

## Streszczenie

Rozprawa doktorska koncentruje się na wykorzystaniu *antycypacyjnych uczących się systemów klasyfikujących ACS* (ang. *Anticipatory Classifier Systems*) w problemach, w których stan przedstawiany jest z wykorzystaniem liczb rzeczywistych. Wiele podobnych badań zostało dotychczas wykonanych dla innych uczących się systemów klasyfikujących LCS (ang. *Learning Classifier Systems*), jednak żaden z nich nie posiadał możliwości budowania wewnętrznej reprezentacji środowiska z użyciem mechanizmu antycypacji. Celem pracy jest przedstawienie możliwych realizacji powyższej integracji z użyciem wybranych testowych problemów. W oparciu o te założenia, sformułowana została hipoteza, zakładająca, że

*Antycypacyjne uczące się systemy klasyfikujące są w stanie poprawnie zbudować wewnętrzną reprezentację modelu rzeczywistoliczbowego środowiska.*

Hipoteza została uprawdopodobniona poprzez osiągnięcie poniższych celów:

**Cel 1 - Zapropowanie modyfikacji wybranych systemów ACS umożliwiającą obsługę danych wejściowych o wartościach rzeczywistych**

Cel został osiągnięty i zweryfikowany na dwa różne sposoby:

1. użycie dozwolonych *przedziałów wartości* jako podstawowej jednostki służącej do reprezentacji danych - zapropowanie autorskiej wersji systemu rvACS (ang. *real-valued Anticipatory Classifier System*),
2. dyskretyzację danych wejściowych - zapropowanie modyfikacji wybranych algorytmów ACS.

Pierwszy sposób inspirowany jest dotychczasowymi osiągnięciami uzyskanymi dla innych systemów uczących. Nowy algorytm rvACS osiąga zadowalające wyniki, jednak poziom złożoności utworzonego rozwiązania, jak i zawilóść interakcji jego wewnętrznych komponentów jest zbyt wysoki.

Drugi sposób pozwala na uniknięcie znaczących modyfikacji w sposobie działania trzech testowanych systemów. Transformacja sygnału wyjściowego środowiska zanim zostanie on przetworzony przez algorytmy uczące, umożliwia zachowanie ich charakterystycznych cech. Dodatkowo, otrzymane rezultaty badań są bardziej obiecujące niż w przypadku zastosowania reprezentacji z wykorzystaniem *przedziałów wartości*.

## **Cel 2 - Zaproponowanie środowisk testowych i metryk do oszacowania wydajności systemów ACS**

Charakterystyka działania algorytmów antycypujących została poddana szczegółowej analizie wykorzystując sześć testowych problemów opisujących aktualny stan za pomocą wartości rzeczywistoliczbowych. Część z nich była już wykorzystywana wcześniej w literaturze podczas badań uczących się systemów klasyfikujących.

Każde z badanych środowisk wykazuje charakterystyczne własności, np:

- problem uczenia nadzorowanego (jednokrokowy) lub uczenia ze wzmocnieniem (wielokrokowy),
- wzajemna zależność poszczególnych atrybutów percepcji,
- rozmiar przestrzeni dostępnych stanów,
- wymagana liczba akcji w celu osiągnięcia nagrody.

Dodatkowo, po raz pierwszy w historii znany problem uczenia ze wzmocnieniem polegający na utrzymywaniu równowagi pręta umieszczonego na wózku został przetestowany przez systemy uczące. Otrzymane wyniki są szczególnie interesujące pod względem wydajności i minimalistycznej postaci otrzymanego rozwiązania.

Działania poszczególnych algorytmów zostały oceniane przez pięć metryk badających różne aspekty działania. Uwagę zwrócono na jakość generowanego rozwiązania, jego rozmiar oraz praktyczne wykorzystanie wiedzy w badanym środowisku.

## **Cel 3 - Zaproponowanie optymalizacji działania systemów ACS w kontekście przetwarzania wejścia rzeczywistoliczbowego**

Wstępne badania wskazały na dwa obszary mające potencjał optymalizacji działania systemów antycypacyjnych w przypadku interakcji z rzeczywistoliczbowymi środowiskami.

Pierwszy, zauważa, że zastosowana reprezentacja wejścia znacząco zwiększa dostępną przestrzeń stanów, co ma znaczący wpływ na działanie każdego badanego algorytmu. Przetestowane zostały cztery strategie, skupiające się na inteligentnym przeszukiwaniu

przestrzeni rozwiązań optymalizując tworzenie wewnętrznego modelu środowiska. Zaproponowane zostało także autorskie rozwiązanie, polegające na optymistycznym generowaniu nowych reguł.

Drugie ograniczenie dotyczy środowisk wielokrokowych, wymagających wykonanie wielu akcji w celu osiągnięcia finalnego stanu. Zaproponowano wymianę komponentu odpowiedzialnego za dystrybucję sygnału nagrody na uśrednioną wersję. Przedstawiono autorską wersję systemu - AACCS2 (ang. *Average Anticipatory Classifier System 2*), wykazującą wzrost wydajności w testowanych problemach.

