


MARKOU DÖNTÉSI FOLYAMATOK

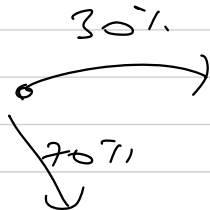
2025-09-22



VOLT

MARKOV-LÁNC:

G + Probability
at each



$$M: \begin{pmatrix} 0,3 & 0,7 \\ & \end{pmatrix}$$

MARKOV DÖNTÉI FOLYAMAT:

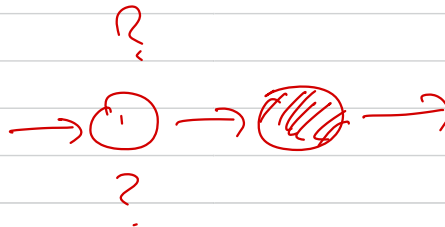
HATÁSOK VAN AZ ÁTMENETI
VÁLTOZÁSOKRA

1 kör kivétel



döntési lehetőségek:

- 50% (+1)
- 50% (0)
- 50% (+2)
- 50% (-1)



MDP

S: állapothalmaz (graf csúcsa)

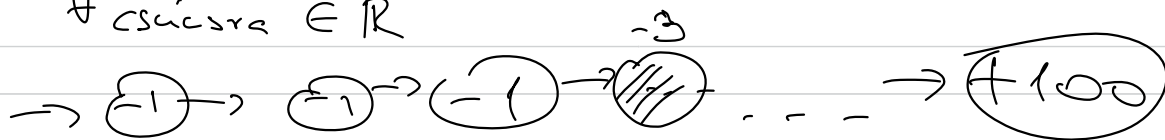
A: (action): \forall csúcsra „taktikai” lehetőségek

$\forall a \in A$: $\text{Prob}(s \rightarrow s' | a)$: ha a **(B)** opció
szint járható:
30%.

POLICY: minden csúcsra mondjuk
meg, mi az általuk

választott action: $a \in A$

REWARD: \forall csúcsra $\in \mathbb{R}$



CÉL: OLTAN STRAT, AMIVEL A VÁRHOZÓ JUTALOM MAXIMÁLIS

PELDÁK: üssék-e 200-as szavát 40/0 -val?

Robot egy labirintusban.

1. AU. Talso a legjobb stratégia megkeresésére.

VAZLAT: Kiszámítjuk, hogy melyik csiúcs mennyire jó.
(utility). Ha ez megvan, választhatunk azt a ^{politikát} ~~dobozt~~,
ami a utility-t maximálja.

$$U(s) = R(s) + \max_a \sum_{s'} P(s \rightarrow s' / a) \cdot U(s')$$

\uparrow \uparrow $\uparrow a$

csúcs reward
(instant) legjobb

$\underbrace{\sum_{s'} P(s \rightarrow s' / a) \cdot U(s')}$ köv. lépésben várható U

U -t ki lehet választani: U_0 : "random" inicializáció

$$U_1(s) := R(s) + \max_a \sum P(s \rightarrow s'_a) U_0(s')$$