

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
Fakulta strojního inženýrství
Ústav automatizace a informatiky

Ing. Vít Ondroušek

**VYUŽITÍ OPAKOVANĚ POSILOVANÉHO UČENÍ
PRO ŘÍZENÍ ČTYŘNOHÉHO ROBOTU**

USING OF REINFORCEMENT LEARNING FOR FOUR
LEGGED ROBOT CONTROL

Zkrácená verze Ph.D. Thesis

Obor: Inženýrská mechanika
Školitel: doc. RNDr. Ing. Tomáš Březina, CSc.

KLÍČOVÁ SLOVA

kráčivý robot, kráčivá chůze, Q-učení, posilované učení

KEYWORDS

walking robot, walking gait, Q-learning, reinforcement learning

OBSAH

1	ÚVOD	5
1.1	Zadání a zaměření práce	5
1.2	Stanovené cíle	5
2	TEORETICKÝ ROZBOR PROBLEMATIKY	5
2.1	Kráčivý pohyb robotů	5
2.2	Přístupy k plánování chůze	6
2.2.1	Přístup založený na modelu	6
2.2.2	Přístup založený na chování	6
2.2.3	Hybridní a ad-hoc přístupy	7
2.3	Architektura řízení využívající DEDS pro plánování chůze	7
2.3.1	Vrstva zpětnovazebných řídicích členů	7
2.3.2	Vrstva zahrnující koordinační mechanismus DEDS	8
2.3.3	Vrstva učení	9
2.4	Postup řešení	10
3	PRAKTICKÁ ČÁST – SOUHRN DOSAŽENÝCH VÝSLEDKŮ	11
3.1	Volba prostředků pro verifikaci	11
3.2	Platforma použitá pro verifikaci návrhu řešení	11
3.3	Sestavení elementárních řídicích členů	12
3.3.1	Řídicí člen postoje	12
3.3.2	Řídicí člen kontaktu	13
3.4	Projekt implementující navržené řešení	14
3.5	Výsledná chůze robotu při komplexních úlohách	15
3.5.1	Úloha rotace robotu	15
3.5.2	Úloha dopředné chůze	17
3.5.3	Úloha dopředné chůze na nakloněné rovině	18
4	ZÁVĚR	19
4.1	Přínos práce	20
4.1.1	Přínos teoretický	20
4.1.2	Přínos praktický	20
4.1.3	Přínos pedagogický	20
	LITERATURA	21
	AUTOROVO CV	25
	ABSTRACT	26

1 ÚVOD

Problematika autonomních mobilních robotů v současné době stále spadá do popředí zájmu vědeckého výzkumu mobilní robotiky. Dle (Ehrenberger, 1993) jsou autonomní roboty třídou technických systémů, které nahrazují mobilní, lokomoční a intelektuální funkce člověka, jejichž základním rysem je cílová činnost bez přímé účasti operátora při jejich řízení, mimo činnost dohlížecí. Hledání možností realizace těchto složitých autonomních systémů má velký potenciál využitelnosti alternativ funkcí v mnoha oborech lidské činnosti. V této disertační práci je teoreticky i prakticky řešena problematika plánování chůze autonomních robotů. Snahou autora je navrhnout a verifikovat jednotku sloužící pro plánování chůze čtyřnohého robotu ve členitém terénu. Jedná se o aktuální problém, který je stále v popředí zájmu vědeckého výzkumu.

1.1 ZADÁNÍ A ZAMĚŘENÍ PRÁCE

Úkolem práce je předložit návrh a simulační ověření adaptivního řídicího systému čtyřnohého robotu, založeného na Q-učení. Prioritou je nalezení (prostřednictvím schématu Q-učení) optimálních simultánních aktivací řídicích členů v různých stavech robotu spolu s vhodnou diskretizací spojitého stavového prostoru. Navrhnutý řídicí systém by měl robotu umožnit řešení tří úloh: kráčivý pohyb v přímém směru, vyhýbání se překážkám a překonání vybraných nouzových stavů robotu.

1.2 STANOVENÉ CÍLE

V rámci státní doktorské zkoušky byly s ohledem na zadání stanoveny tyto cíle disertační práce:

- návrh a implementace množiny elementárních způsobů chování vhodných pro plánování chůze čtyřnohého robotu,
- implementace vhodného koordinačního, resp. plánovacího schématu sloužícího pro propojování těchto chování, např. přístup založený na DEDS,
- dosažení různých stylů chůze čtyřnohého robotu pomocí navrženého řešení,
- verifikační experiment použitého přístupu.

2 TEORETICKÝ ROZBOR PROBLEMATIKY

2.1 KRÁČIVÝ POHYB ROBOTŮ

Kráčivý pohyb jako fenomén je námětem řady vědeckých studií. Nejprve byl předmětem zájmu biologů, kteří sledovali tělesnou stavbu a pohyb zvířat. Naproti tomu technicky zaměřený výzkum se orientuje na hledání abstraktnějšího matematického popisu lokomoce zvířat, který umožňuje zkonstruování biologicky inspirovaných kráčivých strojů (Beer, 1997). Biologická inspirace kráčivých robotů nejčastěji vychází ze savců, plazů, hmyzu aj. Přístupy k výběru a aplikaci poznatků z biologie pro potřeby konstrukce kráčivých robotů se značně liší: může se jednat o

napodobení stavby těla a počtu nohou (tj. návrh konstrukce robotu) (Quinn, 2001), existují práce zabývající se vývojem umělých svalů (Bar-Cohen, 2003), řada podnětů je využívána i při navrhování stylu chůze robotu (Hoff, 1997), rovněž existují snahy napodobit při návrhu řídicí jednotky činnost nervového systému živočichů (např. hmyzu) (Mochida, 1995), (Billard, 2001) atd.

Chůzi lze chápat např. jako koordinovaný způsob zvedání a pokládání noh. Každá chůze je tvořena posloupností tzv. událostí chůze (zvednutí nohy, položení nohy). Jak uvádí (Hardarson, 2002), (McGhee, 1968) existuje u čtyřnohých teoreticky 5040 různých kombinací událostí chůze, ovšem prakticky je jen malé procento z tohoto množství používáno živočichy pro chůzi. A právě tyto způsoby lokomoce jsou významné pro technickou praxi.

2.2 PŘÍSTUPY K PLÁNOVÁNÍ CHŮZE

2.2.1 Přístup založený na modelu

Někdy je tento přístup, např. (MacDonald, 1996), též označován jako geometrický. Jedná se o analytické řešení, které pro plánování chůze využívá model kinematiky a případně i model dynamiky robotu a prostředí. Odvození kinematického a dynamického popisu kráčivého stroje lze nalézt např. v (Řeřucha, 1997).

Výhody

- lze dosáhnout staticky i dynamicky stabilní chůze
- nevyžaduje čas pro učení se chůze za použití robotu

Nevýhody

- dosažitelné výsledky jsou limitovány přesností použitého modelu
- vyžadováno offline zpracování
- nevhodné pro dynamická (proměnná) prostředí

2.2.2 Přístup založený na chování

Návrh řízení v tomto případě vychází z předpokladu, že inteligentní chování robotu může vzniknout na základě interakce malého počtu zpětnovazebných primitiv chování nejnižší úrovně (MacDonald, 1996). Jednotlivá chování jsou většinou realizována pomocí konečných stavových automatů.

Výhody

- není nutné vytvářet model robotu a prostředí
- vhodné i pro strukturovaná, neznámá, dynamická prostředí
- snadná rozšiřitelnost řídicí jednotky (jen při malé změně cílů úlohy)

Nevýhody

- nutnost sestavit velké množství primitiv chování
- absence jednotné architektury určující způsob vzájemného propojení chování
- s rostoucím počtem primitiv chování roste náročnost vzájemného propojení
- změna úlohy může vést k nutnosti zcela nového návrhu řídicí jednotky

2.2.3 Hybridní a ad-hoc přístupy

Kromě těchto dvou výrazných skupin přístupů k plánování chůze robotů lze v odborném diskurzu nalézt celou řadu dalších řešení. Většina těchto přístupů se snaží využít výhody obou přístupů a potlačit jejich nevýhody. Mnoho vědeckých studií hledá inspiraci u biologických systémů. Např. (Ishiguro, 1995) se inspiroje imunitním systémem živočichů. Uvádí přístup založený na chování, který ovšem pro propojení jednotlivých primitiv využívá principy fungování imunitního systému. Obdobně (Mochida, 1995) představuje řešení využívající elementy chování propojené pomocí mechanismů uplatňovaných u smyslového vnímání člověka.

Celá řada prací, např. (Hoff, 1997), (Billard, 2001), (Spence, 2005), se zaměřuje na napodobení tzv. centrálního generátoru vzorů (central pattern generator, dále jen CPG). Jiným zástupcem je přístup založený na paradigmatu DEDS.

