

ŽILINSKÁ UNIVERZITA V ŽILINE  
FAKULTA RIADENIA A INFORMATIKY

AUTOREFERÁT DIZERTAČNEJ PRÁCE  
Študijný odbor: **Aplikovaná informatika**

Žilina, Apríl, 2016

autor : Ing. Michal Chovanec  
vedúci: prof. Ing. Juraj Miček, PhD

ŽILINSKÁ UNIVERZITA V ŽILINE  
FAKULTA RIADENIA A INFORMATIKY

**Ing. Michal Chovanec**

Autoreferát dizertačnej práce  
Aproximácia funkcie ohodnotení v algoritmoch Q-learning neurónovou sieťou

na získanie akademického titulu philosophiae doctor (v skratke PhD.)  
v študijnom programe doktorandského štúdia  
aplikovaná informatika  
v študijnom odbore:  
9.2.9 aplikovaná informatika

Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia forme doktorandského štúdia na katedre technickej kybernetiky, Fakulte riadenia a informatiky Žilinskej univerzity v Žiline

Predkladateľ:  
Ing. Michal Chovanec  
Katedra technickej kybernetiky  
Fakulta riadenia a informatiky  
Žilinská univerzita v Žiline

Školiteľ:  
prof. Ing. Juraj Miček, PhD  
Katedra technickej kybernetiky  
Fakulta riadenia a informatiky  
Žilinská univerzita v Žiline

Oponenti:

Titul, meno a priezvisko :  
Názov pracoviska :

Titul, meno a priezvisko :  
Názov pracoviska :

Titul, meno a priezvisko :  
Názov pracoviska :

Autoreferát bol rozoslaný dňa: .....

Obhajoba dizertačnej práce sa koná dňa ..... o ..... h. pred komisiou pre obhajobu dizertačnej práce schválenu odborovou komisiou v študijnom odbore 9.2.9 aplikovaná informatika, v študijnom programe aplikovaná informatika, vymenovanou dekanom Fakulty riadenia a informatiky Žilinskej univerzity v Žiline dňa .....

prof. Ing. Martin Klimo, PhD. predseda odborovej komisie študijného programu aplikovaná informatika v študijnom odbore 9.2.9 aplikovaná informatika Fakulta riadenia a informatiky Žilinská univerzita Univerzitná 8215/1 010 26 Žilina

## 0.1 Abstrakt

Práca sa zaoberá aproximáciou funkcie ohodnotení konania agenta, v algoritmoch Q-learning. V priestoroch s malým počtom stavov predstavuje vhodné riešenie tabuľka. Pre prípady veľkého počtu stavov je tabuľkové riešenie ťažko vypočítateľné. Je tak nutné použiť aproximáciu. Vhodným kandidátom je neurónová sieť. Tradičné riešenie doprednej siete je však nepoužiteľné z dôvodov nemožnosti takúto sieť učiť. V práci je preto venovaný priestor neurónovej sieti bázických funkcií ktorú už je možné na daný problém trénovať iteráčnými metódami.

## 0.2 Použité metódy

Niektoré tvary bázických funkcií ktoré možno uvažovať pre problém aproximácie

$$f_j^1(s(n), a(n)) = e^{-\sum_{i=1}^{n_s} \beta_{aji}(n)(s_i(n) - \alpha_{aji}(n))^2} \quad (1)$$

$$f_j^2(s(n), a(n)) = \frac{1}{1 + \sum_{i=1}^{n_s} \beta_{aji}(n)(s_i(n) - \alpha_{aji}(n))^2} \quad (2)$$

$$f_j^3(s(n), a(n)) = e^{-\sum_{i=1}^{n_s} \beta_{aji}(n)|s_i(n) - \alpha_{aji}(n)|} \quad (3)$$

kde

$\alpha_{aji}(n) \in \langle -1, 1 \rangle$  určuje polohu maxima funkcie

$\beta_{aji}(n) \in (0, \infty)$  určuje strmosť funkcie.

Aproximovaná funkcia ohodnotení pre  $l$  bázických funkcií je potom

$$Q^x(s(n), a(n)) = \sum_{j=1}^l w(n)_j^x f_j^x(s(n), a(n)) \quad (4)$$

kde  $w(n)_j^x$  sú váhy bázických funkcií.

V práci je ďalej definovaná nová bázická funkcia ktorá kombinuje niekoľko rôznych funkcií, čo vedie na vzťahy

$$P_i(s(n), a(n)) = \begin{cases} r_{ai} & \text{if } s(n) = \alpha_i^1 \\ 0 & \text{inak} \end{cases} \quad (5)$$

$$H_j(s(n), a(n)) = w_{aj} e^{-\beta_{aj} \sum_{i=1}^{n_s} (s_i(n) - \alpha_{aji}^2)^2} \quad (6)$$

$$Q(s(n), a(n)) = \sum_{i=1}^I P_i(s(n), a(n)) + \sum_{j=1}^J H_j(s(n), a(n)) \quad (7)$$

kde

$\alpha_i^1$  sú oblasti kde  $H_j(s(n))$  nadobúda nenulové hodnoty

$\alpha_j^2$  sú oblasti pre ktoré  $f_j(s(n), a(n))$  nadobúda maximum

$r_{ai}$  je hodnota zápornej odmeny  $R(s(n), a(n))$

$w_{aj}$  je váha a zobovedá veľkosti maxima resp. minima pre fukciu

$\beta_{aj}$  je strmosť, a platí  $\beta > 0$

$I$  a  $J$  sú počty bázických funkcií

Označenia  $P$  a  $H$  vznikli z tvaru funkcií : peak a hill. Funkcia bude na ďalších grafoch označená ako Gauss + AT : kombinácia Gaussovej krivky a adaptívnej tabuľky. Mechanizmus učenia zostáva rovnaký ako pre bázické funkcie v predošlej časti. Počet funkcií  $P_i(s(n), a(n))$  bol zvolený 30 a počet funkcií  $H_j(s(n), a(n))$  20. Pre názornosť boli parametre  $r_{ai}$  zvolené záporné a parametre  $\beta_{aj}$  kladné.

Funkcia predstavuje nový tvar bázických funkcií pre aproximovanie funkcie ohodnotení  $Q(s(n), a(n))$ .

## 0.3 Experimentálne výsledky

Úplne znenie práce a cca 15000 súborov s výsledkami je možné nájsť v [2]. Video a zdrojové súbory doplnkového experimentu s robotom je možné nájsť v [1] a [3].

# Literatúra

- [1] Michal Chovanec, Q-learning zdrojové súbory [https://github.com/michalnand/q\\_learning](https://github.com/michalnand/q_learning)
- [2] video robota Motoko Aftermath Michal Chovanec, youtube [https://www.youtube.com/watch?v=8sskJN\\_zuko](https://www.youtube.com/watch?v=8sskJN_zuko)
- [3] Michal Chovanec, Motoko robot zdrojové súbory [https://github.com/michalnand/motoko\\_after\\_math\\_linefollower](https://github.com/michalnand/motoko_after_math_linefollower)