XXXIII JUGOSLOVENSKA KONFERENCIJA ETAN-a, N O V I S A D, 12—17. JUNA 1989.

Renata Zupanc, Tanja Urbančič
Institut "Jožef Stefan"
Jamova 39, 61111 Ljubljana

POSKUS UPRAVLJANJA PREPROSTEGA DINAMICNEGA SISTEMA Z METODAMI UMETNE INTELIGENCE

AN EXPERIMENT IN SIMPLE DYNAMIC SYSTEM CONTROL USING ARTIFICIAL INTELLIGENCE METHODS

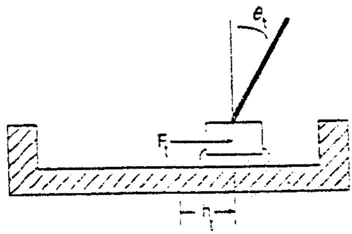
POVZETEK - Reprezentativen primer dinamičnega sistema je invertirano nihalo (vrtljivo vpeta palica na vozičku), ki so ga poskušali krmiliti s klasičnimi analitičnimi metodami, pa tudi z uporabo nekaterih metod umetne inteligence. V referatu je podano upravljanje sistema z uporabo metode preiskovanja prostora stanj, dobljenega iz modela sistema na osnovi intervalske aritmetike. Na generiranih kontrolnih pravilih smo uporabili sistema za induktivno učenje ASISTENT 86 in GINESYS PC. Rezultati kažejo, da dosežemo precej dobro upravljanje že s preprostim modelom. Kritičen je problem razbitja parametrov na intervale. Upravljanje se pri večji globini preiskovanja prostora stanj presenetljivo poslabša - vzrok za to je v večanju napak z globino pri intervalski aritmetiki.

ABSTRACT - The pole-balancing task is often used as an example of the inherently unstable dynamic system. It has been used to demonstrate modern and classical control-engineering techniques. This paper describes an experiment of controlling the pole-cart system. We use a method of searching a state-space, derived from the model based on interval arithmetic. In addition two inductive learning systems, ASSISTANT 86 and GINESYS PC were used to build decision rules from generated control rules. The results show that even with a simple model a rather good control can be achieved. The problem of partitioning of parameter values into intervals is critical. Surprisingly the control gets worse with deeper search because errors due to interval arithmetic grow with increasing search-depth.

1. UVOD

Problem balansiranja palice je primer nestabilnih dinamičnih sistemov, ki se kažejo v številnih primerih balansiranja (npr. pri hoji, usmerjanju raketnega izpuha itd.). Problem ni trivialen in je analitično rešljiv. Med balansiranjem palice in bolj uporabnimi problemi, kot je na primer satelitska kontrola, obstaja precej podobnosti. Zato tudi posvečajo temu problemu toliko pozornosti [5].

Problem voziček-palica je opisan takole: Na ravnem tiru omejene dolžine imamo voziček, na katerem je pritrjena palica, ki je gibljiva le okoli vodoravne osi, pravokotne na tirnice. Na voziček delujemo s silo F konstantne velikosti, ki ji lahko v rednih časovnih presledkih (0.02 s) spremenimo predznak. Voziček je potrebno voditi tako, da ostane na tirnicah in da palica ne pade [4].



Slika 1: Preprosta predstavitev sistema, ki ga kontroliramo

Kvalitativno modeliranje - za razliko od klasičnega analitičnega - opisuje delovanje naprav na intuitiven, človeku razumljiv način, ob uporabi preprostega fizikalnega znanja.

V modelu so kvantitete predstavljene z majhno množico kvalitativnih, simboličnih vrednosti in njihovimi ordinalnimi relacijami. Ni točnih funkcijskih odvisnosti. Zveze so približne (npr. monotonost) in ponavadi v obliki predikatov.

Klasične metode delujejo bolje od kvalitativnih, kadar imamo dosegljiv analitični model. Kjer imamo opravka s problemi, ki so prekompleksni za klasično kontrolo (parametri, ki se spreminjajo s časom, nelinearni sistemi, ali če ne poznamo točnih numeričnih vrednosti parametrov oziroma njihovih medsebojnih odvisnosti - npr. pri fizioloških sistemih), pa se odločimo za preprostejšo, a morda manj zanesljivo kvalitativno kontrolo. Dodatna prednost kvalitativnega pristopa je možnost razlage avtomatsko generiranih odločitev za kontrolne akcije.

