Chapter 11 solutions **Hanna Gábor**

1. *Convert the equation of n-step off-policy TD (7.9) to semi-gradient form. Give accompanying definitions of the return for both the episodic and continuing cases.*

$$w_{t+n} = w_{t+n-1} + \alpha\rho_{t:t+n-1}(G_{t:t+n} - \hat{v}(S_t, w_{t+n-1}))\nabla\hat{v}(S_t, w_{t+n-1})$$

In the episodic case

$$G_{t:t+n} = R_{t+1} + \gamma R_{t+2} + \cdots + \gamma^{t+n-1}R_{t+n} + \hat{v}(S_{t+n}, w_{t+n-1}).$$

In the continuing case

$$G_{t:t+n} = R_{t+1} - \bar{R}_{t+n-1} + R_{t+2} - \bar{R}_{t+n-1} + \cdots + R_{t+n} - \bar{R}_{t+n-1} + \hat{v}(S_{t+n}, w_{t+n-1}).$$