2.3 ARCHITEKTURA ŘÍZENÍ VYUŽÍVAJÍCÍ DEDS PRO PLÁNOVÁNÍ CHŮZE

Tato architektura řízení byla poprvé představena v disertační práci (Huber, 2000). Jedná se o třívrstvou architekturu řízení, která umožňuje autonomní pohyb robotu v neznámém prostředí. Celý systém je vystavěn na malé množině zpětnovazebných řídicích členů, které pracují spojitě v čase, a jsou spojeny se senzorickou soustavou robotu stejně tak jako s jeho akčními členy. Nadřazená vrstva zajišťuje diskretizaci spojitého prostoru na diskrétní události, které je možné popsat formálně. Na základě tohoto formálního popisu pak tato vrstva vyřazuje nepřipustné aktivace řídicích členů, čímž se jednoznačně snižuje velikost stavového prostoru, se kterým pracuje nejvyšší vrstva řídicí jednotky. Ta implementuje vybraný algoritmus učení pro krátkodobé plánování akcí, které povedou ke splnění stanoveného cíle.

2.3.1 Vrstva zpětnovazebných řídicích členů

Nejnižší vrstvu popisované architektury řízení tvoří tzv. báze řízení, $\{\Phi\}$. Jedná se o množinu tvořenou řídicími členy, Φ_i . Instance řídicího členu je vytvořena pomocí propojení přípustných vstupních zdrojů, $\{\sigma\}$, a přípustných výstupních zdrojů, $\{\tau\}$, s řídicím členem, Φ_i . Instance řídicích členů, označované také jako zákony řízení, vytvořené tímto způsobem jsou tedy zobrazováním vstupních zdrojů na zdroje výstupní, které umožňují informacím ze senzorů přímo ovlivňovat akční členy robotu, viz. (1). Pro daný robot pak lze tímto způsobem sestavit množinu řídicích členů, $\{C\}$, podle (2).

$$\sigma_{i\tau}^{\sigma} : \sigma \rightarrow \tau \quad (1)$$

$$\{C\} = \{\Phi_{i\tau_j}^{\sigma_k}\} \quad (2)$$

kde Φ_i je i -tý člen báze řízení, σ_k a τ_j jsou podmnožiny přípustných vstupních a výstupních zdrojů robotu. Při návrhu řídicího členu je proto vždy nutné určit, které vstupní a výstupní zdroje robotu mohou být používány, aby se předešlo vytvoření

nekonzistentních řídicích členů. Jako vstupní a výstupní zdroje mohou být použity např. úhel klonění, klopení a bočení těla robotu nebo jeho pozice x, y, z v prostoru, případně kinematické řetězce noh, formálně označované 0,1,2,3.

2.3.2 Vrstva zahrnující koordinační mechanismus DEDS

Jedná se o prostřední vrstvu popisované architektury řízení, která je vymezena dvěma hlavními cíly. Prvním cílem je na základě formálního popisu jednotlivých členů báze řízení sestavit množinu složených řídicích členů, které jsou schopny generovat složitější chování robotu. Druhým cílem je pak sestavení koordinačního mechanismu, který zahrnuje všechny přípustné stavy robotu a pro každý z nich určuje množinu přípustných aktivací řídicích členů. Neposkytuje však informaci o tom, který z množiny přípustných zákonů řízení se má v daném stavu robotu aktivovat. Vedle těchto hlavních cílů poskytuje prostřední vrstva nezbytné oddělení spojitě činnosti senzorů a aktuátorů od diskrétních formálně popsanych událostí, se kterými pracuje nejvyšší vrstva. Dalším přínosem koordinačního mechanismu je pak snížení velikosti stavového prostoru řešené úlohy a snížení počtu možných aktivací v každém z těchto stavů, což výrazně zkracuje časové nároky algoritmu učení na nalezení vhodné posloupnosti aktivací řídicích členů.

Naproti tomu, u zde popisované architektury řízení, je využíváno složených řídicích členů, které jsou tvořeny užitím hierarchické skladby pomocí operátoru podřízenosti, $<$. Tyto složené řídicí členy zajišťují složitější chování robotu skrze paralelní aktivace základních řídicích členů. Např. v zápisu $\Phi_f = \Phi_g < \Phi_h$ představuje Φ_g podřízený řídicí člen, který nesmí svojí činností ovlivnit činnost nadřazeného řídicího členu Φ_h . Množina řídicích členů, viz. rovnice (2), použitelných pro řešení dané úlohy může být tímto rozšířena o množinu složených řídicích členů:

$$C_m : \left(\dots < \left(\Phi_{i\tau_j}^{\sigma_j} < \Phi_{i\tau_l}^{\sigma_l} \right) \right) \quad (3)$$

kde C_m značí množinu složených řídicích členů, Φ_i značí i -tý člen množiny C a σ, τ jsou příslušné množiny vstupních a výstupních zdrojů.

Jednotlivé základní řídicí členy se obecně mohou snažit o dosažení vzájemně protichůdných cílů. Jejich propojení do složeného řídicího členu by pak způsobilo nekonzistentní chování, při kterém by nebyl splněn ani jeden dílčí cíl řízení. Jelikož je každý základní řídicí člen charakterizován i svým formálním popisem, je možné zcela automaticky vyloučit nekonzistentní složené řídicí členy za použití predikátové logiky a u těch přípustných je možné vyvodit formální popis, opět zcela automaticky. Tímto popisem je myšlena trojice vektorů (S, M, R) , kde vektor S představuje počáteční podmínky, tj. stavy systému, ve kterých může být řídicí člen aktivován. Vektor M popisuje modifikaci stavu systému, tj. jakým způsobem se konfigurace může změnit. Vektor R reprezentuje cíl řízení, tj. o dosažení jakého stavu se řídicí člen snaží.

V DEDS je diskrétní stavový prostor reprezentován funkčními cíly, kterých je dosahováno pomocí přípustné množiny akcí spojitěho řízení. Tím vzniká ucelený

model systému, který popisuje schopnosti robotu v daném prostředí pomocí výsledků chování. Vytvoření tohoto modelu je usnadněno formálními vlastnostmi spojitých řídicích členů a využitím báze řízení. Předložený přístup využívá řídicí členy po dobu procesu konvergence a množinu událostí v podobě aktivací a konvergence řídicích členů. Každý řídicí člen navíc na základě svého cíle rozděluje podřízený prostor fyzikálních stavů na množinu oblastí vlivu a množinu, kterou neovlivňuje. V této souvislosti může být na strategii řízení nahlíženo jako na posloupnost aktivací instancí řídicích členů, jež je charakteristická pro danou úlohu a jež vede systém robotu množinou vyhovujících rovnovážných stavů, přičemž přechody mezi nimi jsou určeny výsledky procesu konvergence. Toto umožňuje, aby bylo celkové chování systému popsáno pomocí DEDS, přičemž diskrétní stavový prostor je z velké části určen na základě atraktorů řídicích členů. Zmíněných vlastností je v této práci využito pro vytvoření modelu systému se stavy reprezentovanými vektorem binárních hodnot (predikátů), p_i , které indikují, který z cílů řízení je či není splněn. Stavový vektor, $q = (p_0, p_1, \mathbf{K}, p_n)$, pak zcela charakterizuje stav, ve kterém se systém nachází.

2.3.3 Vrstva učení

Nejvyšší vrstva popisované architektury řízení zahrnuje vybraný algoritmus učení. Na vstupu této vrstvy je seznam přípustných stavů systému. A pro každý z těchto stavů je k dispozici množina zákonů řízení, které mohou být v daném stavu aktivovány. Cílem této vrstvy je nalezení optimální aktivace v každém ze stavů systému tak, aby byl splněn stanovený cíl řízení, který je popsán stavovým vektorem q , a to pomocí co nejmenšího počtu aktivací dostupných zákonů řízení. Rovněž lze říci, že cílem zvoleného algoritmu učení je nalezení vhodné posloupnosti aktivací jednotlivých instancí řídicích členů, které povedou systém přes množinu přípustných stavů až po stav reprezentující splnění cíle řízení.

V této práci je nejvyšší vrstva řešena s využitím algoritmu Q-učení. Algoritmus Q-učení se řadí do oblasti opakovaně posilovaného strojového učení a poprvé byl popsán ve (Watkins, 1989). Je tedy určen pro řešení úloh, které lze klasifikovat jako Markovský rozhodovací problém. Narozdíl od jiných metod je Q-učení schopné paralelního hledání optimální hodnotící funkce, která určuje ohodnocení pro každou dvojici stav-akce, a optimální strategie řízení, která určuje posloupnost akcí, které vedou k maximální odměně v průběhu času. Z charakteru opakovaně posilovaného učení vyplývá, že oba problémy je algoritmus schopen určit bez vnějšího zásahu učitele a navíc bez znalosti modelu řešené soustavy, v tomto případě čtyřnohého robotu. Učení proto probíhá pouze na základě přímého vlivu jednotlivých akcí na stav systému.

