Operacijska istraživanja

10. predavanje: Dinamičko programiranje i Markovljevi procesi odlučivanja

Sažetak predavanja

- Dinamičko programiranje
  - Općeniti principi
    - (bes)/konačni horizont
- Determinističko dinamičko programiranje
  - Zagonetke
  - Problem naprtnjače
- Stohastičko dinamičko programiranje
- Računalne poteškoće

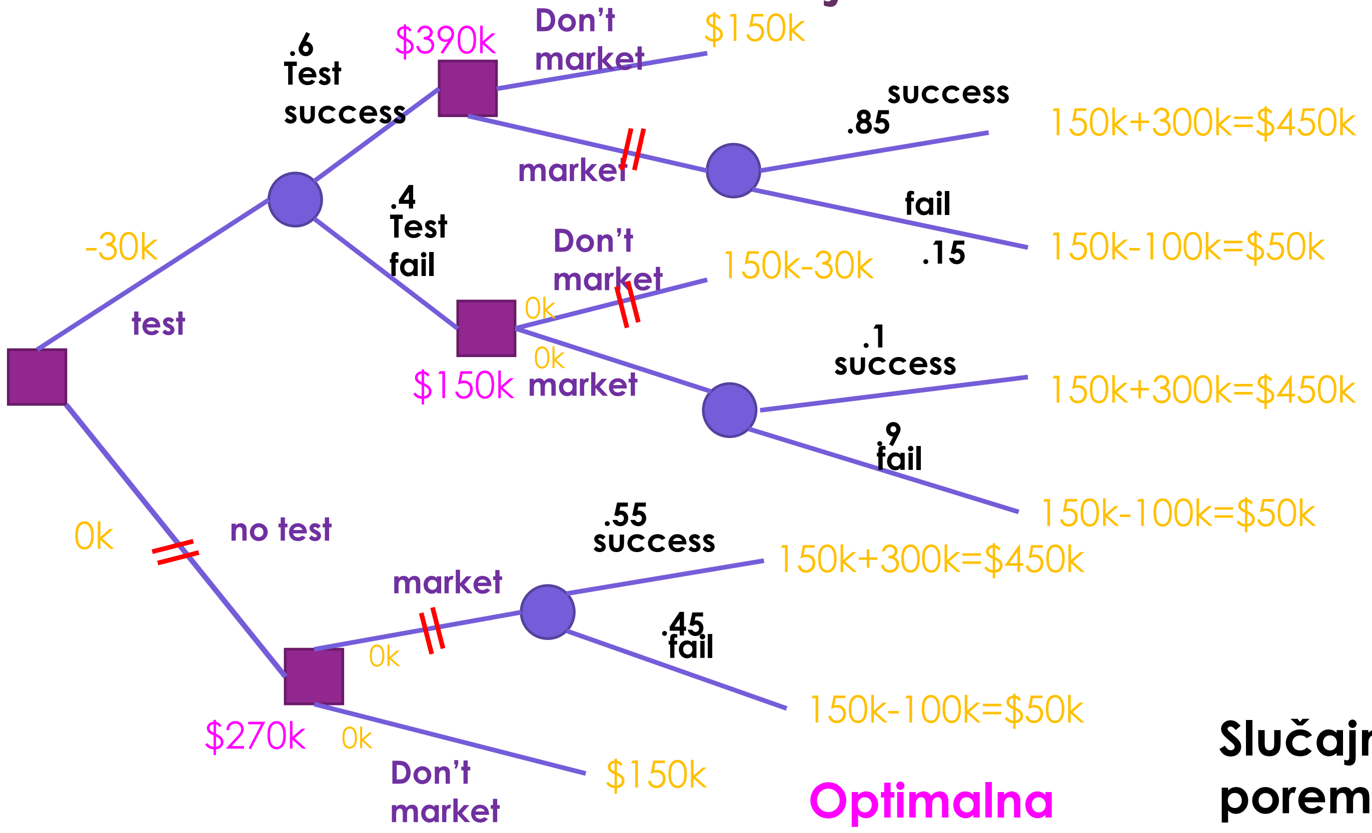
Dinamičko programiranje

- Optimizacijski problem se može podijeliti u niz manjih problema
  - Rekurzija sa memoizacijom
- „Jednostavnije” rješavanje problema
- Rješenja manjih problema se mogu koristiti u većim problemima
  - Korištenje memoriziranih rješenja smanjuje vrijeme izračuna
    - Redukcija u složenosti

