



AGH

**AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W
KRAKOWIE**

**WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI,
INFORMATYKI I INŻYNIERII BIOMEDYCZNEJ**

KATEDRA AUTOMATYKI I ROBOTYKI

Praca dyplomowa magisterska

*Kognitywny system eksploracji środowiska przez
inteligentnego agenta z wykorzystaniem uczenia
motywowanego.*

*Cognitive system of environment exploration by an
intelligent agent using motivated learning*

Autor:

Bartłomiej Styczeń

Kierunek studiów:

Automatyka i robotyka

Opiekun pracy:

dr hab. Adrian Horzyk

Kraków, 2020

Uprzedzony o odpowiedzialności karnej na podstawie art. 115 ust. 1 i 2 ustawy z dnia 4 lutego 1994 r. o prawie autorskim i prawach pokrewnych (t.j. Dz.U. z 2006 r. Nr 90, poz. 631 z późn. zm.): „Kto przywłaszcza sobie autorstwo albo wprowadza w błąd co do autorstwa całości lub części cudzego utworu albo artystycznego wykonania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 3. Tej samej karze podlega, kto rozpowszechnia bez podania nazwiska lub pseudonimu twórcy cudzy utwór w wersji oryginalnej albo w postaci opracowania, artystycznego wykonania albo publicznie zniekształca taki utwór, artystyczne wykonanie, fonogram, wideogram lub nadanie.”, a także uprzedzony o odpowiedzialności dyscyplinarnej na podstawie art. 211 ust. 1 ustawy z dnia 27 lipca 2005 r. Prawo o szkolnictwie wyższym (t.j. Dz. U. z 2012 r. poz. 572, z późn. zm.): „Za naruszenie przepisów obowiązujących w uczelni oraz za czyny uchybiające godności studenta student ponosi odpowiedzialność dyscyplinarną przed komisją dyscyplinarną albo przed sądem koleżeńskim samorządu studenckiego, zwanym dalej «sądem koleżeńskim».”, oświadczam, że niniejszą pracę dyplomową wykonałem(-am) osobiście i samodzielnie i że nie korzystałem(-am) ze źródeł innych niż wymienione w pracy.

*Serdecznie dziękuję ...tu ciąg dalszych
podziękowań np. dla promotora, żony, są-
siada itp.*

Spis treści

1. Wprowadzenie	7
1.1. Cele pracy	8
1.2. Zawartość pracy	9
2. Motywacja.....	11

1. Wprowadzenie

Od początku pierwszych urządzeń elektronicznych, a szczególnie zaprojektowaniu pierwszych komputerów, które mogły wykonywać pewne zbiory instrukcji, naukowcy zastanawiali się jak wykorzystać te moce obliczeniowe. Próbowano zdefiniować pojęcie inteligencji. I tak w roku 1955 termin *sztuczna inteligencja* został wprowadzony jako dyscyplina naukowa. Natomiast w roku 1950 matematyk Alan Turing opracował test nazwany jego nazwiskiem, którego celem jest sprawdzenie czy maszyna, która posiada inteligencję równą lub przewyższającą człowieka.

Test został wprowadzony w artykule [1], w którym Turing zaproponował zastąpienie pytania: *Can machines think?* (ang. *Czy maszyny myślą?*) pewną formą testu. Polega on na prowadzeniu rozmowy przez 2 osoby i jedną maszynę. Zadaniem jednego człowieka i maszyny jest przekonanie sędziego, że są one człowiekiem. Jeżeli sędzia nie jest w stanie zdecydować, oznacza to, że maszyna zdaje test. Ostatecznie pytanie, które zadawał Turing na początku artykułu zostaje zmienione na: *Can machines do what we (as thinking entities) can do?* (ang. *Czy maszyny umieją robić to, co myślące istoty?*).

