Yogesh Haribhau Kulkarni • You
AI Coach (Machine Learning, Natural Language Processin...
now •

Have we reached AGI? That's the Q*uestion.

*Conditions apply

OpenAI drama that's being unfolded seemingly looks out of the world. Doesn't look human. Yes, that's most likely the reason behind it. An algorithm called Q*, which has supposedly achieved AGI and that's supposedly the culprit of the ouster of @Sam Altman.

Picture this: researchers cautioning the board about the potential AGI prowess of Q*. Sam Altman, the CEO at the time, allegedly fell short of full transparency, leading to his dismissal. The town's talking, and it's all about AGI.

What's AGI? AGI (Artificial General Intelligence) is a far more evolved form of AI than the form we see around today, which is called ASI (Artificial Specific Intelligence). ASI can do one task very well, e.g. predicting next move in chess, or translating from English to Hindi, but not both. That's what AGI can do, almost like humans, it can do multiple tasks well.

How to develop AI? These days the most popular approaches are Machine Learning and Deep Learning. They are further divided as Supervised (where inputs and outputs are specified, e.g. classifying if an image is a cat or a dog), Unsupervised (where only inputs are specified, not outputs, e.g. grouping similar customers together) and Reinforcement Learning (where

penalties and rewards are specified, e.g. predicting next move in chess).

ChatGPT that we know, is actually combination of Supervised (transformer architecture , ie. encoder decoder network, e.g. machine translation) and Reinforcement Learning with human feedback.

Q*-Learning is one of the approaches in Reinforcement Learning. It doesn't require prior knowledge of the environment. It learns by doing, adjusting actions based on rewards and penalties. Here, "Q" refers to the quality of actions in a given state. The Q-value of an action in a state represents the expected cumulative reward if that action is taken. Q* specifically emphasizes the absolute optimal policy.

Example:
In a maze, the agent (robot) learns Q-values for each possible action (move left, right, up, down) in each state. During exploration, the agent updates Q-values based on the rewards it receives. Over time, it learns the optimal policy—choosing actions with the highest Q-values in each state—to efficiently navigate the maze and reach the goal.

There are many unknowns, no public validation yet, that's why I said "*" Conditions Apply. But its very clear now, that its not IF AGI is going to come or not, but just, WHEN.

All attempts are being made to make AI a benevolent as possible. But you never know ("NO"?), like the character (who is this guy?) shown below AI can point guns at you.

#qstar #ml #neuralnetwork #ai #artificialintelligence #jamesbond #drno #reinforcementlearning #agi #aritificalgeneralintelligence #samaltman #chatgpt #future