The Game Designer's Playlist

Innovative Games Every Game Designer Needs to Play

Zack Hiwiller

Addison-Wesley

Boston • Columbus • New York • San Francisco • Amsterdam • Cape Town
Dubai • London • Madrid • Milan • Munich • Paris • Montreal • Toronto • Delhi • Mexico City
São Paulo • Sydney • Hong Kong • Seoul • Singapore • Taipei • Tokyo

Contents

CHAPTER 1

FROM SIMPLICITY, THE UNIVERSE

"Man is not born to solve the problem of the universe, but to find out what he has to do; and to restrain himself within the limits of his comprehension."

—Johann Wolfgang von Goethe

In a commencement address in 2005, writer David Foster Wallace relayed this parable: "There are these two young fish swimming along and they happen to meet an older fish swimming the other way, who nods at them and says, 'Morning, boys. How's the water?' [...] [T]he two young fish swim on for a bit and eventually one of them looks over at the other and goes, 'What the hell is water?'" One needs awareness of all the things around us that are so taken for granted that they are invisible. We start our examination of games with one of the most popular games of all time, one that few in the West bother to explore.

A Universe, Conquered

Imagine for a moment that your life to this point took a different path. Imagine instead that since the age of five you were enrolled in a school specifically to train you to be an expert doctor. Your country, small on the world stage but important nonetheless, venerates the skills of doctors fervently, especially their ability to diagnose troublesome cases. Competition among medical students is cutthroat, and only the most rigorous study habits and keenest intuitions can vault a small percentage of students into the ranks of professional diagnosticians. People with lower skill levels still work, of course, but only the best can make a living on just their ability to diagnose. Imagine further that this skill is so venerated that it becomes a bit of a national pastime. Diagnosing is medicine, sure, but it also fills part of the cultural identity—a combination of philosophy and art.

You, after years and years of hard work, have become the world leader in diagnosis. This isn't braggadocio; it is fact. A world ranking system makes this explicit. You spent a few years at the amateur level and have since climbed rank after rank of professionals. You are now a Rank-9 professional and the best in the world. Your services are in high demand, and you have obtained celebrity status. At only 33 years old, you look forward to decades in which to practice your craft at the highest level.

Then you get an odd invitation. A computer company thinks they can make a program that out-intuits the best doctors and they will give you a bunch of money to go head to head. That is fine—many have tried in the past. But the task of diagnosis is just too vast for computers to master. They do fine, certainly. One would take the diagnosis of a specialized computer program over the random guess of a guy on the street, but the art of diagnosis is too nuanced and vast for a computer to compete with the best. Maybe in a generation or two when artificial intelligence (AI) approaches the level of human complexity it will be possible, but right now? No. Easy money.

You accept. The competition is shown live on your country's television networks and streamed to enthusiasts around the world. Expert panelists come on before the competition to repeat what everyone already knows: The human body is too complex. While computers do a good job at approximating human performance, it will be a long time before they can exceed it.

The point of all of this confidence leads to an obvious narrative conclusion. You lose to the computer. In front of your adoring countrymen, in front of the world, your entire worldview must change. Everything for which you've trained your entire life and everything you've believed about the intuitive nature of diagnosis is about to change. Where once you could be confident that your skills had worth, now it is possible with a little infrastructure that anyone could have access to the same (or greater!) power through their phones. What would you do? What would the country do? What would the world do?

To an extent, this is the story of Lee Sedol. In 2016, he, as a Go world champion, was beaten in a best of five series by an AI developed by DeepMind called AlphaGo. And to a shocked community of Go players, it wasn't even close.

In Lee's home country of South Korea, Go is part of the cultural identity. There are Go academies that train hopeful prodigies from their kindergarten days to become professional Go players. A rigid hierarchy of Go castes clearly delineates skill levels, and only four players each year become professionals.

Lee was supposed to win. In interviews before the match, Lee stressed that he was not considering the possibility of losing the five-game series, only that he had to focus on not losing one of the five games. Among online Go communities (at least the English-speaking ones I can find), the outlook was curious confidence. Estimates of Sedol losing one game ranged from 50% to 80%, but odds of Sedol losing three or more of the five-game series tended to be at 10% or lower. Professional gambling outlets had a more even outlook, putting an AlphaGo series victory at around 50%.

