

AlphaStar



**UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH**

Cano Duarte, Yaiza
Monroig Company, Pol
Olivé Fernández Josep, Maria

Índice

Índice	1
Introducción	2
Descripción del producto	2
Métodos y técnicas utilizadas	3
Arquitectura del algoritmo	4
Supervised learning	5
Value and policy evaluation	6
Factor innovador del producto	6
Impacto del producto en la empresa	8
Beneficios	8
Riesgos	8
Posición en el mercado	8
Impacto del producto en el usuario o en la sociedad	9
Beneficios	9
Riesgos	10
Referencias	12

Introducción

La inteligencia artificial ha formado parte de los videojuegos desde sus inicios en 1950 y se utiliza para crear la ilusión de que PNJs (personajes no jugadores) reproducen el comportamiento humano, por ejemplo, dotándolos de la capacidad de generar comportamientos receptivos, adaptativos o inteligentes. El objetivo principal al incorporar la IA a este tipo de entretenimiento era mejorar la experiencia del jugador más allá de lo que se puede conseguir con *machine learning*.

Además, conforme ha ido pasando el tiempo, se ha popularizado la idea de tener oponentes dirigidos por una IA, cosa que aumentaba la dificultad de los niveles, introducía nuevos patrones de movimiento y ocasionaba que los eventos en el juego dependieran de funciones de hash que se ejecutaban según las acciones del jugador. En juegos actuales, técnicas ya implementadas como el *pathfinding* (búsqueda de ruta), utilizado, por ejemplo, por los PNJs para saber hacia dónde dirigirse en una batalla, y los árboles/grafos de decisiones, utilizados, por ejemplo, en aventuras gráficas donde tus decisiones dependen de cómo se desarrollará la historia, están a la orden del día.

Descripción del producto

Como hemos comentado anteriormente, la idea de enfrentar un oponente difícil atrae a la mayoría de los jugadores, incluidos nosotros, así que decidimos realizar un estudio sobre *AlphaStar*, la primera IA que ha sido capaz de derrotar a uno de los mejores jugadores profesionales de *StarCraft II*, obteniendo un resultado de 5-0 con la raza *Protoss*. Este resultado puede no parecer impresionante si se desconoce el auténtico nivel de estrategia, concentración y estrés al cual se está sometido durante una partida. Así pues, comenzaremos explicando cuál es el reto que ofrece este videojuego.

StarCraft II es un videojuego creado por *Blizzard Entertainment* con el cual llevan haciéndose torneos profesionales más de 20 años. El mapa se sitúa en un universo de ciencia ficción y existen diferentes modalidades de juego, el más común es el 1v1 durante 5 partidas. Para empezar, el jugador deberá escoger entre 3 razas, donde cada una de ellas tiene características y habilidades propias. Cada jugador comenzará con un número limitado de unidades, las cuales utilizará para obtener recursos básicos que utilizará para crear más unidades y estructuras que a su vez podrán recolectar nuevos recursos con los que mejorar estructuras y desarrollar tecnologías más sofisticadas, y así durante unas cuantas etapas del juego. El objetivo es acabar obteniendo mejores unidades y estructuras con las que atacar y defenderte de tu oponente hasta que uno de los dos sea invadido por el otro.

Al contrario que otros juegos como el ajedrez, un jugador no obtiene información sobre el otro hasta que las unidades/estructuras oponentes se encuentren dentro del campo de visión

de sus unidades y viceversa; además, no es un juego por turnos, sino que ambos jugadores juegan al mismo tiempo. Todo esto provoca la necesidad de ser capaz de, no solo pensar y ejecutar una estrategia a largo y corto plazo, sino de adaptarse según factores externos como la estrategia del oponente o la situación de la partida.

Así pues, en noviembre de 2016, *DeepMind* junto con *Blizzard* anunciaron que empezarían a trabajar en un proyecto de inteligencia artificial para éste tan conocido juego online. En agosto de 2017 (menos de un año después), se realizó el primer lanzamiento de la API de la IA que ahora conocemos como *AlphaStar* y desde entonces estuvo aprendiendo a cómo jugar hasta que el 12 de diciembre de 2018 obtuvo el resultado de 5-0 contra un jugador profesional que hemos explicado al principio de esta sección.