Općeniti principi

- Elementi dinamičkog programa
  - Faze
  - Stanja
  - Odluke
  - Tranzicije
  - Izravne nagrade
  - Rekurzija za optimalnu funkciju vrijednosti
  - Optimalna strategija
  - Dodatno, ovisno o specifikama
    - Npr. U stohastičkom slučaju, slučajni poremećaji

Stabla odlučivanja



Optimalna funkcija vrijednosti – vrijednost za svako stanje

Slučajni poremećaji – dolaze kod grananja iz čvorova događaja

- **Faze**
- **Stanja** – čvorovi odluka (kvadratići) i terminalne grane
- **Odluke** – grananja iz stanja
- **Tranzicije** – funkcija koja izračunava koje je sljedeće stanje. Sve između stanja.
- **Izravne nagrade** – realiziraju se sa svakom odlukom i terminalnim stanjme
- **Optimalna strategija** – najbolja odluka za svako stanje (osim terminalnog)

Općeniti principi I - elementi

1. Problem se može podijeliti u **faze** k=1…N sa odlukama u svakoj fazi
   - Konačni/beskonačni horizonti problema
2. Skup stanja S_k vezan uz fazu k
   - Stanje $x_k \in S_k$– informacija potrebna za svaku optimalnu odluku
   - Iz trenutnog stanja, optimalna odluka za ostatak problema ne ovisi o prošlosti - **Markovost**
     - **Princip optimalnosti**

Općeniti principi II - elementi

3. Postoji skup C_k mogućih **odluka** vezanih uz svaku fazu k
   - Odluke $u_k \in U_k(x_k) \subset C_k$
   - $U_k(x_k)$ – skup dopuštenih odluka u stanju x_k
4. **Tranzicijska** funkcija f_k definira transformaciju iz trenutnog stanja x_k pod utjecajem trenutne odluke u_k u sljedeće stanje (u narednoj fazi)
   - $f_k: S_k \times C_k \rightarrow S_{k+1}$ (determinističko)
5. Izravna nagrada/trošak dan prema fazama
   - $g_k: S_k \times C_k \rightarrow \mathbb{R}$, k=1,…,N-1
   - $g_N: S_N \rightarrow \mathbb{R}$
   - Ukupni trošak mora biti jednak agregaciji izravnih troškova

Općeniti principi III - elementi

4. Rekurzivna formula koja povezuje trošak tijekom stanja u fazi k sa troškom u fazi k+1

- *(deterministički)*
- $J_k^*(x_k) = min_{u_k}\{g_k(x_k, u_k) + J_{k+1}^*(x_{k+1} = f_k(x_k, u_k))\}$
- *(stohastički)*
- $J_k^*(x_k) = min_{u_k}\mathbb{E}\{g_k(x_k, u_k, w_k) + J_{k+1}^*(x_{k+1} = f_k(x_k, u_k, w_k))\}$
  - (Bellmanove jednadžbe optimalnosti)
- J_k^* se zove optimalna **funkcija vrijednosti**
  - Pod pretpostavkom da je agregacija **suma** i optimizacijski kriterij pod neizvjesnosti očekivana vrijednost

Općeniti principi IV - elementi

5. **optimalna strategija/politika** $\pi^*=\{\mu_1^*,\ldots, \mu_{N-1}^*\}$ koja je sekvenca funkcija $\mu_k^* : S_k \rightarrow C_k$, pravila odluka u svakoj fazi k

- *(determinističko)*
- $\mu_k^*(x_k) = argmin_{u_k}\{g_k(x_k, u_k) + J_{k+1}^*(x_{k+1} = f_k(x_k, u_k))\}$
- *(stohastično)*
- $\mu_k^*(x_k) = argmin_{u_k}\mathbb{E}\,\{g_k(x_k, u_k) + J_{k+1}^*(x_{k+1} = f_k(x_k, u_k, w_k))\}$

Općeniti principi V - elementi

- Problemi sa konačnim horizontom
  - Rješavanje unatražnom indukcijom preko faza(N…1)
- Problemi sa beskonačnim horizontom
  - Valjano definirano za specijalne formulacije
    - Snižene nagrade
    - Prosječne nagrade
  - Procedure za rješavanje:
    - Iterativne procedure
      - LP
      - Iteriranje strategija
      - Iteriranje vrijednosti

Problem I - zagonetke

- Pretpostavimo da imamo 30 šibica na stolu. Počinjem uzimanjem 1, 2, ili 3 šibice. Zatim moj protivnik mora uzeti 1, 2, ili 3 šibice. Nastavljamo na taj način dok se zadnja šibica ne pokupi. Igrač koji pokupi zadnju šibicu je izgubio igru. Kako mogu ja (prvi igrač) biti siguran da ću pobijediti u igri?