Jak bylo uvedeno, algoritmus hledá optimální hodnotící funkci, která každé dvojici stav-akce přiřadí číselnou hodnotu. Jedná se o tzv. Q hodnotu, $Q(x_i, a_i)$, která vyjadřuje, jak moc perspektivní je volba akce řízení, a , ve stavu soustavy, s ,

v časovém kroku, t . Optimální hodnota může být dle Bellmanovy rovnice optimality spočtena následovně:

$$Q^*(x_t, a_t) = r_{t+1} + \gamma \max_{b \in A} Q^*(x_{t+1}, b) \quad (4)$$

kde Q^* je optimální Q hodnota příslušící volbě akce řízení, a_t , ve stavu soustavy, x_t , v časovém kroku, t . Hodnota r_{t+1} určuje okamžitou odměnu soustavy za volbu akce řízení a_t ve stavu x_t . Výsledný stav soustavy po provedení této akce je x_{t+1} . Množina, A , představuje všechny přípustné akce řízení. Diskontní (srážkový) faktor, γ , pro který platí $0 < \gamma \leq 1$, je používán pro snížení vlivu vzdálených odměn na aktuální Q hodnotu. Rovnice (4) určující výpočet optimální hodnotící funkce může být aproximována iterativním výpočtem, jak ukazuje rovnice (5).

$$Q_t(x_{t-1}, a_{t-1}) = (1 - \alpha) Q_{t-1}(x_{t-1}, a_{t-1}) + \alpha \left(r_t + \gamma \max_{b \in A} Q_{t-1}(x_t, b) \right) \quad (5)$$

kde α je parametr učení, určující vliv právě získané okamžité odměny za provedení akce na aktualizovanou Q hodnotu příslušné dvojice stav-akce, $Q_t(x_{t-1}, a_{t-1})$. Tento iterativní způsob výpočtu již umožňuje využití algoritmu Q-učení pro nalezení odhadu optimální hodnotící funkce.

Uvedená rovnice určující způsob výpočtu nové Q hodnoty není příliš vhodná pro použití v nejvyšší vrstvě popsané architektury řízení. Dle rovnice (5) jsou totiž aktualizace prováděny v každém časovém kroku, t . Naproti tomu v architektuře využívající model DEDS je klíčový výskyt událostí aktivace a konvergence používaných zákonů řízení. Přičemž časová náročnost konvergence různých zákonů řízení, resp. instancí řídicích členů, je obecně různá. S uvažováním výše uvedeného je vhodné modifikovat rovnici (5) následovně, (Huber, 2000):

$$Q_k(\mathbf{p}_{k-1}, c_{k-1}) = (1 - \alpha) Q_{k-1}(\mathbf{p}_{k-1}, c_{k-1}) + \alpha \left(r_k + \gamma \max_{c \in C^o} Q_{k-1}(\mathbf{p}_k, c) \right) \quad (6)$$

kde k je diskrétní časový okamžik odpovídající události konvergence ($k-1$) aktivace řídicího členu, c_{k-1} . Dále r_k představuje odměnu získanou během činnosti tohoto řídicího členu a $\mathbf{p}_k$ je predikátový vektor reprezentující stav robotu po ukončení konvergence tohoto členu. Množina C^o je tvořena všemi přípustnými řídicími členy, viz rovnice (2). Tato formulace procesu iterace výrazně snižuje časovou náročnost díky jednodušší struktuře výsledné strategie řízení a nižšímu počtu okamžiků, kdy je prováděn výpočet.

2.4 POSTUP ŘEŠENÍ

Při řešení stanovených cílů bude postupováno následovně. Nejprve je nutné zvolit množinu základních řídicích členů a najít vhodný předpis jejich potenciálové funkce. Pro prvotní zhodnocení správnosti návrhu je zapotřebí posoudit a vykreslit úplný průběh těchto funkcí pro různé výchozí konfigurace robotu.

Dále je nutné implementovat návrh základních řídicích členů ve vhodném simulačním prostředí pro zvolený čtyřnohý robot. Z těchto řídicích členů lze pomocí vhodné algoritmizace sestavit množinu složených řídicích členů a vyvodit automaticky jejich specifikace. Chování robotu generovaná základními i složenými

řídícími členy by měla být simulačně testována. Toto vyžaduje vytvoření modelu robotu ve zvoleném simulačním prostředí. V dalším kroku lze následně sestavit model všech možných chování robotu, reprezentovaný nedeterministickým stavovým automatem.

V následující fázi řešení budou hledány optimální strategie řízení pomocí algoritmu Q-učení. Pro tento účel je nutné vytvořit mechanismus, pomocí něhož budou ukládány a načítány jednotlivé Q hodnoty. Dále je zapotřebí vhodným způsobem reprezentovat nalezené optimální strategie řízení pro dané úlohy. V poslední části řešení je totiž nezbytné zrealizovat monitorovací jednotku, která bude schopná měřit a vyhodnocovat jednotlivé parametry chůze tak, aby bylo možné provést kvalitativní i kvantitativní porovnání těchto stylů.

3 PRAKTICKÁ ČÁST – SOUHRN DOSAŽENÝCH VÝSLEDKŮ

3.1 VOLBA PROSTŘEDKŮ PRO VERIFIKACI

Pro implementaci předkládaného řešení je nezbytně nutné zvolit vhodné programové vybavení, které by umožnilo realizovat jednotlivé kroky popsané v kap. 2.4. V první fázi řešení je přistoupeno k návrhu elementárních řídicích členů. Pro tyto účely lze použít výpočtový software Matlab, který je již tradičně považován za standard v technické praxi. V dalších krocích řešení již musí být posuzováno chování robotu během procesu aktivace a konvergence navržených řídicích členů. Pro zajištění věrohodnosti je přitom vhodné postihnout nejen statické, ale i dynamické účinky působící na robot, jako např. setrvačné síly apod. Pro tyto potřeby byl zvolen fyzikální engine Open Dynamics Engine (dále jen ODE). Pro vykreslení simulované scény byla použita grafická knihovna OpenGL. Požadavek na uložení stavu simulace a její opětovné spuštění byl vyřešen pomocí standardizovaného formátu XODE, což je textový soubor přesně definované struktury ukládaný ve formátu XML.

3.2 PLATFORMA POUŽITÁ PRO VERIFIKACI NÁVRHU ŘEŠENÍ

Předkládané řešení umožňující využití opakovaně posilovaného učení pro řízení robotu klade maximální důraz na dodržení obecnosti a rozšiřitelnosti. Jedná se tedy o univerzální návrh použitelný pro různé kráčivé roboty plnící obecně různé úkoly, jejichž zadání může být měněno bez nutnosti zásadní změny jednotky řízení. Aby však bylo možné doložit vhodnost návrhu a poskytnout porovnání s již publikovanými výsledky v odborné literatuře, je nutné vztáhnout tento obecný návrh ke konkrétní platformě. Jako testovací robot byl zvolen sériově vyráběný model EQ3-KT firmy Lynxmotion, který má tři stupně volnosti na nohu. Je tedy umožněn obecný pohyb těla robotu v prostoru aniž by docházelo ke změně bodů došlapu. Rozměr těla robotu je 240x200x55 mm. Celková hmotnost robotu je 1320 g, z toho 670 g připadá na tělo. Robot je osazen celkem 12ti servopohony Hitech HS-475HB, váha jednoho pohonu činí 40g.

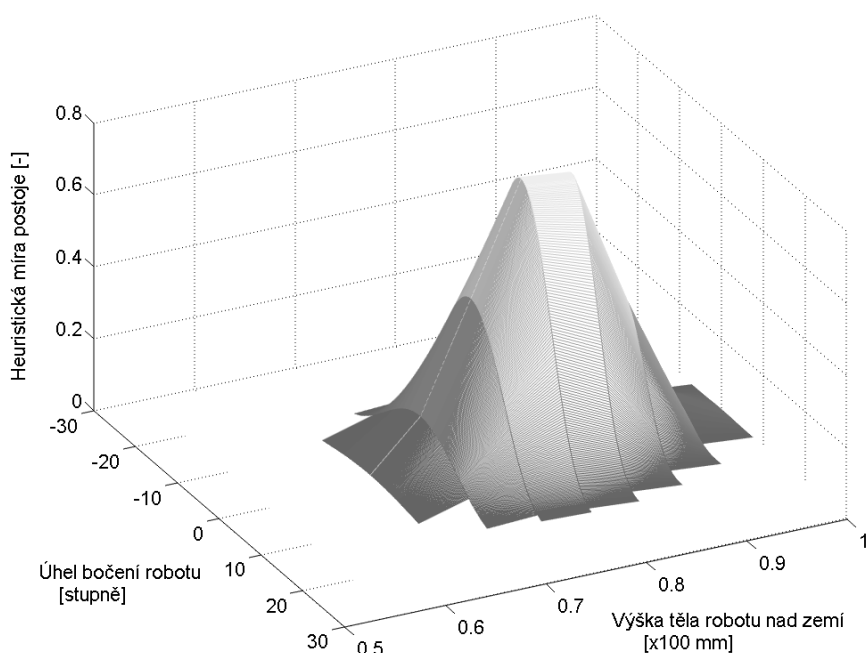
3.3 SESTAVENÍ ELEMENTÁRNÍCH ŘÍDICÍCH ČLENŮ

Základem popsaného návrhu jednotky řízení robotu je báze řízení, obsahující elementární řídicí členy. Pro potřeby plánování chůze robotu je vhodné do této množiny zahrnout řídicí člen postoje, Φ_P , a kontaktu, Φ_K . V práci je navržena modifikace řídicího členu postoje a vlastní řešení řídicího členu kontaktu. Celkem je v disertační práci sestaveno a diskutováno 28 různých instancí těchto dvou řídicích členů. V následujících kapitolách jsou uvedeny vybrané příklady.