#### **Cel 4 - Eksperymentalna analiza przeprowadzonych badań**

W celu zrozumienia natury przeprowadzanych eksperymentów, w każdym przypadku uwzględniany został aspekt czasowy ilustrujący sposób ewolucji wybranych metryk w całym procesie uczenia.

Dodatkowo, porównania pomiędzy poszczególnymi algorytmami bądź wersjami środowisk zostały wykonane z wykorzystaniem technik modelowania probabilistycznego.

#### **Cel 5 - Stworzenie ogólnodostępnej biblioteki programistycznej**

Efektom tezy jest ogólnodostępna biblioteka programistyczna w języku Python implementująca szereg systemów uczących (np. ACS, ACS2, YACS, MACS, X-NCS). Znaczącym osiągnięciem jest pełne odtworzenie dotychczasowych wyników badań wraz z dodaniem możliwości obsługi liczb rzeczywistych. Zaprojektowana ona została zgodnie z aktualnymi trendami, umożliwiając szybką integrację i porównanie się z innymi rodzinami algorytmów.

#### **Słowa kluczowe**

uczące się systemy klasyfikujące; antycypacyjne uczące się systemy klasyfikujące; uczenie ze wzmocnieniem; openai gym, wejście rzeczywistoliczbowe

# Abstract

The thesis focuses on the use of ACSs (*Anticipatory Classifier Systems*) with the input deliberately expressed as real-valued number. Other classifier system families thoroughly investigated such a research area, but no significant contribution was made to systems with explicit prediction capabilities. The aim was to present possible solutions to such integration using several benchmarking problems. Based on this assumption, the following thesis was formulated:

*Anticipatory Classifiers Systems can build the correct internal model of the environment using the real-valued input.*

The thesis was substantiated by achieving the following goals:

## **Goal 1 - Propose modifications towards ACS system capable of handling real-valued input**

This intention was achieved and validated using two independent approaches:

1. changing the rule's attribute representations to incorporate *interval predicates*, resulting in a proprietary system variation named rvACS (*real-valued Anticipatory Classifier System*),
2. implementing the prominent ACS algorithms with the intention of input discretization.

The first approach was based on the actual advancements made for the other systems. The rvACS managed to show promising results. Nevertheless, because of the much richer rule complexity and the cooperation of components, the obtained conclusion is that the nature of rvACS is not aligned with the overall ACS virtues.

The second approach maintains the characteristics and original intentions of three investigated algorithms by affecting only the agent-environment interface layer for discretizing

the input signal. This approach was considered superior to interval formation by obtaining promising results and better rule interpretability.

### **Goal 2 - Propose relevant benchmarking problems and metrics for evaluating ACS performance**

The nature of the performance was carefully reviewed in six problems using the real-valued output as a description of an actual state. Most of them have been used as a toy-problems in LCS (*Learning Classifier Systems*) research, providing different problems of different natures, like:

- single and multiple steps,
- extensive mutual feature interaction,
- vast input space,
- the need for long-action chain building.

Moreover, a famous RL (Reinforcement Learning) *Cart Pole* benchmarking problem was investigated by the LCS for the first time, to our knowledge, obtaining encouraging results with compact knowledge representation.

To highlight performance, five key metrics were selected that investigate the state of the systems from multiple angles. Aspects like the quality of evolving solutions, size, and effective application are emphasized throughout the research.

### **Goal 3 - Propose relevant improvements towards neoteric changes**

Preliminary tests revealed two possible limitations of the system that were further studied.

First, any form of real-valued representation increased the problem's search space, impeding the agent's learning speed. Four techniques probing for the most promising action selection were inspected, especially regarding knowledge acquisition rate. The novel approach, named *Optimistic Initial Quality* was also proposed herein.

The other limitation relates to multistep problems, where certain input discretization might demand the agent to perform a notably larger number of actions to receive the feedback signal. An approach replacing the credit assignment with the undiscounted incentive distribution version resulted in a system named AACSS2 (*Average Anticipatory Classifier System 2*). The method showed performance improvements for specific problems.

**Goal 4 - Conduct an experimental evaluation of intended adjustments**

All the experiments were tried to illustrate the evolution of selected metrics throughout the agent's learning of selected problems.

Moreover, the Bayesian estimation approach emphasized the differences between investigated cases, drawing non-biased conclusions. Often, selected environments were scaled up, where agents' behaviour was challenged with increased complexity.

**Goal 5 - Developing an open-sourced Python Machine Learning framework for evaluating various LCS algorithms**

Significant effort was put to implement the most famous classifier systems algorithms (i.e. ACS, ACS2, YACS, MACS, X-NCS) using the tools used nowadays by data scientist communities. The original performance reported by authors is preserved while adding full support for the real-valued representation of perception. By integrating with the industry standards, all evaluated models share the same code base and operate with a commonly agreed interface enabling effortless benchmarking with other state-of-the-art systems.

Moreover, a conscious endeavour was put to simplify the usage of the developed framework to spark other researchers' interest.

**Keywords**

learning classifier systems; anticipatory classifier systems; reinforcement learning; openai gym, real-valued input