Active Learning and Optimal Predictions

Susanne Still
University of Hawaii, Manoa
Honolulu, HI 96822
sstill@hawaii.edu

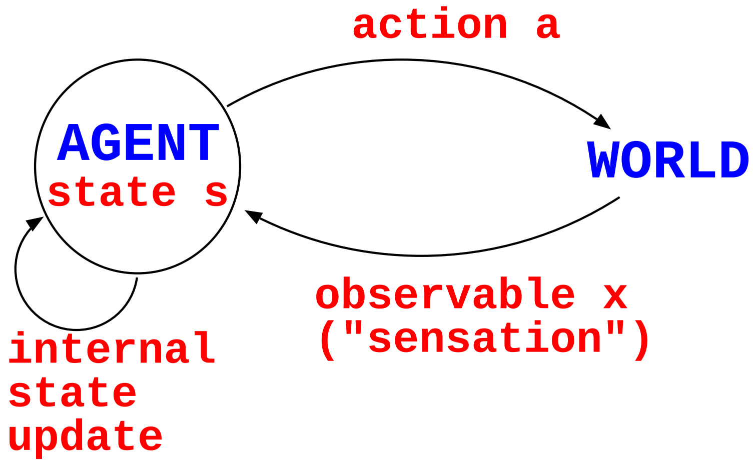
Proposing a new, information theoretic approach to

(inter-) active learning of a physical process by an agent.

*

Thanks: William Bialek.

Problem:



At time t, world produces future data $\mathbf{x}_{\text{future}}$ according to the distribution $Q(\mathbf{x}_{\text{future}}|\mathbf{a}, \mathbf{w})$.
($\mathbf{a}$: action; $\mathbf{w}$: relevant past data, reaching time T into the past).

- Agent updates internal state according to (probabilistic) rule $p(\mathbf{s}_{\text{new}}|\mathbf{v}, \mathbf{s}_{\text{old}})$.
  ($\mathbf{v}$: available observations, reaching time L into the past).
- Agent chooses new action according to (probabilistic) rule $p(\mathbf{a}_{\text{new}}|\mathbf{v}, \mathbf{s}_{\text{old}})$.

Challenge: find the update rules such that
actions and internal representations are chosen optimally.

Intuition: Internal model should make optimal predictions.

Act such that internal representation retains maximal predictive information.

New internal representation should reflect both the the current internal model, and the available observations. **Lossy compression:** Keep only information about the future.

Together, this leads to the **optimization principle**:

$$\max_{\substack{p(\mathbf{s}_{\text{new}}|\{\mathbf{v},\mathbf{s}_{\text{old}}\}) \\ p(\mathbf{a}_{\text{new}}|\{\mathbf{v},\mathbf{s}_{\text{old}}\})}} \mathbf{I}[\mathbf{s}_{\text{new}};\mathbf{x}_{\text{future}}] - \lambda\, \mathbf{I}[\mathbf{s}_{\text{new}};\{\mathbf{v},\mathbf{s}_{\text{old}}\}] - \mu\, \mathbf{I}[\mathbf{a}_{\text{new}};\{\mathbf{v},\mathbf{s}_{\text{old}}\}]$$

General Solution:

$$p_{\text{opt}}(\mathbf{s}_{\text{new}}|\{\mathbf{v},\mathbf{s}_{\text{old}}\}) = \frac{p(\mathbf{s}_{\text{new}})}{Z_s(\mathbf{v},\mathbf{s}_{\text{old}},\lambda)} \times \exp\left[-\frac{1}{\lambda} D_{\text{KL}}\left[p(\mathbf{x}_{\text{future}}|\mathbf{v},\mathbf{s}_{\text{old}}) \| p(\mathbf{x}_{\text{future}}|\mathbf{s}_{\text{new}}) \right] \right]$$

$$p_{\text{opt}}(\mathbf{a}_{\text{new}}|\{\mathbf{v},\mathbf{s}_{\text{old}}\}) = \frac{p(\mathbf{a}_{\text{new}})}{Z_a(\mathbf{v},\mathbf{s}_{\text{old}},\mu)} \times \exp\left[-\frac{1}{\mu} \left(\left\langle D_{\text{KL}}\left[p(\mathbf{x}_{\text{future}}|\mathbf{a}_{\text{new}},\mathbf{v},\mathbf{s}_{\text{old}}) \| p(\mathbf{x}_{\text{future}}|\mathbf{s}_{\text{new}}) \right] \right\rangle_{p(\mathbf{s}_{\text{new}}|\mathbf{v},\mathbf{s}_{\text{old}})} - D_{\text{KL}}\left[p(\mathbf{x}_{\text{future}}|\mathbf{a}_{\text{new}},\mathbf{v},\mathbf{s}_{\text{old}}) \| p(\mathbf{x}_{\text{future}}) \right] \right) \right]$$