3.3.1 Řídicí člen postoje

Hlavním cílem tohoto řídicího členu, Φ_P , je dosažení optimálních poloh noh robotu vzhledem k jejich pracovnímu prostoru. Jako kritérium optimality je použito heuristické míry postoje.

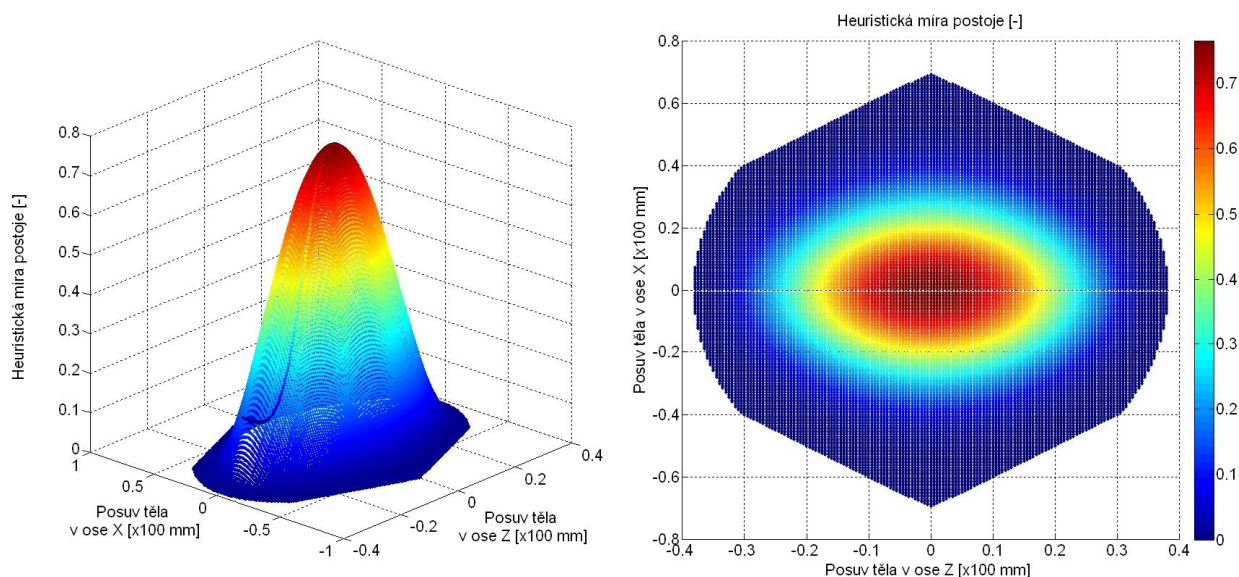
Průběh heuristické míry postoje pro konfigurace robotu s rozdílnou výškou těla nad zemí je uveden na Obr. 1. Je patrné, že existuje pro dané neměnné body došlapu ideální výška těla robotu nad zemí, při které lze docílit maximální hodnoty heuristické míry postoje. Z grafu rovněž plyne, že instance řídicího členu postoje, $\Phi_{P_y}^{0,1,2,3}$, která optimalizuje výšku těla robotu nad zemí prostřednictvím translačního pohybu, bude vyhovovat požadavkům kladeným na tvar plochy funkce.



Obr. 1: Potenciálové funkce řídicího členu postoje pro různé výšky těla nad zemí

Řídicí člen postoje je možné použít také s výstupními zdroji x, z . Příslušná instance pak má tvar $\Phi_{P_{x,z}}^{0,1,2,3}$. Na Obr. 2 je uveden průběh potenciálové funkce tohoto řídicího členu pro výchozí konfiguraci č.1. Graf ukazuje závislost velikosti

heuristické míry postoje na poloze těla robotu, která se mění v rovině těla robotu, tj. ve směru longitudinální a laterální osy těla.



Obr. 2: Potenciálové funkce řídicího členu $\Phi_{P_{x,z}^{0,1,2,3}}$ pro konfiguraci č.1

3.3.2 Řídicí člen kontaktu

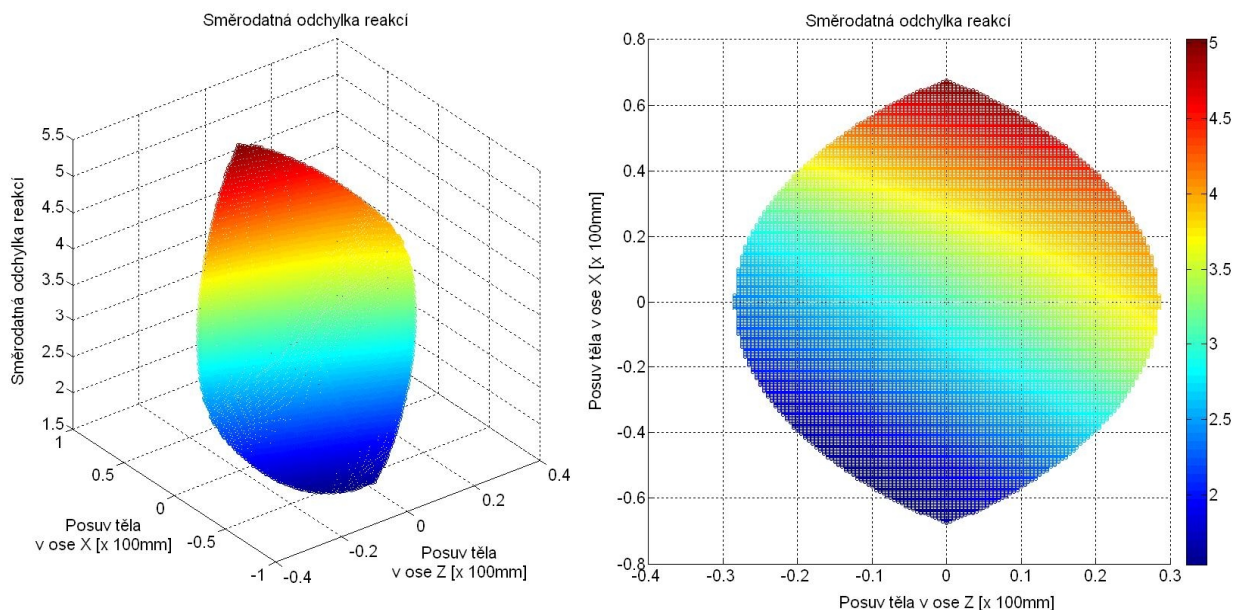
Hlavním cílem řídicího členu kontaktu, Φ_K , je docílení optimálního rozložení sil působících v došlapech. Tohoto stavu se snaží docílit translačním pohybem těla nebo změnou konfigurace vybrané nohy. S ohledem na informace uvedené v teoretickém rozboru bylo přistoupeno k vlastnímu návrhu tohoto řídicího členu.

Cílem řídicího členu kontaktu by mělo být dosažení takové konfigurace, kdy budou jednotlivé síly působící v došlapech rozloženy rovnoměrně. Proto je možné formulovat výpočet potenciálové funkce řídicího členu kontaktu následovně:

$$\Phi_K = \sigma(F_{1..n}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (F_i - \bar{F})^2}, \quad n=3 \quad (7)$$

kde $\sigma(F)$ značí směrodatnou odchylku tří spočtených sil působících v došlapech robotu, F_i jsou jednotlivé síly, $\bar{F}$ je aritmetický průměr těchto sil.