Temat inteligencji był coraz bardziej popularny, ale po kilku latach badań, nie osiągnęto rezultatów. Nastąpiła tzw. zima dla AI (ang. *AI winter*). W trakcie kilku dziesięcioleci nie prowadzono wielu badań w zakresie sztucznej inteligencji. Trwało to do początku lat dziewięćdziesiątych, kiedy to w roku 1998 francuski informatyk Yann LeCun wydał artykuł [2] opisujący konwolucyjne sieci neuronowe, które uczyły bezpośrednio z obrazów, a nie z wcześniej zaprojektowanych cech. Był to początek przełomu w procesie uczenia sztucznych sieci neuronowych, ponieważ pierwszy raz można było przeprowadzić w pełni automatyczne uczenie sieci neuronowej.

Kolejnym przełomem w dziedzinie sztucznej inteligencji była praca Alexa Krizhevskiego, Ilyi Sutskevera i Geoffreya E. Hintona [3]. Ich przełomowa implementacja architektury sieci konwolucyjnej na kartach graficznych umożliwiła stworzenie głębokiej (jak na rok 2012) sieci, która wygrała konkurs ImageNet [4]. Od tego roku tematyka sztucznych sieci neuronowych,

a w szczególności konwolucyjnych sieci neuronowych. Postęp technologiczny w architekturach kart graficznych umożliwił tworzenie coraz większych sieci, które dawały coraz lepsze rezultaty. Powrót do algorytmów uczenia ze wzmocnieniem (ang. *reinforcement learning*) sprawił, że powróciła chęć rozwijania sztucznej inteligencji, a nawet próby regulacji prawnej tych rozwiązań.

Szczególną uwagę należy zwrócić na uczenie ze wzmocnieniem, które przypomina po części silną sztuczną inteligencję (ang. *artificial general intelligence*). Agent mając niewiele informacji o środowisku jest w stanie nauczyć się skomplikowanych zadań bez wcześniejszej wiedzy. Popularnym przykładem jest AlphaGo, który został zaprojektowany przez naukowców z DeepMind do grania w grę Go [5].

Należy jednak pamiętać, że mimo świetnych wyników tych programów komputerowych, to daleko im od silnej sztucznej inteligencji. Są one przystosowane do rozwiązywania jednego problemu, w znanym środowisku. Jak środowisko zostanie zmienione, to wyniki nie są aż tak dobre. AlphaGo nie zacznie nagle autonomicznie sterować samochodem, ponieważ algorytm zna całkowicie inne środowisko. Można zauważyć brak przystosowania do zmieniającego się środowiska.

Jednym z podejść, aby rozwiązać problem dynamicznego środowiska jest umożliwienie agentowi samodzielne decydowanie jakie zadania musi on wykonać, aby rozwiązał główny problem. Uczenie motywowane (ang. *motivated learning*) odnosi się do tego problemu i umożliwia rozwiązywanie złożonych problemów w zmieniających się warunkach działania agenta.

Ta praca jest o uczeniu motywowanym przy eksploracji środowiska przez agenta, który w przypadku tej pracy dyplomowej będzie wykonywana na robocie mobilnym. Zastosowane zostanie środowisko symulacyjne umożliwiające sterowanie takim robotem, odczytywanie różnych parametrów środowiska otaczającego agenta, także tych, których sensory robota, np. kamera, czujniki odległości, nie umożliwiają do odczytania.

1.1. Cele pracy

Celem pracy było zbadanie skuteczności uczenia motywowanego przy eksploracji i poznawaniu wirtualnego środowiska przez inteligentnego agenta kognitywnego. Zakres pracy obejmował przygotowanie wirtualnego środowiska umożliwiającego testowanie zaimplementowanych rozwiązań. Następnie inteligentny agent miał za zadanie poznawanie środowiska i tworzenie nowych skojarzeń i wiedzy poprawiających jego umiejętności oraz stopniowo rozbudowujących hierarchię jego potrzeb i możliwości.

TODO

W ramach pracy należało sprawdzić i porównać skuteczność nowych rozwiązań z zakresu uczenia motywowanego z innymi podejściami, np. uczeniem ze wzmocnieniem.

