



ALPHA STAR



DeepMind

- Atari (2013)
- AlphaGo (2015)
- AlphaZero (2017)
- CTF (2018)
- AlphaStar (2019)



StarCraft 2

(turniejowy 1v1)

- gracze zaczynają z bazą i pewną liczbą robotników, po przeciwnych stronach mapy
- robotnicy zbierają zasoby pozwalające wznosić budynki i rekrutować nowe jednostki
- gracz musi kontrolować dwa aspekty
 - ◆ “**macro**”, czyli ekonomię - tempo zbierania zasobów, rozwijanie technologii, budowanie armii
 - ◆ “**micro**”, czyli wydawanie rozkazów pojedynczym jednostkom
- blizzard stara się utrzymać zasadę “kamień, papier, nożyce”, przez co gracz musi umieć przewidywać plany przeciwnika oraz szybko dostosowywać się do niespodziewanych sytuacji





Wyzwania

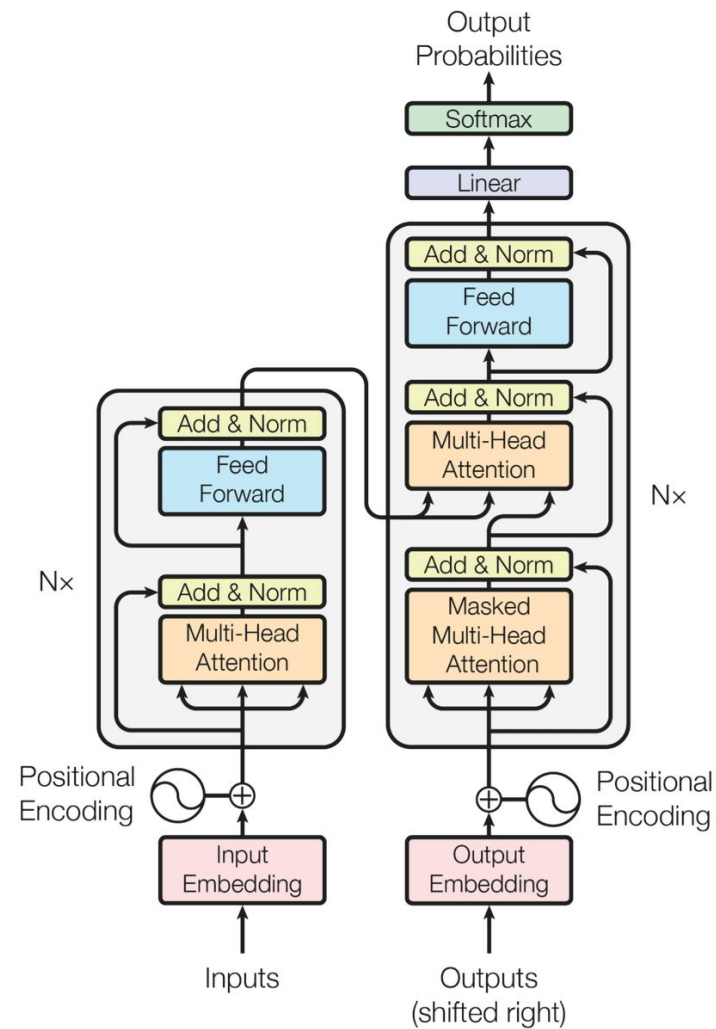
- **Teoria gry:** nie istnieje najlepsza strategia, agent musi znać całą grę zagrań
- **Niedoskonałe informacje:** większość mapy ukryta jest za “mgłą wojny”, agent musi skutecznie prowadzić zwiad
- **Długoterminowe planowanie:** skutki podjętych akcji mogą być widoczne znacznie później
- **Gra czasu rzeczywistego:** akcje muszą być podejmowane natychmiast
- **Ogromna przestrzeń akcji:** Setki jednostek i budynków może wykonywać wiele akcji w dowolnej chwili. Implementacja Deepmind ma 10 do 26 możliwych akcji w każdym momencie.

