CS272 - Theoretical Foundations of Learning, Decisions, and Games

# Lecture 14: Special lecture on Auctions (Annie Ulichney)

10/14

*Lecturer: Nika Haghtalab* *Readings: Blum [2024a], Blum [2024b], Roughgarden [2013]*
*Scribe: Ashley Zhang and Chenran Weng*

## 1 Overview

**Mechanism Design** is the field concerned with designing the rules of a game to induce desirable outcomes when participants act strategically in their own self-interest.

A mechanism typically proceeds in two steps:

1. The designer elicits information or preferences from the participants.
2. Based on the reports received, the mechanism determines an outcome (e.g., allocation and payments).

The fundamental challenge is that the participants' preferences are private and unknown to the designer. The mechanism must therefore rely on participants to *self-report* their preferences, which opens the possibility of strategic misreporting. The key idea in mechanism design is to construct rules so that participants have an incentive to report truthfully, which is formalized through notions of incentive compatibility.

Mechanism design is often described as the "reverse" of game theory.

- **Game Theory:** The rules of the game are fixed. The objective is to analyze strategic interactions and predict the outcome(s).

  Fixed Rules. $\rightarrow$ What are the outcomes?

- **Mechanism Design:** The desired outcome or social choice rule is fixed. The objective is to design the rules that implement this goal when agents behave strategically.

  What rules? $\rightarrow$ Implement desired outcomes.

## 2 Motivating Example: Resource Allocation

Consider a simple resource allocation problem.

- **Setup:** There is a single indivisible item to be allocated among $n$ agents. Each agent $i$ has a private value $v_i \in \mathbb{R}_+$ for the item. This value represents the agent's true willingness to pay.

- **Goal:** The objective is to achieve *efficiency*, that is, to allocate the item to the agent who needs or wants it the most, i.e., the one with the highest value $v_i$.

Let's explore two potential mechanisms to achieve this goal.

## 2.1 Strategy 1: Direct Revelation

A straightforward approach is to simply ask the agents to report their values and give the item to the agent who reports the highest value.

- **Mechanism:** Ask each agent $i$ to report a value $\hat{v}_i$. Allocate the item to agent $j$ such that $\hat{v}_j = \max_i\{\hat{v}_i\}$.

This mechanism fails because there is no incentive for agents to be truthful. Since there is no cost associated with reporting one's value untruthfully, a strategic agent can report an arbitrarily high value in an attempt to receive the item. As a result, the mechanism fails to elicit truthful reports and may end up allocating the item inefficiently.

## 2.2 Strategy 2: First-Price Auction

To introduce a cost for reporting a high value, we can require the winner to pay what they reported.

- **Mechanism:** Ask each agent $i$ to report their willingness to pay, which we'll call a bid $b_i$. Allocate the item to the agent with the highest bid, and charge them the amount they bid.

To analyze behavior in this mechanism, we introduce a standard assumption about agent behavior.

**Assumption 2.1** (Quasi-linear Utility)**.** *The utility $u_i$ of an agent $i$ is their value for the item minus the price $p_i$ they pay.*

$$u_i = v_i - p_i.$$

*If an agent does not receive the item, their utility is 0. Agents will act to maximize their own utility.*

Under this assumption, in the first-price auction, an agent with value $v_i$ has an incentive to bid $b_i < v_i$. By bidding their true value, their utility would be $u_i = v_i - b_i = v_i - v_i = 0$. By bidding lower ("shading their bid"), they can potentially achieve a positive utility if they win. This means the mechanism does not elicit truthful preferences, and the item may not go to the person who values it most.

# 3 Single Item Auctions

The single-item auction is a canonical example in mechanism design. We'll establish the standard model for these auctions and analyze two common auction formats.

## 3.1 The Model

The setup consists of the following components:

- There are $n$ bidders, indexed by $i = 1, \ldots, n$ competing for a single, indivisible item.
- Each bidder $i$ has a private value $v_i \in \mathbb{R}_+$ for the item, representing their true maximum willingness to pay. The value $v_i$ is private information, meaning it is unknown to both the seller and the other bidders.
- If bidder $i$ wins the item and pays a price $p_i$, their utility is $u_i = v_i - p_i$.
- If bidder $i$ loses the auction, they receive no item and pay nothing, so their utility is $u_i = 0$.