Lee lost game one. Then the next day, he lost game two. Then two days later, he lost game three and the series. Game four, Lee won. AlphaGo finished the series winning game five.[1] Lee never won as black, considered the more challenging color.

Journalists made much hay comparing the match to a 1997 match between world Chess champion Garry Kasparov and IBM's Deep Blue AI. Deep Blue won the 1997 series, which was then heralded as a watershed in the human-computer relationship. *Time*'s article about the match at the time had the headline "Can Machines Think?" The similarities are there: the confident human champion, the skeptical enthusiast community, the valuing of human adaptability versus algorithmic power. However, there is a key difference that makes AlphaGo's victory much more interesting, and it has to do with the nature of the game itself.

The Simplicity and Complexity of Go

Go is an ancient game, assumed to be one of the most ancient board games in continuous play. It is a strategy board game for two players. In the form played by Lee and AlphaGo, players sit on opposite sides of a board that has a 19×19 grid. More casual versions of the game use a smaller grid. One player has a pool of black stones, the other white. Players take turns placing stones on the grid's intersections in the hopes of capturing opponent's stones by positioning and increasing the amount of board territory they control.

Key to the game of Go is the concept of *liberty.* A stone has liberty if there is an empty space horizontally or vertically from it on the grid. A stone with no liberty is removed from the

1. While Sedol was a world champion, he was ranked #5 in the world. AlphaGo later beat the world-ranked #1 player in three straight games to much less fanfare in 2017.

board. In the Figure 1.1, white's stone marked #1 has liberties at A and B. Since C is diagonal, it doesn't count as a liberty. White's stone #2 has only one liberty: A. If black plays at A, stone #2 will be captured. In Go terms, stone #2 is in *atari*.[2]

Stones of the same color placed horizontally or vertically become a chain, or essentially one big connected stone. A stone within a chain can have liberty and stay on the board, but the chain must have liberty somewhere. In Figure 1.1, black's stone #3 has no liberty itself, but is connected to the stone to its left, to the stone above it, and through that stone to the stone above and right. That chain of stones acts as one and has only one liberty: D. Thus, the whole chain is in atari.

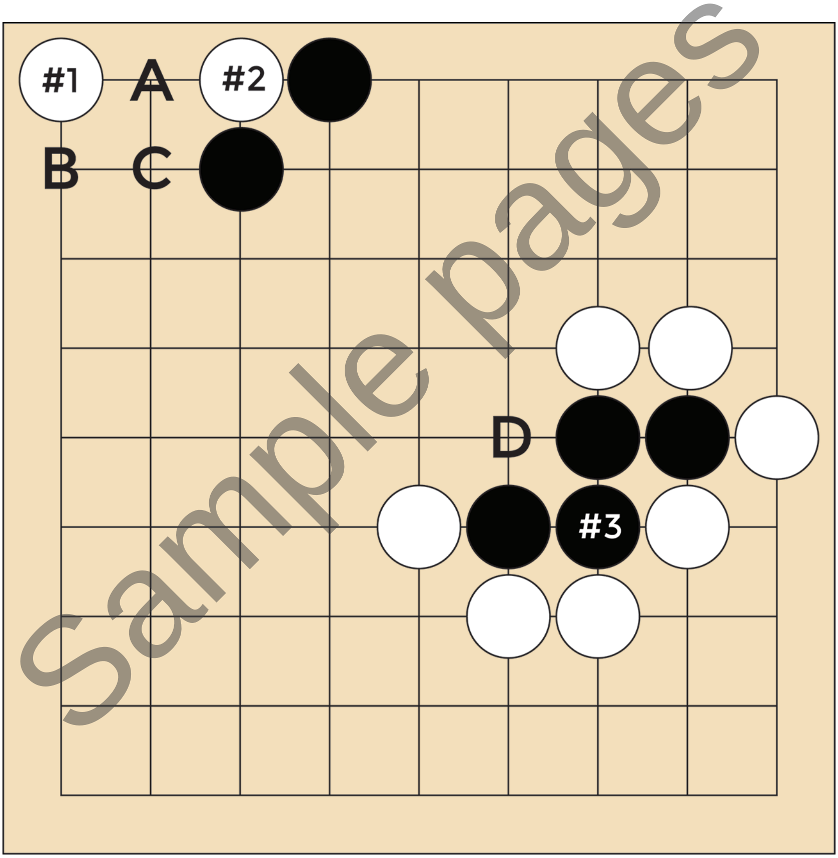

Figure 1.1 A sample Go board

The only other rule of note is that you cannot make a placement that reverts the board back to the state that it was before your last turn, a fix to eliminate infinite loops. Players go back and forth capturing or protecting territory by placing stones. The goal is to hold as much territory as possible.

