Bayesian Reasoning with Constructive Neural Networks

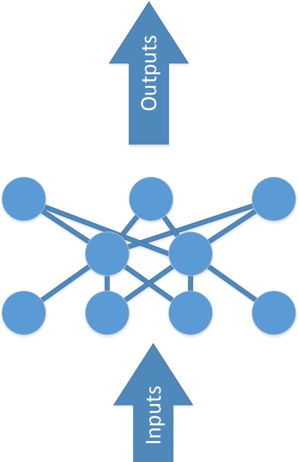
Milad Kharratzadeh[(1)], Thomas Shultz[(2)]

[(1)] Department of Electrical and Computer Engineering, McGill University
[(2)] Department of Psychology & School of Computer Science, McGill University

March 15, 2014
Annual Meeting of the Eastern Psychological Association

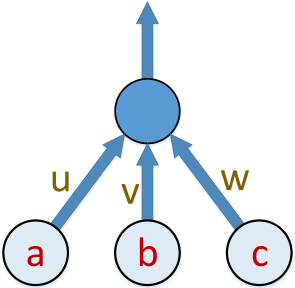
Neural Networks in Brief

- Network of interconnected units and weights
- Conceptually like simplified neurons.
- Goal: learn associations between inputs and outputs (and generalize)



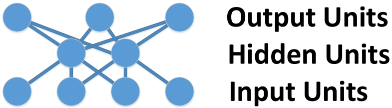
Neural Networks in Brief

- Units are simple mathematical functions:
  - **Input**: weighted sum of outputs of other units,
    $ua + vb + wc$
  - **Output**: a nonlinear function of input
    $f(ua + vb + wc)$
  - Send output to other units doing the same
- Modify weights to reduce the error

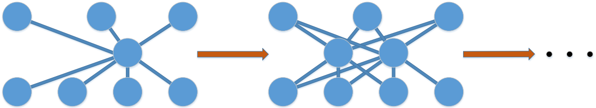


Static VS Constructive Networks

- Static networks:
  - Network's structure fixed in advance (heuristically)
  - Only weights are learned
  - Not psychologically plausible



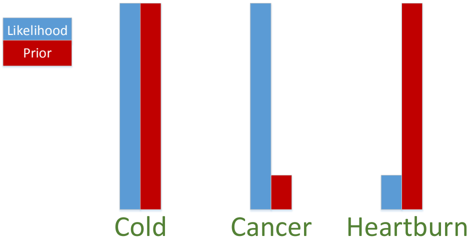
- Constructive networks:
  - Both weights and network's structure are learned
  - Similar to humans' developmental, autonomous learning



- Assume Alice coughs. Most probably, she has ...
  - Cold
  - Heartburn
  - Lung cancer

- Cough a symptom of cold and lung cancer and not heartburn
  $\rightarrow$ Likelihood: observation expectation
- Cold much more common than cancer
  $\rightarrow$ Prior: belief before observation

- Cold high in both



Bayesian Reasoning

- A set of hypotheses to choose from with some priors (e.g., cold, cancer, etc.)
- Degree of belief: Probability, a number between 0 and 1
- Observation (e.g., Alice coughing) $\rightarrow$ likelihood
- Combining prior and likelihood info:
  Bayes' rule: $\textit{posterior belief} \propto \textit{likelihood} \times \textit{prior}$
- Making decision

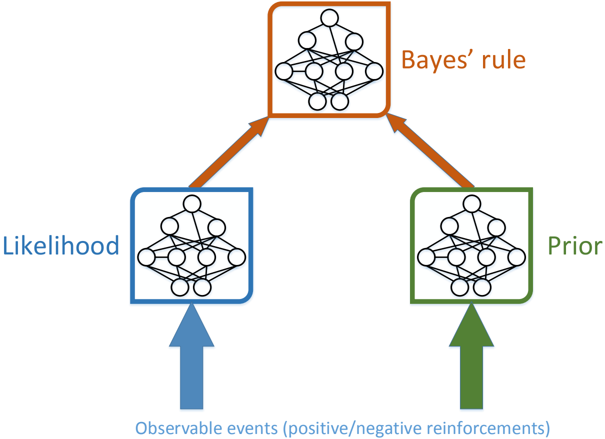
Why use neural networks?

- Challenges with Bayesian approach:
  - Only at Marr's computational level
  - Can be under–constrained
  - Deviation from optimal Bayes' rule (e.g., base–rate neglect)

- Neural nets can help with resolving these issues:
  - At implementation level & psychologically plausible
  - Likelihoods/priors learned from observable events
  - Explaining both Bayesian models and deviations

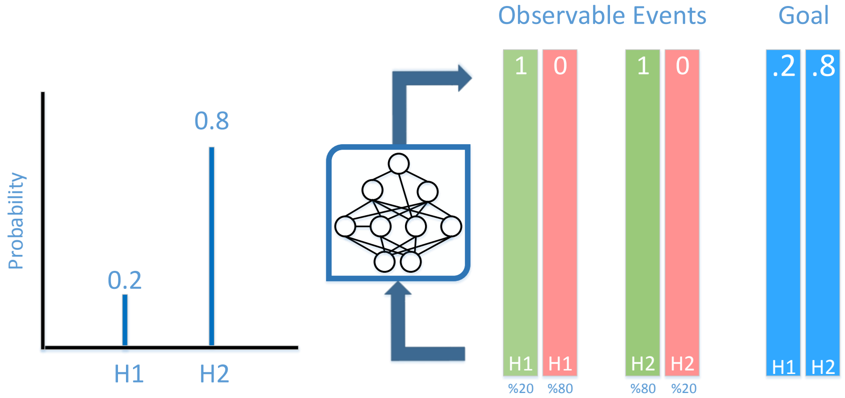
Constructive Neural Net Modeling Bayesian Reasoning

