Select and Sample - A Model of Efficient Neural Inference and Learning

Jacquelyn A. Shelton1, Jörg Bornschein1, Abdul-Saboor Sheikh1, Pietro Berkes2 and Jörg Lücke1

1Frankfurt Institute for Advanced Studies, Germany; 2Volterra Center for Complex Systems, Brandeis University, USA

Highlights

---

**Introduction:**
- Evidence: perception encodes and maintains posterior probability distributions over possible causes of sensory stimuli
- Full posterior representation costly/complex – very high-dimensional, multi-modal, possibly highly highly
- But, the brain can nevertheless perform rapid learning and inference
- Evidence for fast feed-forward processing and recurrent processing

**Goals:**
- Can we find rich representation for very high-dimensional spaces?
- Want: method to combine feed-forward processing and recurrent stages of processing
- Idea: approximate inference and learning with good posterior representation use pre-selection of most relevant latent variables and sample from this selection

**Results:**
- Experiments on image patches with \( H = 1600 \) hidden dimensions
- Method scales well – applicable to high dimensional data while maintaining sampling-based representation of posterior
- All model parameters learnable
- Combined approach formulates pre-selection and sampling as approximations to exact inference in a probabilistic framework for perception

The Setting

---

**Generative model** for sensory data \( \bar{y} = (y_1, \ldots, y_N) \) with hidden causes/objects \( s = (s_1, \ldots, s_N) \) and parameters \( \Theta \):
\[
p(\bar{y} | s, \Theta) = \prod_{n=1}^{N} p(y_n | s_n, \Theta)
\]

Given data set \( Y = (\bar{y}_1, \ldots, \bar{y}_N) \) find maximum likelihood parameters \( \Theta = \text{argmax}_{\Theta} p(Y | \Theta) \) using EM.

**M-step** usually involves a small number of expected values w.r.t. the posterior distribution:
\[
\langle g(\bar{s}) \rangle_{p(\bar{y} | s, \Theta)} = \sum_{s} p(s | \bar{y}, \Theta) g(s)
\]

where \( g(s) \) is usually an elementary function of the hidden variables (e.g. \( g(s) = \delta \) or \( g(s) = s \delta \) for standard sparse coding)

Computation of expectations is usually the computation demanding part


declare references and acknowledgements

References & Acknowledgements

---