2. MODELIRANJE SISTEMA VOZIČEK-PALICA

2.1 Analitični model

Sistem določajo štirje parametri : položaj vozička (X), hitrost vozička (V), naklon palice (Fi) in kotna hitrost palice (FiV); sama dinamika sistema je določena s sistemom dveh diferencialnih enačb drugega reda [1, 5].

2.2 Kvalitativni model

Iz analitičnega modela smo izpeljali verjetnostni kvalitativni model. Vrednosti parametrov so enake, kot jih navaja Anderson v [1]. Le-te v kvalitativnem modelu niso realna števila temveč intervali na realni osi. Sistem je modeliran na treh različnih nivojih podrobnosti. Nivoji se ločijo po tem, na koliko intervalov je razdeljeno dovoljeno območje za posamezni parameter. Model povezuje trenutno stanje sistema (vektor štirih parametrov), akcijo, novo stanje in verjetnost prehoda v novo stanje. Akcija je lahko pozitivna ali negativna (glede na smer sile na voziček).

2.3 Računanje verjetnosti prehodov z intervalsko aritmetiko

Pri nenatančni informaciji (intervalske vrednosti parametrov) je stopnja nedeterminizma zelo velika: istemu kvalitativnemu stanju v trenutku T lahko pri isti akciji sledijo v času T + ΔT različna nova kvalitativna stanja. Stopnjo nedeterminizma lahko zmanjšamo s tem, da za vsako od možnih novih stanj povemo verjetnost, da se bo sistem znašel v njem. Vsakemu prehodu vrednosti parametra je pridružena ocena verjetnosti za to vrednost. Oceno verjetnosti za prehod celotnega stanja (X, V, Fi, FiV) v novo stanje dobimo kot produkt ocen verjetnosti za prehode vrednosti posameznih parametrov, čeprav količine niso neodvisne. Pri izračunu verjetnosti za posamezne parametre predpostavimo v intervalski aritmetiki njihovo enakomerno porazdelitev. Upoštevamo le tiste prehode, katerih verjetnost je večja od 0.01.

Vrednosti posameznih parametrov novega kvalitativnega stanja se torej izračunajo po enačbah dinamike sistema z intervalsko aritmetiko. Za primer si oglejmo računanje vrednosti novega položaja. Izhajamo iz znane Eulerjeve enačbe

$$x(t + Dt) = XN = x(t) + Dt\, \dot{x}(t)$$

Dobiti želimo XN kot vsoto X in DX, pri čemer je DX produkt Dt in V:

$$XN = X + Dt\ V$$

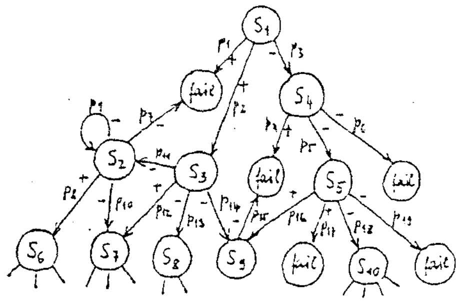
X, V, DX in XN so intervali, Dt pa konstanta. Torej

```
x( X, V, Dt, XN\P) :-
    mult_int_c( V, Dt, DX),        % množenje intervala s konstanto
    plus_int( X, DX, XN\P)         % seštevanje dveh intervalov
    P >= 0.01.                     % verjetnosti zanj več kot 0.01
```

Ker imamo opravka s funkcijami, ki so na danih intervalih monotone, se intervalska aritmetika prevede na dokaj enostavno računanje vrednosti mej intervalov. Poskrbeti je treba, da nova vrednost parametra ni kar poljuben interval, temveč ena od možnih kvalitativnih vrednosti za ta parameter (eden izmed določenih intervalov). Ker lahko več kvalitativnih vrednosti pokriva izračunani interval, je skupna ocena verjetnosti (ki je enaka 1) mednje razdeljena v razmerju dolžin njihovih presekov z izračunanim intervalom (predpostavka enakomerne porazdelitve verjetnosti).

2.4 Prostor stanj

S simulacijo kvalitativnega modela s silo velikosti 1 N in 10 N smo generirali vsa možna naslednja stanja s pripadajočimi verjetnostmi, pri znanem trenutnem stanju in akciji, na izbranem nivoju natančnosti. Na ta način smo dobili 2 x 3 = 6 grafov prehajanja stanj. Vsak graf predstavlja prostor stanj.



