits of memory

Formula lower bounds

- Andreev's function $A: \{0,1\}^{2n} \to \{0,1\}$ is defined by

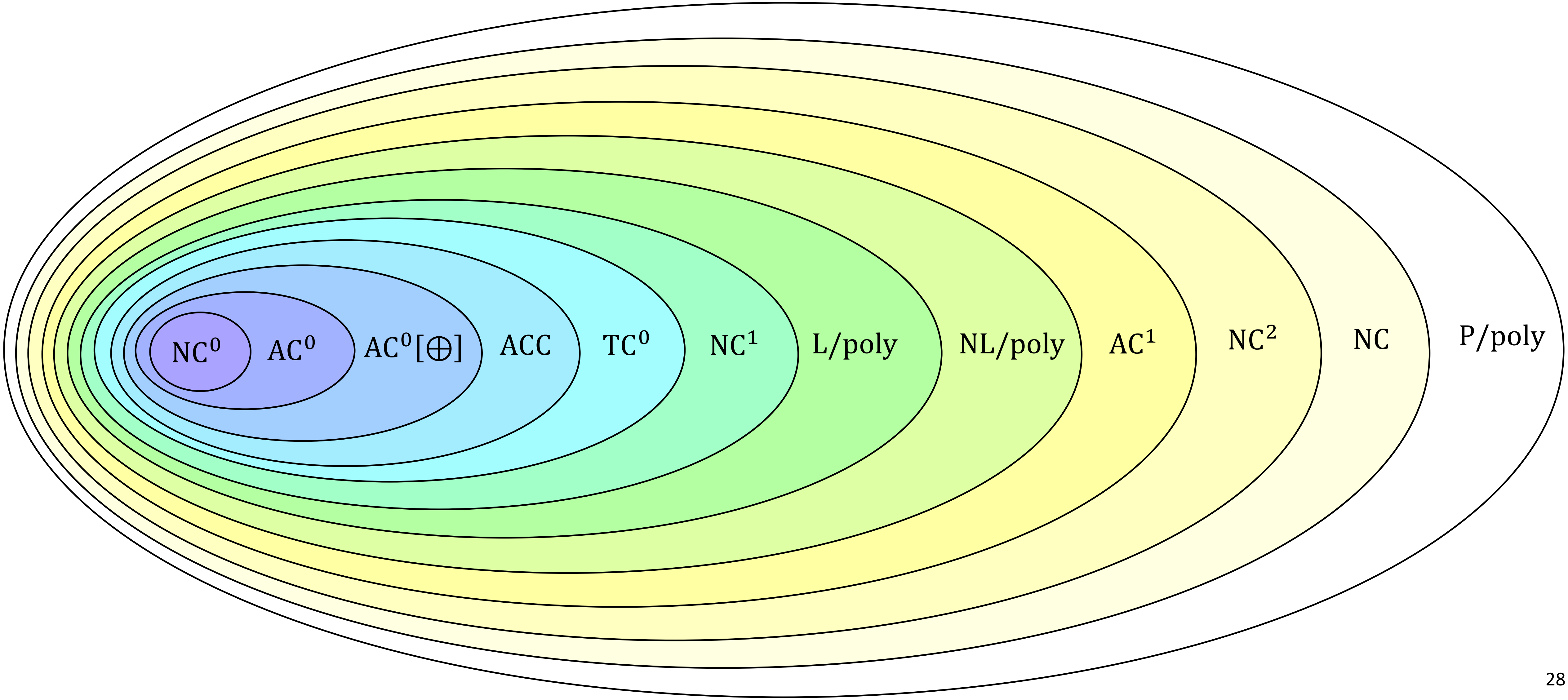
$$A\left(f, x^{(1)}, \dots, x^{(\log n)}\right) = f\left(\mathrm{PARITY}\left(x^{(1)}\right), \dots, \mathrm{PARITY}\left(x^{(\log n)}\right)\right)$$

- **Theorem:** $L(A) \geq \widetilde{\Omega}(n^3)$, where $L(\cdot)$ is De Morgan leafsize
- Proof is based on shrinkage of De Morgan formulas:

$$\mathbb{E}_{\rho \sim R_p}\left[L\left(f|_\rho\right)\right] \leq O\left(p^2 \cdot L(f) + p \cdot \sqrt{L(f)}\right)$$