Solution in the deterministic case ($\lambda \rightarrow 0$, $\mu \rightarrow 0$):

$$p_{\mathsf{opt}}(\mathbf{s}_{\mathsf{new}}|\{\mathbf{v},\mathbf{s}_{\mathsf{old}}\}) \quad \rightarrow \quad \delta_{\mathbf{s}_{\mathsf{new}},\mathbf{s}^*(\mathbf{v},\,\mathbf{s}_{\mathsf{old}})}$$

$$p_{\mathsf{opt}}(\mathbf{a}_{\mathsf{new}}|\{\mathbf{v},\mathbf{s}_{\mathsf{old}}\}) \quad \rightarrow \quad \delta_{\mathbf{a}_{\mathsf{new}},\mathbf{a}^*(\mathbf{v},\,\mathbf{s}_{\mathsf{old}})}$$

with

$$\mathbf{s}^*(\mathbf{v},\,\mathbf{s}_{\mathsf{old}}) = \underset{\mathbf{s}_{\mathsf{new}}}{\mathsf{argmin}} \quad D_{\mathsf{KL}}\Big[\, p(\mathbf{x}_{\mathsf{future}}|\mathbf{v},\mathbf{s}_{\mathsf{old}}) \parallel p(\mathbf{x}_{\mathsf{future}}|\mathbf{s}_{\mathsf{new}}) \,\Big]$$

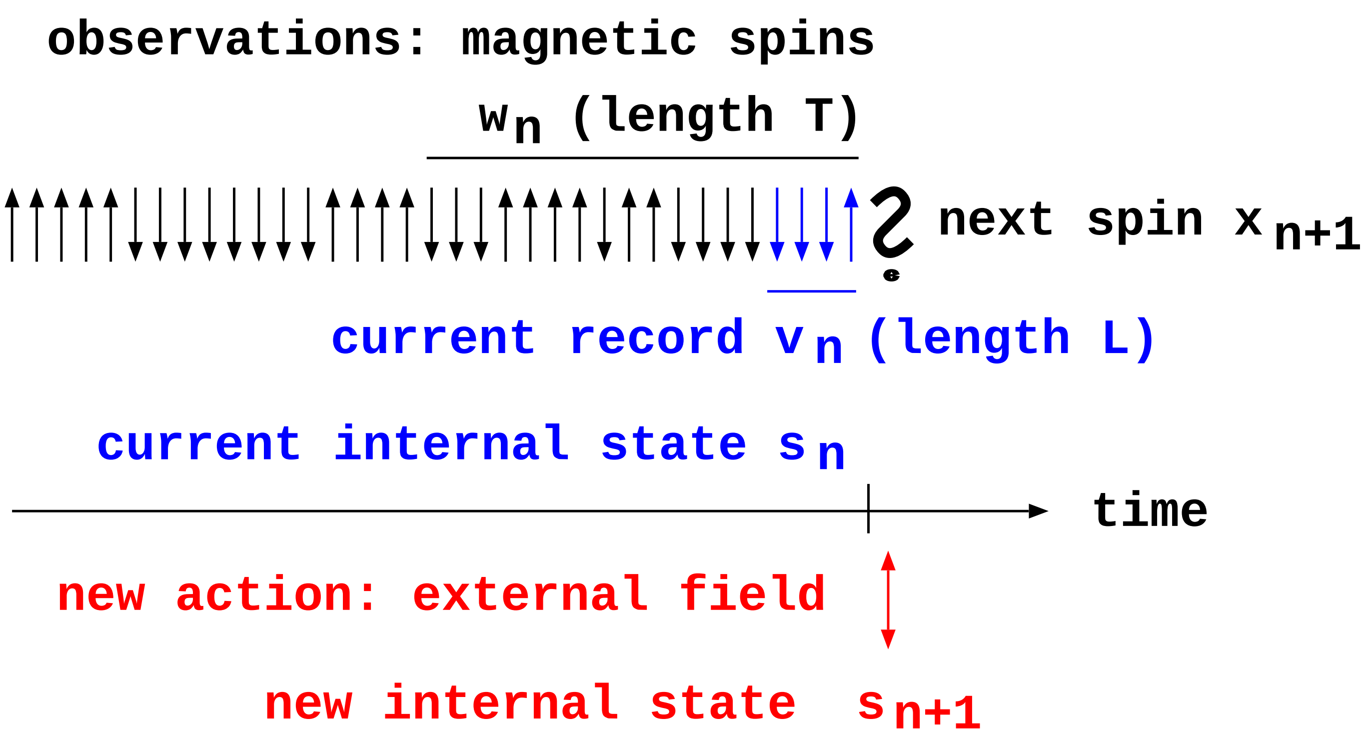
$$\mathbf{a}^*(\mathbf{v},\,\mathbf{s}_{\mathsf{old}}) = \underset{\mathbf{a}_{\mathsf{new}}}{\mathsf{argmin}} \quad D_{\mathsf{KL}}\Big[\, p(\mathbf{x}_{\mathsf{future}}|\mathbf{a}_{\mathsf{new}},\mathbf{v},\mathbf{s}_{\mathsf{old}}) \parallel p(\mathbf{x}_{\mathsf{future}}|\mathbf{s}^*) \,\Big] - D_{\mathsf{KL}}\Big[\, p(\mathbf{x}_{\mathsf{future}}|\mathbf{a}_{\mathsf{new}},\mathbf{v},\mathbf{s}_{\mathsf{old}}) \parallel p(\mathbf{x}_{\mathsf{future}}) \,\Big]$$

