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

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

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 current internal model, and the available observations. **Lossy compression:** Keep only information about the future.

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

$$\max \quad I[s_{\text{new}}; x_{\text{future}}] - \lambda I[s_{\text{new}}; \{v, s_{\text{old}}\}] - \mu I[a_{\text{new}}; \{v, s_{\text{old}}\}]$$

$$p(s_{\text{new}}|\{v, s_{\text{old}}\})$$

$$p(a_{\text{new}}|\{v, s_{\text{old}}\})$$
General Solution:

\[ p_{\text{opt}}(s_{\text{new}}|\{v, s_{\text{old}}\}) = \frac{p(s_{\text{new}})}{Z_s(v, s_{\text{old}}, \lambda)} \times \]
\[ \exp \left[ - \frac{1}{\lambda} D_{KL} \left[ p(x_{\text{future}}|v, s_{\text{old}}) \parallel p(x_{\text{future}}|s_{\text{new}}) \right] \right] \]

\[ p_{\text{opt}}(a_{\text{new}}|\{v, s_{\text{old}}\}) = \frac{p(a_{\text{new}})}{Z_a(v, s_{\text{old}}, \mu)} \times \]
\[ \exp \left[ - \frac{1}{\mu} \left( D_{KL} \left[ p(x_{\text{future}}|a_{\text{new}}, v, s_{\text{old}}) \parallel p(x_{\text{future}}|s_{\text{new}}) \right] \right) p(s_{\text{new}}|v, s_{\text{old}}) \]
\[ - D_{KL} \left[ p(x_{\text{future}}|a_{\text{new}}, v, s_{\text{old}}) \parallel p(x_{\text{future}}) \right] \right) \]
Solution in the deterministic case \((\lambda \to 0, \mu \to 0)\):

\[
p_{\text{opt}}(s_{\text{new}}|\{v, s_{\text{old}}\}) \to \delta_{s_{\text{new}}, s^*(v, s_{\text{old}})}
\]

\[
p_{\text{opt}}(a_{\text{new}}|\{v, s_{\text{old}}\}) \to \delta_{a_{\text{new}}, a^*(v, s_{\text{old}})}
\]

with

\[
s^*(v, s_{\text{old}}) = \arg\min_{s_{\text{new}}} D_{\text{KL}}\left[ p(x_{\text{future}}|v, s_{\text{old}}) \parallel p(x_{\text{future}}|s_{\text{new}}) \right]
\]

\[
a^*(v, s_{\text{old}}) = \arg\min_{a_{\text{new}}} D_{\text{KL}}\left[ p(x_{\text{future}}|a_{\text{new}}, v, s_{\text{old}}) \parallel p(x_{\text{future}}|s^*) \right]
\]

\[
- D_{\text{KL}}\left[ p(x_{\text{future}}|a_{\text{new}}, v, s_{\text{old}}) \parallel p(x_{\text{future}}) \right]
\]
**Example: A binary world of magnetic spins.**

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

Spin interaction:

\[ H(w_n) = \sum_{i=n-T+1}^{n} J_{n-i} x_i \]

**observations: magnetic spins**

\[ w_n \text{ (length T)} \]

**current record** \( v_n \text{ (length L)} \)

**new action: external field**

**new internal state** \( s_{n+1} \)

**current internal state** \( s_n \)
Case $L \geq T$:

Optimal state

$$s^*(v_n, s_n) = \operatorname{argmin}_{s_{n+1}} \left| s_{n+1} - \left[ H(v_n) + a^* \right] \right|$$

Optimal action

Case: small perturbation. Reinforce belief.

$$a^*(v_n, s_n) = \operatorname{argmax}_{a_{n+1}} \left| H(v_n) + a_{n+1} \right|$$

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

$$- \left< \log \left[ \left< Q(x_{n+1} | a^*(v'_n, s'_n), v'_n) \right>_{p(v'_n, s'_n)} \right] \right>_{Q(x_{n+1} | a_{n+1}, v_n)}$$
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, \( \eta \) with variance \( \sigma^2 \)

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

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.