

OpenAI

Why is it losing?

Edward Auttonberry

9/21/2018

On the eleventh of August 2018, Danylo “Dendi” Ishutin, one of the most prominent and popular players in Dota 2’s professional scene stepped onto the stage of Key Arena in Seattle, Washington for his seventh year in competition. Though not actually competing in the tournament this year, he was still present for another challenge that was arguable much more difficult than the tournament that he with his team were the first to conquer. His opponent this time around was not a human, but another computer that two people physically carried onto the stage during the reveal. People’s expectations for this match were not particularly high because the game had a reputation for particularly terrible AI opponents, if they can be called that, and also because the new opponent was just using the same standard opponent scripting API as all of the other terrible bots (OpenAI, 2018). How, realistic, was it to believe that this bot to stand a chance against one of the best players in the world?

Then the game started. The contest was a best two out of three matches. The bot was only one character out of the 114 available, but Ishutin could pick any hero wanted, including the those that carried a specific advantage over the bot’s character. After the character selection phase, the matches all ended before twenty minutes came to pass total. In a stunning display of mechanical skill, the bot absolutely wiped the floor with Ishutin. The professionals at the event that followed him all continued to be crushed by the new virtual opponent. It was not until about a month after the event that the bot became publicly available to the Dota 2 player base, and it was not until about month after that that the bot was successfully triumphed over.

Defeating the OpenAI in a one-on-one match continued to be a goalpost for professional players and the competitive public, but a problem surfaced with the bot: after they beat it the first time, most professionals did not lose again. Even more time passed, and even more people had figured out how to beat OpenAI, and eventually a strategy came out that allowed even casual players to defeat the bot that was standing toe-to-toe with professionals. The next tournament rolled around again in August of 2018 and a new iteration was being presented at the arena this year. OpenAI5 was announced at the tournament, which was a team of 5 OpenAI bots cooperating as a team to play a full game of Dota 2 against another team of five using most of the game’s intended design, such as the full map and item pool. Last year’s exhibition of OpenAI’s flawless triumph over Ishutin caused most to think that the human team, Pain Gaming, was going to lose at least the first game. It was somewhat close, but after identifying the OpenAI5 strategy, and using that information as well as taking advantage of the bots’ weaknesses to climb out of the disadvantage they found themselves in about halfway into the game, Pain Gaming emerged victorious in game one without even having their base invaded, which is a very critical game objective.

This begs the question: how were we able to catch up in just a year? If we look at AlphaGo Zero and the games that it has conquered, there has not been much talk of anybody coming back and reclaiming the domain for mankind from the computers. After only 21 days it defeated Go world champion Ke Jie in 3 out of 3 games and its Elo rating has only increased since then – it is still probably the best Go player in the world and it continues to improve from self-play (Hassabis & Silver, 2017). OpenAI learns through reinforcement learning similarly to AlphaGo, and it plays 180 years of games against itself every day to train itself in its “Gyms,” whereas the most time any player has spent playing Dota 2 is about a year and a half. Considering both these two systems have been training in their respective games for about two years, and the institutions that fund them are not running out of money any time soon, one would expect to see them perform at about the same level in their domains. This is obviously not the case the cause for this misalignment is a topic that deserves thorough investigation, which will be discussed here in detail.

The Dota 2 game space allows continuous movement within a large confined space in two dimensions, with a third independent vertical dimension added that affects gameplay mechanics, such as vision, but cannot be moved through freely. Human players will have no trouble learning to navigate this space once the simple movement controls are learned: click the move-to-location button with the cursor over some space on the map and the game's pathfinding algorithm will do the hard work for you. At the same time, there are other actions a player can take each frame, such as moving to attack a unit, casting a spell, buying from the shop, and many others. This results in a very large search space containing a combination of discrete and continuous layers. OpenAI navigates through this combined space by discretizing the search space into 170,000 possible actions per hero (OpenAI, 2018). The individual LSTM networks produce their decisions on this search space every four game frames, which is about 7 times per second.

This puts the bot at a slight disadvantage compared to the players. In Go, for example, both players, bot or not, are limited to the same discrete playing space. Any Go-playing intelligence will physically be able to make any move that human player can with reasonable decision-making time (a few milliseconds). For OpenAI to be able to make its decisions at the speed that it currently does, it must remove some of its possible play options from its search space by discretizing. This does not affect much of its options except for simple movement. In Dota 2, the position of game artefacts in competitive play is very crucial to certain strategies and game mechanics. For example, there is a very strict cut-off range at which team-aligned non-player character units will attack an opposing player if that player attacks a player on that unit's team. The bot may have to decide whether it should act aggressively towards the player and suffer the damage caused by the con-player characters, whereas the player may be able to fit inside of the middle ground. To try and put this into perspective, the game map's size is 15000 x 15000 game units according to the official Wiki. If we estimate the amount of gameplay decisions in the search space are already discrete and do not affect movement to about 5000 choices, that leaves the search space with 165000 options for movement. If the map is sliced into a grid of 165000 units of equal area, that creates a margin of 1363.63 game units. This limits the possible vertical and horizontal options to about 14 discrete points on either axis for the whole map. From a human player's perspective at the standard zoom level, that would be about one fourth of the horizontal game screen size, though this is somewhat dependent on the resolution of the viewport. With it is simple to see what kind of disadvantage a human player would be at if his opponent could move to any point any point on a continuous 2.5-dimensional space, but he could only move between about 196 different points on the map.

