

Problem wyboru akcji w rozproszonych sterownikach robotów mobilnych

Artykuł poświęcony jest sterowaniu robotów mobilnych. Przeprowadzone są rozważania koncentrujące się wokół zastosowania sterowników o architekturze rozproszonej, tzn. takiej w której w przeciwieństwie do architektury hierarchicznej wszystkie warstwy sterownika działają jednocześnie i mają dostęp zarówno do receptorów jak i do układów wykonawczych robota mobilnego. Powoduje to jednak powstanie tzw. problemu wyboru akcji. Zaprezentowane jest rozwiązanie tego problemu.

Action selection problem in distributed architecture for mobile robot control

Article deals with problems of the mobile robots control. The use of the distributed architecture controllers is discussed. In the case of the distributed architecture, all layers of the controller simultaneously can access sensor's data and can send commands to the effectors. This leads to the action selection problem. Solution to this problem is proposed.

1. WPROWADZENIE

Jednoczesna praca wszystkich warstw sterownika o rozproszonej architekturze prowadzi do powstania tzw. problemu wyboru akcji. Problem ten polega na nieustannym podejmowaniu decyzji: która z jednocześnie działających warstw (lub które z warstw) sterownika powinny w danym momencie mieć wpływ na zachowania sterowanego robota mobilnego. W dalszej części artykułu podane jest nowe rozwiązanie tego problemu z jednoczesnym osiągnięciem pewnych własności, istotnych dla prawidłowego i skutecznego działania takiego sterownika.

Klasyczne już przykłady zastosowań robotów sterowanych przez omawianą tu klasę sterowników stanowią następujące przypadki: inspekcja i ewentualna naprawa urządzeń wewnątrz elektrowni atomowych w rejonach podwyższonego skażenia, inspekcja i wykonywanie podwodnych prac na niebezpiecznych głębokościach i przez długie okresy czasu (np. prace geologiczne, poszukiwanie przedmiotów, badania kosmiczne: sondy międzyplanetarne (np. sondy Pioneer 10, Viking-1,2), badanie powierzchni innych ciał niebieskich (np. Łuńochod-1,2, Sojourner), prace podziemne w kopalniach, w tunelach, transport towarowy i osobowy (np. inteligentne pojazdy samochodowe, budownictwo (np. prace murarskie), transport wewnątrz nowoczesnych fabryk i zakładów, wspomaganie osób niepełnosprawnych (np. inteligentne wózki inwalidzkie, roboty wykonujące niektóre prace domowe), zastosowania w rolnictwie (np. oranie, sianie, opryskiwanie, koszenie, zbieranie owoców, proste prace publiczne (np. koszenie trawy, mycie peronów kolejowych, mycie zewnętrznych okien wieżowców), inspekcja i czyszczenie rur (np. czyszczenie rur kanalizacji pod autostradą), strzeżenie pomieszczeń przed kradzieżą, np. muzeów lub magazynów, zdalne

zwiedzanie muzeów (np. przez sterowanie robota przez Internet). Osobnym rodzajem zastosowań są roboty interwencyjno-inspekcyjne, jak np. SR11 INSPECTOR produkcji PIAP.

2. ZAŁOŻENIA WSTEPNE

Sterownik można uznać za działający prawidłowo jeżeli pozwala on na osiągnięcie koherentnego, racjonalnego i adaptacyjnego zachowania sterowanego robota mobilnego poruszającego się w dynamicznym, nieznanym otoczeniu [1].

Zachowanie koherentne można zdefiniować następująco: w przypadku konfliktów pomiędzy realizacją decyzji kilku warstw sterownika, robot mobilny nie będzie zachowywać się „w sposób niezdecydowany”, czyli nie będzie w zależności od nieznacząco zmiennych stanów otoczenia starać się realizować odmiennych zadań.