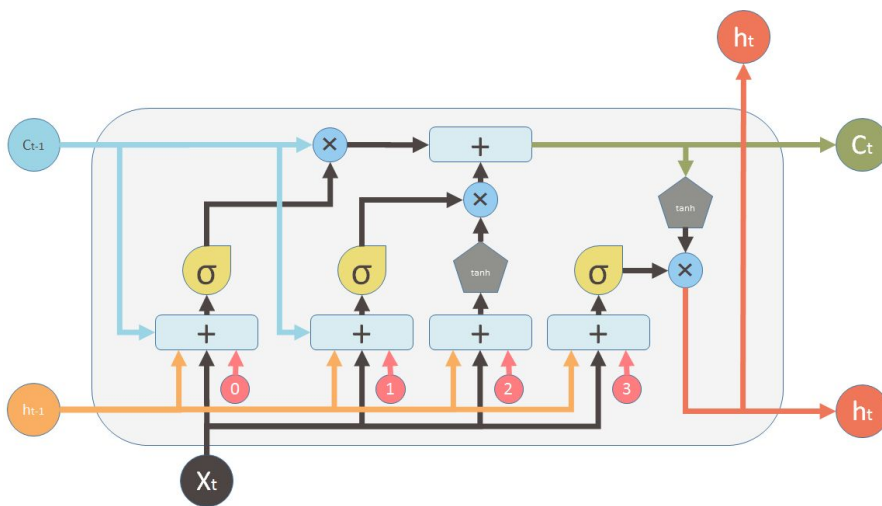


BIG. OOF.

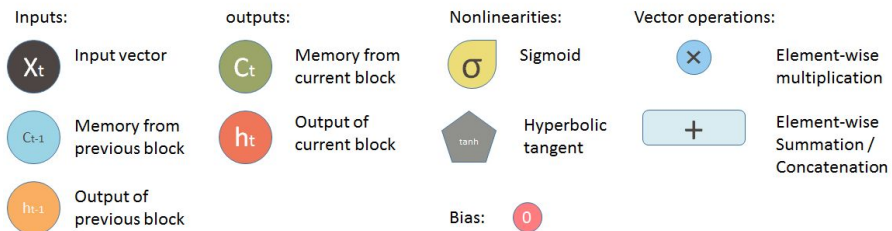
AlphaStar's behaviour is generated by a deep [neural network](#) that receives input data from the raw game interface (a list of units and their properties), and outputs a sequence of instructions that constitute an action within the game. More specifically, the neural network architecture applies a [transformer](#) torso to the units (similar to [relational deep reinforcement learning](#)), combined with a [deep LSTM core](#), an [auto-regressive policy head](#) with a [pointer network](#), and a [centralised value baseline](#). We believe that this advanced model will help with many other challenges in machine learning research that involve long-term sequence modelling and large output spaces such as translation, language modelling and visual representations.