2. Nolen Bushnell named the pioneering game company Atari after this bit of Go terminology.

What seems strange to players familiar with Checkers, Chess, or other abstract strategy games is that the game has no capture that marks the end of the game. Instead, the game continues until both players pass. Practically speaking, the game continues until one player resigns feeling he is in too weak of a position to come back.

We will come back time and again to the comparisons of Go to Chess for a number of reasons. First, Chess is well understood by western audiences. Second, the similarities between Kasparov's loss to Deep Blue and Lee's loss to AlphaGo are significant. Finally, Chess is aesthetically similar enough to Go that it has relevant parallels.

Go is a conceptually simple game: The only weapon in the game is where to place a single stone. Compare this to Chess, in which each piece may have a different means of capture. While there are no "piece types" in Go, a position in Chess matters for both its location and the type of piece there. If you can capture a queen in Chess, it is often a good idea, no matter its location on the board. Go is more egalitarian. Every piece is not of equal worth, but it can certainly look that way on the surface.

Combinatorics illustrates the large problem spaces of both Chess and Go. For a trivially simple game such as "What hand is the penny in?" the "board states" there can only be: {left hand:penny; right hand:nothing} or {left hand:nothing, right hand:penny}.

Tic-tac-toe is a comparatively simple game that actually contains some level of strategy as compared to the pure luck of our penny game. There are nine locations on the board. Each location can have an X, an O, or be empty. Thus, there are 3^9 or 19,683 possibilities. Of course, since some of those boards can never exist because of the rules (all 9 Xs, for example), we can reduce that down to 5,812 legal board states[3] in a 3×3 game of Tic-tac-toe. By this measure, Tic-tac-toe is 2,500 times more complicated than the penny game.

Chess has a board with 64 locations. However, each location is more complex than just white, black, or none. We must consider whether a space has a white pawn, a black knight, or a promoted white queen. Just like we reduced our Tic-tac-toe space to a number of legal boards, the restrictions for legal boards in Chess is more stringent. Not only can we not have 64 white pawns, but we must be careful that a board doesn't put one of the kings in check, among other conditions.

This problem is not trivial. In 1950, mathematician Claude Shannon estimated the number of Chess states as 10^{120}. Later estimates reduced this number to approximately 10^{50} or 100,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000,000. This is considerably larger than our Tic-tac-toe space.

3. From http://brianshourd.com/posts/2012-11-06-tilt-number-of-tic-tac-toe-boards.html.

Go strips the location states back down to three. At each location, there can be a black stone, a white stone, or no stone. This makes the upper boundary of the possible Go boards simple: 3^{361}. But like in our previous examples, we have to remove boards that aren't legal. For instance, any board that contains a stone with no liberty must be removed from consideration. In 2016, researchers were finally able to calculate the number of board states. Rather than give you their exact number, it will be easier to think about a nice round estimate of 10^{170}.

We humans have a hard time thinking about exponentials. How much larger is 10^{170} than 10^{50}? For some reason, when we look at these numbers we want to multiply and say that the Go estimate is 4-ish times larger. On the contrary, it is **10^{120} times larger.** If you had a normal-sized Go board that represented each board state, you could fill the entire observable universe with Go boards and still have plenty of boards to place. If you shrunk each Go board down to a single atom, you still wouldn't even have a chance. The universe would run out of room.

The number of game states is large, sure. But that is just one way to look at a game's complexity. Discussing it more would belabor the point. More than just the vastness of the game space is how interconnected it is. If I choose not to make a move because it will give you a better position next turn, I'm said to be looking one move ahead. If I am able to reject a move because it will open up a good position for you in two moves, I can be said to be looking two moves ahead. Depending on how many possible options I have when it's my turn, each move ahead gets exponentially harder. If I had 10 options for next turn, then I have to consider 10×10 options two turns from now. If a move puts me in a bad spot 10 turns from now, I need to be able to see that one of 10,000,000,000 moves. Even if I was to pare that tree down to "reasonable" states, it would still be massive. A turn in Go can have real lasting effects 100 turns later. Few Chess games reach 100 moves in total.