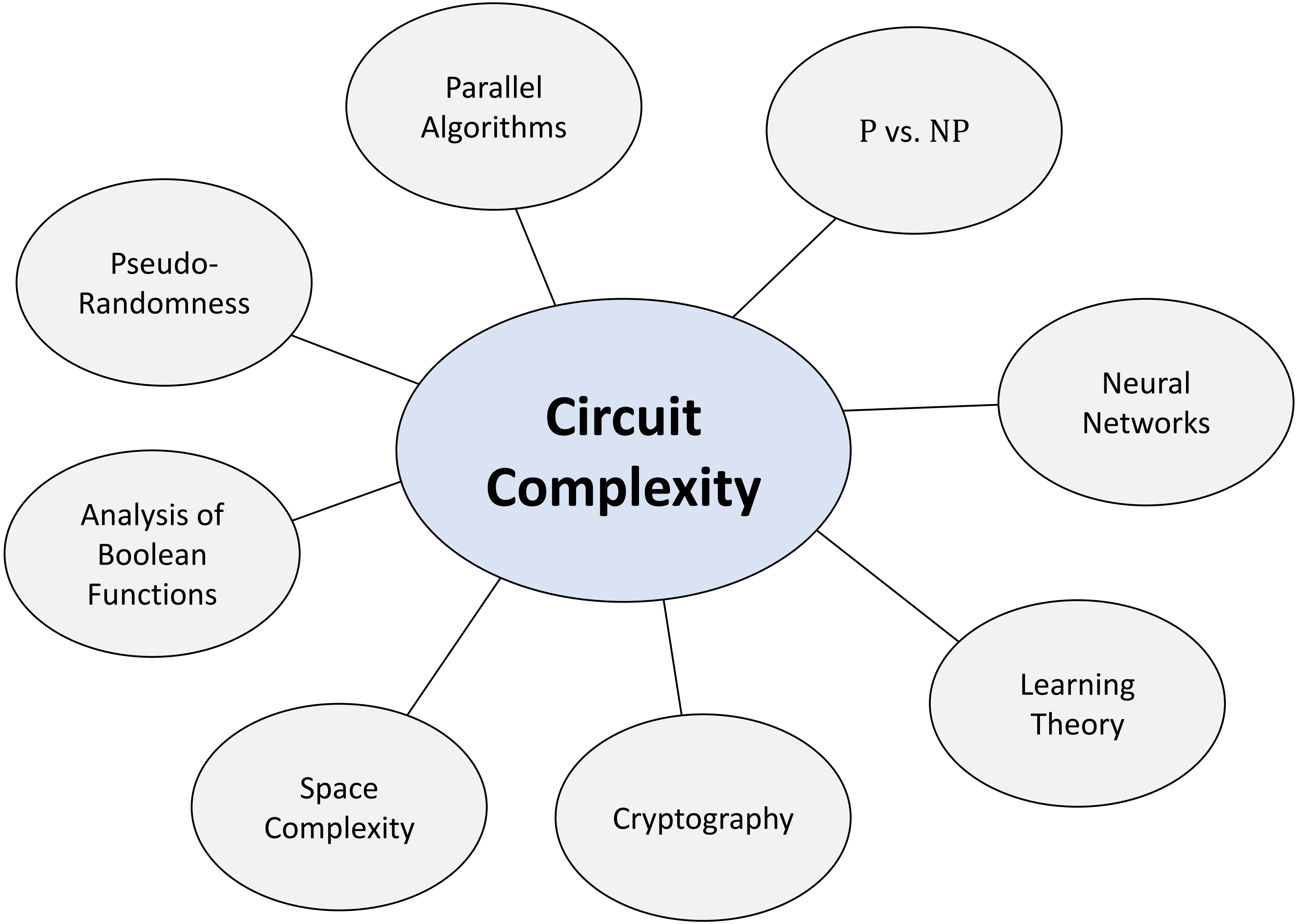
Summary of complexity classes

$\mathrm{NC}^0 \neq \mathrm{AC}^0 \neq \mathrm{AC}^0[\oplus] \neq \mathrm{ACC}$



A few of the many topics we didn't discuss

- Arithmetic circuits ($+$ and $\times$ gates)
- Monotone circuit lower bounds
- Connections between circuit complexity and communication complexity
- (Weak) TC^0 lower bounds


Circuit Complexity
Parallel Algorithms
P vs. NP
Neural Networks
Learning Theory
Cryptography
Space Complexity
Analysis of Boolean Functions
Pseudo-Randomness

Advertisement

- Consider enrolling in my seminar course next quarter!
- Topic: Derandomizing Space-Bounded Computation
  - Is randomness ever necessary for space-efficient computation?
- Less emphasis on exercises, more emphasis on cutting-edge research
  - Will not count as a graduate elective
- Also consider Sasha Razborov's complexity theory course in the spring!

Thank you!

- Being your instructor has been a privilege
- I look forward to reading your expositions
- Please fill out the Graduate Course Feedback Form using My.UChicago (deadline is Sunday, December 15)