TRANSFORMER

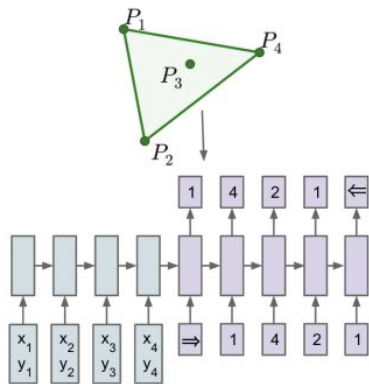




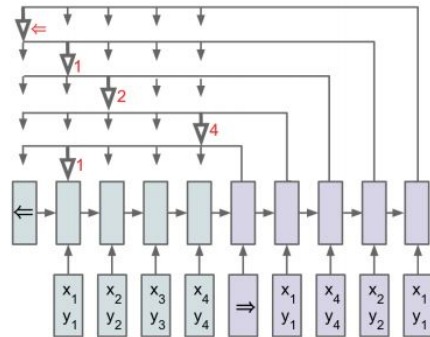
LSTM



POINTER NETWORK



(a) Sequence-to-Sequence

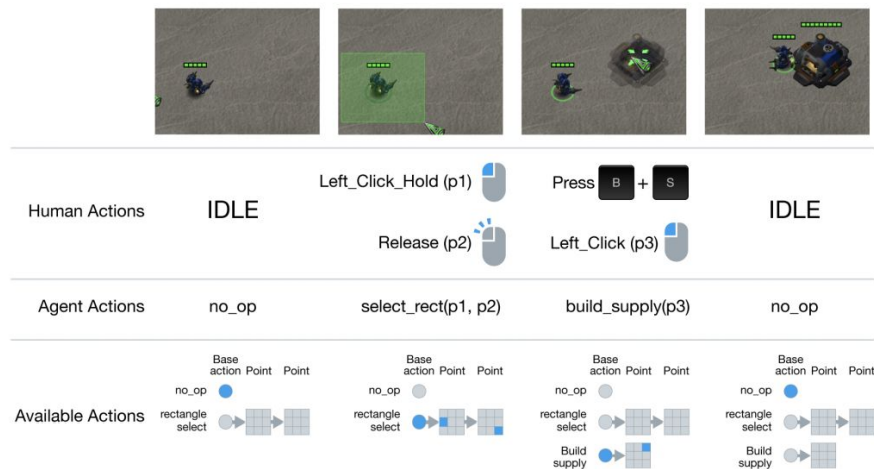


(b) Ptr-Net

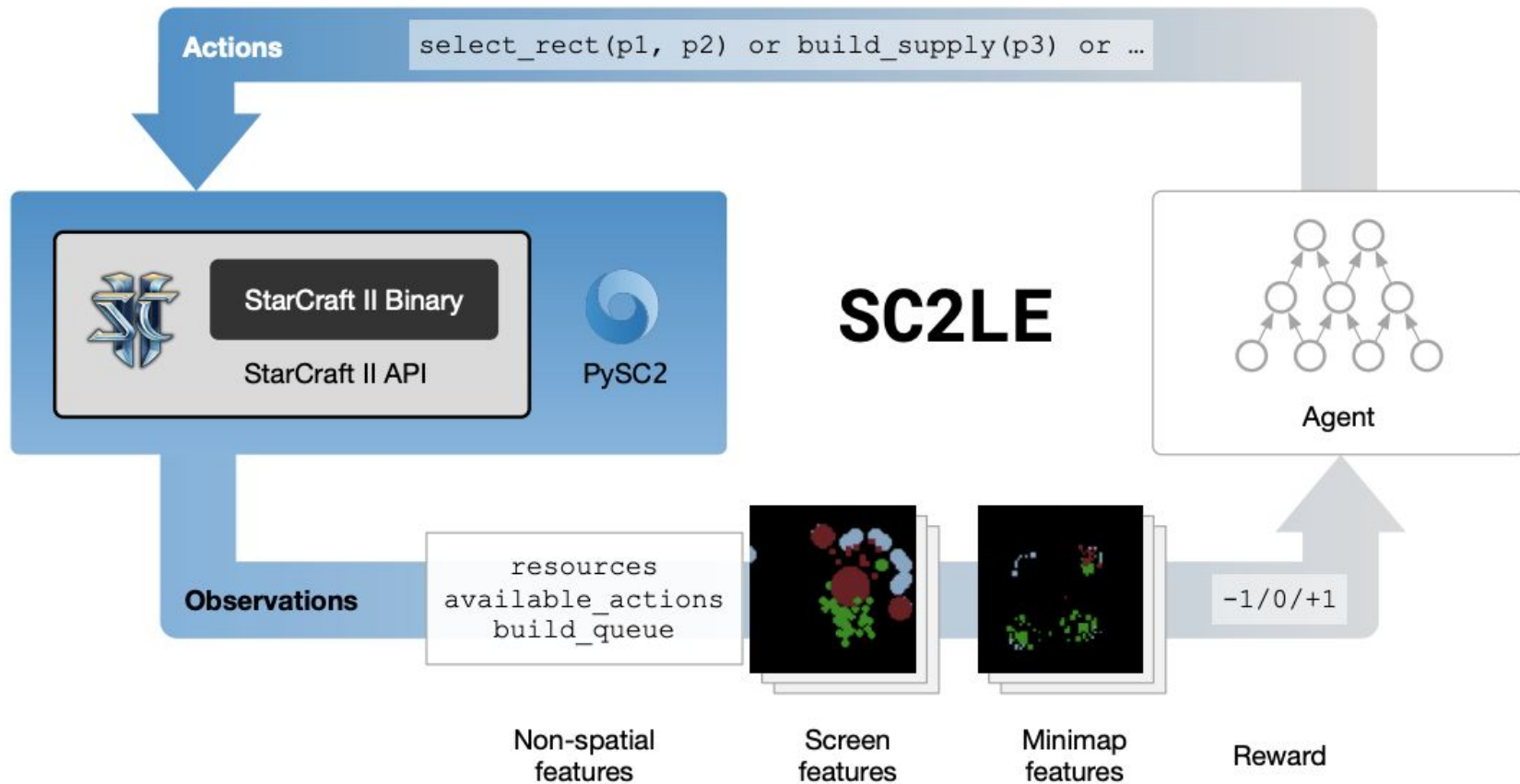
