

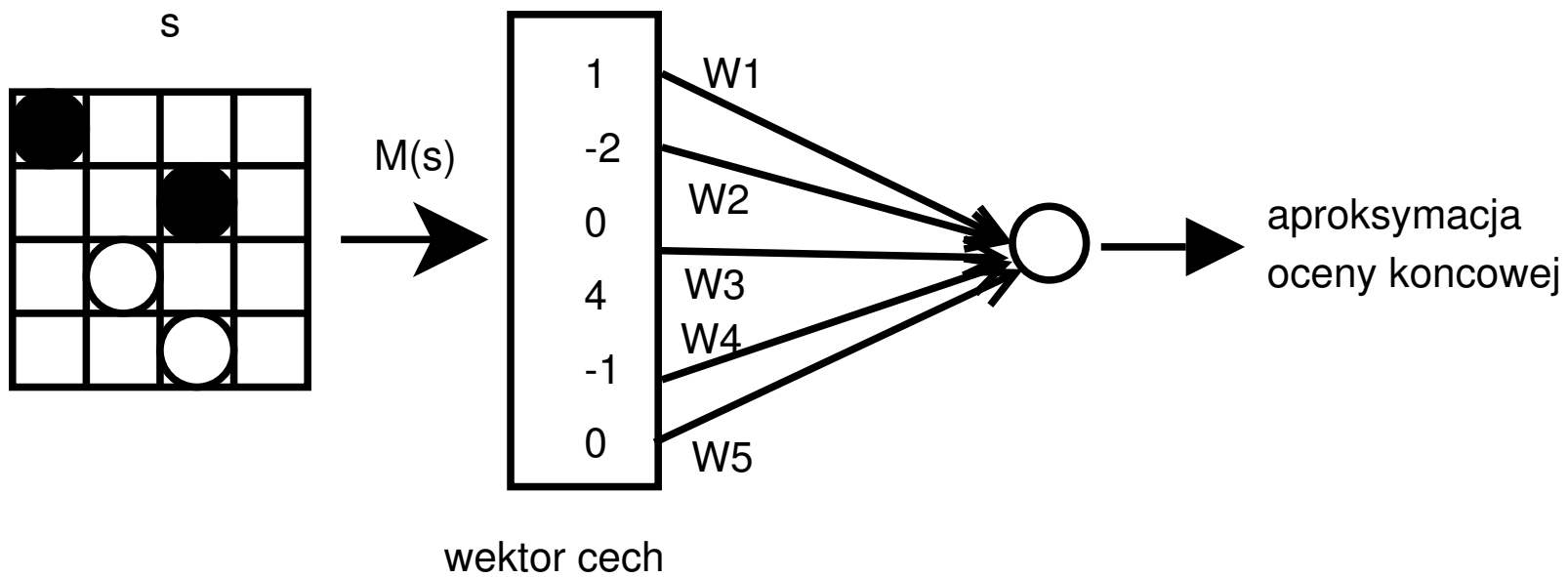
# **Jak nauczyć komputer grać w anty-warcaby**

Daniel Osman

# Spis treści

- **Algorytm Temporal Difference learning**
  - ▷ przypomnienie
  - ▷ funkcja oceny
- **Opis procesu uczenia**
  - ▷ różne podejścia
  - ▷ różne odmiany algorytmu
  - ▷ porównanie z metodą pseudo-ewolucyjną
  - ▷ wyniki

# Schemat



# Proces uczenia. Etap 1

- 10 graczy uczących się
- Wagi zainicjalizowane wartościami losowymi
- 10.000 gier treningowych na 25 losowych przeciwnikach

# Proces uczenia. Etap 2

- 10 graczy uczących się
- Wagi uzyskane z pierwszego etapu
- 10.000 gier treningowych na 25 silniejszych przeciwnikach