We will focus on a common format called a sealed-bid auction.

- Each bidder $i$ privately and simultaneously submits a bid $b_i$ to the seller.
- The seller observes the bids and uses a predetermined set of rules to decide the **allocation** (who gets the item) and the **payment** (how much the winner pays).

## 3.2 Two Fundamental Auction Mechanisms

### 3.2.1 First-Price Auction

In a first-price sealed-bid auction, the rules are:

- **Allocation:** The item is awarded to the bidder who submitted the highest bid. Let $j$ be the winner, so $b_j = \max_i \{b_i\}$.
- **Payment:** The winner $j$ pays the price they bid, $p_j = b_j$.

As noted earlier, a rational bidder will not bid their true value in this auction. If they bid $b_j = v_j$ and win, their utility would be $u_j = v_j - b_j = v_j - v_j = 0$. To get a positive utility, they must bid less than their true value.

### 3.2.2 Second-Price Auction

This auction has a clever twist on the payment rule.

- **Allocation:** As in the first-price auction, the item is awarded to the bidder with the highest bid.
- **Payment:** The winner pays the price of the *second-highest* bid.

This change in the payment rule has a profound effect on bidding strategy: truthful bidding becomes a dominant strategy. We formalize this with the following claim.

**Claim 3.1.** *In a second-price auction, truthful bidding (i.e., setting one's bid $b_i$ equal to one's true value $v_i$) is a dominant strategy.*

*Proof.* Fix any bidder $i$ with a private value $v_i$. We will show that bidding $b_i = v_i$ is optimal for bidder $i$, regardless of the bids submitted by the other players.

Let $B = \max_{j \neq i} b_j$ be the highest bid among all other bidders. Bidder $i$'s outcome depends on how their bid $b_i$ compares to $B$. We analyze three cases.

**Case 1:** $v_i > B$ This is the case where bidder $i$'s true value is higher than anyone else's bid. The socially efficient outcome is for bidder $i$ to win.

- **Truthful bid ($b_i = v_i$):** Since $b_i > B$, bidder $i$ wins the auction and pays $B$. Their utility is $u_i = v_i - B$. Since $v_i > B$, this utility is strictly positive ($u_i > 0$).
- **Overbidding ($b_i > v_i$):** Bidder $i$ still wins (as $b_i > v_i > B$) and pays $B$. The utility is $u_i = v_i - B$. This yields no improvement over bidding truthfully.
- **Underbidding ($b_i < v_i$):**
  - If $B < b_i < v_i$, bidder $i$ still wins and pays $B$. The utility is $u_i = v_i - B$, which is no improvement.
  - If $b_i \leq B$, bidder $i$ loses the auction (or ties and loses, depending on the tie-breaking rule) and their utility is $u_i = 0$. This is strictly worse than the positive utility from bidding truthfully.

In this case, bidding truthfully is the best strategy. Any other bid gives either the same or a strictly worse outcome.

**Case 2:** $v_i < B$ Here, bidder $i$'s value is less than the highest bid from another player. The efficient outcome is for bidder $i$ to lose.

- **Truthful bid ($b_i = v_i$):** Since $b_i < B$, bidder $i$ loses the auction. Their utility is $u_i = 0$.
- **Underbidding ($b_i < v_i$):** Bidder $i$ still loses, and their utility remains $u_i = 0$. There is no change.
- **Overbidding ($b_i > v_i$):**
  - If $v_i < b_i < B$, bidder $i$ still loses. The utility is $u_i = 0$, which is no improvement.
  - If $b_i > B$, bidder $i$ now wins the auction. However, they must pay $B$. Their utility becomes $u_i = v_i - B$. Since we are in the case where $v_i < B$, this utility is strictly negative ($u_i < 0$). This is strictly worse than the zero utility from bidding truthfully.

Truthful bidding again weakly dominates all alternatives.

**Case 3:** $v_i = B$ The bidder's value is exactly equal to the highest competing bid.

- **Truthful bid ($b_i = v_i$):** Bidder $i$ either loses or ties for the highest bid. If they lose, $u_i = 0$. If they win (by tie-breaking), they pay $B = v_i$, so their utility is $u_i = v_i - B = 0$. In all scenarios, the utility is zero.