AUTO-REGRESSIVE POLICY

$$\pi_{\theta}(a|s) = \prod_{l=0}^L \pi_{\theta}(a^l | a^{<l}, s).$$

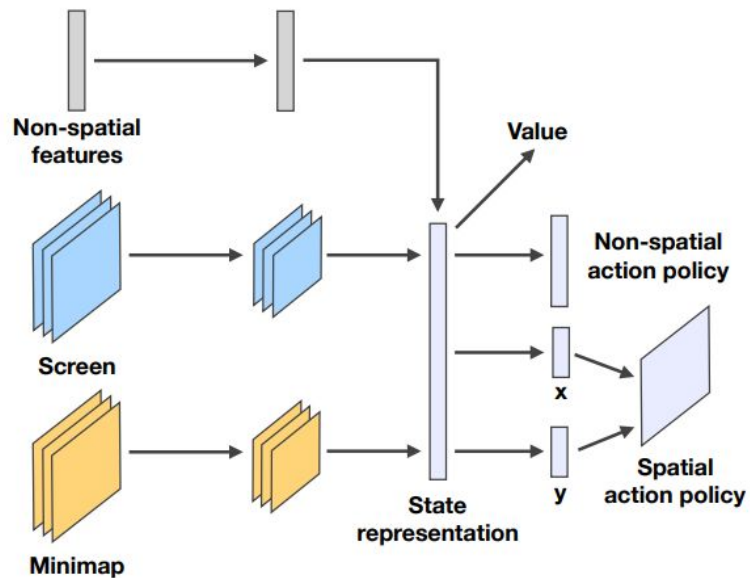
a^0 - funkcja, $a^i, i > 0$ - lista argumentów



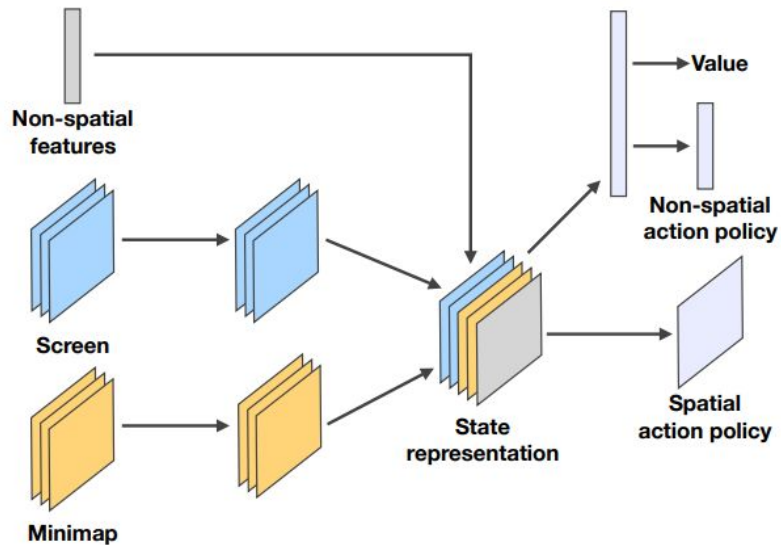
CO JUŻ WIDZIELIŚMY?



CO JUŻ WIDZIELIŚMY?