Slika 2: Del grafa prehajanja stanj

3. GENERIRANJE KONTROLNE STRATEGIJE NA PODLAGI VERJETNOSTNEGA KVALITATIVNEGA MODELA

Na osnovi preiskovanja tako dobljenega grafa smo izpeljali kontrolna pravila tako, da se v vsakem stanju algoritem odloči za akcijo, pri kateri je matematično upanje preko vseh stanj ugodnejše glede na izbrani kriterij. Izbrali smo naslednja kriterija: maksimizirati ocenjeno dobo preživetja (to strategijo imenujemo KAPA) in maksimizirati ocenjeno verjetnost preživetja N korakov (strategija IOTA).

Doba preživetja je ocenjena z rekurzivno funkcijo:

$$\mathrm{kapa}(S,0) = \begin{cases} 0 & \text{če terminal}(S) \\ 1 & \text{sicer} \end{cases}$$

$$\mathrm{kapa}(S,N) = \begin{cases} \max\limits_{A\in\{+,-\}}\Big(1 + \mathrm{SUMA}_{t(S,A,S_i,p_i)}\; p_i \; \mathrm{kapa}(S_i,N-1)\Big) & \text{če non-terminal}(S) \\ 0 & \text{če terminal}(S) \end{cases} \qquad (1)$$

kjer četverka $t(S, A, S_i, p_i)$ označuje i-tega naslednika stanja S z verjetnostjo p_i pri akciji A. Terminalna stanja so stanja, pri katerih vsaj eden izmed parametrov pade izven dovoljenega območja. Vsa ostala stanja so neterminalna stanja.

Verjetnost preživetja N korakov je ocenjena s podobno rekurzivno funkcijo:

$$\mathrm{iota}(S,0) = \begin{cases} 0 & \text{če terminal}(S) \\ 1 & \text{sicer} \end{cases}$$

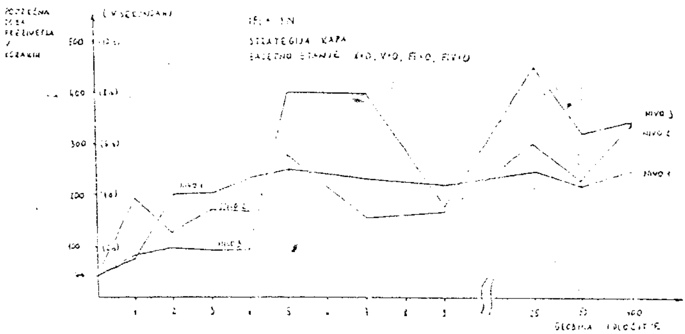
$$\mathrm{iota}(S,N) = \begin{cases} \max\limits_{A\in\{+,-\}} \mathrm{SUMA}_{t(S,A,S_i,p_i)}\; p_i \; \mathrm{iota}(S_i,N-1) & \text{če non-terminal}(S) \\ 0 & \text{če terminal}(S) \end{cases} \qquad (2)$$

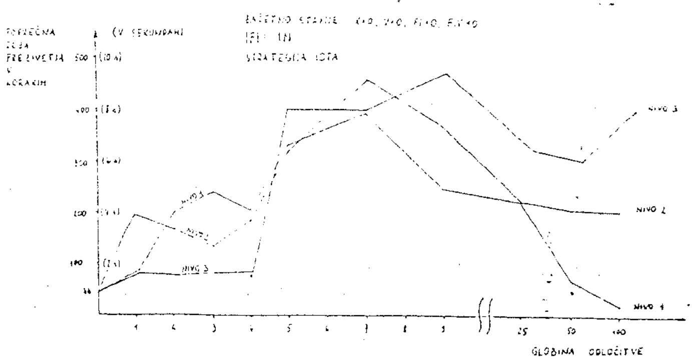
Pri obeh ocenah je pomemben parameter globina preiskovanja grafa. Izbrali smo deset globin: 1, 2, 3, 4, 5, 7, 9, 25, 50 in 100. Z uporabo obeh kriterijev na vseh šestih grafih (dve sili, trije nivoji) dobimo 10 x 2 x 6 = 120 kontrolnih pravil (tabele stanje-akcija).

