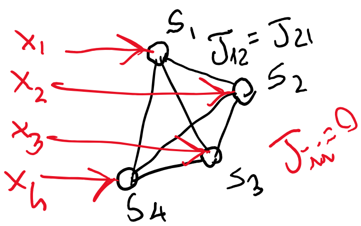
Hopfield and forgetting

Thursday, May 7, 2020 13:49



$$\boxed{s_i(t+1) = f\left(\sum_j^N J_{ij}\, s_j(t) + x_i(t)\right)}$$

f = sign

$\underline{x} = \underline{\xi}^1, \underline{\xi}^2, \dots \underline{\xi}^P$

$J_{ij}(0) \quad J_{ij}(1) \dots J_{ij}(P)$

$\underline{s}(t \to \infty)$

J_{ij}

$\underline{s}(t) = \tilde{\underline{\xi}}^\mu$

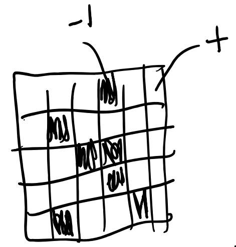
$\underline{s}(t+1) = \underline{\xi}^\mu$

$\underline{\xi}^2 \to \underline{s} \to \underline{\xi}^2$

Cue $\tilde{\underline{\xi}}^2 \to \underline{s} \to \underline{\xi}^2$

$\underline{\xi}^1 \dots \underline{\xi}^P$

$\xi_i^\mu = \pm 1$



$\underline{\xi}^1 = (+ + +, \dots +)$

$\underline{\xi}^2 = (- - -, \dots, -)$

$s_i(t+1) = f\left(\sum_j^N J_{ij}\, s_j(t)\right)$, $J_{ij} = +1$

$\tilde{\underline{\xi}}^1 = (+, -, +, \dots, +)$

$= f\left(\sum_j^N s_j(t)\right)$

$$E(\{s_1 \dots s_N\}) = -\sum_{i,j} s_i J_{ij} s_j = E(t)$$

$$\Rightarrow s_1(t+1) = \text{sign}\Big(\underbrace{\sum_j^N J_{ij}\, s_j(t)}_{h_i(t)}\Big) \Rightarrow E(t+1) \leq E(t)$$

$$\boxed{J_{ij} = J_{ji}} \quad \boxed{J_{ii} = 0}$$

$$\Delta E = E(t+1) - E(t) = -\sum_{i,j} s_i(t+1)\, J_{ij}\, s_j(t+1) + \sum_{i,j} s_i(t)\, J_{ij}\, s_j(t)$$

$$= -\sum_{j=2}^N s_1(t+1)\, J_{1j}\, s_j(t+1) - \sum_{i=2}^N s_i(t+1)\, J_{i1}\, s_1(t+1) +$$

$+ \quad (t) \quad (t) \quad (t) \quad (t)$

$J_{ij} = J_{ji}$

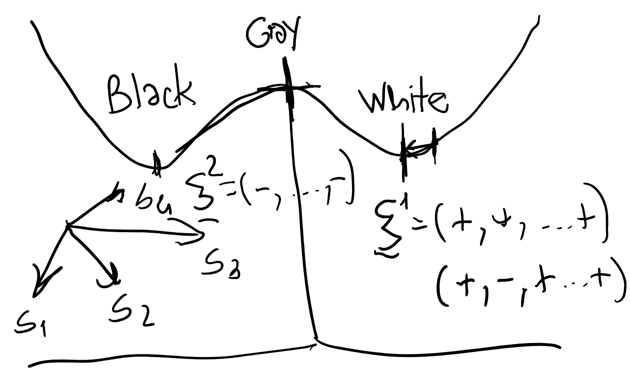
$$\Delta E = -2\left(s_1(t+1) - s_1(t)\right)\underbrace{\left(\sum_{j=2}^{N} J_{1j} s_j(t)\right)}_{h_1(t)}$$