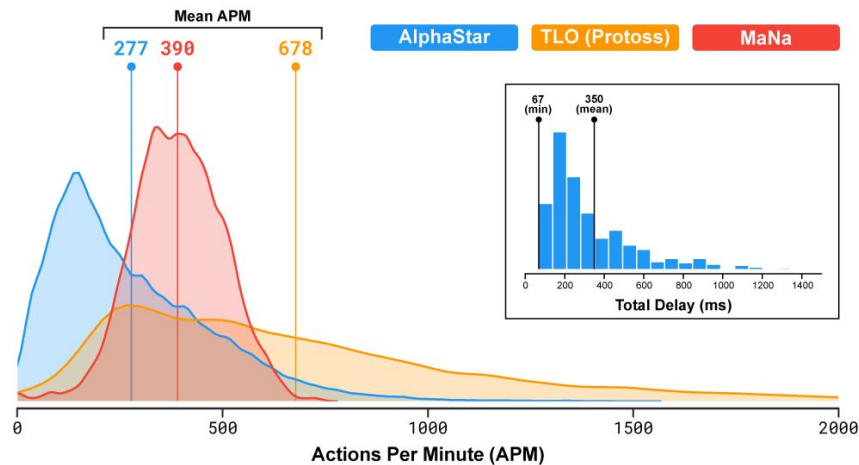
(a) Atari-net



(b) FullyConv

Cel: gra agenta ma przypominać człowieka

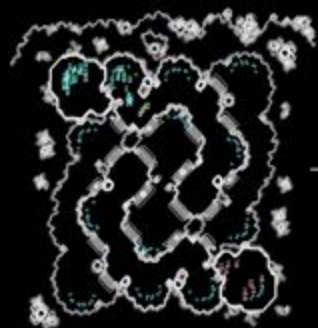
- ograniczony APM
- opóźnienie symulujące czas reakcji
- agent musi obserwować fragment mapy, którego dotyczy akcja