# Różne podejścia

**LB** - uczenie następuje po każdej grze

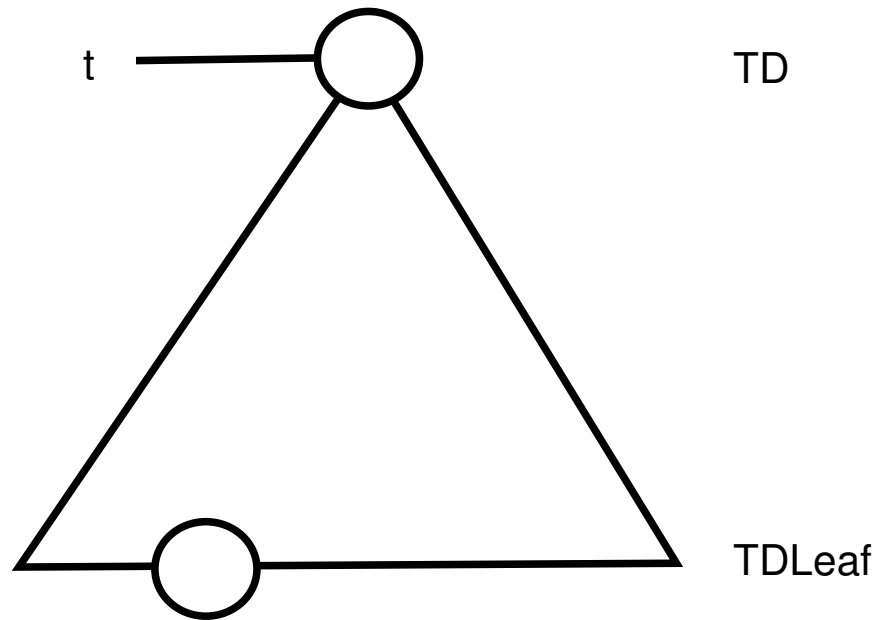
**LL** - uczenie tylko po grach przegranych albo zremisowanych

**L3** - LL + koncentrowanie się na silniejszych przeciwnikach

# Faza testująca

- Co 250 gier treningowych gracz uczący się rozgrywa 100 gier z przeciwnikami testującymi
- 2 grupy przeciwników testujących
  - ▷ losowi
  - ▷ silni