$$\boxed{\Delta E = -2\,\Delta s_1 \cdot h_1(t)}$$

1. $\Delta s_1 = 0 \Rightarrow \Delta E = 0$
2. $\Delta s_1 \neq 0 \Rightarrow \underline{s_1(t+1)} = \text{sign}(h_1(t)) = \underline{-s_1(t)}$

$$\Delta s_1 = \text{sign}(h_1) - (-\text{sign}(h_1)) = 2\,\text{sign}(h_1)$$

$$\Delta E = -2\,\text{sign}(h_1)\cdot h_1 = -2|h_1| < 0$$



$\underline{\xi}^1 = \{+,-,+,\dots +\}$

$\underline{\xi}^2 = \{-,+,-,\dots -\}$

$$s_i' = \underline{\overset{\pm}{\xi_i^1} \cdot \overset{\pm 1}{s_i}}$$

$(+,+,+,\dots +)$

$(- - - \quad -)$

$s_i = \xi_i^1$

$s_i' = \xi_i^1 \xi_i^1 = (\xi_i^1)^2 = +1$

$$s_i'(t+1) = \text{sign}\left(\sum_{j=1}^{N} \overset{+1}{J'_{ij}} s_j'(t)\right)$$

$$\cancel{\xi_i^1 \cdot \xi_i^1}\; s_i(t+1) = \text{sign}\left(\sum_{j=1}^{N} \xi_j^1 s_j(t)\right)\cdot \overset{\pm 1}{\xi_i^1}$$

$+1$

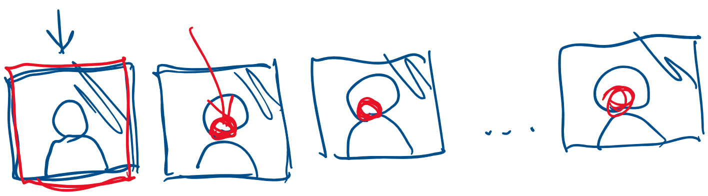
$$s_i(t+1) = \mathrm{sign}\left(\sum_j^N \xi_i^1 \xi_j^1 s_j(t)\right)$$

$$J_{ij} = \xi_i^1 \xi_j^1 \qquad \xi_i^2 = -\xi_i^1$$

$$\rightarrow \quad \underline{J_{ij} = \sum_\mu^p \xi_i^\mu \xi_j^\mu}$$

ξ_i^μ Random uncorrelated $\rightarrow \langle \xi_i^\mu \xi_j^\nu \rangle = 0$

$\forall \mu \neq \nu,\ i \neq j$



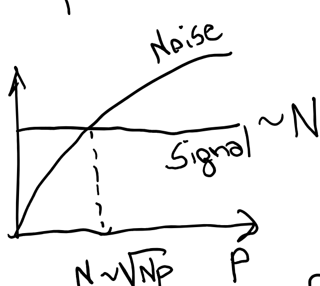
$$h_i = \sum_{j\neq i} J_{ij} \xi_j^1 = \sum_{j\neq i} \sum_\mu^p \xi_i^\mu \xi_j^\mu \xi_j^1 =$$

Signal

$$= \sum_j \xi_i^1 \underbrace{\xi_j^1 \xi_j^1}_{+1} + \sum_{j\neq i}^{N-1} \sum_{\mu>1}^{p-1} \overbrace{\xi_i^\mu \xi_j^\mu}^{\pm 1} \xi_j^1$$

$\sim Np$ Noise

$$\underline{N \xi_i^1} + \eta \rightarrow (0, \sqrt{Np})$$



$p \sim N$

$$\boxed{p \sim N}$$

Amit Gutfreund Sompolinsky

$$p = \underline{0.14} N$$

$p > 0.14N \Rightarrow$ Blackout catastrophe

$$J_{ij} = \sum_{\mu}^{p} \xi_i^\mu \xi_j^\mu \begin{cases} +p \\ -p \end{cases} \qquad p \sim N$$

$0.14N$

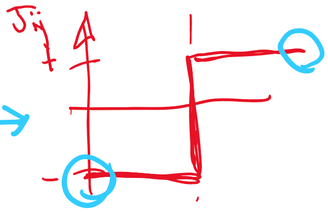
$$J_{ij} = \mathrm{sign}\left(\sum_{\mu}^{p} \xi_i^\mu \xi_j^\mu\right) = \pm 1 \quad \Rightarrow p \sim N \qquad p = 0.08N$$

$$\boxed{J_{ij} = \pm 1}$$

M

$$\mathrm{Signal}^1 = \sum_i^N \xi_i^1 h_i^1 = \sum_{i,j} \xi_i^1 J_{ij} \xi_j^1 =$$

$$= \sum_{i,j} \underbrace{(\xi_i^1 \xi_j^1)}_{\Delta J_{ij}^1 = \pm 1} \cdot \underbrace{J_{ij}}_{= \pm 1}$$



$\Delta J_{ij}^1 = +1$

$$qM \cdot (1-q)(1-q) = qM(1-q)^{p-1}$$

$\boxed{\xi^1}$ ξ^2 ξ^3

$J_{ij} = \xi_i^1 \xi_j^1$

$\Delta J_{ij}^1 = +1$

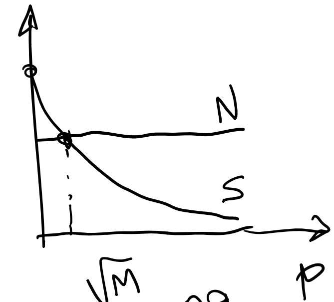
$$qM(1-q)^p = qM e^{\log(1-q)^p} = qM e^{p\log(1-q)} =$$

$$\boxed{\simeq qM e^{-pq}}$$

$$\boxed{\simeq q M e^{-pq}}$$

Signal $\sim q M e^{-pq}$

Noise $\sim \sqrt{M}$

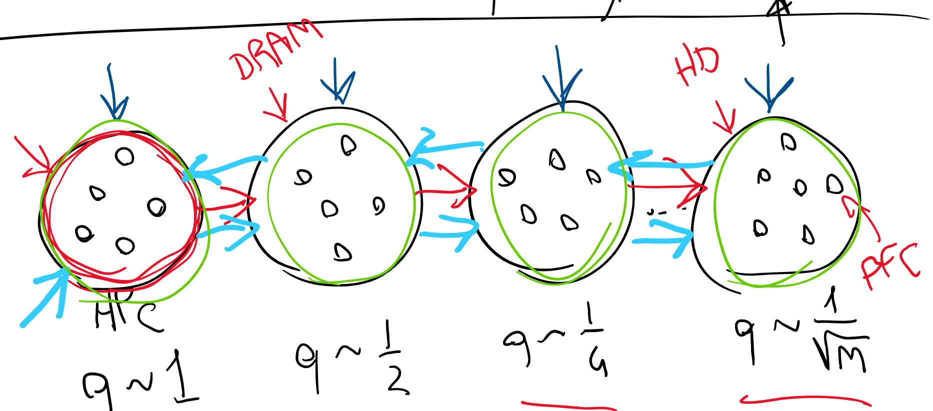


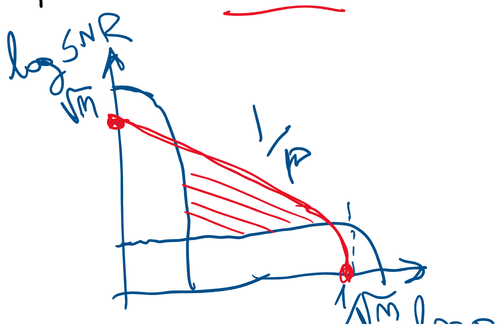
$$\sqrt{M} \sim q M e^{-pq}$$

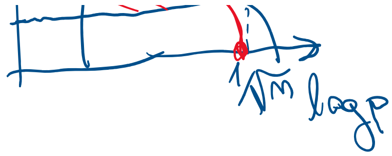
$$\boxed{p \sim N} \rightarrow \boxed{p \sim \frac{1}{q} \log q\sqrt{M}}$$