Instance řídicího členu kontaktu, $\Phi_{K_{x,z}^{0,2,3}}$, se snaží o minimalizaci směrodatné odchylky sil působících v došlapech noh 0,2,3 translačním pohybem těla robotu. Na Obr. 3. je tato funkce ukázána pro výchozí konfiguraci robotu č. 11. Průběh lze považovat za přijatelný, neboť funkce má právě jedno globální maximum a je klesající na celém definičním oboru. Především z pravé části Obr. 3 vyplývá, že tento řídicí člen bude při své činnosti posouvat střed těla robotu (resp. průmět středu těla do roviny kolmé na směr tíhové síly) ke středu trojúhelníku, jehož vrcholy tvoří body došlapu noh 0,2,3 (resp. průměty těchto bodů do roviny kolmé na směr tíhové síly). Řídicí člen se tak zároveň snaží o dosažení trojúhelníkové statické stability, přičemž podpůrný trojúhelník je tvořen nohami robotu, které daná instance používá jako vstupní zdroje.



Obr. 3: Potenciálová funkce řídicího členu $\Phi_{K_{x,z}}^{0,2,3}$ pro zvolenou konfiguraci

3.4 PROJEKT IMPLEMENTUJÍCÍ NAVRŽENÉ ŘEŠENÍ

Předchozí kapitola se zabývala návrhem báze řízení tvořící základ řídicí jednotky robotu. Po dokončení této první fáze, bylo nutné implementovat pomocí zvolených prostředků všechny tři vrstvy architektury založené na paradigmatu DEDS, jejíž teoretický rozbor je uveden v kap. 2.3 a to včetně algoritmu Q-učení. Pro tento účel byl sestaven komplexní software projekt, jehož hlavním účelem je:

- ověření funkčnosti koordinačního, resp. plánovacího schématu
- ověření funkčnosti vytvořeného modelu robotu
- zjištění výsledných elementů chování robotu pomocí navržených řídicích členů
- nalezení vhodného schématu paralelních aktivací kompozitních řídicích členů
- nalezení parametrů algoritmu Q-učení pro řešení zadaných úloh robotu
- posouzení výsledných chování robotu v rámci řešených úloh, viz, kap. 3.5

Pro dosažení těchto cílů bylo nezbytné:

- algoritmizovat chování všech navržených instancí základních řídicích členů
- propojit činnost těchto zákonů řízení s modelem robotu
- zajistit automatické sestavení kompozitních řídicích členů
- implementovat koordinační mechanismus DEDS pomocí stavového automatu
- implementovat algoritmus Q-učení a proces učení propojit s podřízenou vrstvou
- zajistit grafické zobrazení chování robotu
- zajistit ukládání a načítání simulace a jejich parametrů

Všechny tyto dílčí cíle a požadavky byly splněny a jsou do projektu zapracovány. Veškeré prezentované výsledky byly dosaženy pomocí tohoto software, který je součástí předkládaného řešení.

3.5 VÝSLEDNÁ CHŮZE ROBOTU PŘI KOMPLEXNÍCH ÚLOHÁCH

Po té co byla ukázána elementární chování robotu generovaná elementárními řídicími členy a také některá složitější chování robotu, realizovaná kompozitními řídicími členy, bylo přistoupeno k návrhu řešení vybraných komplexních úloh robotu. Jedná se především o tyto úlohy:

- rotace robotu změnou úhlu bočení
- chůze robotu v longitudinálním směru po vodorovné rovině
- chůze robotu v longitudinálním směru po nakloněné rovině

Pro každou z těchto úloh byla navržena vlastní reprezentace stavu robotu, zvolena množina základních řídicích členů a zadány omezující podmínky úlohy. Tyto údaje byly použity jako vstup vytvořené aplikace, pomocí níž byla vytvořena množina kompozitních řídicích členů, model chování robotu pro danou úlohu a nalezena vhodná posloupnost aktivací jednotlivých zákonů řízení prostřednictvím algoritmu Q-učení. V žádné z následujících úloh nemá algoritmus učení k dispozici znalost o tom, zda se robot nachází na vodorovné nebo nakloněné rovině, ani zda jsou v okolí robotu umístěny překážky.

3.5.1 Úloha rotace robotu

Cílem této úlohy bylo nalezení posloupnosti aktivací dostupných zákonů řízení, které by vedly k rotaci robotu kolem svislé osy na vodorovné rovině.

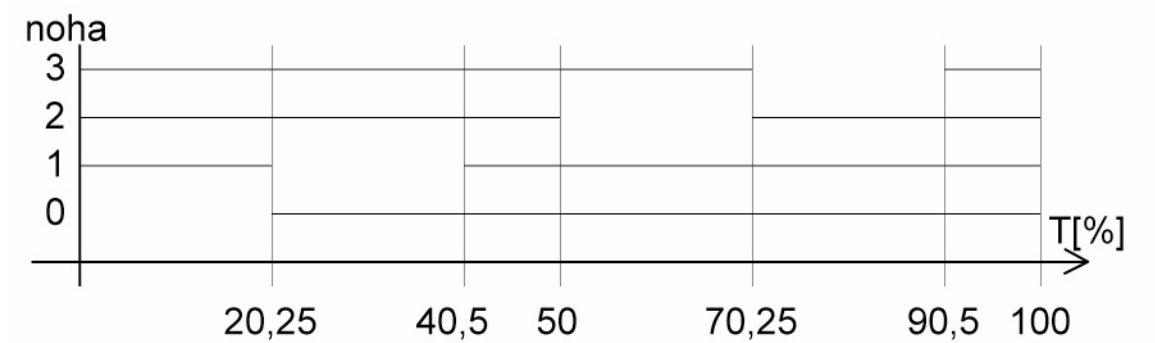
Pomocí zvolených parametrů učení byla nalezena výsledná chůze robotu, která zajišťuje rotaci robotu. Jedná se o chůzi:

- cyklickou: pohyb noh je realizován v pravidelných časových intervalech
- otáčivou: libovolný bod těla rotuje kolem osy těla
- regulární: každá noha má stejný činitel využití
- plazivou: v každém okamžiku je v přenosové fázi maximálně jedna noha
- nespojitou: k pohybu těla dochází pouze v některých fázích chůze

Činitel využití β_{cv} je u jednotlivých noh shodný. Průměrná hodnota tohoto ukazatele z pěti cyklů je 79,75%. Relativní fáze, Φ_{ri} , noh 1,2,3 vůči referenční noze 0 nabývá hodnot:

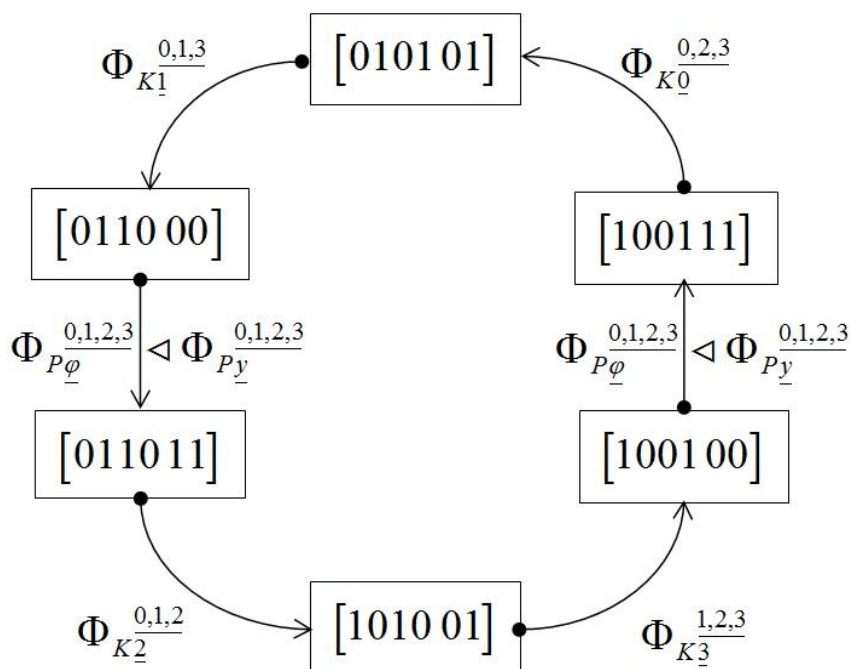
$$\Phi_{ri} = \{0, 29.75, 70.25\} \quad (8)$$

Graf výsledné chůze, ukazující střídání oporné a posuvné fáze noh v průběhu jednoho cyklu, je ukázán na Obr. 4. Na ose X jsou vyneseny procentuální hodnoty z doby trvání jednoho cyklu, T . K rotaci těla dochází dvakrát během tohoto cyklu a to vždy, když se všechny nohy nachází v oporné fázi. Celkový čas, během kterého je manipulováno tělem robotu, činí 19% doby trvání cyklu.