Riješ.: onaj tko se nađe na izboru sa 1 preostalom šibicom je izgubio. Unatražno, 1. igrač uvijek može isforsirati da jedan „ciklus" potroši **4 šibice**, što znači da 1.igrač želi ostavljati 2. igraču neki broj 1+4*k, što kulminira sa 29. Dakle, igrač 1 počinje sa uzimanjem jedne šibice i ima zajamčenu pobjedu forsirajući cikluse od 4 šibice od tad nadalje.

Problem II – alokacija resursa

- Ograničeni resursi (budžet) alocirani preko nekoliko aktivnosti (ulaganja, prilike)
  - LP – pretpostavke
    1. Alokacije su ne-negativni brojevi
    2. Ostvarena korist iz aktivnosti je **proporcionalna** dodijeljenim resursima
    3. Ukupna korisnost iz aktivnosti je suma korisnosti individualnih aktivnosti
  - DP – pretpostavke
    1. Alokacije aktivnostima su članovi **konačnih skupova**
    2. Ukupna korisnost aktivnosti je **agregacija** korisnosti iz individualnih aktivnosti

Problem II – alokacija resursa

- Finco ima \$6,000 za investiranje, itri investicije za odabir. Ako je s_j dolara (u tisućama) investirano u investicije j, tada je ostvaren netto povrat (u tisućama) $r_j(s_j)$, gdje su $r_j(s_j)$ sljedeći:

LP
- Cjelobrojne alokacije
- Proporcionalnost

- $r_1(s_1) = 7s_1 + 2, \ (s_1 > 0)$
- $r_2(s_2) = 3s_2 + 7, \ (s_2 > 0)$
- $r_3(s_3) = 4s_3 + 5, \ (s_3 > 0)$
- $r_1(0) = r_2(0) = r_3(0) = 0$

MILP
- Proporcionalnost

- Količina uložena u svako ulaganje mora biti u cjelobrojnim tisućama \$. Kako Finco treba alocirati \$6,000 da ostvari najveći netto povrat?

Problem II – alokacija resursa

$\max\{r_1(s_1) + r_2(s_2) + r_3(s_3)\}$

s.t. $s_1 + s_2 + s_3 = 6$

s_j ne-negativni integer (j = 1, 2, 3)

Faze k – Alokacija u investicije 1,2,3: $k \in \{1,2,3\}$

Stanja x_k - Dostupan novac za investicije (u $1000-ama) $\forall k$, $x_k \in \{0,1,2,3,4,5,6\}$

Odluke u_k- Alokacije u investicije k, sa ostatkom za ostala ulaganja

Transition function f_k : $x_{k+1}=f_k(x_k,u_k)=x_k-u_k$

Izravna nagrada $g_k(x_k,u_k)=r_k(u_k)$

Rekurzija za optimalnu funkciju vrijednosti

$$J_k^*(x_k) = max_{u_k}\{r_t(u_k) + J_{k+1}^*\big(x_{k+1} = f_k(x_k, u_k)\big)\}, \forall k$$

$$J_4(x) = 0$$

Optimalna funkcija vrijednosti **u ovom slučaju** može biti pohranjena u $|S_k| \times |\{1,2,3\}|$ tablicu

Problem II – alokacija resursa

Stanje x_k	$J_3^*(x_3)$; $\mu_3^*(x_3)$	$r_2(u_2)$	$J_2^*(x_2)$; $\mu_2^*(x_2)$
0	0; (0)	0	
1	9; (1)	10	
2	13; (2)	13	
3	17; (3)	16	
4	21; (4)	19	
5	25; (5)	22	
6	29; (6)	25	

$$J_2^*(x_2) = max_{u_2}\{r_2(u_2) + J_3^*(x_3 = x_2 - u_2)\}$$

Koristimo samo zadnja dva stupca r_2 i J^*_3 za izračun novog stupca!