AlphaStar fue creado originalmente con la idea de parecer un humano cuando jugara ya que se quería realizar un estudio a ciegas donde jugadores humanos se enfrentaran a ella sin darse cuenta de que estaban jugando contra un IA. Con este objetivo en mente, si se diera el caso en el que AlphaStar fuera capaz de jugar a una velocidad y con una precisión superhumanas arruinaría completamente dicho experimento.

Con dichas declaraciones, podríamos pensar que el prestigio y la repercusión de la victoria explicada anteriormente no se debe a que el torneo fuera entre un humano y un superordenador... ¿o sí?

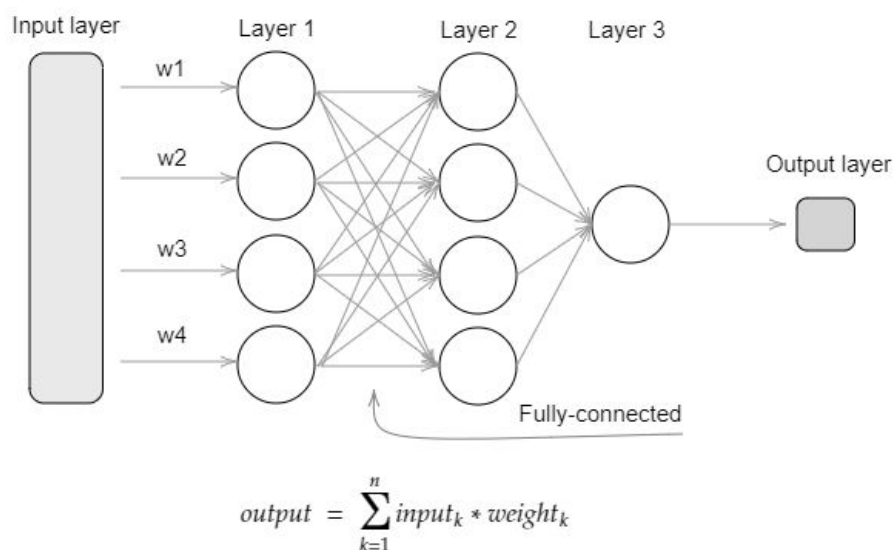
Métodos y técnicas utilizadas

Entorno del algoritmo

El entorno de trabajo del algoritmo utilizado es muy similar al de una persona ya que es el mismo juego en tiempo real, al jugador no humano se le proporciona una vista de una cámara la cual le da un campo limitado, esta cámara al igual que los jugadores humanos la podrá mover y será decisión suya saber hacia dónde es mejor enfocar la atención. AlphaStar tiene la ventaja de que puede hacer un mejor seguimiento de los objetos fuera de su campo de visión sin embargo no puede interactuar con ellos si no mueve la cámara en la posición adecuada. Otra ventaja que tiene el agente inteligente (en este contexto usaremos el término agente, algoritmo y AlphaStar de manera intercambiable) es que tiene más flexibilidad a la hora de actuar y seleccionar cosas ya que no lo tiene que hacer de manera manual como lo haría un humano. Sin embargo, a pesar de estas ventajas, no se muestra ninguna mejoría en el juego. Otra ventaja que podría tener el agente es que teóricamente podría hacer muchas más acciones en un instante de tiempo en comparación a una persona, es por eso que se le limita a un cierto número. Por último el algoritmo también tiene una latencia de reacción ya que al igual que antes debe adecuarse a las condiciones de juego humanas.