TODO

Pierwszym etapem był przegląd literatury związanej z tematem uczenia motywowanego oraz uczenia ze wzmocnieniem, którego celem było zgłębienie wiedzy na ten temat w celu zrozumienia różnic i możliwych ograniczeń powyższych rozwiązań.

Bardzo ważnym etapem był wybór środowiska symulacyjnego, aby móc szybko testować zaimplementowane rozwiązania. Takie podejście umożliwia bezpieczne sprawdzenie algorytmów, w takim środowisku wystarczy jeden przycisk, aby przerwać całe zadanie bez zagrożenia dla ludzi w środowisku oraz samego robota. Ważne, aby wybrane środowisko umożliwiło dużą ingerencję w środowisko. Tworzenie dynamicznego świata jest podstawą, aby można było testować skuteczność algorytmów uczenia motywowanego.

Większość algorytmów zaimplementowano w językach C++/Python, tak aby wysoka logika mogła być szybko prototypowana w języku jakim jest Python, natomiast wszelkie obliczenia, które wymagają dużej wydajności zostały zaimplementowane w języku C++.

Istotnym aspektem tej pracy dyplomowej jest analiza porównawcza z innymi znanymi i ugruntowanymi algorytmami do eksploracji środowiska, np. SLAM (ang. *simultaneous localization and mapping*). Porównanie osiągnięć agenta wykorzystującego algorytmy uczenia motywowanego z podobnym agentem korzystającym z innych podejść umożliwi sprawdzenie czy umożliwienie tworzenia własnych wewnętrznych celów sprawia, że agent dostosowuje się do zmiennego środowiska.

1.2. Zawartość pracy

Ta część będzie napisana jak cała reszta pracy zostanie ukończona. . .

2. Motywacja...

Motywacja to coś bardzo ważnego a mi teraz tego trochę brak :D.

Motivated learning for computational intelligence - notatki.

Agent określany jako wiara – pożądania – intencja (ang. *belief – desire – intention*) jest traktowany jako odnośnik do architektury zastosowanej dla agentów uczenia motywowanego. Ból abstrakcyjny powstaje, jeżeli agent nie jest w stanie wykonać akcji, która obniży wartość bólu pierwotnego. Ból abstrakcyjny nie jest stymulowany przez sensor wartości fizycznej (np. głód, zmęczenie), ale jako ból generowany wewnętrznie. Bóle tworzą hierarchię wszerz i włąb, tworząc struktury grafowe. W takich strukturach teoretycznie mogą pojawiać się zamknięte ścieżki, np. brak jedzenia może sprawić, żeby kupić więcej jedzenia, a do tego potrzebne są pieniądze, które można zdobyć sprzedając jedzenie. Należy zastosować mechanizm wykrywający i zapobiegający takim sytuacjom.

Algorytm tworzenia celów (ang. *goal creation system*)

1. Wybierz dominujący ból stosując regułę zwycięzca bierze wszystko (ang. *winner takes all*) spośród konkurujących ośrodków bólu.
 - Jeżeli żaden z bóli nie przekracza wcześniej zdefiniowanego progu, czekaj aż któryś z nich przekroczy ten próg.
2. Jako aktualny cel wybierz zmniejszenie dominującego bólu.
 - aktualny cel motywuje agenta do działania.
3. Wybierz wcześniej nauczoną akcję, która z najwyższym prawdopodobieństwem spełni aktualny cel.
 - Jeżeli nie ma żadnego, idź do punktu 6.
4. Sprawdź czy wybrana czynność może być wykonana w aktualnym środowisku. Jeśli nie, idź do punktu 3.
5. Wykonaj akcję.