Cognitive theorists debate how many "moves ahead" the human brain is realistically able to hold. There are real limits to the amount of information we can consider. Computers, while faster thinkers than us, have real limits as well. And much like humans, they use the past as a shorthand.

More Than Thinking, Learning

In the late 18th century, audiences marveled at the "Automaton Chess Player," also called the "Mechanical Turk." It was purportedly a mechanical man that could play expert-level Chess. It played against Catherine the Great, Benjamin Franklin, and Napoleon Bonaparte. Charles Babbage, who would later be considered the father of the computer, lost to the Turk twice. Edgar Allen Poe wrote about the machine, purporting it to be a hoax because a "pure" machine should always win.

Poe was right. The Turk was a hoax. Reportedly, during its match with Napoleon it swept the pieces off the table in a rage after Napoleon tested it with a series of illegal moves. That should have been a hint to its mechanisms. The trick was that a grand master Chess player controlled the movements of the automaton using magnets and strings while hidden beneath the Chess board. The clever part of the con was how the human player could be hidden when the mechanisms underneath were opened and examined.

For a long time, the pursuit of a machine that could beat a human was focused on isolated computation. If the best human could look seven moves ahead and a machine eight, then the machine should be able to capitalize on that advantage. However, it turns out that looking farther ahead isn't how high-level players actually play the game well.

Psychologists William Chase and Herbert Simon published a study in the 1970s in which both Chess novices and Chess experts were given patterns of Chess pieces on boards to memorize and recall. Some of the boards were examples from real matches; others were random assortments of pieces. Chess experts were able to recall the patterns from real matches better than the random patterns, whereas novices had equal trouble with both. This lead to the conclusion that at least part of a Chess player's expertise is the ability to compare a current situation with one they have encountered before.

If an expert player plays 10 games of Chess a day every year for 40 years, she could end up with a library of experience of 150,000 games. Her recall would not be perfect, but this could provide a broad range of experience from which to say, "I remember a game like this; I made this move and lost," or "I remember another game like this; I made this move and won." Comparing the two experiences, she could then make the latter move.

The ability to draw on past experiences was essential to one of Kasparov's early victories over Deep Blue. He was able to switch styles on the fly while Deep Blue was using the same hard-coded strategy of what makes for a "good" move. In Kasparov–Deep Blue I, it was not relevant that Deep Blue could evaluate 200 million positions in a single turn if the criteria to determine which of these moves was best was flawed.

AlphaGo benefitted from 20 years of technological advancement in the speed of computers since Deep Blue's victory. But as we've seen, Go itself is exponentially more complicated, and brute force alone would not be the key to winning the game. Instead, AlphaGo does what Kasparov did to succeed: It uses what it knows from the past to determine what the value of a move in the present game might be.

Let's use a trivial Chess example. You have two pawns, a rook, and your king; the opponent has one queen, two pawns, a rook, and her king. Who is more likely to win? Your opponent has more pieces, and those pieces are worth more using standard Chess scoring. But what if your pawns are positioned in a way that you can mate the opponent in the next move? Then we need to revise what the opponent's queen is worth in that particular situation. We do that using heuristics. The algorithms that do the evaluation of specific moves are called *value networks.* The networks that determine what to do given the value of the game state are called "policy networks."

The AlphaGo team "trained" its policy algorithms by showing it 30 million examples played by real Go masters to suggest what real masters would do in numerous situations. This would be sufficient if the goal of AlphaGo was to play like a human master, but the goal was to be better than the human masters. So AlphaGo uses what it knows from that database to play the current game against itself a vast number of times, refining its value network—the rules that identify what a particular position is worth—to generate a recommended move.

Here, then, is the relevant difference between Deep Blue and AlphaGo. We told Deep Blue what a good move looked like and it found the best move it could, given time. AlphaGo *tells itself* what a good move looks like given the circumstances and finds the best move it can, given that. And each time it plays, it has more information to guide better and better decisions.

Because what AlphaGo does is now firmly in its black box, we don't really know what the criteria are that guide its search at any given time. We don't tell it that a queen is nine times as valuable as a pawn. We just show it enough games and give it enough time for testing, and it figures that out. In Game 2, Move 37 of the AlphaGo–Lee match, AlphaGo chose a move that its own predictive model, based on its library of professional Go games, identified as having a probability of 1 in 10,000 of being the best move. That is, a professional Go player would almost never consider the move. Lee, shocked by the move, rose from his chair and left the room for a few minutes. AlphaGo went on to win that game.