Problem II – alokacija resursa

Stanje x_k	$J_3^*(x_3)$; $\mu_3^*(x_3)$	$r_2(u_2)$	$J_2^*(x_2)$; $\mu_2^*(x_2)$
0	0; (0)	0	0; (0)
1	9; (1)	10	10; (1)
2	13; (2)	13	19; (1)
3	17; (3)	16	23; (1)
4	21; (4)	19	27; (1)
5	25; (5)	22	31; (1)
6	29; (6)	25	

$$J_2^*(x_2) = max_{u_2}\{r_2(u_2) + J_3^*(x_3 = x_2 - u_2)\}$$

$$J_2^*(x_2 = 6) = max_{u_2|\, u_2+x_3=6}\{r_2(u_2) + J_3^*(x_3 = 6 - u_2)\}$$

Problem II – alokacija resursa

Stanje x_k	$J_3^*(x_3)$; $\mu_3^*(x_3)$	$r_2(u_2)$	$J_2^*(x_2)$; $\mu_2^*(x_2)$
0	0; (0)	0	0; (0)
1	9; (1)	10	10; (1)
2	13; (2)	13	19; (1)
3	17; (3)	16	23; (1)
4	21; (4)	19	27; (1)
5	25; (5)	22	31; (1)
6	29; (6)	25	35; (1)

$$J_2^*(x_2 = 6) = max \begin{Bmatrix} r_2(u_2 = 0) + J_3^*(x_3 = 6), \\ r_2(u_2 = 1) + J_3^*(x_3 = 5), \\ r_2(u_2 = 2) + J_3^*(x_3 = 4), \\ r_2(u_2 = 3) + J_3^*(x_3 = 3), \\ r_2(u_2 = 4) + J_3^*(x_3 = 2), \\ r_2(u_2 = 5) + J_3^*(x_3 = 1), \\ r_2(u_2 = 6) + J_3^*(x_3 = 0) \end{Bmatrix}$$

Najbolje kombinacija 10+25=35, za u_2=1

Problem II – alokacija resursa

Stanje x_k	$J_3^*(x_3)$; $\mu_3^*(x_3)$	$r_2(u_2)$	$J_2^*(x_2)$; $\mu_2^*(x_2)$	$r_1(u_1)$	$J_1^*(x_1)$; $\mu_1^*(x_1)$
0	0; (0)	0	0; (0)	0	
1	9; (1)	10	10; (1)	9	
2	13; (2)	13	19; (1)	16	
3	17; (3)	16	23; (1)	23	
4	21; (4)	19	27; (1)	30	
5	25; (5)	22	31; (1)	37	
6	29; (6)	25	35; (1)	44	

$$J_1^*(x_1) = max_{u_1}\{r_1(u_1) + J_2^*(x_2 = x_1 - u_1)\}$$

Koristimo samo zadnja dva stupca r_1 i J^*_2 za izračun novog stupca!

Problem II – alokacija resursa

Stanje x_k	$J_3^*(x_3)$; $\mu_3^*(x_3)$	$r_2(u_2)$	$J_2^*(x_2)$; $\mu_2^*(x_2)$	$r_1(u_1)$	$J_1^*(x_1)$; $\mu_1^*(x_1)$
0	0; (0)	0	0; (0)	0	0; (0)
1	9; (1)	10	10; (1)	9	10; (0)
2	13; (2)	13	19; (1)	16	19; (0)
3	17; (3)	16	23; (1)	23	28; (1)
4	21; (4)	19	27; (1)	30	35; (2)
5	25; (5)	22	31; (1)	37	42; (3)
6	29; (6)	25	35; (1)	44	49; (4)

$$J_1^*(x_1) = max_{u_1}\{r_1(u_1) + J_2^*(x_2 = x_1 - u_1)\}$$

Problem II – alokacija resursa

Stanje x_k	$J_3^*(x_3)$; $\mu_3^*(x_3)$	$r_2(u_2)$	$J_2^*(x_2)$; $\mu_2^*(x_2)$	$r_1(u_1)$	$J_1^*(x_1)$; $\mu_1^*(x_1)$
0	0; (0)	0	0; (0)	0	0; (0)
1	9; (1)	10	10; (1)	9	10; (0)
2	13; (2)	13	19; (1)	16	19; (0)
3	17; (3)	16	23; (1)	23	28; (1)
4	21; (4)	19	27; (1)	30	35; (2)
5	25; (5)	22	31; (1)	37	42; (3)
6	29; (6)	25	35; (1)	44	49; (4)

