



$$\begin{aligned}
 v_{\pi}(s) &\doteq \mathbb{E}_{\pi}[G_t \mid S_t = s] \\
 &= \mathbb{E}_{\pi}[R_{t+1} + \gamma G_{t+1} \mid S_t = s] \\
 &= \mathbb{E}_{\pi}[R_{t+1} + \gamma v_{\pi}(S_{t+1}) \mid S_t = s]
 \end{aligned}$$

(from (3.9))

(4.3)

	L	M	R	Discount value	1	iter	S1	S2	S3	S2 meer waard dan S1
L	0	1	0			1	0	0	0	ONWAAR
M	0,5	0	0,5			2	-0,1	-0,1	0	ONWAAR
R	0	0	0			3	-0,2	-0,2	0	ONWAAR
						4	-0,3	-0,3	0	ONWAAR
						5	-0,4	-0,4	0	ONWAAR
						6	-0,5	-0,5	0	ONWAAR
						7	-0,6	-0,6	0	ONWAAR
						8	-0,7	-0,7	0	ONWAAR
						9	-0,8	-0,8	0	ONWAAR
						10	-0,9	-0,9	0	ONWAAR
						11	-1	-1	0	ONWAAR
						12	-1,1	-1	0	WAAR
						13	-1,1	-1	0	WAAR
						14	-1,1	-1	0	WAAR

Conclusie; Ik stop met rekenen na de 12de iteratie omdat dan het patroon stopt en de waardens optimaal zijn