Występowanie zachowania racjonalnego [2] można wykazać dowodząc występowanie takich własności jak:

- *Kompatybilność* – robot mobilny nie może na skutek nieracjonalnego sterowania próbować wykonywania dwóch wzajemnie sprzecznych akcji.
- *Wspólna jednostka miary wagi wykonania danej czynności* - (ang. common currency): jeżeli dochodzi do konfliktu pomiędzy różnymi warstwami sterownika, czyli starają się one wykonać sprzeczne czynności (co prowadzi do utraty kompatybilności), to aby możliwy był racjonalny wybór którą z czynności wykonać, musi istnieć jakaś wspólna dla wszystkich warstw jednostka miary określająca jak wykonanie danej czynności jest aktualnie ważne.
- *Konsekwencja* - (ang. consistency): kwestia racjonalności zachowań robota mobilnego może być ujęta również następująco: jeżeli będąc dwukrotnie w tej samej sytuacji (rozumianej jako stan otoczenia i stan wewnętrzny sterownika) robot mobilny podejmie za każdym razem różne działania, to nie może być uznany za racjonalny.
- *Przechodność* - (ang. transitivity): wybór akcji jest dokonywany na podstawie wspólnej dla wszystkich warstw jednostki miary wagi wykonania danej czynności. Jeżeli wybrana zostanie czynność A, która zostanie uznana za ważniejszą niż czynność B (czyli: $A > B$) i jeżeli czynność B będzie uznana za ważniejszą od C (czyli: $B > C$), to w przypadku wyboru pomiędzy A i C musi być wybrana czynność A (czyli musi zająć: $A > B > C$).

Zachowanie koherentne jest warunkiem koniecznym dla zaistnienia zachowania racjonalnego.

Zachowanie adaptacyjne to zmiany zachowania stosownie do zmiennych stanów otoczenia, szybsze reagowanie sterownika przy powtarzających się stanach otoczenia. Adaptacji mogą podlegać warstwy sterujące lub sposób organizacji ich współpracy, czyli metoda wyboru akcji.

3. BUDOWA STEROWNIKA ROZPROSZONEGO DZIAŁAJĄCEGO WEDŁUG NOWEJ METODY WYBORU AKCJI

Proponowany sterownik jest rozproszony (ang. distributed), przez co rozumie się jednoczesną i asynchroniczną pracę wszystkich jego warstw.

Dane z receptorów są przetwarzane w modułach wstępnego przetwarzania danych (modułach WPD) stanowiących równocześnie pewnego rodzaju specjalizowane filtry. Następnie dane te są wykorzystywane przez poszczególne warstwy sterownika do generowania stosownych

sterowań, oraz używane przez tzw. neurony decyzyjne w procesie wyboru akcji. Moduły WPD zbierają przez receptory informacje zarówno o stanie otoczenia jak i o stanie wewnętrznym robota mobilnego. W niektórych przypadkach moduły WPD mogą wywierać na siebie wzajemny wpływ, np. tłumienie.

O tym, w jaki sposób stopnie swobody sterowanego robota mobilnego zostaną udostępnione warstwom sterownika, decyduje tzw. mechanizm decyzyjny.

Akcje podejmowane przez robota mobilnego wpływają na zmianę stanu jego otoczenia, stanowiąc rodzaj pętli sprzężenia zwrotnego dla sterownika i będąc tym samym jedynym torem przekazywania informacji pomiędzy warstwami.

Wymagane jest aby w każdym wybranym momencie możliwe było określenie stanu wewnętrznego robota mobilnego i stanu jego otoczenia, oraz stanów do których dążą poszczególne warstwy. Dzięki temu możliwe jest zdefiniowanie metryki opisującej odległość pomiędzy stanem aktualnym i pożądanym. Nie jest wymagane aby przestrzenie stanów były dyskretne (mogą być ciągłe).