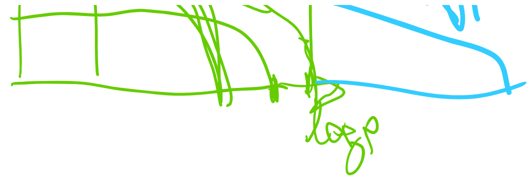
$$\frac{S}{N} \sim q\sqrt{M}\, e^{-pq}$$

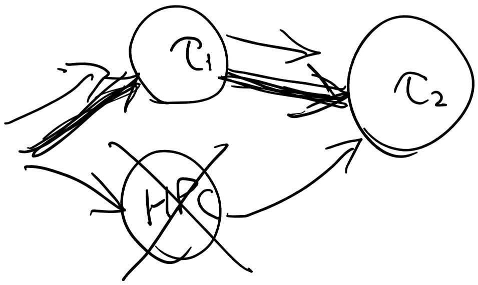
	p	S_0/N_0
Slow : $q \ll 1$ $\quad q \sim \frac{1}{\sqrt{M}}$	$p \sim \sqrt{M}$	$\sim O(1)$
Fast : $q \sim 1$	$p \sim \log M$	$\sqrt{M}$

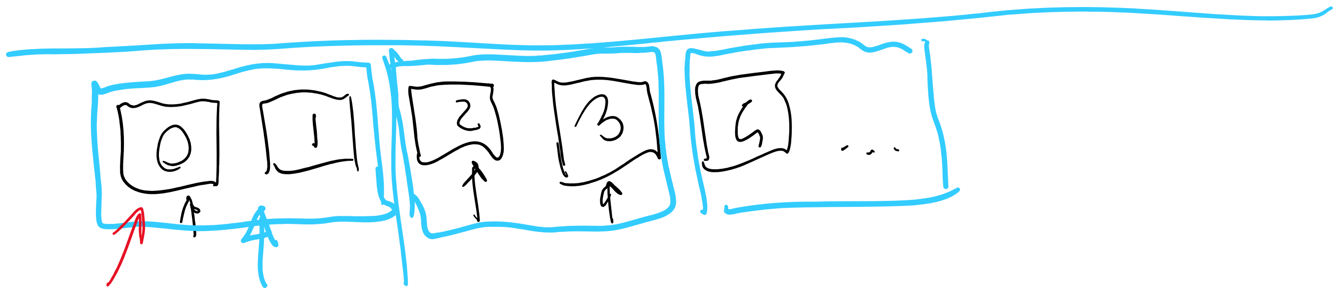


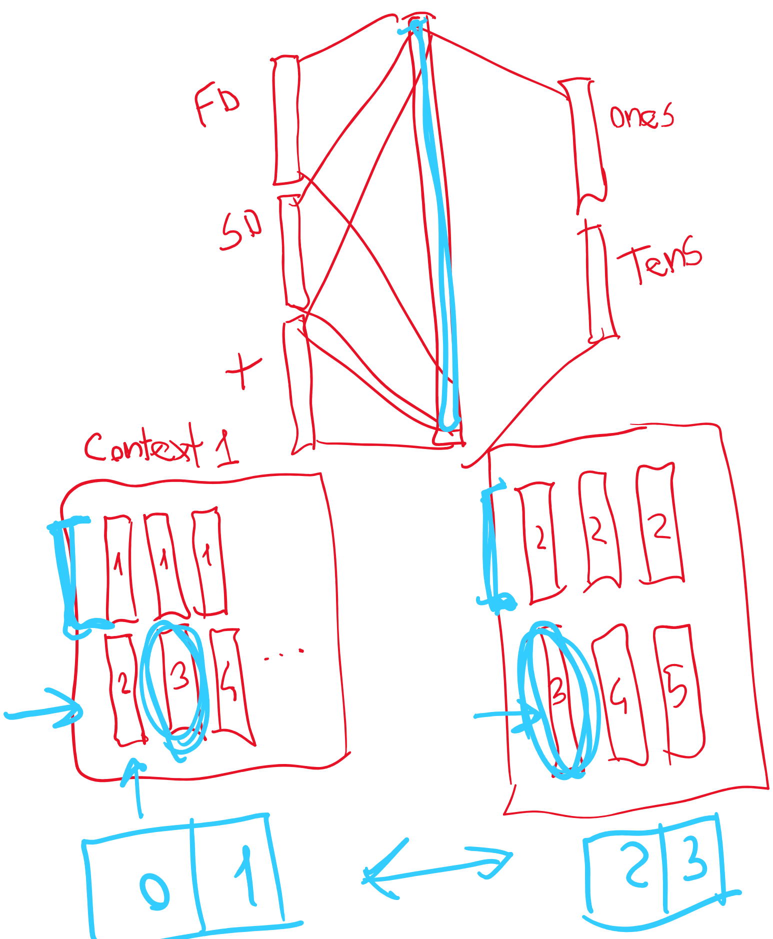




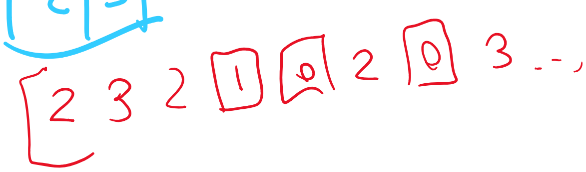






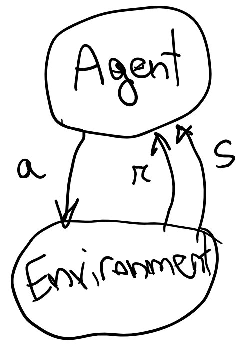


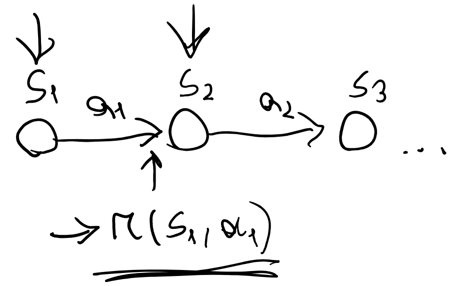
- Replay activity
- Sym.