Obr. 4: Graf chůze v úloze rotace

Ve výsledné chůzi jsou aktivovány především čtyři instance základního řídicího členu kontaktu a dvě kompozitní instance zajišťující rotaci robotu a změnu výšky těla nad povrchem. Konkrétní posloupnost aktivací a přechody mezi stavy robotu jsou ukázány na Obr. 5. Uvedené přechody mezi stavy nastávají v případě rotace robotu na vodorovné rovině nejčastěji. Pokud by ovšem některý z aktivovaných zákonů řízení nebyl úspěšný, např. vlivem nepříznivé konfigurace robotu, výskytem překážky apod., došlo by k přechodu soustavy do stavu, který není ve schématu uveden. V tomto případě by ovšem činnost robotu nebyla zastavena, ale došlo by k aktivaci nejlépe ohodnoceného zákona řízení pro daný nový stav. K takovýmto událostem docházelo především v počátečních stádiích procesu učení, kdy byly testovány i velmi nepříznivé posloupnosti aktivací vedoucí robot do konfigurace, kdy se např. více noh nacházelo ve svých mezních polohách.



Obr. 5: Přechody mezi stavy robotu v úloze rotace

3.5.2 Úloha dopředné chůze

Cílem této úlohy bylo nalezení optimální posloupnosti aktivací zákonů řízení, která zajistí chůzi robotu ve směru jeho longitudinální osy, tj. osy $X^{(2)}$, po vodorovné rovině.

Uvedené parametry učení vedly k nalezení strategie, která zajišťuje chůzi robotu ve směru longitudinální osy. Výslednou chůzi robotu je možné považovat za cyklickou. Hlavní cyklus, který zajišťuje chůzi je ovšem několikrát narušen a to v okamžiku, kdy některý z aktivovaných řídicích členů nedosáhne svého cílového stavu. V následujícím textu bude popsán tento hlavní cyklus. Možná porušení tohoto hlavního cyklu a příčiny neúspěchu konvergence řídicích členů budou diskutovány v následující kapitole.

Přechody mezi jednotlivými stavy robotu v průběhu zmíněného cyklu jsou ukázány na Obr. 7. K posuvu robotu v požadovaném směru dochází dvakrát během jednoho cyklu, a to pod vlivem aktivací $\Phi_{Py}^{0,1,2,3} < \Phi_{K_{x,z}}^{1,2,3}$ a $\Phi_{Py}^{0,1,2,3} < \Phi_{K_{x,z}}^{0,1,2}$. V těchto případech je rovněž optimalizován postoj robotu změnou výšky těla nad podložkou. Mezi těmito hlavními posuvy těla nastává fáze chůze, kterou je možné označit jako přípravnou. V těchto částech cyklu jsou totiž měněny konfigurace noh levé, resp. pravé, strany robotu. Zatímco přední nohy jsou nastavovány pomocí řídicího členu postoje, zadní nohy jsou nastavovány pomocí řídicího členu kontaktu. Přední nohy robotu, tj. 1 a 2, nemohou být nastaveny instancí řídicího členu kontaktu, neboť nebyly zařazeny do výchozí množiny C . Naopak zadní nohy robotu, tj. 0 a 3, nejsou nastavovány řídicím členem postoje, protože výsledná konfigurace je nevýhodná pro dosažení dostatečně velkého posuvu, který je motivován udělováním odměny, r .

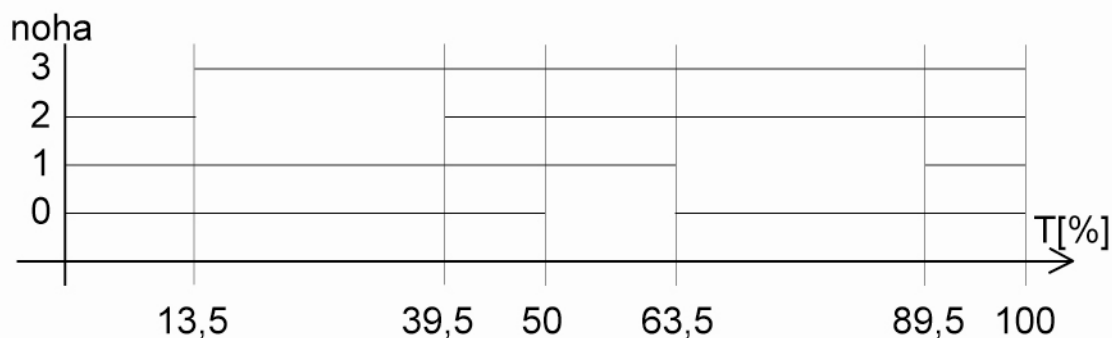
Pokud dochází k úspěšným konverencím řídicích členů uvedených na Obr. 7, pak robot vykonává chůzi:

- cyklickou: pohyb noh je realizován v pravidelných časových intervalech
- vlnitou: pohyb noh začíná zadní nohou a pokračuje v dopředném směru, laterálně spárované nohy jsou fázově posunuty přesně o polovinu cyklu kroku
- plazivou: v každém okamžiku je v přenosové fázi maximálně jedna noha
- nespojitou: k pohybu těla dochází pouze v některých fázích chůze

Činitel využití pro nohy 0,3 nabývá hodnoty $\beta_{cv} = 86,5\%$ a pro nohy 1,2 je $\beta_{cv} = 74\%$. Hodnoty tohoto ukazatele se liší, protože pohyb předních noh je realizován řídicím členem postoje, který manipuluje nohou jiným způsobem než řídicí člen kontaktu, který je používán pro změnu konfigurace noh zadních. Relativní fáze, Φ_{ri} , noh 1,2,3 vůči referenční noze 0 nabývá hodnot:

$$\Phi_{ri} = \{0, 40, 26.5\} \quad (9)$$

Graf výsledné chůze, ukazující střídání oporné a posuvné fáze noh v průběhu jednoho cyklu, je ukázán na Obr. 6. Na ose X jsou vyneseny procentuální hodnoty z doby trvání jednoho cyklu, T . K posuvu těla dochází dvakrát během tohoto cyklu a to vždy, když se všechny nohy nachází v oporné fázi. Celkový čas, během kterého je manipulováno tělem robotu, činí 21% doby trvání cyklu.



Obr. 6: Graf chůze v úloze dopředné chůze po rovině

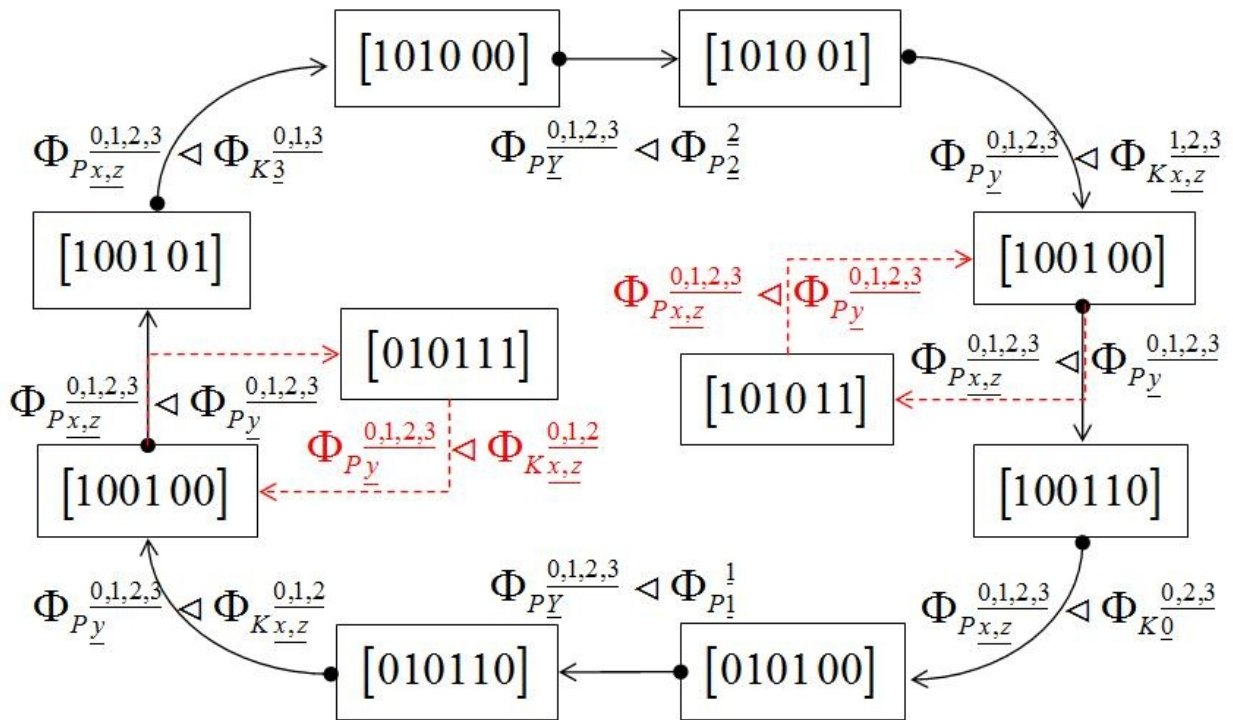
3.5.3 Úloha dopředné chůze na nakloněné rovině

Cílem této úlohy je ověřit schopnost robotu pohybovat se po nakloněné rovině v longitudinálním směru. Pro potřeby experimentálního ověření byl robot umístěn na nakloněnou rovinu, která je pootočená o úhel stoupání 5° kolem osy $Z^{(1)}$. Pro ověření schopností řídicích členů a učícího algoritmu byla jako výchozí strategie použita chůze naučená při řešení úlohy chůze na vodorovné rovině.