OpenAI's documentation does not explain in detail how this issue is currently being addressed in practice, but it is likely using a combination of clicking on different points and moving short distances between them to reach specific locations. However, this also raises another issue, because characters in the game do not usually move in a straight line from point A to point B due to obstacles and map boundaries. It must navigate around these artefacts to be able to reach certain locations. When OpenAI is trained, it trains using the current standard version of the map. Dota 2 is a constantly changing game with constantly changing rules. One of the rules that gets changed the most is the layout of the map, which sometimes occurs every few weeks. When the bot is trained, it must relearn the map, because even the smallest changes in the positioning of some obstacles can cause the decisions that would have previously had a very high chance of producing a good reward to provide a very unfortunate and unexpected consequence instead, due to resulting changes in how the game's pathfinder would navigate the player to some locations. The decisions that neural networks make on a certain path get weaker every time the network must retrain or relearn the results of those decisions (Harvey, 1996). Because Dota 2, unlike more classical games like chess, is constantly changing the rules to play the game by, this would suggest that every time a minor game update is pushed, the bot's decision-making skills get weaker until it starts

from scratch again. These effects are not just limited to changes in the map affecting the discretized search space for movement. This also applies to minor changes in the already discrete portions of the game's search space, such as spell damage, unit movement speed, character statistics, purchasing cost on items.

These effects all reflect the OpenAI's ability to make decisions with short-term consequences but seeing as how the average Dota 2 game lasts for 45 minutes, the bot's ability to make these decisions for the long term is also very crucial. There is a fundamental difference between Dota 2 as well as games like it, and games of pure strategy like chess and Go. Dota 2 is much more akin to cooperative athletic sports games in the sense that a player must first develop the technical ability to perform in the game with maximum efficiency, and then develop and execute the strategies that best target the opponent's weaknesses. Consider soccer. For teams of young or inexperienced players, the team with the best kickers and the defense with the best stance will be the team that wins, which contrasts with high-skill teams where these skills are the bare minimum for every player and are expected to be perfect in each, so that only the team with the best strategic foresight and execution will score the most points. This is true of the skill bracket in Dota 2 as well, and the reason why only very skilled players can defeat the OpenAI – only the players that can match the bot in technical ability and resource efficiency are able to strategize against the bot. This is also the reason why professionals so consistently defeat the bot. OpenAI has not yet learned to strategize. According to professional player Park “March” Tae Won, who played against OpenAI5 with his team before its debut and served as commentator for the debut match, OpenAI5 has only recently adopted a certain strategy that is very popular on the professional scene right now, but it is very easy to adapt to the bots' strategy after the first game (Channel, 2018, at 6:16). This means that currently OpenAI is unable to develop strategies as game progresses, it is only able to use the exact strategies that it has learned from practicing.

The OpenAI is an implementation of LSTM networks, meaning that they can efficiently persist short-term information (Olah, 2015). Being such would allow OpenAI to do analysis of game trends in real time. These trends could include net-worth and experience gaps, player/team movement, and small game details from which projections can be induced, such as “someone not on my team has been repeatedly taking this power up that I want to get.” The bots could apply these trends to make slight deviations in the learned strategies to maximize effect, so the example given previously may, after some more training, result in the decision of “I should get to the power up earlier so that I can get it first when it spawns and maybe the guy who has been taking it will show up for me to kill him.” On the professional scene, the most common complaint about the bots is that their strategy never changes much, and there is a lot of certainty about how the bots will try to play the game. Small real-time changes to the OpenAI strategy will prevent these players from being able to predict the bots and ultimately increase performance in terms of success rate for the bots.

When talking about making these trends, the important resource for developing these trends is the data required to plot and analyze them. However, there is a fundamental limit on what data OpenAI has access to in Dota 2 and how specific it is. OpenAI plays the game by utilizing the official bot scripting API developed for the game to see bots “better” than the ones available by default (OpenAI, 2018). The API's full function is to allow any algorithm written to automate playing Dota 2 to have access to information that any normal player would have access to, which is true to a point. With this API, the bots have access to basic game information: health of units, number of buildings standing, gold, is character currently moving, locations of hazards, etc. However, there is simply too much information for Valve Software to try to make available through an API for them to do so. This means that bots do not see everything that humans see, they only see what is available to them through code. The do not see things

like direction and proximity of projectiles, particle effects, areas with spells casted on them, etc. The way OpenAI current adapts to this disadvantage is by inferring this information based on what they can see, i.e. the bot can see its health drop slowly and that it is affected by a particular status associated with an area-of-effect spell, and therefore it tries to leave it (OpenAI, 2018). This presents a side-effect of the environment that OpenAI operates in. This is that that bot cannot react to some environment changes until it has already directly affected their status in the game. In the video of the match versus Pain Gaming, we can see the bots react instantly to the human players appearing from nowhere, before they even have a chance to activate spells and start fights. However, they do not react to the turret that one of the humans can lay down until that turret has already started firing at them.