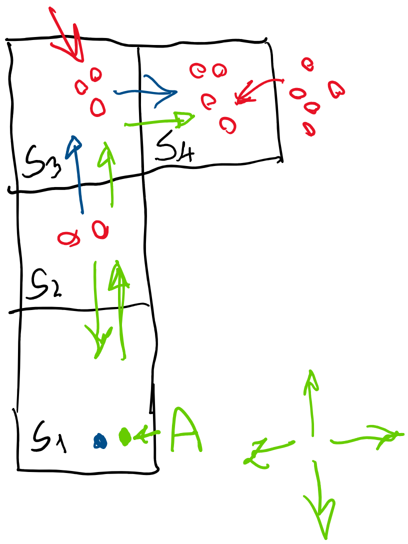


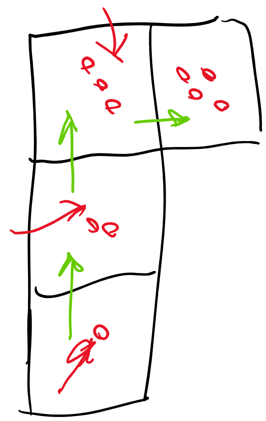
RL Agent

RL









$\pi : S \to A$

$\pi(s_t) = a_t \qquad \gamma < 1$

$$V^{\pi}(s_t) = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ \dots =$$

$$= \sum_{1}^{\infty} r_k \gamma^{k-1}$$

$$\pi^* = \underset{\pi}{\operatorname{argmax}} V^{\pi}(s) \qquad \forall s$$

$$V^{\pi^*}(s) = V^*(s)$$

$\hat{V}(s) \to V^*(s) \qquad s_{t+1} = f(s_t, a_t)$

$\hat{V}(s) \rightarrow V^*(s)$ $\qquad$ $s_{t+1} = f(s_t, a_t)$

$s_t,\ a_t$ randomly $\qquad$ $s_t \xrightarrow{r_t} s_{t+1}$

$$\hat{V}(s_t) \leftarrow \underline{r_t + \gamma \hat{V}(s_{t+1})}$$

$$V^*(s_t) = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots$$
$$= r_t + \gamma \left(r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots \right)$$
$$= r_t + \gamma V^*(s_{t+1})$$

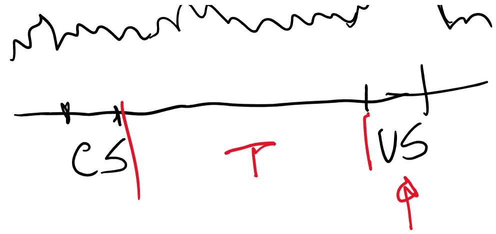
$$\hat{V}(s_t) = (1-\alpha)\hat{V}(s_t) + \alpha\left(r_t + \gamma \hat{V}(s_{t+1})\right)$$

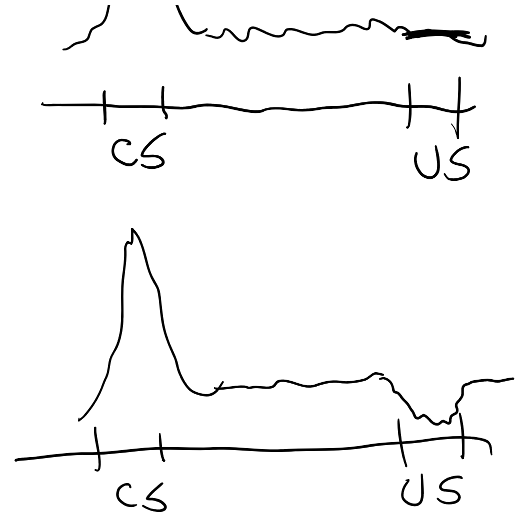
$$\rightarrow \quad \Delta \hat{V}(s_t) = \alpha \underbrace{\left(r_t + \gamma \hat{V}(s_{t+1}) - \hat{V}(s_t)\right)}_{\text{RPE} \quad \delta}$$