Example: A binary world of magnetic spins.

$$Q(x_{n+1}|a_{n+1}, \mathbf{w}_n) = \frac{e^{x_{n+1}(H(\mathbf{w}_n)+a_{n+1})}}{2\cosh\left[H(\mathbf{w}_n) + a_{n+1}\right]}$$

Spin interaction:

$$H(\mathbf{w}_n) = \sum_{i=n-T+1}^{n} J_{n-i} x_i$$



Case $L \geq T$:

Optimal state

$$s^*(\mathbf{v}_n, s_n) \;=\; \underset{s_{n+1}}{\text{argmin}} \; \Big| \, s_{n+1} - \Big[\, H(\mathbf{v}_n) + a^* \, \Big] \, \Big|$$

Optimal action

Case: small perturbation. Reinforce belief.

$$a^*(\mathbf{v}_n, s_n) \;=\; \underset{a_{n+1}}{\text{argmax}} \; \Big| \, H(\mathbf{v}_n) + a_{n+1} \, \Big|$$

Case: arbitrary large external control. Trade off between reinforcing belief and not creating too much order.

At the same time have to maximize the entropy-like term

$$- \Big\langle \, \log \Big[\, \Big\langle \, Q(x_{n+1} | a^*(\mathbf{v}'_n, s'_n), \mathbf{v}'_n) \, \Big\rangle_{p(\mathbf{v}'_n, s'_n)} \, \Big] \, \Big\rangle_{Q(x_{n+1} | a_{n+1}, \mathbf{v}_n)}$$

Case $L \geq T$:

Much uglier equations, but qualitatively similar result.

Main difference: The internal model implicitly contains an average over all possible pasts that are not explicitly available to the learner, but implicitly relevant for the correlations in the world.

Example 2

Functional relationship between action (input to world; query) and observable (output).

$x_n = f(a_n) + \eta$ Small gaussian noise, η with variance σ^2

$$Q(x_{n+1}|a_{n+1}) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x_{n+1} - f(a_{n+1}))^2}$$

for invertible functions f

1. The next sample, a_{n+1}, will be chosen where the slope of f is largest.

2. Both, the optimal action, and the optimal state, are chosen such that $\log[p(a_{n+1})]$ and $\log[p(s_{n+1})]$, respectively, are minimized, i.e. the learner tries to find responses to pasts which are as unique as possible.