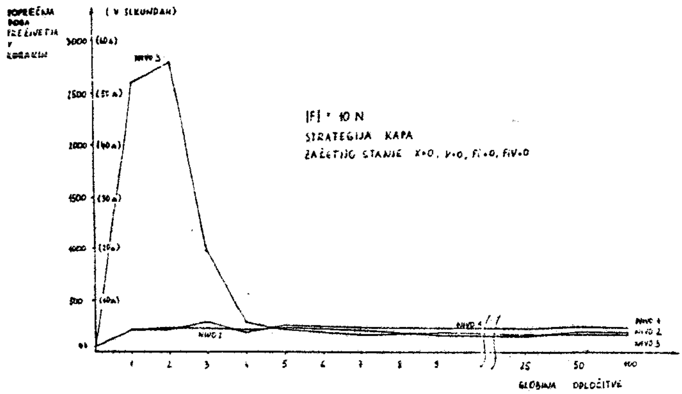
4. REZULTATI TESTIRANJA KONTROLNIH PRAVIL

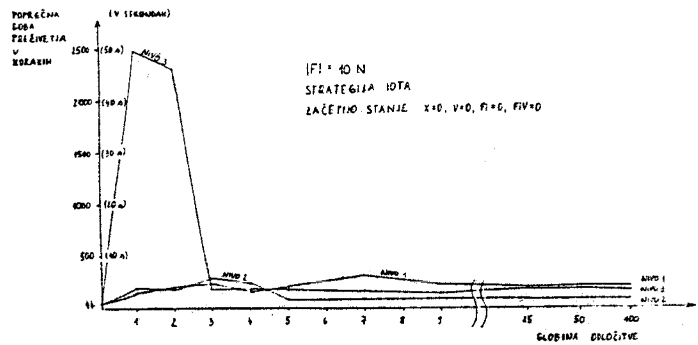
Namesto dejanske naprave smo uporabili numerični simulator, ki uporablja eksaktne analitične enačbe dinamike sistema. Vsa pravila so testirana s silo 1 N na fiksnem naboru 10 izbranih začetnih stanj. Za vsako začetno stanje se izračunata tudi poprečna doba preživetja in standardna deviacija.

Poprečno dobo preživetja v odvisnosti od globine preiskovanja grafa prikazujejo diagrami na sliki 3a, b, c, d:









Slika 3a, b, c, d: Poprečna doba preživetja v odvisnosti od globine odločitve

Iz diagramov smo ugotovili naslednje:

- **Izbira intervalov:** pravilna izbira števila intervalov in njihovih mej za posamezni parameter je kritična pri sili 10 N in majhni globini preiskovanja; krmiljenje na nivoju tri preživi 10 do 16 krat dlje kot na nivoju ena ali dva; kontrola na nivoju ena se ne razlikuje signifikantno od kontrole na nivoju dva - pri obeh se kvaliteta kontrole zelo spreminja z globino preiskovanja.
- **Stabilnost kvalitete upravljanja:** Pri sili 10 N se kvaliteta upravljanja pri gledanju več kot štiri korake naprej stabilizira in je približno dvakrat boljša od naključne strategije.
- **IOTA : KAPA:** Razlike med strategijama IOTA in KAPA pri sili 10 N niso velike
- **Pretiravanje s silo:** dobljena s silo 10 N so primerna tudi za kontrolo s silo 1 N (na nivoju tri pri velikih globinah preiskovanja celo bolj kot kontrolna pravila dobljena s silo 1 N), kar kaže na določeno robustnost sistema.
- **Maksimum:** Pri globini pet do osem je pri strategiji IOTA opazen maksimum kvalitete obnašanja (najbolj izrazit za nivo ena; na nivoju tri komaj opazen). Velika globina preiskovanja grafa ima za posledico kvečjemu slabše rezultate pri kontroli. Zakaj ?
  * napake se pri intervalski aritmetiki z globino večajo (intervali povečujejo napake pri izračunavanju verjetnosti),
  * potrebno je sprejeti kompromis med napakami pri izračunavanju (intervalska aritmetika) in globino preiskovanja grafa (na nivoju tri je kvaliteta upravljanja na globinah 1 in 2 znatno boljša - faktor 10 do 15 - od kvalitete upravljanja na večjih globinah). Verjetnosti v grafu prehodov so brezpogojne, pri gledanju v globino pa bi morali upoštevati pogojne verjetnosti (zgodovino).

