$$
\begin{aligned}
x(t) &\in X = \{x_1, x_2, \ldots, x_n\} \\
u(t) &\in U = \{u_1, u_2, \ldots, u_m\} \\
t &\in T = \{t_0, t_0 + 1, \ldots, t_\mathrm{f}\}
\end{aligned}
$$

$$
\begin{aligned}
\phi &:\ X \times U \to X & x(t+1) &= \phi(x(t), u(t)) \\
\ell &:\ X \times U \to \mathbb{R} & &\ell(x(t), u(t)) \\
\ell_\mathrm{f} &:\ X \to \mathbb{R} & &\ell_\mathrm{f}(x(t_\mathrm{f}))
\end{aligned}
$$

Let $x(t_0) = x_0 \in X$ be given, and consider the optimal control problem:

$$
J(t_0, x_0; u(\cdot)) = \sum_{\tau=t_0}^{t_\mathrm{f}-1} \ell(x(\tau), u(\tau)) + \ell_\mathrm{f}(x(t_\mathrm{f}))
$$

$$
\inf_{\substack{u(\tau)\in U \\ \tau \in T}} J(t_0, x_0; u(\cdot))
$$

Define the cost-to-go:

$$
V :\ T \times X \to \mathbb{R} \qquad V(t, x) = \inf_{\substack{u(\tau)\in U \\ \tau \in \{t, t+1, \ldots, t_\mathrm{f}\}}} J(t, x; u(\cdot))
$$

Note that computing the cost-to-go $V(t_0, x_0)$ from the initial state x_0 at the initial time t_0 essentially amounts to minimize the cost $J(t_0, x_0; u(\cdot))$.

If $t = t_\mathrm{f}$:

$$\begin{aligned} V(t_\mathrm{f}, x) &= \inf_{u(t_\mathrm{f})\in U} J(t_\mathrm{f}, x; u(t_\mathrm{f})) \\ &= \inf_{u(t_\mathrm{f})\in U} \underbrace{\ell_\mathrm{f}(x)}_{\substack{\text{independent} \\ \text{of } u(t_\mathrm{f})}} = \ell_\mathrm{f}(x) \end{aligned}$$

If $t = t_\mathrm{f} - 1$:

$$\begin{aligned} V(t_\mathrm{f}-1, x) &= \inf_{u(t_\mathrm{f}-1), u(t_\mathrm{f})\in U} J(t_\mathrm{f}-1, x; u(\cdot)) \\ &= \inf_{u(t_\mathrm{f}-1), u(t_\mathrm{f})\in U} \{\underbrace{\ell(x(t_\mathrm{f}-1), u(t_\mathrm{f}-1))}_{\text{independent of } u(t_\mathrm{f})} + \underbrace{\ell_\mathrm{f}(x(t_\mathrm{f}))}_{\substack{\text{depend on both} \\ u(t_\mathrm{f}-1) \text{ and } u(t_\mathrm{f})}} \} \\ &= \inf_{u(t_\mathrm{f}-1)} \{\ell(x, u(t_\mathrm{f}-1)) + \underbrace{\inf_{u(t_\mathrm{f})} \ell_\mathrm{f}(x(t_\mathrm{f}))}_{= V(t_\mathrm{f}, x(t_\mathrm{f})) = V(t_\mathrm{f}, \phi(x, u(t_\mathrm{f}-1)))} \} \\ &= \inf_{u(t_\mathrm{f}-1)\in U} \{\ell(x, u(t_\mathrm{f}-1)) + V(t_\mathrm{f}, \phi(x, u(t_f - 1)))\} \\ &= \inf_{u\in U} \{\ell(x, u) + V(t_\mathrm{f}, \phi(x, u))\} \end{aligned}$$

For $t < t_\mathrm{f}$:

$$\begin{aligned} V(t, x) &= \inf_{\substack{u(\tau)\in U \\ \tau\in\{t, t+1, \ldots, t_\mathrm{f}\}}} J(t, x; u(\cdot)) \\ &= \inf_{\substack{u(\tau)\in U \\ \tau\in\{t, t+1, \ldots, t_\mathrm{f}\}}} \{\sum_{\tau=t}^{t_\mathrm{f}-1} \ell(x(\tau), u(\tau)) + \ell_\mathrm{f}(x(t_\mathrm{f}))\} \\ &= \inf_{\substack{u(\tau)\in U \\ \tau\in\{t, t+1, \ldots, t_\mathrm{f}\}}} \{ \underbrace{\ell(x, u(t))}_{\substack{\text{independent of } u(\tau), \\ \tau \in \{t+1, t+2, \ldots, t_\mathrm{f}\}}} + \underbrace{\sum_{\tau=t+1}^{t_\mathrm{f}-1} \ell(x(\tau), u(\tau)) + \ell_\mathrm{f}(x(t_\mathrm{f}))}_{\text{depend on all } u(\tau),\ \tau \in \{t, t+1, \ldots, t_\mathrm{f}\}} \} \\ &= \inf_{u(t)\in U} \{\ell(x, u(t)) + \underbrace{\inf_{\substack{u(\tau)\in U \\ \tau\in\{t+1, t+2, \ldots, t_\mathrm{f}\}}} \{\sum_{\tau=t+1}^{t_\mathrm{f}-1} \ell(x(\tau), u(\tau)) + \ell_\mathrm{f}(x(t_\mathrm{f}))\}}_{= V(t+1, x(t+1)) = V(t+1, \phi(x, u(t)))} \} \\ &= \inf_{u(t)\in U} \{\ell(x, u(t)) + V(t+1, \phi(x, u(t)))\} \\ &= \inf_{u\in U} \{\ell(x, u) + V(t+1, \phi(x, u))\} \end{aligned}$$

Bellman equation:

$$\begin{aligned} V(t_\mathrm{f}, x) &= \ell_\mathrm{f}(x) && \text{for all } x \in X \\ V(t, x) &= \inf_{u \in U} \{\ell(x,u) + V(t+1, \phi(x,u))\} && \text{for all } x \in X \text{ and all } t \in \{t_0, t_{0+1}, \dots, t_\mathrm{f} - 1\} \end{aligned}$$

Let us suppose that the cost-to-go V has been determined. For a given state x at time t, the optimal input $u(t)$ is given as

$$u(t) = \arg\min_{u \in U} \{\ell(x,u) + V(t+1, \phi(x,u))\}$$

This inspires the implantation of the optimal control in a state feed back form:

$$u(t) = u(x(t)) = \arg\min_{u \in U} \{ \underbrace{\ell(x(t),u) + V(t+1, \phi(x(t),u))}_{\text{computed using the measured state } x(t)} \}$$

$$x(t+1) = \phi(x(t), u(t)) \qquad x(t_0) = x_0 \in X$$