

keter-
gan-
tun-
gan
ke-
bi-
jakan
op-
ti-
mal
pada
dis-
tribusi
state
awal
Reward

Per-
gan-
tung
pada
State
atau
State-
Aksi.
 $r(s, a)$ *ataur*(s).

(s)
(s, a)
Kebijakan
Op-
ti-
mal
Per-
gan-
tung
pada
Dis-
tribusi
State
Awal.

π^*
 π
*ek-
spek-
tasi*
dis-
tribusi
state
awal

$\pi^*(p) = \arg \max_{\pi} E_{S_0 \sim p} [V^{\pi}(S_0)],$

$V^{\pi}(S_0)$
 π
 S_0
Contoh:
Dis-
tribusi
Awal
Uni-
form
VS.
Non-
uniform.

*un-
form*

$p_{uniform}(s) = \frac{1}{|S|},$

\hat{S}
 S^*
 $\pi^{uniform}$
berbeda
 π_p^*

$p = [1, 0, 0, \dots]$ *ataup* $= [0, 1, 0, \dots].$

p
 s_1
penekanan
Penamaan
Ke-
bi-
jakan
Op-
ti-
mal
Sesuai
Dis-
tribusi
Awal.

$\pi_0^*(\text{untuk distribusi awal tertentu}), \pi_p^*(\text{untuk distribusi awal } p).$

p
 π_p^*
Interpretasi
Ke-
selu-
ruhan.
MDP
tidak

(s)
 (s, a)
Kebijakan
op-
ti-
mal
 $\pi^* = \arg \max_{\pi} E_{S_0 \sim p} [V^{\pi}(S_0)]$.

p
~~*uni-*~~
~~*form*~~
~~*non-*~~
~~*uniform*~~
 π^*
Inti
perbe-
daan
 p