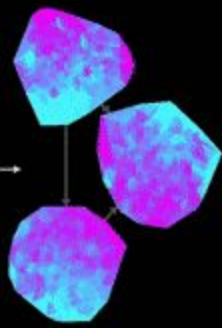




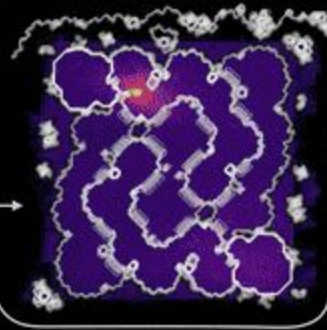
Raw Observations



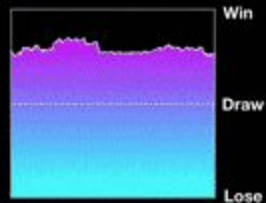
Neural Network Activations



Considered Location



Outcome Prediction



Considered Build/Train

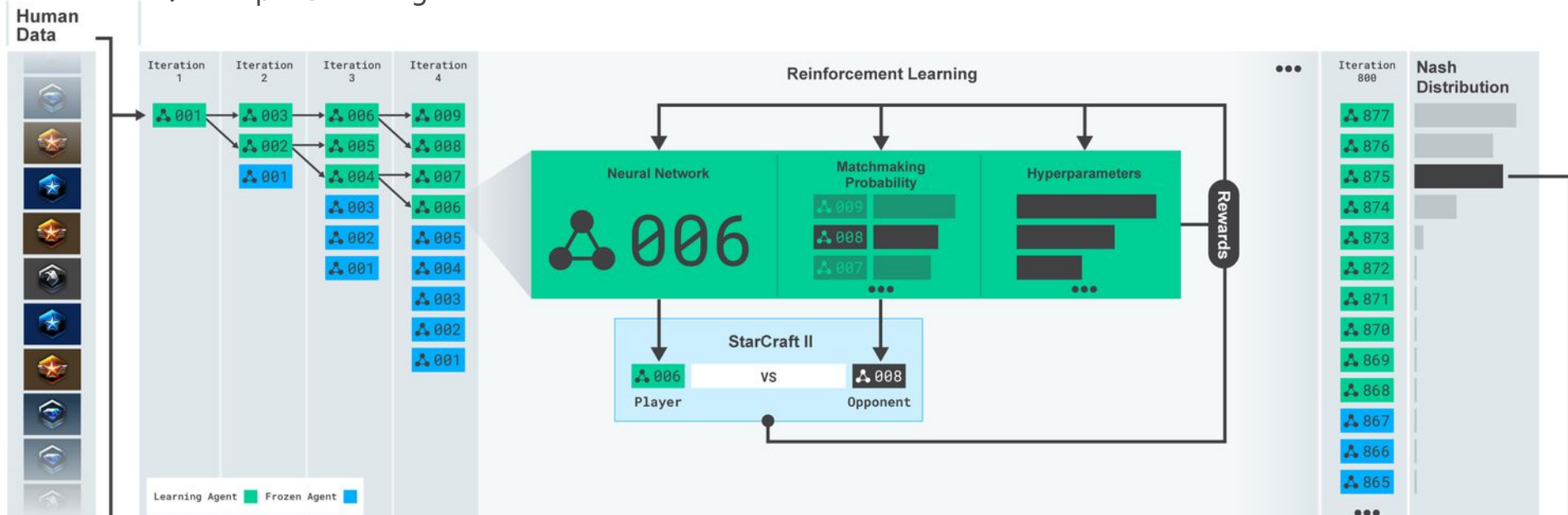


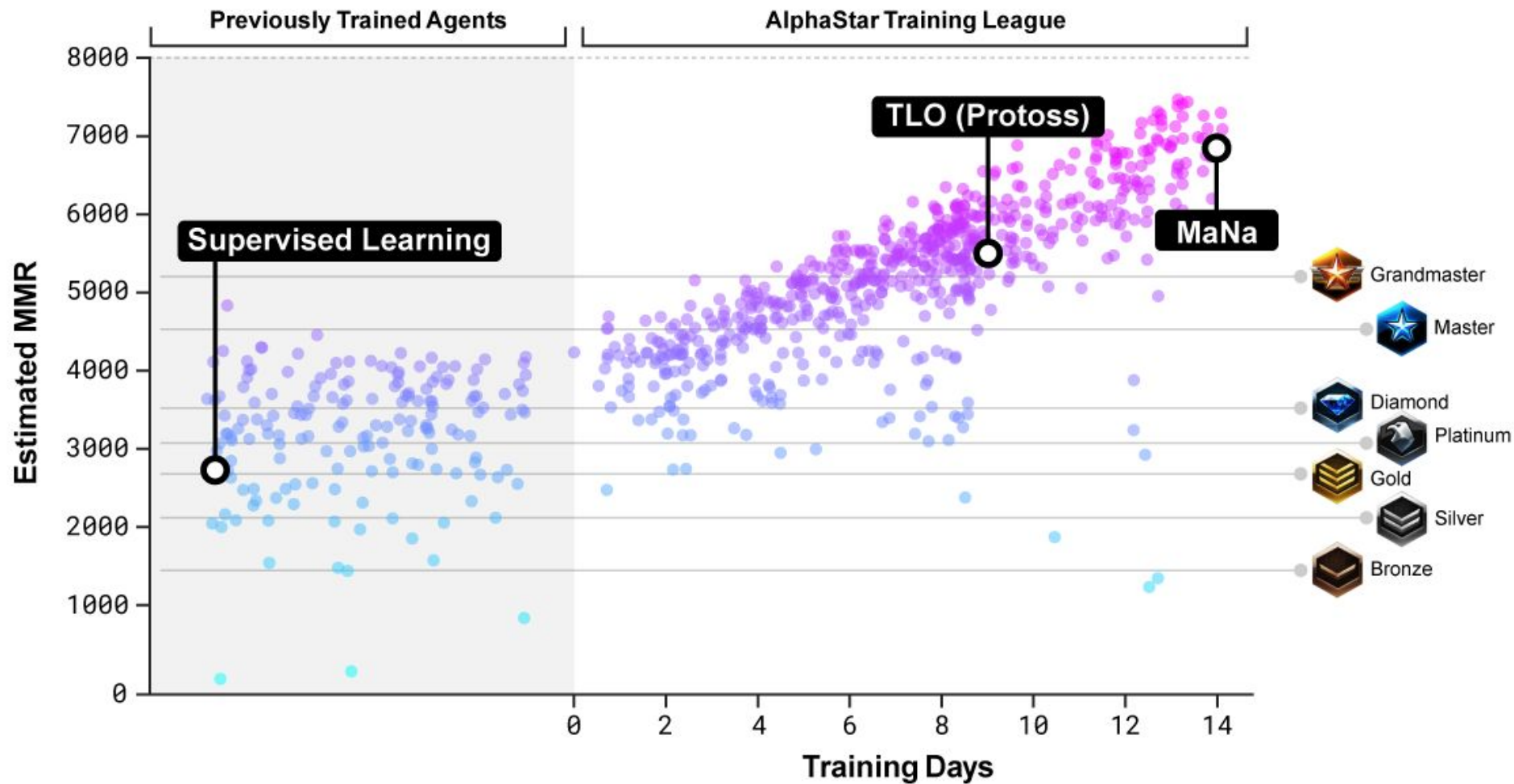
14:32
Galaxy LE

	AlphaStar	LiquidTLO
SCORE	177 / 100	147 / 110
SUPPLY	945 / 1015	335 / 1000
MINERALS	758 / 872	442 / 1030
GAS	64	61
WORKERS	113	86
ARMY	940	1377
PRODUCTION		

Trening

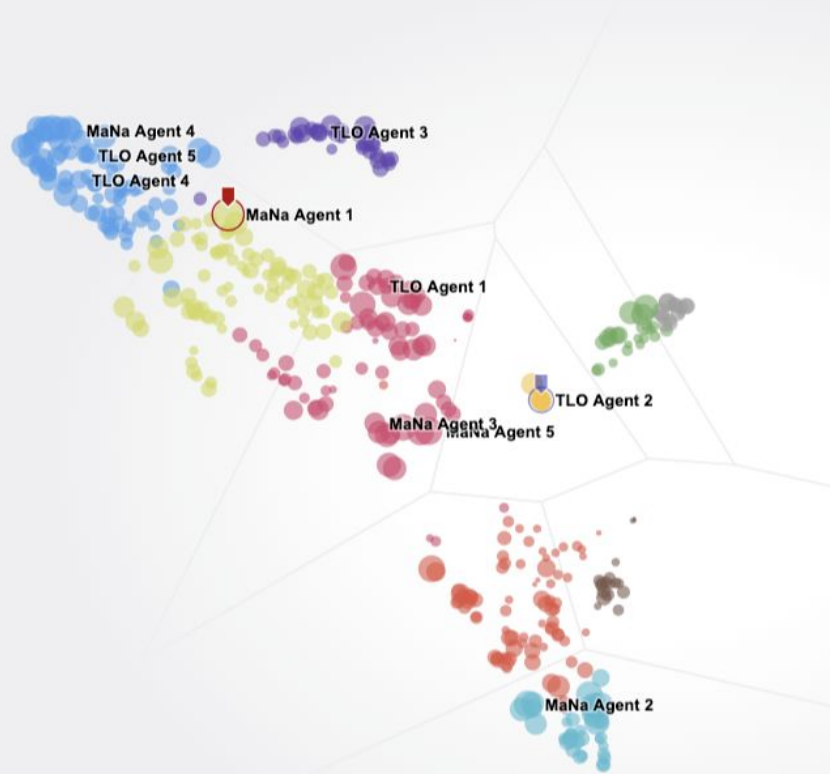
- Supervised Learning
 - ◆ używając anonimowych replayów rozgrywek udostępnionych przez Blizzard
- Reinforcement Learning
 - ◆ AlphaStar League





AlphaStar League Strategy Map

Coloured by strategy clusters



MaNa Agent 1

MMR 7468

TLO Agent 2

6549 MMR



Win rate of MaNa Agent 1 against TLO Agent 2

86%

Unit average count per game



Tech tree build probability



Aby zapewnić różnorodność w lidze, każdy agent ma własny cel. Może nim być na przykład pokonanie konkretnego typu lub dystrybucji agentów, przy użyciu tylko określonych jednostek. Cele są dostosowywane podczas uczenia.



AlphaStar trenowany jest na skalowalnej, rozproszonej konfiguracji wykorzystującej Google TPU v3. W ciągu 14 dni działania Ligi, przy użyciu 16 TPU na agenta, każdy z nich zyskał doświadczenie warte 200 lat rzeczywistej gry.

Źródła

- blogpost <https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/>
- poprzedni paper o StarCraftcie <https://arxiv.org/abs/1708.04782>
- Pointer Network <https://papers.nips.cc/paper/5866-pointer-networks.pdf>
- Transformer <https://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf>