1 Symbol

1.1 Constant

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9$$

$$\alpha, \beta, \gamma, \delta, \epsilon(\varepsilon), \zeta, \eta, \theta(\vartheta), \iota, \kappa, \lambda, \mu, \nu, \xi, o, \pi, \rho(\varrho), \sigma, \tau, \upsilon, \pi(\varphi), \chi, \psi, \omega$$

$$\mathbb{1}, \mathcal{N}, \mathcal{R}$$

1.2 Scalar

$$a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q, r, s, t, u, v, w, x, y, z$$

1.3 Vector

$$\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}, \mathbf{f}, \mathbf{g}, \mathbf{h}, \mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{l}, \mathbf{m}, \mathbf{n}, \mathbf{o}, \mathbf{p}, \mathbf{q}, \mathbf{r}, \mathbf{s}, \mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}$$

1.4 Matrix

$$\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H}, \mathbf{I}, \mathbf{J}, \mathbf{K}, \mathbf{L}, \mathbf{M}, \mathbf{N}, \mathbf{O}, \mathbf{P}, \mathbf{Q}, \mathbf{R}, \mathbf{S}, \mathbf{T}, \mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}$$

1.5 Tensor

$$\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{D}, \mathsf{E}, \mathsf{F}, \mathsf{G}, \mathsf{H}, \mathsf{I}, \mathsf{J}, \mathsf{K}, \mathsf{L}, \mathsf{M}, \mathsf{N}, \mathsf{O}, \mathsf{P}, \mathsf{Q}, \mathsf{R}, \mathsf{S}, \mathsf{T}, \mathsf{U}, \mathsf{V}, \mathsf{W}, \mathsf{X}, \mathsf{Y}, \mathsf{Z}$$

1.6 Set

$$\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}, \mathcal{F}, \mathcal{G}, \mathcal{H}, \mathcal{I}, \mathcal{J}, \mathcal{K}, \mathcal{L}, \mathcal{M}, \mathcal{N}, \mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$$

2 Statistics

2.1 Probability

$$x \sim \mathcal{X} \tag{sample}$$

$$x \doteq \mathcal{X} \tag{define}$$

$$x \leftarrow \mathcal{X} \tag{generate}$$

$$x \leftsquigarrow \mathcal{X} \tag{shuffle}$$

$$p(\alpha)\,, p(a)\,, p(\mathbf{a})\,, p(\mathbf{A})\,, p(\mathsf{A})\,, p(\mathcal{A})$$

$$p\big(\alpha \mid \beta\big)\,, p\big(a \mid b\big)\,, p\big(\mathbf{a} \mid \mathbf{b}\big)\,, p\big(\mathbf{A} \mid \mathbf{B}\big)\,, p\big(\mathsf{A} \mid \mathsf{B}\big)\,, p\big(\mathcal{A} \mid \mathcal{B}\big)$$

3 Indexing

$$\operatorname*{argmax}_{a \in \mathcal{A}} \; p(a)$$
$$\operatorname*{argmin}_{x \in \mathcal{X}} \; p\big(x \mid y\big)$$

4 Distribution

$$\sigma(x)$$

$$\frac{\exp\big(p(x)\big)}{\sum_{x' \in \mathcal{X}} \exp\big(p(x')\big)}$$

5 Neural Networks

5.1 Activation

$$\max(\mathbf{0}, \mathbf{x})$$

$$\tanh(\mathbf{x})$$

$$\{0, 1\}$$
$$\{a, \ldots, z\}$$

$$\mathbf{x} \oplus \mathbf{y} \quad (\ 26.85016\text{pt})$$
$$\mathbf{x} \ominus \mathbf{y} \quad (\ 26.85016\text{pt})$$
$$\mathbf{x} \odot \mathbf{y} \quad (\ 26.85016\text{pt})$$
$$\mathbf{x} \oslash \mathbf{y} \quad (\ 26.85016\text{pt})$$

$$x \mathbin{\bar{*}} y \tag{1}$$

$$\mathbf{A}^{\top}$$
$$\mathbf{A}^{\text{-}1}$$
$$\mathbf{A}^{\text{-}\top}$$

$$\mathbb{1}\left[x\right]$$

$$\begin{aligned}
\mathbf{i}_t &= \sigma(\mathbf{W}_i \cdot \mathbf{x}_t + \mathbf{b}_i) \\
\mathbf{f}_t &= \sigma\Big(\mathbf{W}_f \cdot \mathbf{x}_t + \mathbf{b}_f\Big) \\
\mathbf{o}_t &= \sigma(\mathbf{W}_o \cdot \mathbf{x}_t + \mathbf{b}_o) \\
\mathbf{g}_t &= \tanh\big(\mathbf{W}_g \cdot \mathbf{x}_t + \mathbf{b}_g\big) \\
\mathbf{c}_t &= \mathbf{i}_t \odot \mathbf{f}_t
\end{aligned}$$