Wyboru akcji dokonuje mechanizm decyzyjny. Składa się on z modułów WPD, neuronów decyzyjnych i arbitra. Arbitr jest to element sterownika umieszczony pomiędzy wszystkimi jego warstwami a efektorami sterowanego robota mobilnego. Jego zadaniem jest ostateczne przydzielanie warstwom sterownika dostępu do napędów poszczególnych stopni swobody sterowanego robota mobilnego. Każda taka warstwa realizuje swoje zadanie, inne niż pozostałe warstwy. Przy rozwiązywaniu problemu konstrukcji mechanizmu decyzyjnego wzorowano się na działaniu niektórych organizmów żywych. Wykorzystano elementy techniki sztucznych sieci neuronowych. Mechanizm decyzyjny zachowuje się jak sieć neuronowa o ilości tzw. neuronów decyzyjnych odpowiadającej ilości warstw w sterowniku. Każdy taki neuron posiada tylko jedno wejście i jedno wyjście. Sygnały na wejściach neuronów to sygnały pochodzące z odpowiadających warstwom modułów WPD. Sygnały na wyjściach neuronów decydują o tym która warstwa będzie w danym momencie sterować zachowaniem robota mobilnego. Sygnały wyjściowe są wynikiem iloczynu sygnałów wejściowych i odpowiadających im wag. Jako warstwa sterująca w danym momencie wybierana jest przez arbitra ta, o największym sygnale wyjściowym (pobudzeniu) przypisanego jej neuronu. W danym momencie jako warstwa działająca wybierana jest przez arbitra ta, która ma największe pobudzenie.

Wartości wag zmieniają się bezustannie. W danym momencie wynikają głównie z historii działania robota mobilnego lecz również w pewnym stopniu z bieżącej sytuacji. Wartości sygnałów z modułów WPD, głównie wynikają ze stanu otoczenia, są jednak również pośrednio uzależnione od historii poprzez wpływ na nie stanu wewnętrznego sterowanego robota mobilnego. Wagi te składają się z dwóch elementów: **część stała** - przypisana wstępnie danej warstwie, **część adaptacyjna** - wynik bezustannego procesu adaptacji mechanizmu decyzyjnego.

Moduły WPD przekazują do neuronów decyzyjnych sygnały, będące złożeniem sygnałów reprezentujących sytuację zewnętrzną, zwanych dalej **bodźcami zewnętrznymi** (odległości w przestrzeni stanów otoczenia pomiędzy bieżącymi stanami otoczenia, a stanami zamierzonym), oraz sygnałów reprezentujących sytuację wewnętrzną, zwanych dalej **imperatywami wewnętrznymi** (odległości w przestrzeni stanów wewnętrznych pomiędzy

bieżącymi stanami wewnętrznymi, a stanami do których zaistnienia dane warstwa sterująca dąży).

W części doświadczalnej pracy zbadano pięć metod uczenia neuronów decyzyjnych (zmian wartości części adaptacyjnych wag neuronów decyzyjnych). Metody te są opisane wzorami poniższymi wzorami [3, 4].

Metoda klasyczna Hebba

$$w_i(k) = w_i(k-1) + \eta x_i(k) y_i(k) \quad (3.1)$$

gdzie:

- $w_i(k)$ - nowa waga otrzymana w k -tym kroku uczenia
- k - krok uczenia numer k
- i - numer neuronu decyzyjnego
- $w_i(k-1)$ - poprzednia waga
- η - współczynnik uczenia
- x - sygnał wejściowy do neuronu decyzyjnego (sygnał z modułu WPD)
- y - pobudzenie neuronu ($y = x * w$)

Metoda Hebba z zapominaniem

$$w_i(k) = w_i(k-1) + \eta x_i(k) y_i(k) - \gamma w_i(k-1) y_i(k) \quad (3.2)$$

gdzie:

- γ - tzw. współczynnik zapominania

Metoda „instar”

$$w_i(k) = w_i(k-1) + \eta (x_i(k) - w_i(k-1)) \quad (3.3)$$

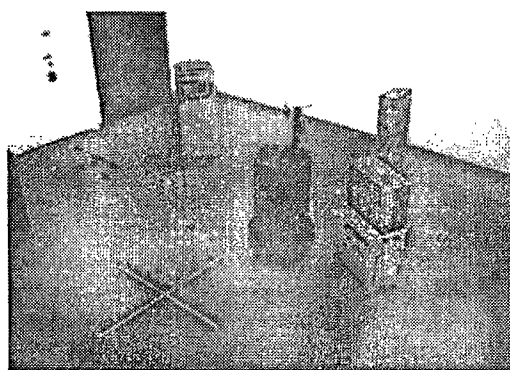
Metoda „outstar”

$$w_i(k) = w_i(k-1) + \eta (y_i(k) - w_i(k-1)) \quad (3.4)$$

Metoda Oji

$$w_i(k) = w_i(k-1) + \eta y_i(k) [x_i(k) - w_i(k-1) y_i(k)] \quad (3.5)$$

4. CZĘŚĆ EKSPERYMENTALNA



Rys. 4.1. Przykładowe doświadczenie na torze badawczym.

Wszystkie doświadczenia zostały przeprowadzone w laboratorium, z wykorzystaniem robota mobilnego Nomad 200 i programu sterownika do realizacji przykładowego zadania (rys. 4.1.). Przykładowe zadanie polega na pobieraniu „przedmiotów” z „magazynu” i sukcesywnym przewożeniu ich do „domu”. Jednocześnie robot musi unikać zderzeń z przeszkodami i dbać o właściwy stan naładowania swych akumulatorów. Jeśli akumulatory rozładują się do pewnego poziomu, robot powinien dotrzeć do „stacji ładowania akumulatorów”. Robot taki jako złożony system.

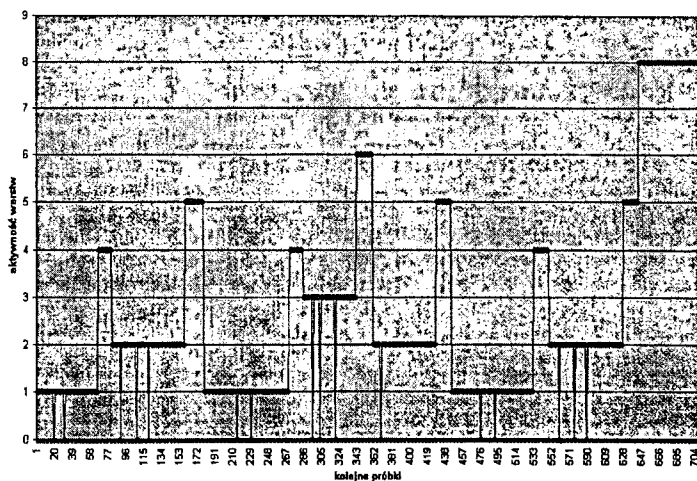
mechatroniczny „zużywa się”, więc powinien monitorować stopień zużycia ważniejszych części mechanicznych. Jeśli ten stopień zużycia staje się odpowiednio duży, robot powinien udać się do „serwisu” i tam wyłączyć się, aby umożliwić „przeгляд”. W tej sytuacji przykładowy sterownik składa się z dziewięciu warstw.

W ramach badań wykonano ponad 60 eksperymentów, zbierając około 65 Mb danych na różnych trasach robota, przy różnych ustawieniach parametrów sterownika, w różnych warunkach, przy różnych metodach uczenia mechanizmu decyzyjnego. Porównano działanie sterownika przy wykorzystaniu pięciu różnych metod uczenia. Otrzymane wyniki wszystkich eksperymentów wykorzystano do analizy zachowań robota mobilnego sterowanego za pomocą sterownika o architekturze rozproszonej z wykorzystaniem proponowanej metody wyboru akcji [1].