- Any other bid $b_i \neq v_i$ results in either losing (utility 0) or winning and paying $B$, which also results in utility $u_i = v_i - B = 0$.

The best possible utility is 0, which is achieved by bidding truthfully.

**Conclusion:** Across all possible scenarios for the other players' bids, bidding $b_i = v_i$ yields a utility that is at least as high as, and sometimes strictly higher than, the utility from any other possible bid $b_i' \neq v_i$. Therefore, truthful bidding is a dominant strategy. □

# 4 Desirable Properties of Mechanisms

Now that we have seen some examples, let's formalize the properties we want a mechanism to have. The designer's goal is to specify rules that lead to desirable outcomes even when participants act strategically to maximize their own utility.

## 4.1 Dominant Strategy Incentive Compatibility (DSIC)

A mechanism is **Dominant Strategy Incentive Compatible (DSIC)** if reporting truthfully is a dominant strategy for every agent. This means that, for each participant, truthful reporting maximizes their utility regardless of what any other agent reports.

Let the mechanism be defined by an allocation rule $f$ that determines the outcome (e.g., who receives the item) and a payment rule $p$ that specifies the transfers (e.g., how much each participant pays). Let $v_i$ be the true private value of agent $i$, $b_i$ be their reported value (or bid), and $b_{-i}$ be the vector of reports from all agents other than $i$. Agent $i$'s utility is:

$$u_i(b_i, b_{-i}) = v_i(f(b_i, b_{-i})) - p_i(b_i, b_{-i}).$$

**Definition 4.1** (Dominant-Strategy Incentive Compatibility)**.** A mechanism $(f, p)$ is **DSIC** if for every agent $i$, for every possible true value $v_i$, for every possible false report $b_i \neq v_i$, and for all possible reports $b_{-i}$ from the other agents, the following inequality holds:

$$v_i(f(v_i, b_{-i})) - p_i(v_i, b_{-i}) \geq v_i(f(b_i, b_{-i})) - p_i(b_i, b_{-i}).$$

In this inequality, the left-hand side represents agent $i$'s utility when reporting truthfully, while the right-hand side is their utility from misreporting. This inequality requires that truth-telling weakly dominate every possible deviation. As shown earlier, the second-price (Vickrey) auction is a DSIC mechanism.

## 4.2 Economic Efficiency

A second desirable property of a mechanism is efficiency. A mechanism is efficient if it maximizes total social welfare, that is, the sum of all agents' values for the chosen outcome.

In the single-item auction setting, an allocation can be represented by a vector $x = (x_1, \dots, x_n)$, where $x_i$ is an indicator variable for agent $i$ receiving the item:

$$x_i = \mathbb{I}\{\text{agent } i \text{ wins}\}.$$

The **social welfare** of an allocation $x$ is

$$W(x; v) = \sum_{i=1}^{n} v_i x_i.$$

A mechanism is called **efficient** if its allocation rule always selects the outcome that maximizes welfare:

$$f(v) \in \arg\max_x W(x; v).$$

In the single-item auction case, efficiency means allocating the item to the person with the highest valuation. Finally, a mechanism should ensure that agents are willing to participate voluntarily.

**Definition 4.2** (Individual Rationality (IR))**.** A mechanism is **individually rational (IR)** if for every agent $i$ and for every possible vector of true values $v = (v_1, \dots, v_n)$, their utility is non-negative when everyone reports truthfully:

$$v_i(f(v)) - p_i(v) \geq 0.$$

Here, $f(v)$ is the allocation function. This condition guarantees that participation never makes an agent worse off than sitting out of the mechanism. It is also referred to as a participation constraint.

# 5 The Vickrey-Clarke-Groves (VCG) Mechanism

The **Vickrey-Clarke-Groves (VCG)** mechanism is a general method for constructing mechanisms that are both DSIC and efficient. It generalizes the logic of the second-price (Vickrey) auction to a broad class of settings.

## 5.1 The General Setting

The VCG framework is defined more broadly:

- There are $n$ players (or agents), indexed by $i = \{1, \dots, n\}$.
- There is a finite set of possible outcomes, $\mathcal{A}$.