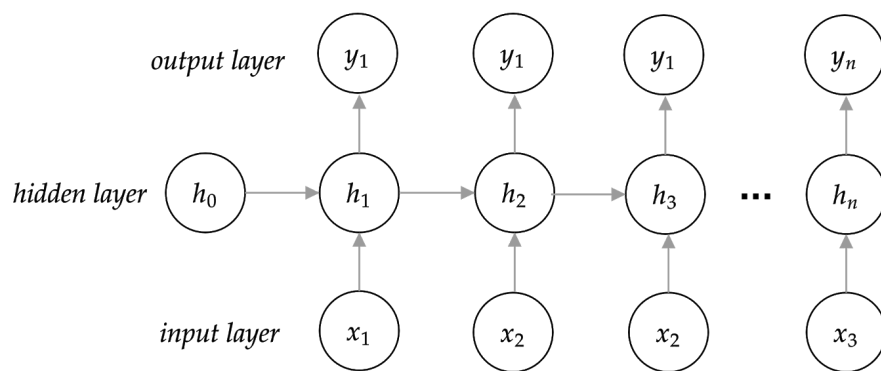
Arquitectura del algoritmo

La base del funcionamiento del AlphaStar se basa en una subrama de *machine learning* llamada *deep learning*; esta rama se basa principalmente en la utilización de redes neuronales artificiales, estas están basadas vagamente en las redes neuronales del cerebro humano ya que las neuronas de los seres humanos generan una interacción entre ellas inmensamente compleja, tanto que hoy en día aun no se sabe por completo su funcionamiento. Volviendo al tema anterior, las redes neuronales artificiales(RNA) se pueden representar con un grafo (aunque a nivel computacional se representan como matrices). La idea de una neurona artificial es que dada una serie de entradas estas se multiplican por una serie de pesos que determinan el valor, importancia o potencia que se le tiene que dar a una entrada específica al pasar por esa neurona, al final de todo se suman todos los valores lo que da lugar a la salida. La siguiente imagen muestra un ejemplo visual de su funcionamiento. Como se puede ver una neurona sola simplemente representa una ecuación lineal muy sencilla, la potencia de las redes neuronales en sí viene dada por la cantidad de conexiones que hay entre ellas y la cantidad de capas que haya. A las redes neuronales con más de 3 capas se les llaman redes neuronales profundas y son las que son capaces de resolver todos estos problemas actualmente. Las RNA tienen muchas otras cualidades, defectos y propiedades; sin embargo, no es necesario para la explicación de la arquitectura del AlphaStar.



La arquitectura del modelo neuronal está hecho a base de un tipo de RNA recurrentes, específicamente una subespecie que contiene células llamadas *Long Short Term Memory* , estas células tienen la característica de poder “almacenar” la memoria de eventos pasados a base de recuperarlos o desecharlos dependiendo de su importancia de los eventos actuales, su funcionamiento es muy similar al de un registro que es capaz de retroalimentarse con su propia salida. Las redes neuronales al igual que todas las técnicas de *machine learning* se basan en entrenar al agente para que vaya mejorando, y trabaje mejor, el entrenamiento de las

RNA es especialmente difícil e inestable, por su dificultad está catalogado como un *Black Box* donde no se sabe porque a veces funciona y a veces no, lo que dificulta la reproducibilidad de los experimentos. Por esta razón, AlphaStar se basa en una multitud de técnicas tanto nuevas como ya existentes. Cada una de estas técnicas están explicada a continuación.



Multi-agent learning

El entrenamiento del algoritmo utilizada un entrenamiento multiagente, que consiste en poner múltiples agentes a jugar entre ellos para que aprendan entre ellos y mejoren concurrentemente. El método multi agente solo se utiliza durante el aprendizaje de AlphaStar ya que en el momento de la evaluación solo se necesita uno.

Supervised learning

El aprendizaje supervisado se basa en entrenar al modelo neuronal a base de darle unos datos de entrada (input), hacer una predicción de acción (output), generar el error que ha generado el modelo a partir de una métrica de error como MSE (mean squared error), en este caso la métrica que se usa es específica del juego llamada *MMR score*. Después de generar el error, se calcula el gradiente de todas las neuronas para averiguar el error de cada una de ellas, este cálculo se usa una técnica llamada *reverse mode automatic differentiation*. Por último se aplica una regla de mejora para mejorar la capacidad de la red, en este caso se usa específicamente un optimizador Adam (*adaptive moment estimation*). Para conseguir el error entre la predicción y el resultado verdadero (labels) se usan datos de partidas reales de jugadores profesionales.

Reinforcement learning

La técnica *Reinforcement Learning* es un tipo de aprendizaje que permite al ordenador aprender por sí mismo, ya sea jugando con sí mismo o basándose en partidas reales. Principalmente se intenta imitar el aprendizaje que existe en la naturaleza en la que se le da una recompensa al agente inteligente, ya sea negativa o positiva, con el fin de fomentar las acciones que este prenda. El algoritmo utilizado se basa en el paradigma *Actor-Critic*.