- Jeśli ta akcja *obniżyła* wartość dominującego bólu:
 - (a) Zwiększ wartości wag zależności pomiędzy aktualnym bólem a akcją jaka została wykonana i zwiększ wartość wag odpowiadających abstrakcyjnemu bólowi powiązanego z tą akcją.
 - (b) Idź do punktu 1.
 - Jeśli ta akcja *nie obniżyła* wartości dominującego bólu.
 - (a) Zmniejsz wartości wag zależności pomiędzy aktualnym bólem a akcją jaka została wykonana i zmniejsz wartość wag odpowiadających abstrakcyjnemu bólowi powiązanego z tą akcją.
 - (b) Idź do punktu 3.
6. Wykonaj eksplorację przestrzeni akcji mającą na celu spełnienie celu.
- Jeśli nowa akcja *zmniejszyła* wartość dominującego bólu.
 - (a) Zwiększ wartości wag zależności pomiędzy aktualnym bólem a akcją jaka została wykonana i stwórz nowy abstrakcyjny ból związany z niemożliwością wykonania tej akcji.
 - (b) Idź do punktu 1.
 - Jeśli nowa akcja *nie zmniejszyła* wartości dominującego bólu, idź do punktu 6.

Uczenie motywowane może być zastosowane wspólnie z uczeniem opartym na ciekawości (ang. *curiosity based learning*) – ciekawość poinformuje agenta o nowych odkryciach, podczas gdy uczenie motywowane skupi się na poszukiwaniu konkretnych celów, które nie są podane przez twórcę (tak jak w uczeniu ze wzmocnieniem).

Porównanie uczenia motywowanego do uczenia ze wzmocnieniem.

Tabela 2.1. Porównanie uczenia ze wzmocnieniem do uczenia motywowanego. Źródło: Motivated learning for computational intelligence

Uczenie ze wzmocnieniem	Uczenie motywowane
Jedna funkcja kosztu (określona zewnętrznie)	Wiele funkcji kosztu (tworzone przez agenta)
Mierzalne nagrody – może być zoptymalizowane	Wewnętrzne nagrody (niemierzalne) – nie może być zoptymalizowane
Możliwy do przewidzenia	Niemożliwy do przewidzenia
Zadania zdefiniowane przez badacza	Agent sam stwarza sobie cele
Maksymalizacja nagrody – potencjalnie niestabilne przy optymalizacji	Algorytm minimax – stabilne
Brak wewnętrznych motywacji i tworzenia celów	Wewnętrzne motywacje i tworzenie celów
Zawsze aktywne	Wykonuje akcje, kiedy jest taka potrzeba albo agent uzna to za konieczne
Wysiłek uczenia wzrasta wraz ze zwiększaniem skomplikowania środowiska	Agent uczy się lepiej w skomplikowanych środowiskach

Motivated Learning for the Development of Autonomous Systems

Motywacja agenta powinna pojawić się w wyniku procesu rozwojowego (Pfeifer i Bongard 2006). Zostało to zaobserwowane u ludzi i twierdzi się, że jest to wynikiem nagradzania akcji, które choć trochę przekraczają możliwości człowieka. Ludzie mają wewnętrzną potrzebę zadawania pytań: "Dlaczego?" i "Jak?" w celu zrozumienia otaczającego świata (środowiska).

Na podstawie reguły ciekawości został zaproponowany inteligentny, adaptacyjny system ciekawości (ang. *intelligent adaptive curiosity*).

Bibliografia

- [1] Alan Turing. „Computing Machinery and Intelligence”. W: *Mind* 236 (1950), s. 433–460. DOI: <https://doi.org/10.1093/mind/LIX.236.433>.
- [2] Yann LeCun. „Gradient-based learning applied to document recognition”. W: (1998). DOI: <http://yann.lecun.com/exdb/publis/pdf/lecun-99.pdf>.
- [3] Geoffreya E. Hintona Alexa Krizhevskiego Ilyi Sutskevera. „ImageNet Classification with Deep Convolutional Neural Networks”. W: (2012). DOI: <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- [4] Stanford Vision Lab. *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)*. URL: <http://www.image-net.org/challenges/LSVRC/>. dostęp: 30.08.2020.
- [5] DeepMind. *AlphaGo*. URL: <https://deepmind.com/research/case-studies/alphago-the-story-so-far>. dostęp: 30.08.2020.