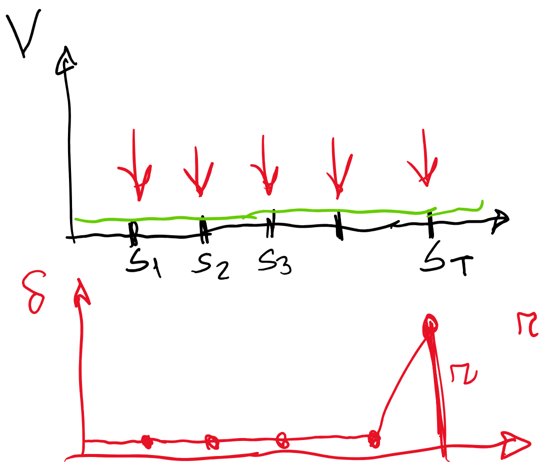
s_t = terminal

$$\Delta \hat{V}(s_t) = \alpha(\overbrace{r_t - \hat{V}(s_t)}^{\delta})$$

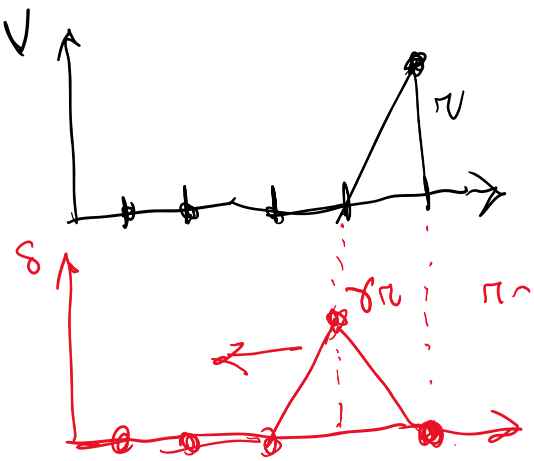
$$\Delta \hat{V}(s_t) = \alpha\left(\underline{\gamma \hat{V}(s_{t+1})} - \underline{\hat{V}(s_t)}\right)$$



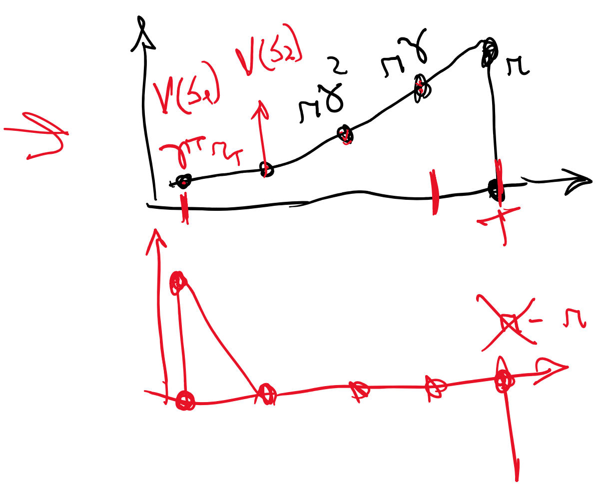


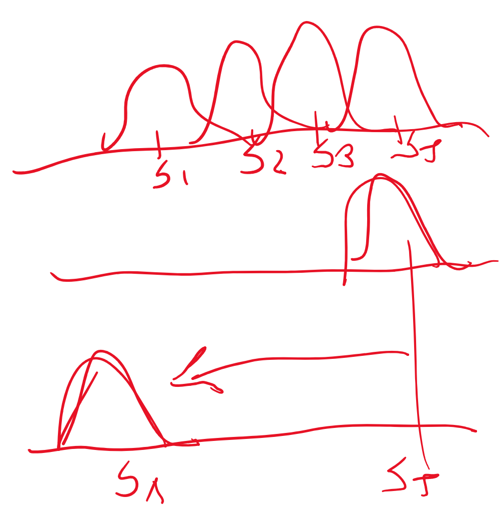


$$r + \cancel{\gamma \hat{V}(s_{t+1})} - \hat{V}(s_t)$$



$r - r = 0$





$$V^*(s_1) = r_1 + \gamma r_2 + \gamma^2 r_3 + \dots \gamma^T r_T$$

$$V^*(s_1) = \underbrace{r_1 + \gamma r_2 + \gamma r_3}_{=0} + \dots 0 \; r_T$$

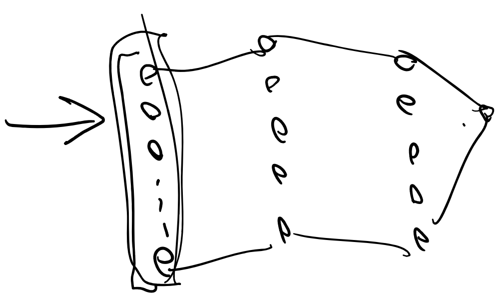
$$V^*(s_2) = r_2 + \gamma r_3 + \dots \gamma^{T-1} r_T$$

$$V^*(s_T) = r$$

$$\underline{V^*(s), \quad s_{t+1} = f(s_t, a_t)}$$

$$\boxed{Q(s,a)} = r(s,a) + \gamma V^*(f(s,a))$$

$$\gamma \max_{a'} Q(f(s,a), a')$$



$Q(s,a)$