$$\begin{aligned}
&-\log\Big(p\big(y_t \mid \mathbf{h}_{t-1}, y_{t-1}; \theta\big)\Big) \\
&\epsilon \sim \mathcal{U}(a, b) \\
&\epsilon \sim \mathcal{N}\Big(0, (\sigma/t)^2\Big) \\
&\epsilon \sim \mathcal{B}(1, p) \\
&\epsilon \sim \mathcal{B}\Big(1, (\sigma/t)^2\Big)
\end{aligned}$$

$$\max_{x \in \mathcal{X}} f(x)$$
$$\min_{x \in \mathcal{X}} f(x)$$
$$\sum_{x \in \mathcal{X}} f(x)$$
$$\frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} f(x)$$

$$\big\{ f(x) \mid x \in \mathcal{X} \big\}$$

$$\frac{\partial f(x)}{\partial x}$$
$$\frac{\partial^2 f(x)}{\partial x^2}$$
$$\frac{\partial^2 f(x)}{\partial x \, \partial y}$$

$$\mathbb{E}_{s_t \sim E, a_i \sim \pi} \left[R(s_t, a_t) \right] \tag{2}$$

$$\mathbf{y} = \mathbf{W}_l^{[1,3]} * \mathbf{x} + \mathbf{b}_l \tag{3}$$

6 Reinforcement Learning

$$\nabla_{\mathbf{x}} f(\mathbf{x})$$
$$\nabla_{\mathbf{x}}^2 f(\mathbf{x})$$

$$\mathbb{E}\left[\frac{\partial f(x)}{\partial x}\right], y^*, y' \tag{4}$$

$$\tilde{\mathbf{y}} \tag{5}$$

$$\left(x\right), \left[x\right], \left\{x\right\} \tag{6}$$

7 Decoding

$$\hat{y}_t = \underset{y' \in \mathcal{V}}{\operatorname{argmax}} \; p\big(\mathbf{h}_{t-1}, y_{t-1}\big) \tag{7}$$

$$\hat{y}_t = \underset{y' \in \mathcal{V}}{\operatorname{argmax}} \; p\big(\mathbf{h}_{t-1}, \hat{y}_{t-1}\big) \tag{8}$$

$$\hat{y}_t = \underset{y' \in \mathcal{V}}{\operatorname{argmax}} \; p\big(\mathbf{h}_{t-1} + \epsilon, y_{t-1}\big) \tag{9}$$
$$\epsilon \sim \mathcal{N}(0, 1) \tag{10}$$

$$\hat{y}_t = \underset{y' \in \mathcal{V}}{\operatorname{argmax}} \; p\big(\mathbf{h}_{t-1} + \mathbf{a}, y_{t-1}\big) \tag{11}$$
$$\mathbf{a} = \pi\big(\mathbf{a} \mid \mathbf{x}; \theta_\pi\big) \tag{actor}$$

$$S_t \mapsto G_t \tag{Monte Carlo}$$
$$S_t \mapsto \mathbb{E}\Big[R_{t+1} + \gamma \hat{v}\big(S_{t+1}\big)\Big] \tag{Dynamic Programming}$$
$$S_t \mapsto R_{t+1} + \gamma \hat{v}\big(S_{t+1}\big) \tag{TD(0)}$$
$$S_t \mapsto R_{t+1} + \gamma R_{t+2} + \dots \gamma^n \hat{v}\big(S_{t+n}\big) \tag{n-step TD(0)}$$

$$\overline{\mathrm{VE}}(\theta_v) \doteq \mathbb{E}_{s\sim\mu}\left[v_\pi(s) - \hat{v}_\pi(s;\theta_v)\right] \qquad \text{(mean squared value error)}$$

$$a'a' \tag{12}$$

$$^{\epsilon}\mathcal{C}(\mathbf{x}) \tag{13}$$

$$\begin{aligned}
\ell_\pi &= \mathbb{E}_{\mathbf{s}_t\sim\mu(s),a_t\sim\pi}\left[-\big(r_t - \hat{v}(\mathbf{s}_t;\theta_{\hat{v}})\big)\cdot\log\Big(\pi\big(a_t \mid \mathbf{s}_t;\theta_\pi\big)\Big)\right] \\
\ell_{\hat{v}} &= \mathbb{E}_{\mathbf{s}_t\sim\mu(s),a_t\sim\pi}\left[\big(r_t - \hat{v}(\mathbf{s}_t;\theta_{\hat{v}})\big)^2\right]
\end{aligned}$$

