Possibly Useful Formulas

Inverse Kinematics: $\mathscr{C}_{\text{obs}} = \{\mathbf{c} : \exists \mathbf{b} : \phi(\mathbf{b},\mathbf{c}) \in \mathscr{W}_{\text{obs}}\}$

Admissible: $\hat{h}(n) \leq h(n)$

Consistent: $\hat{h}(n) - \hat{h}(m) \leq h(n,m)$

Unification: $S : \{\mathscr{V}_P, \mathscr{V}_Q\} \rightarrow \{\mathscr{V}_Q, \mathscr{C}\}$ such that $S(P) = S(Q) = U$

Value Iteration: $u_i(s) = r(s) + \gamma \max_a \sum_{s'} P(s'|s,a) u_{i-1}(s')$

Policy Evaluation: $u_\pi(s) = r(s) + \gamma \sum_{s'} P(s'|s,\pi(s)) u_\pi(s')$

Policy Improvement: $\pi_{i+1}(s) = \operatorname{argmax}_a \sum_{s'} P(s'|s,a) u_{\pi_i}(s')$

Q-Learning: $q(s,a) = r(s) + \gamma \sum_{s'} P(s'|s,a) \max_{a'} q(s',a')$

TD Learning: $q(s_t,a_t) += \eta \left(r_t + \gamma \max_a q(s_{t+1},a) - q(s_t,a_t) \right)$

SARSA: $q(s_t,a_t) += \eta \left(r_t + \gamma q(s_{t+1},a_{t+1}) - q(s_t,a_t) \right)$

Actor-Critic: $\mathscr{L}_{\text{critic}} = \frac{1}{2}\left(q(s,a) - q_{\text{local}}(s,a)\right)^2, \;\; \mathscr{L}_{\text{actor}} = -\sum_a \pi_a(s) q(s,a)$

REINFORCE: $\Delta w_{i,j} = \eta (r-b) \sum_{t=1}^{T} \frac{\partial \log \pi_{a_t}(s_t)}{\partial w_{i,j}}$

Pinhole Camera: $\frac{x'}{x} = \frac{y'}{y} = \frac{-f}{z}$

Convolution: $w[k,l] * x[k,l] = \sum_i \sum_j w[k-i,l-j] x[i,j]$

Backprop: : $\frac{\partial \mathscr{L}}{\partial w[i,j]} = \sum_k \sum_l \frac{\partial \mathscr{L}}{\partial y[k,l]} \frac{\partial y[k,l]}{\partial w[i,j]}$

Max-Pooling: $z[m,n] = \max_{\substack{(m-1)p+1 \leq k \leq mp \\ (n-1)p+1 \leq l \leq np}} y[k,l]$

Backprop: $\frac{\partial \mathscr{L}}{\partial y[k,l]} = \frac{\partial \mathscr{L}}{\partial z[m,n]}$ if $y[k,l] = \max_{\substack{(m-1)p+1 \leq i \leq mp \\ (n-1)p+1 \leq j \leq np}} y[i,j]$