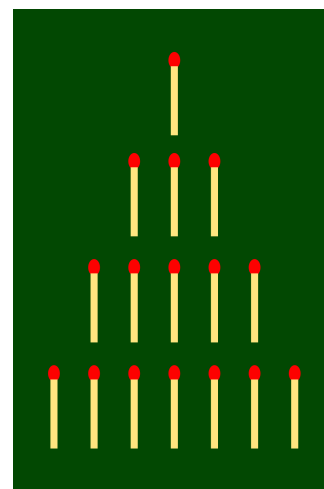
Value and policy evaluation

Cada situación en la que está sometida el agente/jugador se puede representar con un estado, a su vez todas las posibles situaciones se pueden representar con múltiples estados, dependiendo del juego el número de estados es pequeño o en el caso de AlphaStar es extremadamente grande. Para juegos con espacio de estados pequeños se utilizan *Tabular methods* estos se basan en almacenar una tabla para representar los estados en la cual cada estado tiene un *value*, el cual representa que tan bueno es este respecto al resto. Los métodos tabulares son capaces de resolver problemas de inteligencia artificial de manera exacta en la mayoría de los casos, sin embargo cuando el espacio de búsqueda es tan grande se utilizan métodos aproximados tales como el uso de redes neuronales. Para mejorar la estrategia de juego del agente, este juega contra sí mismo actualiza este *value* de manera aproximada. A su vez este agente utiliza una política de aprendizaje llamada *Off-Policy*. Esto significa que las acciones que en realidad hay dos agentes, uno que toma las decisiones y que se encarga de explorar el espacio de búsqueda y otro agente *greedy* al cual se le actualiza el *value*. De esta manera nos aseguramos de explorar todo el espacio de búsqueda y a la vez tener otra política que nos permita hacer las mejores jugadas.

Factor innovador del producto

AlphaStar no es ni mucho menos la primera IA que se usa en el mundo de los videojuegos, y precisamente, para entender porque esta IA representa una gran innovación en su sector es importante fijarnos en sus antecesores y orígenes.

Como ya comentamos anteriormente, los videojuegos han sido un campo de investigación en inteligencia artificial desde sus mismísimos inicios, y es que, un videojuego no deja de ser una perfecta representación de un mundo en miniatura, que proporciona un entorno controlado y ordenado a una IA para aprender. Una de las primeras IA's de las que se tiene constancia data de 1951, el juego en cuestión, *Nim*, tiene un origen incierto aunque se rumorea que surgió en China. Aun así, la primera referencia escrita en Europa se remonta a inicios del siglo 16. Este simple juego, consiste en remover una serie de objetos de una estructura piramidal, en la que dependiendo de la versión, el jugador que coje el ultimo objeto restante gana o pierde según las normas correspondientes. En cada turno, un jugador solo puede coger uno o más objetos de una misma fila, y deberá esperar a su siguiente turno para poder tocar otra. Hay que recalcar, que en el momento en el momento de su



desarrollo y publicación, esta versión computarizada de *Nim* era tecnología de lo más avanzada.

Paralelamente, en 1951, Christopher Strachey creó un programa para jugar a las damas y Dietrich Prinz uno para el ajedrez, que también fueron de los primeros programas jamás programados, y con unas IA extremadamente simples. Pero no fue hasta principios de los sesenta, que un programa de ajedrez, de Arthur Samuel, consiguió las habilidades suficientes como para desafiar a un rival amateur decente. Este camino culmina en la ampliamente conocida victoria de Deep Blue, de IBM, frente a Kasparov en 1997.

Veinte años más tarde que la versión computarizada de *Nim*, también salió al mercado *Pong* de la mano de Atari, en su formato recreativo ampliamente conocido. Este era capaz de jugar e incluso ganar regularmente a jugadores con mucha experiencia y habilidad en el juego.

Aun así, la mayoría de los primeros videojuegos desarrollados, se basaron más en la lógica discreta y la competición entre dos jugadores.

Es evidente que AlphaStar representa un enorme salto respecto a sus ya lejanos antecesores, pero comparándose con su competencia actual, ¿qué es lo que hace a esta IA realmente interesante?

Recientemente, y con el gran *boom* de las redes neuronales, son muchas las empresas que se han aventurado a desarrollar IA's, no solo con la esperanza de que fueran capaces de jugar a ciertos videojuegos, sino con la intención de que estas pudieran igualar o incluso superar a los mejores jugadores de cada uno. Y muchas de ellas han tenido un éxito significativo, en videojuegos como Mario, Atari, Quake III Arena Capture the Flag, y Dota 2, pero hasta ahora, ninguna otra había sido capaz de jugar a StarCraft con la maestría que ha demostrado AlphaStar. La mayoría de resultados obtenidos previamente en este videojuego, implican técnicas que podrían tacharse de chapuceras. Algunos ejemplos serían la necesidad de dotar a la IA de capacidades sobrehumanas (como por ejemplo para ejecutar más acciones por minuto de las que serían posibles para cualquier persona), imponer restricciones significativas en las normas del juego o incluso reduciendo la complejidad del mapa en el que se jugaba la partida. Y incluso con todas estas facilidades, ninguna de las IA rivales de AlphaStar había nunca conseguido siquiera acercarse al nivel de los jugadores más profesionales del videojuego.