- Each agent $i$ has a private valuation function $v_i : \mathcal{A} \to \mathbb{R}$, which specifies their value for each possible outcome.
- For an outcome $a \in \mathcal{A}$ and a payment $p_i$, the utility for player $i$ is $u_i(a, p_i) = v_i(a) - p_i$.

Let $v = (v_1, \ldots, v_n)$ denote the vector of all players' valuation functions. The social welfare for a given outcome $a$ is the sum of all players' values for that outcome:

$$W(a; v) := \sum_{j=1}^{n} v_j(a).$$

The maximum possible social welfare for a given set of valuations $V$ is given by:

$$W^*(v) := \max_{a \in \mathcal{A}} W(a; v).$$

## 5.2 The VCG Mechanism

VCG is a direct revelation mechanism: each agent reports a valuation function $\hat{v}_i$. Let $\hat{v} = (\hat{v}_1, \ldots, \hat{v}_n)$ be the vector of reported valuations. computes the outcome (allocation) and payments based on the reported profile.

1. **Allocation Rule:** Choose the outcome $a^*$ that maximizes the social welfare based on the reported valuations.
$$a^* = f(\hat{v}) = \arg\max_{a \in \mathcal{A}} \sum_{i=1}^{n} \hat{v}_i(a).$$
2. **Payment Rule:** Each player $i$ pays an amount equal to the "harm" their presence imposes on the welfare of the other players.
$$p_i(\hat{v}) = \left( \max_{a \in \mathcal{A}} \sum_{j \neq i} \hat{v}_j(a) \right) - \left( \sum_{j \neq i} \hat{v}_j(a^*) \right).$$
The first term is the maximum possible welfare for everyone else if player $i$ did not play. The second term is the actual welfare everyone else gets in the chosen outcome $a^*$.

It can be shown that the VCG mechanism is both DSIC and efficient. The second-price auction is a special case of VCG applied to the single-item allocation problem.

## 5.3 Truthfulness in VCG

**Claim 5.1.** *The VCG mechanism is DSIC.*

*Proof.* Fix any agent $i$ with true valuation $v_i$, and fix the reported valuations of all other players, $v_{-i}$. We will show that player $i$'s utility is maximized by reporting $\hat{v}_i = v_i$.

Let's first calculate the utility for player $i$ if they are truthful.

$$\begin{aligned} u_i(\text{truthful}) &= v_i(a^*) - p_i(v) \\ &= v_i(a^*) - \left[\max_{a \in \mathcal{A}} \sum_{j \neq i} v_j(a) - \sum_{j \neq i} v_j(a^*)\right] \\ &= \left(v_i(a^*) + \sum_{j \neq i} v_j(a^*)\right) - \max_{a \in \mathcal{A}} \sum_{j \neq i} v_j(a) \\ &= \sum_{j=1}^{n} v_j(a^*) - \max_{a \in \mathcal{A}} \sum_{j \neq i} v_j(a) \\ &= \max_{a \in \mathcal{A}} \sum_{j=1}^{n} v_j(a) - \left(\max_{a \in \mathcal{A}} \sum_{j \neq i} v_j(a)\right). \end{aligned} \tag{1}$$

Then, what is player $i$'s utility if they lie and report a different valuation $\hat{v}_i \neq v_i$? Let the reported profile be $v' = (\hat{v}_i, v_{-i})$.

$$\begin{aligned} u_i(\text{dishonest}) &= v_i(f(v')) - p_i(v') \\ &= \sum_{j=1}^{n} v_j(f(v')) - \max_{a \in \mathcal{A}} \sum_{j \neq i} v_j(a). \end{aligned} \tag{2}$$

Now we compare the utility from being truthful versus being dishonest.

$$(1) - (2) = \max_{a \in \mathcal{A}} \sum_{j=1}^{n} v_j(a) - \sum_{j=1}^{n} v_j(f(v')) \geq 0.$$

This implies that:

$$u_i(\text{truthful}) \geq u_i(\text{dishonest}).$$

Hence, truthful reporting is a dominant strategy. □

*Remark* 5.2. VCG mechanisms are also efficient. The allocation rule $f(v)$ explicitly maximizes total welfare. Moreover, under nonnegative valuations, they are individually rational and involve no positive transfers (i.e., the mechanism will never pay agents).