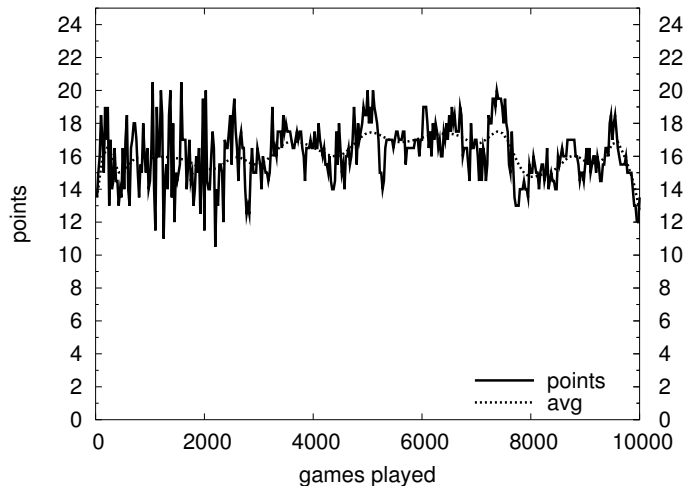
# Algorytm TDLeaf



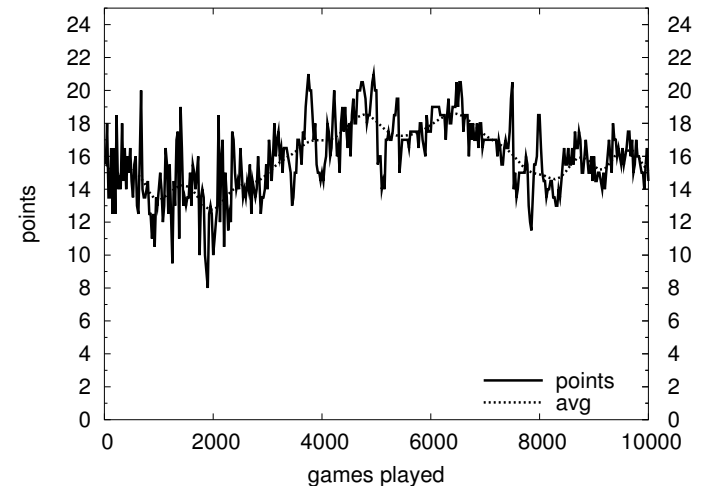


# Algorytm pseudo-ewolucyjny

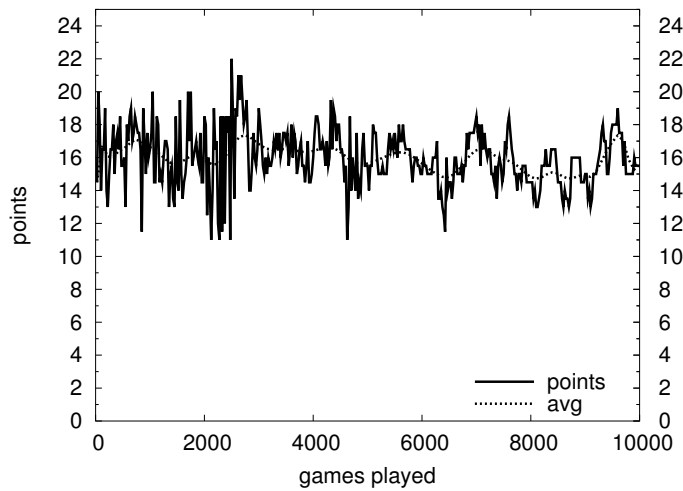
- 1 przeciwnik
  - ▷ mutacja przeciwnika raz na dwie gry
  - ▷ po każdej grze gracze zamieniają się miejscami
- Krzyżowanie
  - ▷ jeżeli przeciwnik wygrywa obie gry albo wygrywa i remisuje to wagi gracza uczącego się przesuwają się o 5% w stronę wektora wag przeciwnika