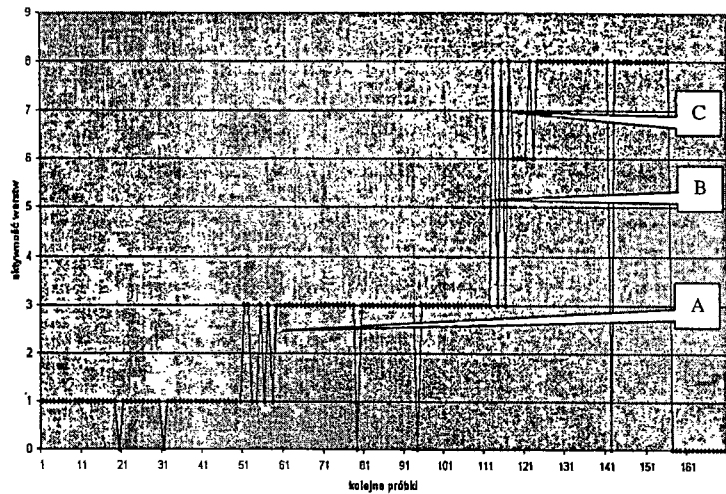
4.1. Koherencja

Koherencja proponowanego sterownika wyposażonego w mechanizm decyzyjny została udowodniona na gruncie doświadczeń z rzeczywistym robotem mobilnym jako obiektem sterowania. Koherencję można badać bezpośrednio przy użyciu tzw. wykresów aktywności warstw. Rys. 4.2. przedstawia przebieg doświadczenia w którym nie doszło do utraty koherencji, a rys. 4.3 – przypadek w którym doszło do utraty koherencji (na skutek celowych działań). Na tych wykresach oś pozioma odwzorowuje upływ czasu, a oś pionowa to numer warstwy aktualnie sterującej zachowaniem systemu autonomicznego.

Wszystkie doświadczenia, nie licząc tych w których celowo wprowadzono zaburzenia funkcjonowania mechanizmu decyzyjnego przez zmianę niektórych jego parametrów lub metody uczenia, pokazują że stworzona metoda wyboru akcji za pomocą mechanizmu decyzyjnego pozwala na osiągnięcie zachowania koherentnego. Osiągnięto zachowanie koherencji niezależnie od wykorzystanego toru doświadczalnego (na trasach różniących się rozmieszczeniem przeszkód i celów), od niektórych parametrów sterownika i od gwałtownych zmian stanu otoczenia.



Rys.4.2. Koherentna aktywność warstw (2 próbki ≈ 1 sekunda).



Rys.4.3. Utrata koherencji: punkty A, B i C (2 próbki \approx 1 sekunda).

4.2. Racjonalność

Racjonalne zachowanie sterownika jest częściowo zapewnione przez odpowiednie zaprojektowanie modułów WPD. Odpowiednia ich konstrukcja pozwala wstępnie ustalić które warstwy sterownika powinny zadziałać przy wystąpieniu danego stanu otoczenia oraz danego stanu wewnętrznego. Jednak o wyborze warstw sterujących decydują nie tylko moduły WPD, gdyż te mogą pobudzać do pracy wiele warstw w tym samym czasie. Wybór akcji wynika w głównej mierze z pracy mechanizmu decyzyjnego. Przy podejmowaniu decyzji o tym, która warstwa w danym momencie będzie sterować zachowaniem robota mobilnego, brane są pod uwagę zarówno okoliczności zewnętrzne (stan otoczenia), wewnętrzne (stan wewnętrzny), jak i dotychczasowa historia działań (wagi adaptacyjne). Zbudowany system sterowania należy więc do klasy tzw. systemów autonomicznych usytuowanych (uwarunkowanych, ang. situated embodied agents). Zachowanie takich systemów nie bazuje na żadnej rozbudowanej wewnętrznej reprezentacji stanu otoczenia, modelach tego otoczenia ani działaniach o naturze symbolicznej (jak np. w systemach ekspertowych). Główną rolę w działaniach takich systemów są uwarunkowania i powiązania typu sensoryczno-motorycznego, lecz nie są to jedynie systemy reaktywne (działające w wyniku współpracy prostych odruchów). Systemy usytuowane są adaptacyjne, posiadają pewną historię. W rezultacie, w identycznych sytuacjach (rozumianych jedynie jako stany otoczenia) nie muszą zachować się one zawsze tak samo. Ich zachowanie zależy od ich przeszłości i stanów wewnętrznych.