DeepMind has already created an AI that can play Atari games with only the raw images from the screen as input. The AI is not told the rules of Space Invaders, only that the goal is to maximize the score. Off it goes, playing millions of games, bettering itself, without ever really knowing why Earth needs to be saved from those invaders. While I write this, DeepMind is working on solving StarCraft II using techniques similar to those used by AlphaGo.

To emphasize the speed of developments in the AI world, in the time between writing the first draft of this chapter and my first revisions of it, DeepMind created a derivative AI that taught itself Chess and Shogi (a Japanese cousin of Chess) that could beat the reigning best AIs. And it accomplished this after only being told the rules and simulating games with itself for less than 24 hours.

Play to Win

In October 1990, newspapers across the country ran a panel of Gary Larson's famous *The Far Side* comic called "Hopeful parents." In it, a child plays video games, rapt, while his proud 1950s-style parents look on from behind, sharing a thought balloon of a "Help Wanted" ad in the far future of September 2005. The ad is loaded with pleas for game players: "Do you laugh in the face of Killer Goombas? Call us. $80,000/yr plus a free house."

Philosophers have engaged with the question of humor for centuries, but one explanation that is popular is the benign violation theory. It posits that what we find funny is a violation of expectations with a benign outcome. For instance, seeing someone unexpectedly getting hit with a pie is funny because it is unexpected and has no real consequence. Larson's "Hopeful parents" panel was funny in 1990 because it was unexpected and benign. It was unexpected because who expected someone to get paid to play video games? It is benign because if they are wrong, it isn't like the little boy's friends are all out learning plumbing and accountancy at that age (and even if they were—it is just a cartoon).

But Larson's joke isn't that unexpected anymore. In September 2005, the date the cartoon looks forward to, I was employed as a game designer and my parents had the cartoon hanging in their home. We were stretching the joke a little—being a game designer is about playing games as much as being a chef is about eating food. Skip forward a few years and the unexpected aspect of the joke loses even more focus.

Market intelligence firm Newzoo estimates that eSports will be a $1.4 billion field with an audience of more than 300 million people by 2020. Defense of the Ancients 2 (or Dota2) player Saahil "UNiVeRsE" Arora is reported to have made $2.7 million in tournament prizes (not counting sponsorship revenues). And like any popular sport, betting exists on the periphery. Numerous sites let you bet on popular games like Defense of the Ancients 2, Counter-Strike: Go (not the same Go we covered above, thankfully), or StarCraft II.

The 2006 World Chess Championship was marred by controversy when the manager of one player accused the other of excessive and frequent bathroom breaks, the implication being that when he got into a tough spot, he'd head to the bathroom and receive assistance from a Chess program. As evidence, the manager showed that upon returning from the bathroom, the player's moves corresponded with the recommended move from a popular Chess program 78% of the time.

Whether there was cheating at the event is not particularly relevant here. Samsung received a patent in 2016 for augmented-reality contact lenses that can overlay real vision with computer-generated information. What happens to human Chess competition when those are readily available? Do you ban contact lenses in general? What if the equivalent becomes implantable?

Doping is a constant concern in all sports. Yet doping alone won't make for a great athlete. It simply separates the already elite. A DeepMind-powered AI wiring directly into your vision could allow anyone to compete with the best. When DeepMind's StarCraft II AI is complete, how will eSports audiences know what they are watching is real human performance and not AI driven or AI augmented? What will that do to the viewing experience? If DeepMind can use its techniques on games as wildly different as StarCraft II, Space Invaders, and Go and master them all in a decade, how long will it be before it can at least help with every type of human game?

While Deep Blue was a true supercomputer, consumer-level hardware can now run Chess programs that can beat every player on Earth. Human-versus-human Chess championships are now quaint junior varsity leagues in comparison to the AI-versus-AI tournaments.

This can go one of two ways. The first is the optimistic approach. Long ago, we developed machines that can transport us at great speeds and great distances, yet we still have sprinting and distance running. Players aren't using Segways during soccer matches. There is a burden on professional leagues to adhere to a strict set of performance limitations to uphold the sanctity of the game and, while there are some challenges, it works fairly well. The fact that elite cyclists have a high incidence of doping has done nothing to dissuade bikers from hitting the road in cycling contests all over the world. Chess tournaments draw steady participation in this post Deep-Blue-in-your-pocket world. As much of a cultural touchstone as Go is in east Asia, it would be hard to imagine a 2,500-year streak of popularity melted by recent developments.