7.1 Q Learning

$$A \leftarrow \pi\big(\cdot \mid S\big) \tag{14}$$
$$S', R \leftarrow \mathcal{E}(A) \tag{15}$$
$$A' \leftarrow \underset{A'\in\mathcal{A}}{\operatorname{argmax}}\ \pi\Big(\cdot \mid S'\Big) \tag{16}$$
$$S \to R + \gamma Q\Big(S', A'\Big) \tag{17}$$

$$\ell = \mathbb{E}_{A\sim\pi(\cdot\mid S),S',R\leftarrow\mathcal{E}(A)}, \tag{18}$$

7.2 Sarsa

$$A \leftarrow \pi\big(\cdot \mid S\big) \tag{19}$$
$$S', R \leftarrow \mathcal{E}(A) \tag{20}$$
$$A' \leftarrow \pi\Big(\cdot \mid S'\Big) \tag{21}$$
$$S \to R + \gamma Q\Big(S', A'\Big) \tag{22}$$

7.3 TD(λ)

$$\begin{aligned}
A &\leftarrow \pi\big(\cdot \mid S\big) \\
S', R &\leftarrow \mathcal{E}(A) \\
\mathbf{z} &\leftarrow \gamma\lambda\mathbf{z} + \nabla_{\mathbf{w}}\,\hat{v}(S;\mathbf{w}) \\
S &\to R + \gamma Q(S;\mathbf{w})
\end{aligned}$$

7.4 Expected Sarsa

$$V(S_t) \leftarrow V(S_t) + \alpha \left[G_t - V(S_t)\right] \tag{23}$$

$$V(S_t) \leftarrow V(S_t) + \alpha \, \mathbb{E}_\pi \left[R_{t+1} + \gamma V\big(S_{t+1}\big) - V(S_t)\right] \tag{24}$$

$$\begin{aligned} G_t &= R_t + \gamma G_{t+1} \\ &= R_t + \gamma R_{t+1} + \gamma^2 G_{t+2} \\ &= \sum_{k=t}^{T} \gamma^{k-t} R_k \end{aligned}$$

$$V(S_t) \leftarrow V(S_t) + \alpha \left[R_{t+1} + \gamma V\big(S_{t+1}\big) - V(S_t)\right] \tag{25}$$

8 Exercise 6.1

$$V_{t+1}(S_t) \leftarrow V_t(S_t) + \sigma \Big[R_{t+1} + \gamma V_t\big(S_{t+1}\big) - V_t(S_t)\Big]$$

$$\delta_t \doteq R_{t+1} + \gamma V_t\big(S_{t+1}\big) - V_t(S_t)$$

$$\psi_{t+1} \doteq V_{t+1}\big(S_{t+1}\big) - V_t\big(S_{t+1}\big)$$

$$\begin{aligned}
G_t - V_t(S_t) &= R_{t+1} + \gamma G_{t+1} - V_t(S_t) + \gamma V_{t+1}\big(S_{t+1}\big) - \gamma V_{t+1}\big(S_{t+1}\big) + \gamma V_t\big(S_{t+1}\big) - \gamma V_t\big(S_{t+1}\big) \\
&= \delta_t + \gamma \Big(G_{t+1} - V_{t+1}\big(S_{t+1}\big)\Big) + \gamma \psi_{t+1} \\
&= \delta_t + \gamma \delta_{t+1} + \gamma^2 \Big(G_{t+2} - V_{t+2}\big(S_{t+2}\big)\Big) + \gamma \psi_{t+1} + \gamma^2 \psi_{t+2} \\
&= \delta_t + \gamma \delta_{t+1} + \cdots + \gamma^{T-t-1} \delta_{T-1} + \gamma^{T-t} \left(G_T - V_T(S_T)\right) + \gamma \psi_{t+1} + \gamma^2 \psi_{t+2} \cdots + \gamma^{T-t} \psi_T \\
&= \delta_t + \gamma \delta_{t+1} + \cdots + \gamma^{T-t-1} \delta_{T-1} + \gamma^{T-t} \left(0 - 0\right) + \gamma \psi_{t+1} + \gamma^2 \psi_{t+2} \cdots + \gamma^{T-t} \psi_T \\
&= \sum_{k=t}^{T-1} \gamma^{k-t} \delta_k + \sum_{k=t}^{T-1} \gamma^{k-t+1} \psi_{k+1} \\
&= \sum_{k=t}^{T-1} \gamma^{k-t} \delta_k + \sum_{k=t}^{T-1} \gamma^{k-t+1} \Big(V_{k+1}\big(S_{k+1}\big) - V_k\big(S_{k+1}\big)\Big) \\
&= \sum_{k=t}^{T-1} \gamma^{k-t} \bigg(\delta_k + \gamma \Big(V_{k+1}\big(S_{k+1}\big) - V_k\big(S_{k+1}\big)\Big)\bigg)
\end{aligned}$$

$$\mathfrak{E} \tag{26}$$