While this is not entirely at the fault of the OpenAI's development, it still presents a significant challenge in its path to fully surpassing professional players. One of two things must occur to get over this speed bump: either the Dota team at Valve must make more and more specific information available through the botting API, or OpenAI is going to have to rely on a lot of physical resources to be able to render Dota 2 games and play them based on feedback from computer vision. In any case, OpenAI can not improve at much better rates without the ability to obtain all feedback from the game.

This includes feedback from the bots themselves. As stated before, each bot in an OpenAI5 team is an independent LSTM network that operates based solely on game input to make decisions. The bots do not communicate with each other in any capacity (OpenAI, 2018). This is another point in the strategizing discussion that human players have over OpenAI. When skilled Dota 2 players are on a team together and working towards the goal of destroying the enemy base, they are constantly communicating with each other about what they are doing. Information that is critical to the execution of a short-term strategy is made available to teammates like "I am about to start this fight." This information lets the human players know to get ready for a fight with the other team. OpenAI only uses information about the current decisions of its teammates to make decisions, but it has no idea what its teammates are going to do in the future. In the Pain Gaming match, the OpenAI5 team often cast critical spells that do not work together at the same time. Most of the time this would still result in the bots achieving their goal for the fight, but in some cases, it caused them to lose their advantage for the scope of the fight and lose some teammates afterwards. Interactive inter-player cooperation between human and machine players has shown potential increase in human's ability to complete some games successfully (Swiechowski, Merrick, Mandziuk, & Abbass, 2015). The machine provides cooperative output to assist the human avatar in completing the game objective. This process can be reworked so that the machine output is interpreted by another machine. In this case, it would be output from one instance of OpenAI being intercepted by the other cooperative instances of OpenAI. Each network would then be able to communicate with each other to deliver intent, which it is currently unable to do.

There is a lot of ground for OpenAI to cover. Only about a year ago, there was a lot of support behind it because it was beating the best of people at a different kind of game from chess or go. It was performing short-term strategizing, faking players out, hiding, and generally playing like a good player would. Many were disappointed to see how OpenAI5 performed this year, losing its debut match to the last place team of the tournament where it had beaten one of the best without error only a year before, and losing to other non-professional teams in test runs before the debut. There are many flaws with the way OpenAI plays the game currently, some of which are being worked on, other which cannot be worked on until some developments come along.

For OpenAI to see success in the next iteration it must overcome the roadblocks that have been presented. The discretized search space for movement become denser to allow for better movement and better adaptation to changing environments. Information throughout the game must be stored and

analyzed through all stages of the game to allow for alterations in strategy and potentially learning long-term strategizing. OpenAI must have the ability to see everything that a human would see that the current botting API does not provide. Finally, OpenAI must be able to communicate with the other instances of itself that it acts cooperatively with to be able to execute the most efficient team plays. AlphaGo is undeniably way ahead of OpenAI in its domain. There could be any number of reasons for this. Maybe AlphaGo is simply a much better general game playing machine, or maybe OpenAI is learning to play a much harder game. Either way, one or both is going to be better than humans at all of our games at some point.

References

- Channel, O. D. (2018, August 2018). Humans vs OpenAI5 | The International 2018 Main Event. Vancouver, British Columbia, Canada: Valve Software. Retrieved from <https://www.youtube.com/watch?v=nGhkCuQloXI>
- Harvey, I. (1996). *Relearning and Evolution in Neural Networks*. University of Sussex, School of Cognitive and Computing Sciences, Brighton. Retrieved from <http://users.sussex.ac.uk/~inmanh/Relearn.pdf>
- Hassabis, D., & Silver, D. (2017, October 18). *AlphaGo Zero: Learning from scratch*. Retrieved from deepmind.com: <https://deepmind.com/blog/alphago-zero-learning-scratch/>
- Olah, C. (2015, August 27). *Understanding LSTM Networks*. Retrieved from colah's blog: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- OpenAI. (2018). *OpenAI Five*. Retrieved from openai.com: <https://blog.openai.com/openai-five/>
- Swiechowski, M., Merrick, K., Mandziuk, J., & Abbass, H. (2015). Human-Machine Cooperation in General Game Playing. *Advances in Computer-Human Interactions*, 96-100.