Analizując racjonalność zachowań robotów mobilnych trzeba pamiętać o zjawisku tzw. ograniczonego zachowania racjonalnego (ang. bounded rationality) [5]. Przykładem występowania tego zjawiska może być sytuacja, w której robot omija przeszkodę z wykorzystaniem zbyt dużej ilości manewrów omijania przeszkód. Możliwości utrzymania zachowania racjonalnego są ograniczone ze względu na:

- a) ograniczone możliwości przetwarzania danych (np. ograniczony rozmiar pamięci, ograniczona szybkość procesora komputerowego, itp.),

- b) ograniczone możliwości pozyskania danych o stanie otoczenia wynikające ze skończonej liczby receptorów (np. czujników odległości) o skończonych dokładnościach, zasięgach, itp.,
- c) ograniczoną wiedzę o regułach postępowania w danej sytuacji (przewidywanie konsekwencji swych działań na podstawie skończonej bazy wiedzy),
- d) ograniczone możliwości realizacji komend sterowania obiektu - nie jest możliwe sterowanie z błędem zerowym.

W badaniach eksperymentalnych spotkano się z nieracjonalnym zachowaniem robota. Nieracjonalność zachowania polegała na przykład na: pomijaniu wykonywania zadań do wykonania w danym celu, brak koherencji zachowań (nie można mówić o zachowaniu racjonalnym bez koherencji), omijaniu nieistniejących przeszkód, taranowaniu przeszkód, krążeniu tylko wokół jednego celu, realizacji tylko wybranej części zadań z całkowitym pominięciem pozostałych, zachowaniu wynikającym jedynie ze stanów wewnętrznych – bez uwzględnienia stanu otoczenia. Żadne z tych zjawisk nie występuje w przypadku sterownika z mechanizmem uczonego metodą „instar”, z odpowiednio dobranymi parametrami.

Racjonalność opisywanego tu sterownika należy rozważyć również w świetle przyjętej na wstępie definicji.

Kompatybilność: W związku z tym, że zachowanie robota mobilnego jest wynikiem wyboru jednej z warstw sterownika jako aktualnie kontrolującej jego ruchy, nie występuje tu możliwość zaistnienia konfliktu w postaci próby jednoczesnego wykonania dwóch wzajemnie sprzecznych czynności, np. jednoczesna jazda do przodu i do tyłu. Konflikty takie nie występują również w obrębie poszczególnych warstw. W rezultacie kompatybilność jest zawsze zachowana.

Wspólna jednostka miary wagi wykonania danej czynności: Wspólną jednostką miary są pobudzenia warstw sterownika, które są następnie porównywane przez arbitra. Sygnały te są jednoznacznie zdefiniowane i w rezultacie postulat ten jest spełniony.

Konsekwencja: Konsekwencja działań sterownika jest tu zawsze zachowana. Wszystkie zachowania robota mobilnego są rezultatem sterowania wynikającego jedynie ze stanów wewnętrznych (wliczając w to wagi adaptacyjne) i stanów otoczenia robota. Na przebieg sterowania nie mają wpływu żadne inne czynniki. W rezultacie zawsze w przypadku takich samych stanów otoczenia i wewnętrznych sterownika, zachowanie robota mobilnego jest takie samo.

Przechodność: Pobudzenia warstw sterownika są wyrażone liczbowo. Każdą taką wartość można porównywać z inną, zgodnie z podstawowymi regułami matematyki. W ten sposób postulat przechodności jest spełniony z definicji.

Robot mobilny wyposażony w sterownik rozproszony z proponowanym mechanizmem wyboru akcji jest więc racjonalny.

4.3. Adaptacja

W związku z istnieniem mechanizmu decyzyjnego, którego działanie opiera się m.in. na dynamicznym przydzielaniu warstwom wag decyzyjnych, opisywany sterownik wykazuje cechy adaptacyjne. Dowodem tego są wyniki badań przeprowadzonych eksperymentalnych. Pokazują one, że sterownik dostosowuje się do zmiennych stanów otoczenia i jest w stanie - bez żadnych regulacji ze strony człowieka - sterować ruchem robota na różnych trasach.