Even largely technical sports draw a line on what is acceptable. The "formula" in Formula One racing refers to the set of rules for what teams are and are not allowed to include in their car in a given season. Formula One racing is extremely popular despite the reality that it isn't the pinnacle of what is possible in terms of automobile speed; it is the pinnacle of what is possible given a distinct set of rules. These guidelines will help us define the role of sports and games in an increasingly augmented world.

The pessimistic approach ranges from economic dystopia to Armageddon. In 2016, billionaire Elon Musk said that increases in AI and automation will soon make most current jobs obsolete. He should know; his Tesla company is one of the companies developing self-driving trucks that aim to put the world's 20 million truckers in a new career. A 2016 McKinsey study examined the automatability of over 800 careers. It found that 51% of the U.S. economy's activities were susceptible to automation. I, for one, welcome our new robot overlords.

Futurists have gone further. Many estimate that based on past increases in computing power, we should have the ability to digitally simulate the human brain sometime in the next 15 years. After that, who knows? In the 1950s, the great computer scientist Alan Turing said that when machines become more complex than the human brain, the machines would be in control. Others have stated that once machines are artificially intelligent in the general sense, that they would be able to build smarter and smarter machines that we cannot yet imagine. In *Darwin Among the Machines,* George Dyson wrote, "In the game of life and evolution, there are three players at the table: human beings, nature, and machines. I am firmly on the side of nature. But nature, I suspect, is on the side of the machines." I. J. Good, the mathematician and contemporary of Turing, was one of the people responsible for helping popularize Go in the West. He also believed that a smarter-than-human machine would eventually lead to the extinction of mankind.[4]

Summary

We've gotten into some interesting topics. And while they are certainly game related, what do they have to teach game designers? Let's come back for a moment from the world of Godlike machines to the realm of simple game design.

If you exclude basics like the term definitions and end state, Go really has only three rules: adjacent stones of the same color are considered one stone, stones that have no liberty are removed, and you cannot recreate a former board position. A fourth rule can be employed that allows a weaker player to start with additional stones, but this is optional. No other game wrestles such a great amount of complexity of play from such a simple rule set.

Go has no weird rules or edge cases. Even Chess has *en passant* capturing and *castling,* special asterisks that flummox new players for their rare application. Go remains beautiful because of its combination of simplicity and depth.

When designing a game, you will be constantly bombarded by your subconscious with possible additional ideas. Wouldn't it be cool if my main character could fly? Or walk on the ceiling? Wouldn't it be great if my card game had one extra card type? Or my turns had one more phase? So many possibilities could be created! However, there are also bugs and unintended consequences that creep in with new features. Games with more complex rulesets are harder to teach to new players and harder to develop and test. If Go can be described in a paragraph, played for three millennia, and considered the pinnacle of AI development, then does your game *really* need that extra feature?

I've never created a game as simple and as elegant as Go; almost no one has. But understanding Go helps a designer to understand the interplay between rules, systems, and play experience in a way that is difficult to put into words. It is everything essential about game design freeze-dried and preserved for eternity: decision making, aesthetics, tactics, strategy, psychology, philosophy, risk, reward, intuition, and mathematics. It is the essential start for any game designer playlist.

4. A really great takedown of this position is at the following link, but is too out-of-scope to cover here: https://backchannel.com/the-myth-of-a-superhuman-ai-59282b686c62

Additionally, and while it seems strange to talk about this given the quintessentially analog nature of the games discussed, AI is a large part of most digital games. Understanding the complexity of how an algorithm plays a simple game like Tic-tac-toe, Nim, or Battleship is a good first step. Understanding how to tackle more complex games like Othello or Backgammon is next. When you consider the complexity of acting humanlike in a game as simple as Checkers, it will help clarify the algorithms required in any other design you may create.

GAMES COVERED

Playlist Game #1: Go

Designer: Unknown

Why: Simply put, Go is one of the most sublime games ever created. A designer that cannot appreciate the game's depth based on the dynamics created by simple rules is a designer that will have trouble understanding complex interactions in even relatively simple game systems.