$x_3=x_2-u_2$

$x_2=x_1-u_1$

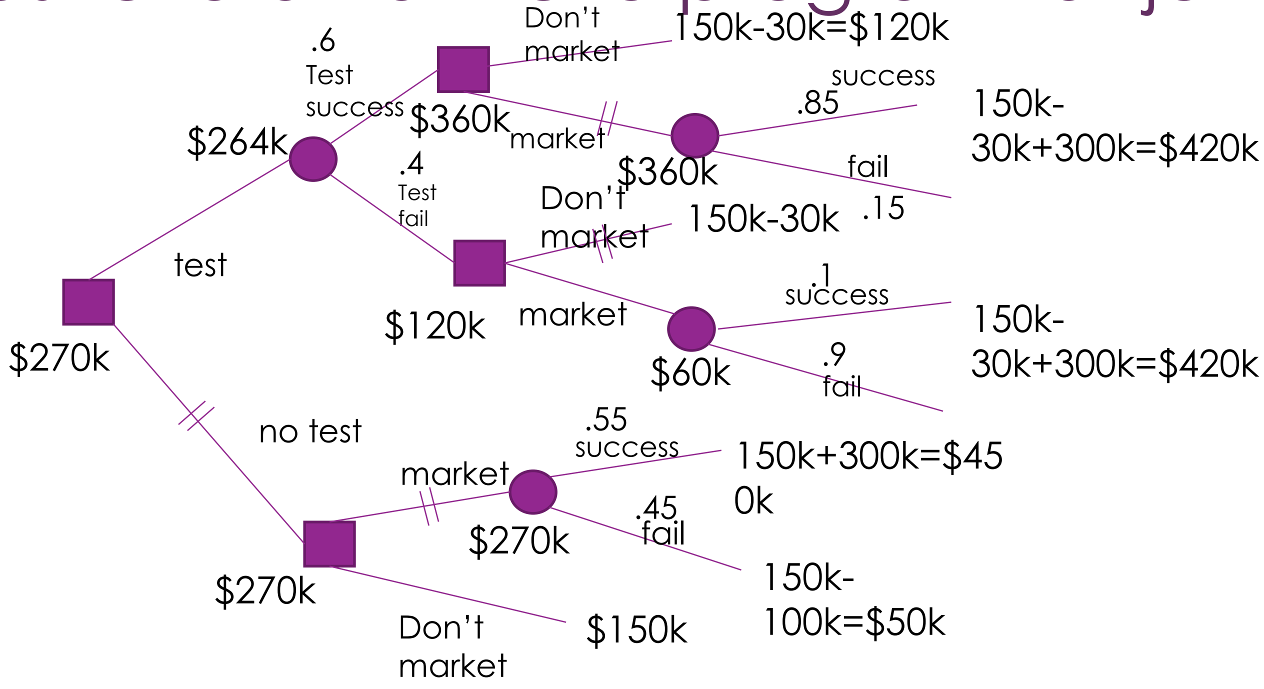
- Vrijednost optimalnog rješenja je \$49,000 za uloženih \$6,000
- Optimalna alokacija jest $u_1=4, u_2=1, u_3=1$ (vrijednosti u zagradama u tablici)

Stohastičko dinamičko programiranje

- Pristup za rješavanje Markovljevih procesa odlučivanja (MDP)
  - Kontrola stohastičkih procesa u diskretnom vremenu
  - Stohastička tranzicija između stanja

- Konačni horizont
  - Unatražna indukcija – unatražno DP
- Beskonačni horizont
  - LP
  - Iteriranje vrijednosti
  - Iteriranje strategija

Stohastičko dinamičko programiranje



$$J_k^*(x_k) = min_{u_k} \mathbb{E}\{g_k(x_k, u_k, w_k) + J_{k+1}^*(x_{k+1} = f_k(x_k, u_k, w_k))\}$$

- Konačni horizont – unatražna indukcija
  - Stablo odlučivanja
    - Čvorovi odluke ⇔ stanja
    - Bridovi iz čvorova događaja u odluke ⇔ stohastičke tranzicije
    - Očekivana vrijednost se računa u čvorovima događaja
    - Minimizacija se računa u čvorovima odluka

Računalne poteškoće

$$J_k^*(x_k) = min_{u_k} \mathbb{E}\,\{g_k(x_k, u_k, w_k) + J_{k+1}^*(x_{k+1} = f_k(x_k, u_k, w_k))\}$$

- Sve egzaktne procedure DP-a pate od 3 prokletstva **dimenzionalnosti**:

1. U prostoru odluka
    - Funkcija vrijednosti – **izračun** i **pohrana** za sva stanja
2. U prostoru odluka
    - Optimizacija preko svih odluka
3. U prostoru slučajnih poremećaja
    - Izračun očekivane vrijednosti

- Realistični problemi imaju velike dimenzionalnosti sva tri prostora