5. ODLOČITVENA PRAVILA, DOBLJENA S SISTEMOMA ZA INDUKTIVNO UČENJE ASISTENT 86 IN GINESYS PC

Sistema za induktivno učenje, ASISTENT 86 in GINESYS PC, smo uporabili na nekaj izbranih tabelah stanje-akcija iz poglavja tri. Vsak element tabele stanje-akcija je učni primer. Iz primerov ASISTENT 86 zgradi odločitveno drevo, GINESYS PC pa odločitveno pravilo, sestavljeno iz več IF-THEN pravil [2, 3].

Analiza odločitvenih dreves in pravil pa pokaže:

- **Kompleksnost:** Odločitvena drevesa so bolj kompleksna od pravil, ki jih generira GINESYS PC, in tudi bolj asimetrična.
- **Enostavni in sestavljeni pogoji:**
  * pri vrhu drevesa so pokrite očitne situacije, ki jih je potrebno neposredno krmiliti. Enako velja za pravila na vrhu odločitvenega pravila, ki so si podobna. Primer:

    (FiV = Pos) ==> Pos_Act

    Če se palica giblje v pozitivni smeri, izvrši pozitivno akcijo.
  * pri sili 10 N v pravilih nastopajo skoraj izključno enostavni pogoji, pri sili 1 N pa imamo večinoma konjunktivno sestavljene pogoje. Taki sestavljeni pogoji pokrivajo bolj zapletene situacije, na primer:

    (Fi = Near_Pos) & (FiV = Neg) ==> Neg_Act

    Če je palica malo nagnjena od navpičnice in se giblje proti njej, jo je potrebno zaustaviti.

- **Krmiljenje položaja:** Položaj vozička je mogoče krmiliti šele pri veliki globini preiskovanja prostora stanj. Parameter X nastopa le pri dnu dreves oz. na dnu hierarhije pravil.
- **Simetrija pravil:** Zgornji del hierarhije pravil je simetričen: za vsako pravilo obstaja "zrcalno" pravilo, kjer v pog... nastopajo parametri z nasprotno predznačenimi kvalitativnimi vrednostmi in ki pokrivajo nasprotno predznačeno akcijo kot prvotno pravilo.

6. ZAKLJUČEK IN RAZPRAVA

Raziskovanje v smeri preiskovanja prostora stanj v globino je koristno, ker lahko iz poskusov povzamemo več zanimivih ugotovitev. Obstaja kompromis med globino preiskovanja in napaka i zaradi intervalske aritmetike. Precej dobro upravljanje doseže že s preprostim modelom. Položaj vozička postane pomemb šele pri večji globini preiskovanja prostora stanj. Pretir anje s silo je koristno (generirana pravila s silo velikosti 10, upravljanje s silo velikosti 1).

Za nadaljnje delo bi bilo zanimivo preučiti vpeljavo aposteriornih verjetnosti prehodov stanj, ki bi jih dobili z večkratnim krmiljenjem, in jih primerjati z apriornimi. Verjetno se da dosti pridobiti z opustitvijo predpostavke enakomerne porazdelitve verjetnosti po intervalih ter z upoštevanjem pogo nih verjetnosti oziroma zgodovine, pa tudi s planiranjem in učen em na podlagi izkušenj. Raziskave v teh smereh so že v teku.

LITERATURA

[1] Anderson, C. W. (1987) Strategy Learning with Multilayer Connectionist Representations, v Langley, P. (ur.), Proceedings of the 4th International Workshop on Machine Learning, Morgan Kaufmann

[2] Karalić, A. (1988) Implementacija sistema GINESYS PC, diplomska naloga, Fakulteta za elektrotehniko, Ljubljana

[3] Kononenko, I. (1985) Razvoj sistema za induktivno učenje ASISTENT, magistrsko delo, Fakulteta za elektrotehniko, Ljubljana

[4] Michie, D. & Chambers, R. A. (1971) Boxes: An Experiment In Adaptive Control, Machine Intelligence 2, Edinburgh University Press, stran 137-152

[5] Sammut, C. (1988) Experimental Results from an Evaluation of Algorithms that Learn to Control Dynamic Systems, Proceedings of 5th International Machine Learning Conference, Ann Arbor, Michigan, Morgan Kaufmann