Esta versión contrasta con el éxito rotundo de AlphaStar en StarCraft II, la cual mediante el uso de una red neuronal profunda entrenada directamente de datos extraídos del propio videojuego, nos ha mostrado resultados nunca vistos.

Conseguir tal habilidad en el que se considera el más difícil de todos los videojuegos de estrategia en tiempo real (RTS - Real Time Strategy), es un auténtico hito.

Impacto del producto en la empresa

Beneficios

El momento en el que se lanzó la primera versión de AlphaStar fue un factor muy importante para alcanzar el éxito ya que superó las expectativas de la gente que lo esperaba desde que fue anunciada y que especulaban su salida entre 1 año y 2 más tarde. El prestigio que ganó la empresa se consolidó aún más cuando tan solo un año después AlphaStar consiguió ganar una partida de 5 rondas contra un jugador profesional.

AlphaStar es uno de los múltiples proyectos de DeepMind, y su principal función es ejercer de sandbox conjuntamente con algunos de los otros proyectos de la empresa. Estos proyectos tienen como función principal servir de entreno para las IA de forma que después estas puedan resolver otros problemas del mundo real.

Como ya comentamos en su momento, entrenar a una IA en base a un videojuego tienen múltiples ventajas, ya que estos juegos no son más que representaciones simples y limitadas de un mundo con sus propias normas, en definitiva, un mundo menos caótico y con mucho menos ruido que el nuestro.

Proyectos como AlphaStar o AlphaGo son de vital importancia para preparar el campo sobre el que las IA algún día podrían ayudarnos a resolver problemas cruciales, y por lo tanto, representan un gran activo para la empresa.

Riesgos

Al igual que cualquier producto, AlphaStar puede tener inconvenientes y estos vienen dados principalmente por los usuarios. La rapidez con la que la IA aprendió a jugar y a ganar a jugadores profesionales se ha convertido en un arma de doble filo, ha hecho que la mayoría de usuarios tanto que juegan habitualmente como los que solo ven los torneos, hayan expresado su expectación de que AlphaStar sea capaz de derrotar a los jugadores profesionales de Starcraft II al menos el 50% de las partidas para 2021.

Posición en el mercado

Uno de los beneficios más importantes al igual que cualquier otro producto es el dinero, AlphaStar representa un producto que beneficia tanto a la empresa que lo produce, y a los productores de StarCraft, ya que proporcionan un medio de publicidad muy extenso para ambos partidarios.

Por ejemplo, desde que AlphaStar ganó al jugador profesional MaNa, otros jugadores profesionales Europeos se han animado a jugar contra AlphaStar después de que DeepMind anunciara que habían realizado cambios respecto a la interacción de esta IA con la cámara y la velocidad con la que ejecutaba acciones (ambas dos controversias que tuvieron una notable repercusión y que explicamos en el apartado de riesgos del impacto del producto en la sociedad).

Impacto del producto en el usuario o en la sociedad

Beneficios

Uno de los objetivos determinantes de empresas tales como DeepMind y OpenAI, es conseguir lo que se conoce como *AGI* o Inteligencia General Artificial. A diferencia de muchos sistemas de inteligencia artificial que solo son expertos en algún dominio de conocimiento específico (e.g. Sistemas basados en el conocimiento), o que solo resuelven un problema específico lo más optimizado posible (e.g. A*, Búsqueda local, etc.), los sistemas de inteligencia artificial general se centran en poder trabajar en cualquier ontología, además de esto, son capaces de aprender y de mejorar con el tiempo.

