MPD tuple :

$\langle S, A, T, R \rangle$

S: states $s_t \in \{ HN, HE, HW, LN, LE, LW \}$

A: actions $\{ GN, GE, GW \}$

T: transitions

$P_{GN}(s_{t+1} \mid s_t)$:

	HN	HE	HW	LN	LE	LW
HN	0.66	0	0	0.34	0	0
HE	0	1	0	0	0	0
HW	0	0	1	0	0	1
LN	0	0	0	1	0	0
LE	0	0.34	0	0	0.66	0
LW	0	0	0.33	0	0	0.67

$P_{GE}(s_{t+1} \mid s_t)$:

	HN	HE	HW	LN	LE	LW
HN	1	0	0	0	0	0
HE	0	0.66	0	0	0.34	0
HW	0	0	1	0	0	0
LN	0.33	0	0	0.67	0	0
LE	0	0	0	0	1	0
LW	0	0	0.36	0	0	0.64

$P_{GW}(s_{t+1} \mid s_t)$:

	HN	HE	HW	LN	LE	LW
HN	1	0	0	0	0	0
HE	0	1	0	0	0	0
HW	0	0	0.68	0	0	0.32
LN	0.33	0	0	0.67	0	0
LE	0	0.34	0	0	0.66	0
LW	0	0	0	0	0	1

R: rewards , we define costs instead :

$c(GN) = ?$, $c(GE) = ?$, $c(GW) = ?$

DEFINIR COSTES