Dodatkowym przejawem adaptacji jest szybsze reagowanie sterownika przy powtarzających się stanach otoczenia. W przykładowym zadaniu wynika to z budowy modułu WPD warstwy omijania przeszkód. Imperatyw wewnętrzny i_{weW} zawiera zmienną pomocniczą która opisuje częstotliwość zderzeń z przeszkodami. Im są one częstsze, tym większa jest wartość i_{weW} w razie wykrycia przeszkody w strefie bezpieczeństwa. Kiedy i_{weW} rośnie, rośnie też sygnał wejściowy do neuronu decyzyjnego, a w rezultacie szybciej rośnie waga adaptacyjna warstwy omijania przeszkód. Im waga ta jest większa, tym szybciej rośnie pobudzenie tej warstwy (jest to przykład zachowania autokatalitycznego). W rezultacie warstwa omijania przeszkód stanie się szybciej warstwą wybraną przez arbitra.

5. WNIOSKI

Jednym z ciekawszych wniosków tej pracy jest spostrzeżenie, iż przy budowaniu coraz bardziej złożonych systemów sterujących zaciera się (wydawałoby się wyraźna) granica pomiędzy architekturą hierarchiczną, a architekturą rozproszoną (równoległą). Istnieje punkt widzenia, z którego każda z warstw sterownika (składająca się m.in. z czujników, modułu WPD, warstwy sterującej, niskiego poziomu sterowników efektorów) jest strukturą hierarchiczną. W końcowym efekcie można na całość spojrzeć jak na strukturę hybrydową – rozproszony system złożony z podsystemów o charakterze hierarchicznym.

Dalsze prace nad zaproponowanym rozwiązaniem wyboru akcji w sterownikach rozproszonych mogą iść w kilku kierunkach. Jednym z nich mogą być próby zastosowania techniki algorytmów genetycznych. W szczególności technika ta mogłaby znaleźć zastosowanie w doborze parametrów sterownika. Dobór tych parametrów, oddzielnie dla każdej warstwy, można by przeprowadzić za pomocą metod algorytmów genetycznych. Ciekawym jawi się także zbadanie możliwości struktury zagnieżdżonej. Wtedy grupy warstw np. prostych odruchów, składających się na złożone zachowania, wyposażone mogłyby być w swoje oddzielne mechanizmy decyzyjne. Na drugim, wyższym poziomie organizacji – złożone zachowania rywalizowałyby pomiędzy sobą, również za pomocą mechanizmu decyzyjnego.

Najważniejszym krokiem, który może być wykonany w przyszłych pracach, może być zbadanie działania proponowanego mechanizmu wyboru akcji zaimplementowanego na komputerach o przetwarzaniu równoległym. W omawianych badaniach równoległa praca warstw sterownika była jedynie symulowana.

LITERATURA

- [1] Szykarczyk P., „Zastosowanie równoległej architektury warstwowej i problem wyboru akcji w sterowaniu systemów autonomicznych”, wyd. Wojskowej Akademii Technicznej, Warszawa, 1999, rozprawa doktorska
- [2] McFarland D., Bösser Th., „*Intelligent Behavior in Animals and Robots*”, A Bradford Book, The MIT Press, Cambridge, Massachusetts, London, England, 1993
- [3] Osowski S., „*Sieci neuronowe w ujęciu algorytmicznym*”, Wydawnictwa Naukowo-Techniczne, Warszawa, 1996
- [4] Korbicz J., Obuchowicz A., Uciński D., „*Sztuczne sieci neuronowe, podstawy i zastosowania*”, Akademicka Oficyna Wydawnicza PLJ, wyd. I, Warszawa 1994
- [5] Simon, H.A., „*Bounded rationality*”, in: J. Eatwell, M. Millgate & P. Newman (eds.): *The New Palgrave: A Dictionary of Economics*. London and Basingstoke: Macmillan