

Markovovské rozhodovacie procesy

S - množina stavov

A - množina akcií

$P_{sa}(s')$ - pravdepodobnosť, že ak v stave s urobíme akciu a , tak skončíme v stave s'

$R(s)$ - odmena za dosiahnutie stavu s

γ - diskontný faktor

Odmena pri postupnosti stavov $s_0, s_1, s_2, \dots, s_k$:

$$R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \dots + \gamma^k R(s_k)$$

Rovnice pre počítanie hodnotovej funkcie $V^\pi(s)$:

$$V^\pi(s) = R(s) + \gamma \sum_{s'} P_{s, \pi(s)}(s') V^\pi(s')$$