Utilizarea Reinforcement Learning în AlphaGo

Poponet Tiberiu-Sergiu

Grupa 30231

AlphaGo este un program de calculator dezvoltat de compania DeepMind Technologies, o subsidiară a Alphabet Inc. (compania-mamă a Google). Acesta a devenit faimos în 2016 când a învins campionul mondial de Go Lee Sedol marcatând un progres semnificativ în domeniul inteligenței artificiale. Un aspect cheie al succesului AlphaGo este utilizarea Învățării Profunde Reinforcement (RL), o paradigmă de învățare care implică luarea deciziilor bazate pe interacțiunea agentului cu mediul înconjurător și pe feedbackul obținut.

AlphaGo utilizează o abordare hibridă, combinând tehnici de învățare supervizată și RL. RL este esențial în faza de evaluare a mutărilor și îmbunătățirea jocului, iar această abordare a fost detaliată într-un articol semnificativ publicat de cercetătorii DeepMind.

În cadrul AlphaGo, RL este folosit în principal pentru antrenarea rețelei de joc (policy network) și a rețelei de evaluare a pozițiilor (value network). Agenții de RL sunt instruiți să joace Go împotriva lor înșiși și să optimizeze politica de joc și evaluarea pozițiilor în funcție de rezultatele obținute. Procesul de antrenare a implicat milioane de jocuri simulare, permițând AlphaGo să dobândească o înțelegere profundă a strategiilor și tacticii din Go.

Policy network este responsabil pentru generarea de mutări în timpul jocului. Acesta este instruit folosind metode de RL pentru a maximiza șansele de a obține victoria într-un joc simulat. RL îi permite să ajusteze strategiile în funcție de succesul obținut în partidele anterioare.

Value network evaluază pozițiile din joc și furnizează o estimare a șanselor de victorie ale jucătorului într-o anumită stare. RL este utilizat pentru a antrena value network pentru a oferi estimări precise, ceea ce îmbunătățește strategiile de joc ale AlphaGo.

Articolul de referință pentru acest referat este "Mastering the game of Go with deep neural networks and tree search" publicat în revista Nature în 2016 de către Silver et al. Acest articol oferă o perspectivă detaliată asupra metodelor utilizate în AlphaGo și modului în care RL contribuie la performanța sa remarcabilă în jocul de Go.

În concluzie, AlphaGo demonstrează eficiența Învățării Profunde Reinforcement în abordarea problemelor complexe precum jocul de Go. Integrarea acestei paradigme în dezvoltarea programelor de calculator a dus la progrese semnificative în domeniul inteligenței artificiale. Articolul oferă o perspectivă detaliată asupra modului în care RL a fost implementat în AlphaGo, iar această abordare a deschis calea către explorarea și înțelegerea mai profundă a inteligenței artificiale și a învățării automate.