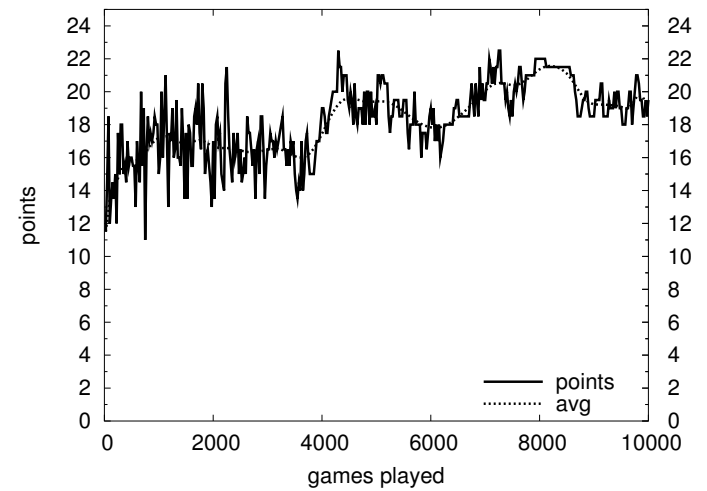
(a) gracz 1



(b) gracz 3

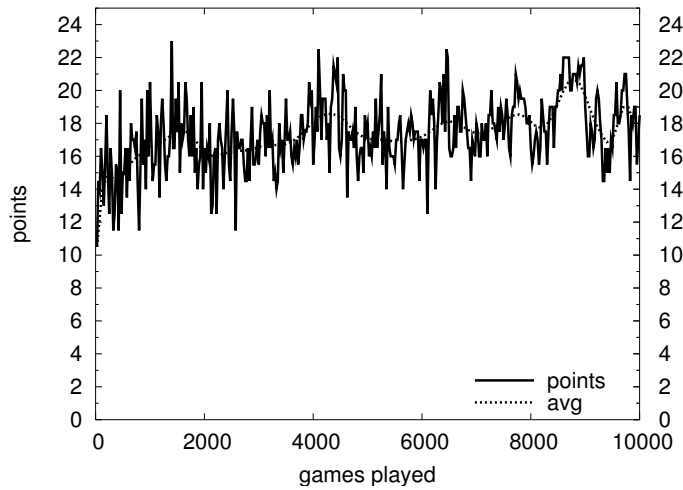


(c) gracz 5

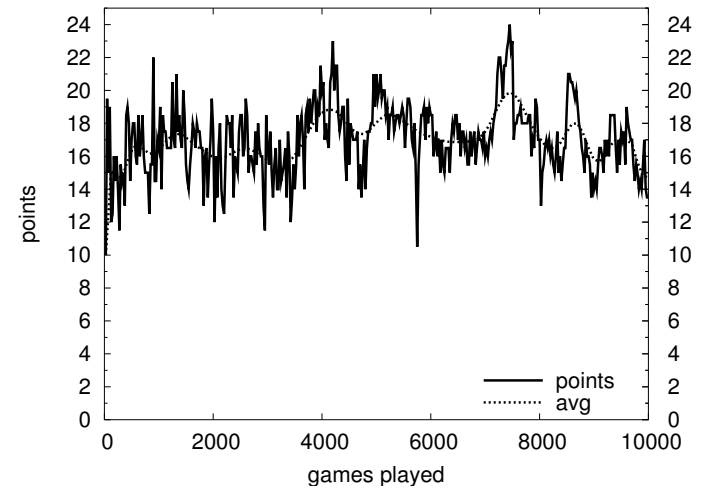


(d) gracz 7

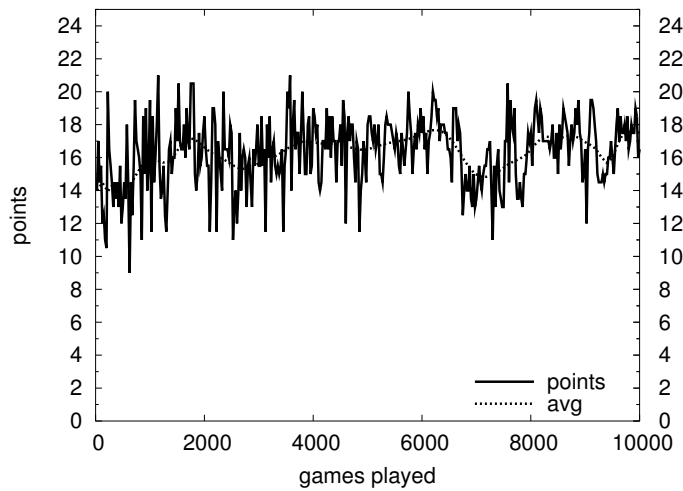
Postępy w treningu. Etap 1. LB



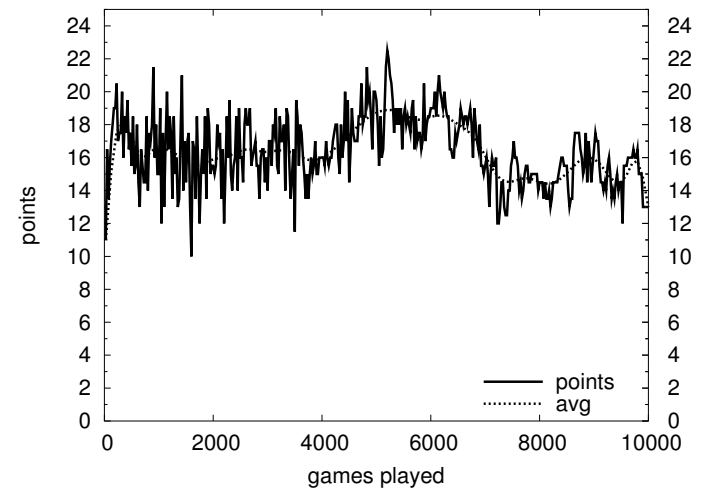
(e) gracz 1



(f) gracz 3



(g) gracz 5



(h) gracz 7

Postępy w treningu. Etap 1. LL

## Trening. Etap 1.

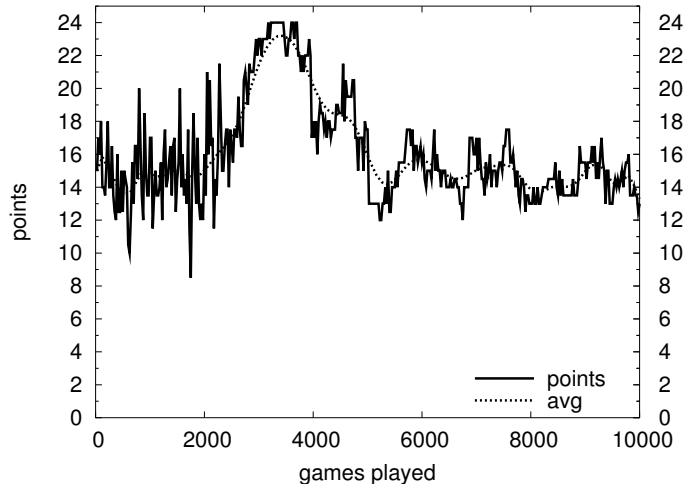
gry	LL	LB	EVO
1 - 2,500	64.9%	64.7%	64.5%
2,501 - 5,000	70.5%	70.2%	63.5%
5,001 - 7,500	72.0%	72.2%	64.8%
7,501 - 10,000	70.8%	69.1%	63.1%