El sistema de inteligencia general más conocido somos por supuesto nosotros. El ser humano es capaz de aprender y aprender infinitas de problemas distintos. Este es el objetivo de estas empresas, conseguir una inteligencia similar a la nuestra, o al menos que pueda ayudarnos en muchos campos a la vez. Aunque este sea el objetivo, aún estamos lejos de poder alcanzarlo, sin embargo poco a poco vamos consiguiendo generalizar la inteligencia artificial, y AlphaStar es un ejemplo de esto. AlphaStar vive en un propio mundo miniatura, con sus propias reglas y leyes, el mundo de StarCraft. AlphaStar es capaz de aprender todo lo que existe dentro de su juego, el siguiente paso sería agrandar este mundo. Otros sistemas como AlphaZero ya son capaces de jugar diversos juegos a la vez, sin necesidad de darles información, es decir lo han aprendido todo ellos mismos.

En definitiva, la inteligencia general artificial sería un gran avance a la sociedad actual, con AlphaStar somos capaces de echar un pequeño vistazo a lo que sería la inteligencia artificial del futuro.

Por otro lado, este producto provoca una gran emoción en la comunidad de jugadores, principalmente la de StarCraft, en la cual todo jugador desearía poder competir contra AlphaStar. El hecho de que una inteligencia artificial sea capaz de ganarle a un humano en un juego de estas características está más normalizado que antes, aun así, este hecho provoca una

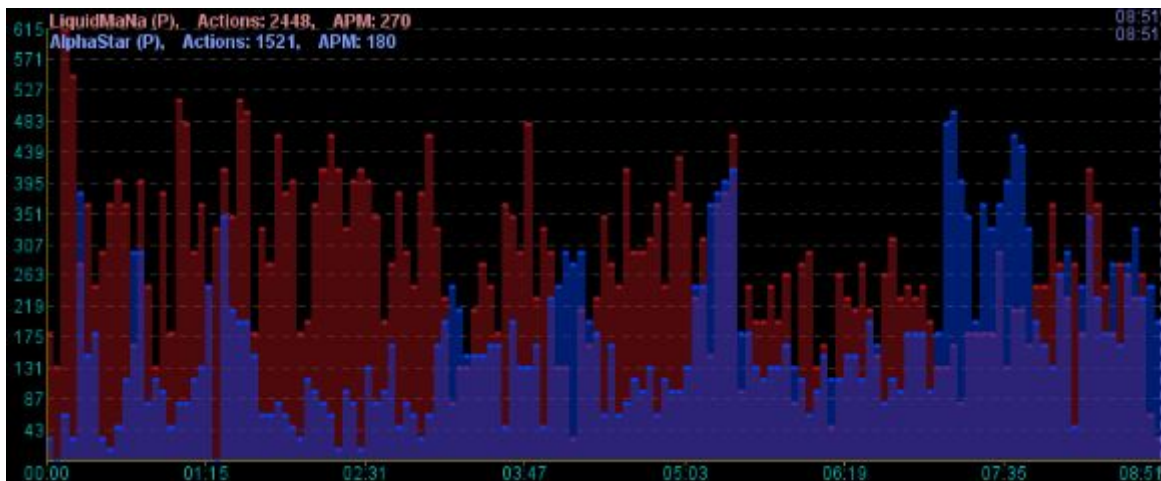
sensación extraña de emoción, incertidumbre y puede que miedo, ya que nos abre una multitud de posibilidades.

Riesgos

Existe una gran controversia respecto a la velocidad a la que juega AlphaStar, aunque DeepMind desde un primer momento quería que la velocidad de esta IA fuera parecida a la de un jugador humano, lo cierto es que no ha sido posible y este ha sido un factor que ha influenciado en los resultados de las partidas.

Starcraft II es un juego que premia a la habilidad de micro gestionar varias cosas a la vez y de dar muchas órdenes o comandos en un corto periodo de tiempo. Como ya hemos explicado, los jugadores deben construir sus bases, gestionar recursos, estar atentos al mapa, investigar para obtener mejor tecnología, instruir tropas, combatir... Y todo esto de manera simultánea. Y, de nuevo, a pesar que DeepMind impuso restricciones y límites en cuanto a cuan rápido AlphaStar podía reaccionar o cuantas órdenes podía dar en un tiempo determinado. Mucha gente creyó que dicha IA actuaba, en ciertos momentos, a más velocidad y con más precisión de lo que podría un ser humano.

La siguiente gráfica muestra las APM (acciones por minuto) del jugador profesional MaNa (en rojo) y las de la IA AlphaStar (en azul) en su segunda partida en un intervalo de 5 segundos (eje Y) durante los primeros minutos de partida (eje X).