## 5.4 The Second-Price Auction as a Special Case

We can also show that the familiar second-price auction is simply a special case of the more general VCG mechanism. To see this, we define the components of the VCG framework for the single-item auction setting:

- **Set of Outcomes $\mathcal{A}$:** An outcome is the allocation of the single item. We can represent the outcome where player $i$ wins with the basis vector $e_i$, which is a vector of zeros with a 1 in the $i$-th position. So, $\mathcal{A} = \{e_1, e_2, \ldots, e_n\}$.
- **Valuation Functions $v_i$:** Player $i$'s value for the outcome $e_j$ (where player $j$ wins) is their private value $v_i$ if they are the winner ($i = j$), and 0 otherwise.

$$v_i(e_j) = \begin{cases} v_i & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Under this specification, the allocation rule assigns the item to the highest bidder $i = \arg\max_j v_j$ and the payment rule yields $p_i = \max_{j \neq i} v_j$ for the winner and $p_i = 0$ for all other agents. This is precisely the second-price auction from earlier.

# 6 The Revelation Principle

The Revelation Principle is a foundational concept in mechanism design. It states that if a certain social goal can be achieved by *any* mechanism, it can also be achieved by a simple, direct, and truthful one.

**Principle 1** (The Revelation Principle)**.** *For every mechanism $M$ in which every player has a dominant strategy, there exists an equivalent direct-revelation, DSIC mechanism $M'$.*

# 7 Single-Parameter Environments

While the VCG mechanism is powerful and general, many important problems have a simpler structure. We now focus on a common and important class of problems known as single-parameter environments.

## 7.1 Definition of the Environment

A mechanism design problem is in a **single-parameter environment** if it satisfies the following conditions:

- There are $n$ agents. Each agent $i$ has a private value $v_i \in \mathbb{R}_+$ that can be summarized by a single parameter.
- The mechanism chooses an allocation $x(b) = (x_1(b), \ldots, x_n(b)) \in \mathcal{X}$ where $\mathcal{X}$ is the set of all feasible allocations.
- Each agent's utility is quasi-linear:

  $$u_i(b) = v_i x_i(b) - p_i(b),$$

  where $p_i(b)$ is the payment determined by the mechanism.

In this setting, we often focus on designing the allocation rule first, and then determining if a payment rule exists to make it incentive-compatible. This leads to two important definitions.

**Definition 7.1** (Implementable Allocation)**.** An allocation rule $x$ is said to be **implementable** if there exists a payment rule $p$ such that the complete mechanism $(x, p)$ is DSIC.

**Definition 7.2** (Monotone Allocation)**.** Let player $i$'s reported bid be $b_i$, and the vector of other players' bids be $b_{-i}$. An allocation rule $x$ is **monotone** if for every player $i$ and any fixed bids $b_{-i}$ from the other players, the allocation to player $i$, $x_i(b_i, b_{-i})$, is non-decreasing in their bid $b_i$.

## 7.2 Myerson's Lemma

In Myerson [1981], Myerson characterizes the DSIC mechanisms in single-parameter environments.

**Theorem 7.3** (Myerson, 1981)**.** *For a single-parameter environment:*

1. *An allocation rule $x$ is implementable if and only if it is monotone.*

2. *If an allocation rule $x$ is monotone, there exists a unique payment rule $p$ (up to an additive constant) such that the mechanism $(x, p)$ is DSIC.*

3. *This unique payment rule $p$ has an explicit form*

$$p_i(b_i, b_{-i}) = b_i x_i(b_i, b_{-i}) - \int_0^{b_i} x_i(z, b_{-i}) dz.$$

# References


Noga Alon, Omri Ben-Eliezer, Yuval Dagan, Shay Moran, Moni Naor, and Eylon Yogev. Adversarial laws of large numbers and optimal regret in online classification. In *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2021, page 447–455, New York, NY, USA, 2021. Association for Computing Machinery. ISBN 9781450380539. doi: 10.1145/3406325.3451041. URL https://doi.org/10.1145/3406325.3451041.

Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund, and Robert E Schapire. The nonstochastic multiarmed bandit problem. *SIAM journal on computing*, 32(1):48–77, 2002.

Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *Annual Conference Computational Learning Theory*, 2009. URL https://api.semanticscholar.org/CorpusID:9403043.

Claude Berge. *Topological Spaces*. Oliver and Boyd, London, 1963.

Avrim Blum. Mechanism design i: Auctions, examples, and vcg, 2024a. URL https://home.ttic.edu/~avrim/AGT24/Lecture%209%20-%20Mechanism%20Design1.pdf. TTIC 31260 Algorithmic Game Theory, Lecture 9, April 17, 2024.

Avrim Blum. Mechanism design ii: more vcg and the revelation principle, 2024b. URL https://home.ttic.edu/~avrim/AGT24/Lecture%2010%20-%20Mechanism%20Design2.pdf. TTIC 31260 Algorithmic Game Theory, Lecture 10, April 22, 2024.

Noam Brown and Tuomas Sandholm. Superhuman ai for multiplayer poker. *Science*, 365(6456): 885–890, 2019.

Constantinos Daskalakis, Paul W. Goldberg, and Christos H. Papadimitriou. The Complexity of Computing a Nash Equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.

Yuan Deng, Jon Schneider, and Balasubramanian Sivan. Strategizing against no-regret learners. *Advances in neural information processing systems*, 32, 2019.

Martin Grötschel, László Lovász, and Alexander Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2. Springer Science & Business Media, 2012.

Nika Haghtalab, Thodoris Lykouris, Sloan Nietert, and Alexander Wei. Learning in stackelberg games with non-myopic agents. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 917–918, 2022.

Nika Haghtalab, Chara Podimata, and Kunhe Yang. Calibrated stackelberg games: Learning optimal commitments against calibrated agents. *Advances in Neural Information Processing Systems*, 36:61645–61677, 2023.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis with adaptive adversaries. *J. ACM*, 71(3), June 2024a. ISSN 0004-5411. doi: 10.1145/3656638. URL https://doi.org/10.1145/3656638.

Nika Haghtalab, Tim Roughgarden, and Abhishek Shetty. Smoothed analysis with adaptive adversaries. *Journal of the ACM*, 71(3):1–34, 2024b.

Steve Hanneke. The optimal sample complexity of pac learning. *Journal of Machine Learning Research*, 17(38):1–15, 2016. URL http://jmlr.org/papers/v17/15-389.html.

David Haussler. Decision theoretic generalizations of the pac model for neural net and other learning applications. *Information and Computation*, 100(1):78–150, 1992. ISSN 0890-5401. doi: https://doi.org/10.1016/0890-5401(92)90010-D. URL https://www.sciencedirect.com/science/article/pii/089054019290010D.

Elad Hazan and Tomer Koren. The computational power of optimization in online learning. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 128–141, 2016.

Adam Kalai and Santosh Vempala. Efficient algorithms for online decision problems. *Journal of Computer and System Sciences*, 71(3):291–307, 2005.

Yin Tat Lee, Aaron Sidford, and Santosh S Vempala. Efficient convex optimization with membership oracles. In *Conference On Learning Theory*, pages 1292–1294. PMLR, 2018.

N. Littlestone and M.K. Warmuth. The weighted majority algorithm. *Information and Computation*, 108(2):212–261, 1994. ISSN 0890-5401. doi: https://doi.org/10.1006/inco.1994.1009. URL `https://www.sciencedirect.com/science/article/pii/S0890540184710091`.

Nick Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Mach. Learn.*, 2(4):285–318, April 1988. ISSN 0885-6125. doi: 10.1023/A:1022869011914. URL `https://doi.org/10.1023/A:1022869011914`.

Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

Jr. Nash, John F. Equilibrium points in N-person games. 36(1):48–49. ISSN 0027-8424. doi: 10.1073/pnas.36.1.48. URL `https://doi.org/10.1073/pnas.36.1.48`.

Tim Roughgarden. Cs364a: Lecture notes on algorithmic game theory (fall 2013), 2013. URL `https://timroughgarden.org/f13/f13.pdf`.

David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017.

Leslie G. Valiant. A theory of the learnable. *Commun. ACM*, 27:1134–1142, 1984. URL `https://api.semanticscholar.org/CorpusID:59712`.