Při snaze uplatňovat již naučenou strategii aktivací zákonů řízení se ukázalo, že některé instance sice plní svůj cíl řízení, ale výsledná konfigurace robotu je odlišná než konfigurace při řešení předchozí úlohy. Jedná se především o instance $\Phi_{K \frac{0,1,2}{x,z}}$ a $\Phi_{K \frac{1,2,3}{x,z}}$, které se snaží rovnoměrně rozložit síly v došlapech předních noh a jedné z noh zadních. Cíl aktivace je splněn, protože je dosaženo stabilního postavení na nohách, které jsou použity jako vstupní zdroje instance, ale koncová hodnota směrodatné odchylky reakcí, dle rov. (7), je větší než v předchozí úloze. To znamená, že i konfigurace robotu po dokončení procesu konvergence je odlišná. Následná aktivace řídicího členu vede na jiný než očekávaný stav, čímž dochází k opuštění hlavního cyklu. Nové přechody mezi stavy jsou znázorněny na Obr. 7 červenou barvou.

Popsaná situace, kdy není splněn cíl aktivovaného zákona řízení, může nastat i v případě chůze na vodorovné rovině ovšem s nižší pravděpodobností. Čím bude mít rovina větší úhel stoupání, anebo čím více překážek, resp. nerovností terénu, se bude v prostředí vyskytovat, tím větší bude pravděpodobnost výskytu přechodu do jiných než naznačených stavů.

K obdobné situaci může dojít i v případě jiných než dvou výše zmíněných instancí. Výslednou chůzi proto nelze považovat za cyklickou a nelze ji hodnotit pomocí kvantitativních kritérií. Vizualně lze posoudit výslednou chůzi pomocí videozáznamu přiloženého k originálnímu výtisku disertační práce. Popsané dodatečné aktivace řídicích členů odpovídají opakované snaze robotu o docílení stabilního postavení na nohách 0,1,2 (resp. 1,2,3), což lze ve videozáznamu chůze pozorovat. Výslednou chůzi je možné označit jako nespojitou a plazivou.



Obr. 7: Přechody mezi stavy robotu v úloze dopředné chůze po nakloněné rovině

4 ZÁVĚR

Uvedené verifikační experimenty prokázaly schopnost učícího algoritmu nalézt vyhovující vzájemně odlišné styly chůze při řešení různých úloh. Rovněž prokázaly vhodnost diskretizace spojitého stavového prostoru a ověřily funkčnost koordinačního schématu architektury DEDS. Experiment provedený na nakloněné rovině prokázal schopnost navrženého řešení adaptovat se v případě změny okolního prostředí. Představené řešení může být dále vylepšeno zjištěním pravděpodobností přechodů pod vlivem aktivací zákonů řízení. Tímto způsobem je možné odstranit nedeterminismus stavového automatu prostřední vrstvy řídicí jednotky. Díky tomu je následně možné použít i offline Q-učení a urychlit tak proces učení.

Další výhodou představuje možnost přechodu do módu řízení pomocí oprátora. Tohoto způsobu lze snadnou dosáhnout vypnutím nejvyšší vrstvy architektury řízení, která provádí naučený výběr z množiny dostupných zákonů řízení. Výběr akce řízení je ponechán operátorovi, přičemž aktivace zákonů řízení, které by mohli vést do nepřijatelného stavu robotu, jsou automaticky vyřazovány. Operátor proto může vybírat pouze z množiny přípustných zákonů řízení.

Přestože je předložené řešení řídicí jednotky vztaženo ke konkrétní robotické platformě, popsany způsob návrhu je univerzální a může být s minimem nutných změn upraven pro jiný typ kráčivého robotu. Navíc předložené řešení může najít uplatnění nejen v mobilní robotice, ale také například při řešení pohybu průmyslových manipulátorů.

4.1 PŘÍNOS PRÁCE

4.1.1 Příklad teoretický

Hlavní teoretický přínos této práce spočívá ve vytvoření robustní architektury řízení kráčivého robotu, která umožňuje využití strojového učení pro nalezení různých stylů chůze. Za přínos lze dále považovat:

- zavedení vhodné diskretizace spojitého prostoru umožňující použití algoritmu Q-učení pro řízení kráčivého robotu
- návrh algoritmy simulace simultánních aktivací řídicích členů
- návrh řídicího členu kontaktu schopného činnosti ve strukturovaném prostředí
- zajištění reaktivního chování robotu na neznámé překážky v jeho bezprostřední blízkosti

4.1.2 Příklad praktický

Za praktické přínosy této práce lze považovat:

- vytvoření komplexního software projektu, který umožňuje provádět dynamické simulace robotických soustav v různých prostředích
- popsání způsobu řízení robotu může být snadno využito i pro řízení jiných mechatronických soustav
- uvedení průběhu funkcí použitých řídicích členů pro různé konfigurace robotu
- vyhodnocení naučených stylů chůze pomocí standardizovaných ukazatelů

4.1.3 Příklad pedagogický

- v rámci řešení této disertační práce bylo vytvořeno několik bakalářských prací zabývajících se mobilní robotikou [33], [35]
- zveřejněné videozáznamy chůze robotu ukazují poutavou formou činnost strojového učení, čehož může být využito pro motivaci zájmu studentů o algoritmy umělé inteligence
- vytvořený software projekt umožní studentům se zájmem o robotiku jednoduchým způsobem navrhnout simulace vlastních robotických systémů

LITERATURA

- ALEXANDER R.M.: The gaits of bipedal and quadrupedal animals, *Int. J. Robotics Research* 3(2): 49-59. 1984.
- BAR-COHEN Y.: Biologically Inspired Intelligent Robots Using Artificial Muscles, *International Conference on MEMS, NANO and Smart Systems (ICMENS'03)*, 2003.
- BEER R.D., QUINN R.D., CHIEL H.J., RITZMANN R.E.: Biologically inspired approaches to robotics: What can we learn from insects?, *Communications of the ACM*, Vol.40 No.3, March 1997.
- BILLARD A., IJSPEERT A.: Biologically Inspired Neural Controllers for Motor Control in a Quadruped Robot, *Int. Joint Conf. on Neural Networks (IJCNN'00)*, 2000.
- BILLARD A.: Learning Motor Skills By Imitation: A Biologically Inspired Robotic Model, *Int. journal Cybernetics and Systems*, 32:155-193, 2001.
- BOR K., BŘEZINA T.: Návrh a implementace kompozitního řídicího členu chůze 4-nohého robotu, *Diplomová práce, Vysoké Učení Technické v Brně, Fakulta Strojního Inženýrství*, 2004.
- BROOKS R.A.: A robot that walks; emergent behaviors from carefully evolved network. *Neural Computation*, 1(2): 355-363, 1989.
- CONNOLLY Ch., GRUPEN R.: Applications of Harmonic Functions to Robotics, *Univ. of Massachusetts, Amherst*, 1992.
- EHRENBERGER Z, Kolíbal Z.: *Průmyslové Roboty III*, Nakladatelství VUT Brno, Brno, 1993.
- GRUPEN R.A.: A Developmental Organization for Robot Behavior, *University of Massachusetts, Amherst*, 2004.
- HILDEBRAND M.: Symmetrical gaits of horses. *Science*, 1965.
- HARDARSSON F.: Stability analysis and synthesis of statically balanced walking for quadruped robots, *Doctoral Thesis, Royal Institute of Technology, Stockholm*, 2002
- HOFF J., BEKEY G.A.: A cerebellar Approach to Adaptive Locomotion for Legged Robots, *Proc. Int. Symposium on Computational Intelligence in Robotics and Automation (CIRA'97)*, 1997.
- HUBER M., GRUPEN R.: A Hybrid Discrete Event Dynamic Systems Approach to Robot Control, *Technical Report #96-43, Univ. of Massachusetts*, 1996.
- HUBER M.: A Hybrid Architecture for Adaptive Robot Control. *Ph.D. Dissertation, Univ. of Massachusetts, Amherst*, 2000.