## Faza testująca - losowi przeciwnicy. Etap 1.

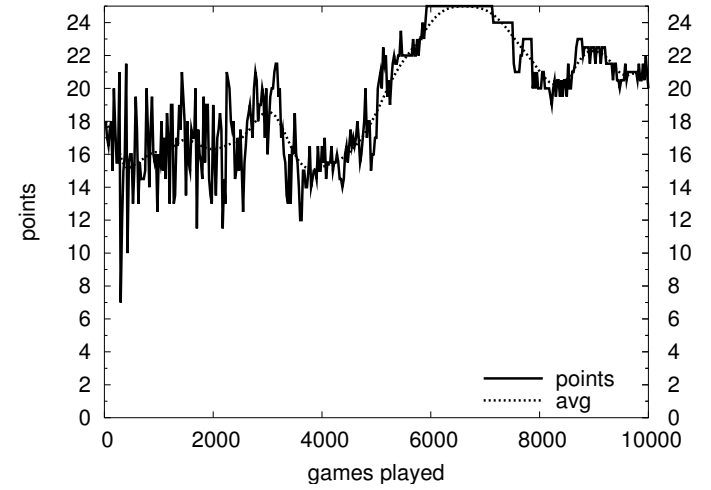
gry	LL	LB	EVO
1 - 2,500	66.8%	69.9%	51.7%
2,501 - 5,000	70.1%	70.9%	55.7%
5,001 - 7,500	71.7%	70.0%	57.2%
7,501 - 10,000	70.7%	70.7%	59.4%

## Faza testująca - silni przeciwnicy. Etap 1.

gry	LL	LB	EVO
1 - 2,500	44.2%	43.9%	30.1%
2,501 - 5,000	51.6%	43.2%	32.9%
5,001 - 7,500	49.2%	43.5%	35.3%
7,501 - 10,000	50.0%	44.5%	36.8%



(i) TDLeaf. Gracz 8 LL



(j) TD. Gracz 1 LL

Najlepsi gracze TDLeaf i TD

## Trening. Etap 2.

gry	TDLeaf+LL	TDLeaf+LB	TD+LL	TD+LB	TD+L3	EVO
1 - 2,500	61.3%	59.0%	61.8%	52.9%	56.7%	65.3%
2,501 - 5,000	66.3%	65.6%	64.3%	54.9%	45.5%	65.7%
5,001 - 7,500	60.1%	63.3%	71.9%	51.8%	40.8%	62.2%
7,501 - 10,000	60.1%	64.6%	62.8%	52.3%	37.5%	70.1%

## Faza testująca - losowi przeciwnicy. Etap 2.

gry	TDLeaf+LL	TDLeaf+LB	TD+LL	TD+LB	TD+L3	EVO
1 - 2,500	71.8%	71.9%	72.0%	69.5%	73.5%	
2,501 - 5,000	70.6%	72.5%	74.4%	71.2%	74.9%	
5,001 - 7,500	70.4%	72.4%	72.5%	72.1%	72.8%	
7,501 - 10,000	70.2%	71.4%	73.1%	73.2%	74.4%	

## Faza testująca - silni przeciwnicy. Etap 2.

gry	TDLeaf+LL	TDLeaf+LB	TD+LL	TD+LB	TD+L3	EVO
1 - 2,500	51.7%	50.4%	53.3%	43.3%	53.3%	
2,501 - 5,000	52.2%	55.0%	56.7%	44.8%	57.4%	
5,001 - 7,500	50.6%	56.1%	54.0%	47.9%	56.7%	
7,501 - 10,000	50.8%	56.0%	54.8%	48.2%	56.1%	

# Pytania