- A constructive neural network formed of three modules:
  - Representing priors
  - Representing likelihoods
  - Applying Bayes' rule



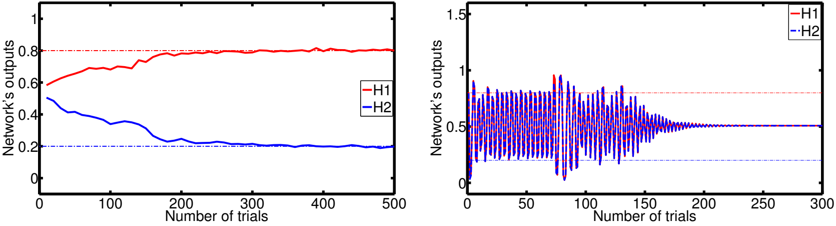
Probability Matching

- First two modules perform **Probability Matching**:
  - **Input:** hypotheses
  - **Output:** degrees of belief (probability)
  - NO access to actual probabilities
  - Only positively (1) or negatively (0) reinforced instances



- H1: $1, 0, 0, 0, 1, 0, 0, 0, \ldots \rightarrow 0.2$
- H2: $0, 1, 1, 1, 0, 1, 1, 1, \ldots \rightarrow 0.8$

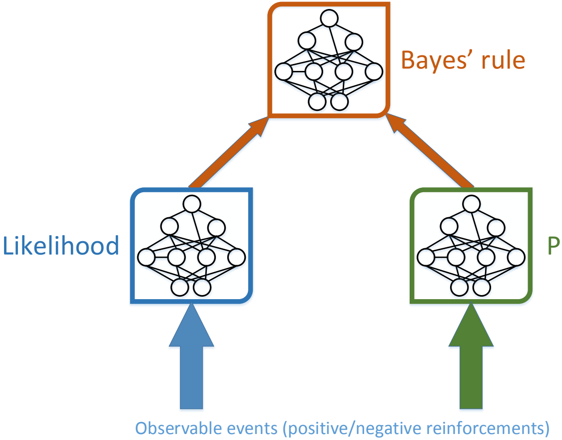
- Can neural nets successfully learn the probabilities?
  — Static networks: Not always
  — Constructive networks: Yes



(a) Constructive Neural Net (b) Static Neural Net

Bayes' Rule Module

- **Input**: likelihoods/priors from probability matching module
- Applying Bayes' rule
- **Output**: posterior beliefs



Base–rate Neglect

- Not taking full account of priors
- Basing decision only on likelihoods (assuming equal priors)

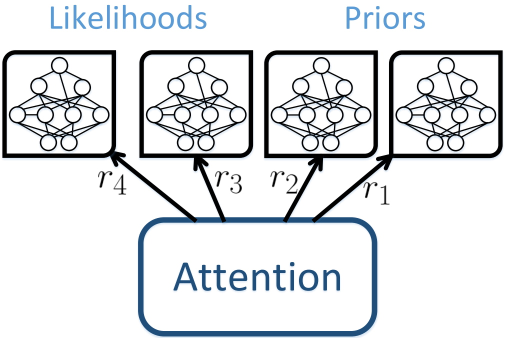
- **Example**: Tom is an opera buff who enjoys touring art museums when on vacation. Which situation is more likely?
  - Tom plays trumpet for a major symphony orchestra.
  - Tom is a farmer.

- **Reasons**: deliberate neglect, failure to recall, long–term synaptic decay, etc.

Modeling Base–rate Neglect as Weight Disruption

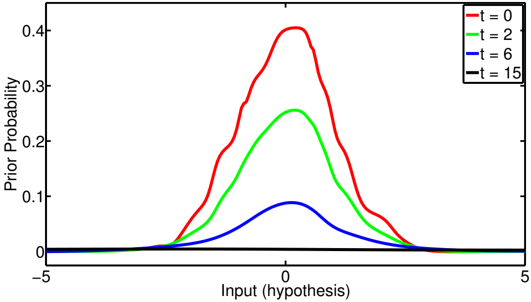
- **Attention module** applying **weight disruption**
- $w_{\text{new}} = r \times w_{\text{old}}, \quad 0 \leq r \leq 1$
- Likelihoods → $r = 1$
- Priors → $0 \leq r < 1$

- Effects of attention, memory indexing, and relevance
- Long–term synaptic decay



Base–rate Neglect Results

- $r = 0.8^t$, i.e., higher t → more weight disruption
- More disrupted weights → more equal priors → more base–rate neglect
- **Conclusion**: base-rate neglect (a deviation) can be explained in the same framework as Bayes' rule



Conclusion

- Neural net model of Bayesian Reasoning
- Realistic inputs
- Autonomous learning through constructive neural net
- Representing probabilities in neural circuitry
- Unifying Bayesian accounts AND deviations

- Bayesian models and neural nets can be viewed as being at different and complementary levels

Questions?

NSERC CRSNG

McGill

Social Sciences and Humanities Research Council of Canada
Conseil de recherches en sciences humaines du Canada
Canada