



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 | 7 |
| 1.2. Zawartość pracy..... | 7 |
| 2. Motywacja... .. | 9 |

1. Wprowadzenie

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 poniższej pracy jest zastosowanie uczenia motywowanego do eksploracji nieznanego środowiska przez agenta (robota mobilnego) i porównanie osiągnięć takiego agenta w porównaniu z podobnym agentem wykorzystującym algorytmy uczenia ze wzmocnieniem, np. przy jednoczesnej lokalizacji i mapowaniu SLAM (ang. *simultaneous localization and mapping*).

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*).