- JEFFERSON A., Coelho Jr., Grupen R.: Effective Multifingered Grasp Synthesis, Technical Report, Univ. of Massachusetts, Amherst, 1994.
- ISHIGURO A., WATANABE Y., UCHIKAWA Y.: An Immunological Approach to Dynamic Behavior Control for Autonomous Mobile Robots, Proc. Int. Conf. on Intelligent Robots and Systems (IROS '95), 1995.
- JEONG K., YANG T., OH J.: A study on the support pattern of a quadruped walking robot for aperiodic motion, In. Proc. IROS, str. 308-313, Pittsburgh, PA, 1995.
- KOPINEC M.: Návrh a realizace senzorické soustavy náklonu kráčivého mobilního robota, diplomová práce, Ústav mechaniky těles, mechatroniky a biomechaniky, Fakulta strojního inženýrství, VUT v Brně, 2004.
- KRČEK P., DVOŘÁK J.: Plánování Cesty Autonomního Lokomočního Robotu na Základě Strojového Učení, disertační práce, Vysoké Učení Technické v Brně, Brno, 2010.
- KREJSA J., VĚCHET S.: Rapidly Exploring Random Trees Used for Mobile Robots Path Planning. Inženýrská Mechanika 12 (4) (2005) pp. 231-238, 2005
- MACDONALD W.S.: Legged Locomotion over Irregular Terrain using the Control Basis Approach, Master Project, University Of Massachusetts, Amherst, 1996.
- MACDONALD W.: Design and implementation of four-legged walking robot, BS thesis, ECE dept., MIT, Amherst, 1994.
- MCGHEE R.B., Frank A. A.: On the stability properties of quadruped creeping gaits, J. Mathematical Biosciences, 1968.
- MARÍK, V. a kol.: Umělá inteligence 1, Academia, Praha, 1993.
- MOCHIDA T., Ishiguro A., Aoki T., Uchikawa Y.: Behavior arbitration for autonomous mobile robots using emotion mechanisms, Proc. Int. Conf. on Intelligent Robots and Systems (IROS '95), 1995.
- ONDROUŠEK V.: Návrh a implementace koordinačního mechanismu řídících členů chůze čtyřnohého robota, diplomová práce, Vysoké Učení Technické v Brně, Fakulta Strojního Inženýrství, 2004.
- PETTERSSON L., Jansson K., Rehbinder H., Wikander J.: Behavior-based Control of a Four Legged Walking Robot, Department of Machine Design, Royal Institute of Technology, Sweden, 1998.
- QUINN R. D., Nelson G. M., Bachmann R. J., Kingsley D. A., Offi J., Ritzmann R. E.: Insect Designs for Improved Robot Mobility, 4th International Conference On Climbing and Walking Robots, Berns and Dillmann eds., 69-76, 2001.

- RAI L., Kang S.J., Rim K.: Dynamic Skill-Based Software Architecture for Locomotion Control in Shape Reconfigurable Mobile Robots, Proc. IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems and 2nd Int. Workshop on Collaborative Computing, Integration, and Assurance (SEUS-WCCIA'06), 2006.
- RIDDERSTRÖM CH.: Legged locomotion: Balance, control and tools - from equation to action, Royal Institute of Technology, Stockholm, 2003.
- ŘEŘUCHA V.: Inteligentní řízení krácejícího robota, Habilitační práce, Vojenská Akademie v Brně, 1997.
- SERIŠ R.: Využití ODE pro Sestavení Dynamického Modelu Čtyřnohého Robotu, bakalářská práce, VUT v Brně, Brno, 2008.
- SHAVLIK J., DIETTERICH, T.: Readings in Machine Learning, San Francisco, Morgan Kaufmann, pp. 267-283, 1990.
- SCHREIBER P.: Návrh Automatického Generátoru Prostředí Pro Mobilí Robot, bakalářská práce, VUT v Brně, Brno, 2008.
- SPENNEBERG D. a kol., Aramies: A Four-Legged Climbing and Walking Robot, Faculty of Mathematics and Computer Science, University of Bremen, 2005.
- THRUN S., Wolfram B., Dieter F.: A Real-Time Algorithm for Mobile Robot Mapping With Applications to Multi-Robot and 3D Mapping, proc. on IEEE International Conference on Robotics and Automation, San Francisco, 2000.
- U.S.ARMAY: Logistical vehicle off-road mobility. Fort Eustis, Va.: U.S. Army, Transportation Combat Developments Agency, feb 1967. (Project TCCO 62-5).
- VĚCHET S., KREJSA J., GREPL R.: How to build a robot with no money, merkur, lego and old stepper motor, proc. of Engineering Mechanics 2005, Praha, Institute of Thermomechanics, Theoretical and Applied Mechanics, Academy of Sciences of Czech Republic, 2005, pp. 335-336, ISBN: 80-85918-93-5.
- VĚCHET S., KREJSA J., ONDROUŠEK V.: The Development of Autonomous Racing Robot Bender, časopis Inženýrská Mechanika, s. 277-287, 2007.
- WETTERGREEN D., PANGELS H., BARES J.: Behavior-based gait execution for the Dante II walking robot, Proc. Int. Conf. on Intelligent Robots and Systems (IROS '95), 1995.
- WATKINS, C. J. C. H.: Learning from delayed rewards, PhD thesis, University of Cambridge, Cambridge, England, 1989.
- WITTEN, I. H.: An adaptive optimal controller for discrete-time Markov environments, Information and Control, 34:286-295, Academic Press, Londýn, 1977, ISSN 0019-9958.

WOLFRAM B., a kol.: Experiences with an interactive museum tour-guide robot, J. of Artificial Intelligence, 1999.

Internetové zdroje:

DVOŘÁK J.: Přednášky předmětu Expertní systémy, dostupné z www.uai.fme.vutbr.cz/~jdvorak/vyuka/es/es.htm, cit. dne 5.3.2011.

NILSSON J.: Introduction to Machine Learning, Stanford University, 2005. Dostupné z <http://robotics.stanford.edu/~nilsson/MLBOOK.pdf>

Open Dynamics Engine, dozstupné z www.ode.org, citováno dne 20.5.2011.

Robot EQ3-KT, fa Lynxmotion, dostupné z www.lynxmotion.com, cit. dne 12.10.2008

FREESE M.: Virtual Robot Experimentation Platform, dostupné z www.v-rep.eu, cit. dne. 5.3.2011.

WOERGOETTER F., Porr B.: Reinforcement learning, Scholarpedia, 3(3):1448, 2008, dostupné z www.scholarpedia.org/article/Reinforcement_learning, cit. dne 8.3.2011

PALČÁK F.: Možnosti využití programu ADAMS pro studium, výuku i výzkum, odborná přednáška, FSI VUT, 21.3.2011, dostupné z mechatronika.fme.vutbr.cz/podporaodbornehovzdelavaniclenutymu/prednaskydetail/4

AUTOROVO CV

Osobní údaje

Ing. Vít Ondroušek, narozen 27.4.1981 v Brně,

Pracoviště

Ústav automatizace a informatiky
Fakulta strojního inženýrství
Vysoké učení technické v Brně
Technická 2896/2
616 69 Brno

Funkce na pracovišti

technický pracovník

Vzdělání a akademická kvalifikace

1999-2004: Vysoké Učení Technické v Brně, Ústav Automatizace a Informatiky,
1995-1999: Gymnázium Vídeňská 47, obor informatika

Přehled zaměstnání

2008-dosud: technický pracovník, Ústav automatizace a informatiky, Fakulta strojního inženýrství, Vysoké učení technické v Brně

Pedagogická činnost

Výuka cvičení předmětů: Matematické základy informatiky, Algoritmy umělé inteligence, Informatika 1. Vedení bakalářských a diplomových prací.

Vědeckovýzkumná činnost

2005-2011: Simulační modelování mechatronických soustav. Výzkumný záměr MSM 0021630518

2005-2008: Simulace mechanické funkce vybraných prvků lidského těla, KONTAKT 1P05ME789

2004: Výzkum a vývoj mechatronických soustav. Výzkumný záměr MSM 262100024

ABSTRACT

The Ph.D. thesis is focused on using the reinforcement learning for four legged robot control. The main aim is to create an adaptive control system of the walking robot, which will be able to plan the walking gait through Q-learning algorithm. This aim is achieved using the design of the complex three layered architecture, which is based on the DEDS paradigm. The small set of elementary reactive behaviors forms the basis of proposed solution. The set of composite control laws is designed using simultaneous activations of these behaviors. Both types of controllers are able to operate on the plain terrain as well as on the rugged one. The model of all possible behaviors, that can be achieved using activations of mentioned controllers, is designed using an appropriate discretization of the continuous state space. This model is used by the Q-learning algorithm for finding the optimal strategies of robot control. The capabilities of the control unit are shown on solving three complex tasks: rotation of the robot, walking of the robot in the straight line and the walking on the inclined plane. These tasks are solved using the spatial dynamic simulations of the four legged robot with three degrees of freedom on each leg. Resulting walking gaits are evaluated using the quantitative standardized indicators. The video files, which show acting of elementary and composite controllers as well as the resulting walking gaits of the robot, are integral part of this thesis.