A primera vista puede parecer que DeepMind sí que restringió la velocidad de AlphaStar a un nivel humano. Pero cuando diferentes usuarios empezaron a indagar más en profundidad en las acciones que habían hecho, descubrieron que las primeras acciones del jugador profesional simplemente eran las de cambiar de punto de vista entre los diferentes grupos de patrullas que iba creando (acciones innecesarias pero que realizan muchos jugadores para “calentar” los dedos con los que juegan toda la partida). Así pues, no le dieron mucha importancia a este elevado número de APM y cambiaron de foco, apuntando ahora al pico de APM de AlphaStar que se encuentra aproximadamente en el minuto 7 y en el cual se aprecia

una enorme diferencia respecto a los de MaNa. En ese minuto, ambos jugadores entraron en combate y parece ser que AlphaStar superó los APM máximos que había realizado hasta el momento para eliminar la mayoría de fuerzas enemigas y sacar una ventaja clave que iría agrandando a lo largo de la partida. Finalmente acabó ganando AlphaStar y, no nos queremos extender mucho en esto ya que es algo difícil de comprender para personas que no saben nada de este juego, pero mucha gente aseguró que dicho combate del principio había sentenciado el resultado de la partida y que no habría habido manera humana de remontarla.

Relacionado con lo que acabamos de contar, existe otra controversia aunque menos popular, sobre la diferencia en cómo interactúa un humano con la cámara, a como lo hace la IA. El jugador tiene que mover tropas por el mapa para poder ver lo que hay mientras que AlphaStar tiene visión de todo el mapa, aunque en teoría lo tiene como con niebla para que no pueda sacar información sin mover tropas. El hecho es que AlphaStar no gasta prácticamente acciones en mover tropas por el mapa para obtener visión y, aunque se sabe que esto genera una clara ventaja para la IA, no se sabe cuán de grande es.

El hecho de porqué explicamos estas controversias en riesgos del impacto en la sociedad es porque han sido los mismos usuarios quienes han realizado las investigaciones y quienes han tenido sentimientos encontrados con esta IA. ¿Se sentirán engañados por las declaraciones de DeepMind? ¿Les dará igual que AlphaStar supere los límites humanos? ¿Querrían jugar contra ella o se sentirían frustrados hasta el punto de enemistarse con el juego o sus creadores? ¿Se puede humanizar una IA?

Referencias

- [1] Archimedes' Lab, *Online Game by Archimedes' Lab* recuperado de https://www.archimedes-lab.org/game_nim/play_nim_game.html
- [2] Charles C. Margossian (14 de enero de 2019) *A Review of Automatic Differentiation and its Efficient Implementation*. *arXiv:1811.05031v2*
- [3] DeepMind (25 de enero 2019) *AlphaStar Resources* recuperado de <https://deepmind.com/research/open-source/alphastar-resources>
- [4] Michael Graziano (17 de septiembre de 2019) *Rethinking Consciousness: A Scientific Theory of Subjective Experience*. W. W. Norton & Company
- [5] Oriol Vinyals, Igor Babuschkin, David Silver (30 de octubre 2019) *Grandmaster level in StarCraft II using multi-agent reinforcement learning* recuperado de https://www.nature.com/articles/s41586-019-1724-z.epdf?author_access_token=IZH3nqPYtWJXfDA10W0CNNRgN0jAjWel9jnR3ZoTv0PSZcPzJFGNAZhOIk4deBCKzKm70KfinloafEF1bCCXL6IIHHgKaDkaTkBcTEv7aT-wqDoG1VeO9-wO3GEoAMF9bAOt7mJ0RWQnRVMbyfgH9A%3D%3D
- [6] Richard S. Sutton, Andrew G. Barto (19 de octubre de 2018) *Reinforcement Learning: An Introduction 2nd Edition*. A Bradford Book
- [7] The AlphaStar Team (30 de octubre 2019) *AlphaStar: Grandmaster level in StarCraft II using multi-agent reinforcement learning* recuperado de <https://deepmind.com/blog/article/AlphaStar-Grandmaster-level-in-StarCraft-II-using-multi-agent-reinforcement-learning>
- [8] Wikipedia, *Artificial intelligence in video games* recuperado de https://en.wikipedia.org/wiki/Artificial_intelligence_in_video_games
- [9] AI Impacts, *The unexpected difficulty of comparing AlphaStar to humans* recuperado de <https://aiimpacts.org/the-unexpected-difficulty-of-comparing-alphastar-to-humans/>