

UCZENIE ZE WZMOCNIENIEM - WSPÓŁCZESNE WYZWANIA

Uczenie maszynowe dzieli się na trzy obszary, uczenia z nadzorem (supervised learning), uczenia nienadzorowanego (unsupervised learning) i uczenia ze wzmocnieniem (reinforcement learning RL). Yann LeCun jeden z ojców współczesnej rewolucji w uczeniu maszynowym nazwał RL „wiśnią na torcie”. Pomimo, że w swoim wystąpieniu stwierdził, że wiemy jak „zrobić tę wiśnię” wciąż jest wiele podstawowych problemów, które muszą być rozwiązane. Na przykład niedawne przełomowe osiągnięcia w grach (DQN, GO, Dota2) wciąż operują w sterylnych środowiskach i wymagają ogromnej mocy obliczeniowej. Celem tego projektu jest zbadanie niektórych z tych ograniczeń. W tym celu chcemy stworzyć na Uniwersytecie Warszawskim grupę badawczą prowadzącą badania w dziedzinie RL na światowym poziomie

Cel prowadzonych badań/hipoteza badawcza. Planujemy badań następujące tematy: **uczenie hierarchiczne (hierarchical learning), uczenie stopniowe (curriculum learning), uczenie na poziomie meta i pamięć (meta-learning /memory), metody modelowe i bezmodelowe w RL (model-free and model-based methods).**

Uczenie hierarchiczne wciąż jest słabo zbadaną dziedziną. W niedawnym artykule, w którym przedstawiliśmy nową metodę uczenia bazując na której planujemy zbadać dwa ważne aspekty uczenia hierarchicznego. Pierwsza to kompatybilność makro-akcji a druga to automatyzacja wykrywania struktury hierarchicznej. W obu przypadkach proponujemy nowe algorytmy, których efektywność będzie zbadana podczas projektu.

Uczenie stopniowe Proponujemy nową metodę generowania programu. Założmy, że chcemy rozwiązać zadania, które jest jednym z wielu podobnych, ale różniących się poziomem trudności zadań. Nasz algorytm w każdej fazie nauki wybiera główne zadanie, w sposób, który ma za zadanie maksymalizować prędkość problemu uczenia. Pozostałe zadania są trenowane jednocześnie, co ma dodatkowo przyspieszyć tempo gromadzenia wiedzy. W projekcie zaimplementujemy i przeprowadzimy testy naszej metody.

Uczenie na poziomie meta i pamięć Celem naszych badań będzie zbadanie użyteczności różnych architektur w uczeniu na meta poziomie. Zbadamy ich możliwość pamiętania i do wykonywania algorytmów szybkiego uczenia RL na semantycznie spójnych dziedzinach.

Metody modelowe i bezmodelowe w RL Głównym celem badawczym jest stworzenie algorytmu łączącego uczenie modelowe i bezmodelowe. Podzieliliśmy to zadanie na trzy części. Pierwsza jest poświęcona automatycznemu wykrywaniu niskowymiarowych, lokalnych modeli RL. Druga część będzie poświęcona dostosowania algorytmu MCTS, która będzie dostosowana do takich modeli. Przepuszczamy, że kluczowe będzie unikanie wielokrotnych wizyt w tym samym (podobnym). Ostatecznym celem jest połączenie obydwu technik w jedną metodę.

Metoda badawcza/metodyka. Standardową metodą stosowaną w nowoczesnych badaniach nasz sztuczną inteligencją jest przeprowadzanie eksperymentów. W projekcie prezentujemy kilka nowych metod, do każdej z nich proponujemy eksperyment, który zweryfikuje ich użyteczność. Aby skutecznie wprowadzić ten plan w życie planujemy zakupić komputery wyposażone w GPU i stworzyć zespół badawczy.

Wpływ spodziewanych rezultatów na rozwój nauki. RL jest jedną z ważniejszych obszarów współczesnych badań nad sztuczną inteligencją. Badania zaproponowane w projekcie są nowatorskie i mam y nadzieję, że przyczynią się do rozwoju dziedziny.

Uczenie hierarchiczne W społeczności badawczej panuje przekonanie, że hierarchiczność może być kluczowym elementem prawdziwie inteligentnych systemów. Zbadanie kompatybilności makro-akcji może przyczynić się do opanowania technik potrzebnych użyciu hierarchiczności. Drugie zadanie dotyczy zbadania możliwości automatycznego wykrywania struktury hierarchicznej. Jest to kluczowe zagadnienie w tej stosunkowo mało zbadanej dziedzinie i każdy wynik może przyczynić się do istotnego postępu.

Uczenie stopniowe Znalezienie ogólnej metody generowania programu szkolenia dla agenta byłoby znaczącym postępem dla RL. Mamy nadzieję, że nasz algorytm będzie stanowił postęp w tym kierunku.

Uczenie na poziomie meta i pamięć Uczenie na poziomie meta jest oparte o założenie, że rozwiązujemy problem z pewnej klasy, w której są pewne zasady (być może niejawne). Wykrywanie takich zasad może znacząco poprawić efektywność algorytmów RL. Ma to szczególne znaczenie dla praktycznych zastosowań, które często są zdefiniowane przez problemy zawierające wiele niejawnej wiedzy o świecie.

Metody modelowe i bezmodelowe w RL Znalezienie metod łączenia metod modelowych i bezmodelowych może mieć znaczący wpływ na dziedzinę RL. Ważny postęp w tym zakresie zostały poczyniony przez klasę algorytmów GPS. W projekcie chcielibyśmy powtórzyć ten sukces dla problemów dyskretnych. Metody takie mogą znacząco zmniejszyć ilość danych potrzebnych do treningu.