



- Predictive coding model of Rao and Ballard.

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## └ Introduction

## └ Introduction

1. Prior predictions are compared to stimuli and the model parameters are updated considering prediction errors, features corresponding to receptive fields in the the primary sensory cortex are learned.

- Predictive coding model of Rao and Ballard.
- Free-energy model of Friston.

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## └ Introduction

## └ Introduction

1. Prior predictions are compared to stimuli and the model parameters are updated considering prediction errors, features corresponding to receptive fields in the the primary sensory cortex are learned.
2. Weight stimuli by their noise, learn features using their covariance, implement attentional modulation changing the variance of attended features.

- Predictive coding model of Rao and Ballard.
- Free-energy model of Friston.

# Introduction

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## └ Introduction

### └ Introduction

- Predictive coding model of Rao and Ballard.
- Free-energy model of Friston.
- Hebbian plasticity.

1. Prior predictions are compared to stimuli and the model parameters are updated considering prediction errors, features corresponding to receptive fields in the the primary sensory cortex are learned.
2. Weight stimuli by their noise, learn features using their covariance, implement attentional modulation changing the variance of attended features.
3. Synaptic strenght is changed proportionally to activities of pre-synaptic and post-synaptic neurons.

- Predictive coding model of Rao and Ballard.
- Free-energy model of Friston.
- Hebbian plasticity.

# Introduction

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## └ Introduction

## └ Introduction

- Predictive coding model of Rao and Ballard.
- Free-energy model of Friston.
- Hebbian plasticity.
- Free energy minimization.

- Predictive coding model of Rao and Ballard.
- Free-energy model of Friston.
- Hebbian plasticity.
- Free energy minimization.

1. Prior predictions are compared to stimuli and the model parameters are updated considering prediction errors, features corresponding to receptive fields in the the primary sensory cortex are learned.
2. Weight stimuli by their noise, learn features using their covariance, implement attentional modulation changing the variance of attended features.
3. Synaptic strenght is changed proportionally to activities of pre-synaptic and post-synaptic neurons.
4. Minimization of free energy can be seen as the base of many theories of perception.

## Working hypotheses

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## └ Introduction

- Working hypotheses

1. The state of a neuron is determined only by the synaptic weight and the state of its input neurons.

- Local computation.

① Local computation.

## Working hypotheses

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## └ Introduction

- Working hypotheses

1. The state of a neuron is determined only by the synaptic weight and the state of its input neurons.
2. Synaptic plasticity depends only on the activities of pre-synaptic and post-synaptic neurons.

- ④ Local computation.
- ⑤ Local plasticity.

## Working hypotheses

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## └ Introduction

- Working hypotheses

- ① Local computation.
- ② Local plasticity.
- ③ Basic neuronal computation.

- 1 Local computation.
- 2 Local plasticity.
- 3 Basic neuronal computation.

1. The state of a neuron is determined only by the synaptic weight and the state of its input neurons.
2. Synaptic plasticity depends only on the activities of pre-synaptic and post-synaptic neurons.
3. The state of a neuron is the result of the application of a monotonic function to the linear combination of states and synaptic weights of input neurons.



## Single variable model

- Feature is a scalar variable  $v \in \Omega_v$ .
- Stimulus is a scalar variable  $u \in \Omega_u$ .

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

- Single variable model

- └ Single variable model

1. The model describes the inference of a single variable from a single sensory input.

- Feature is a scalar variable  $v \in \Omega_v$ .
- Stimulus is a scalar variable  $u \in \Omega_u$ .

## Single variable model

- Feature is a scalar variable  $v \in \Omega_v$ .
- Stimulus is a scalar variable  $u \in \Omega_u$ .
- Relation between feature and stimulus is a differentiable function  $g : \Omega_v \rightarrow \Omega_u$ .

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

- Single variable model

- └ Single variable model

1. The model describes the inference of a single variable from a single sensory input.
2. In general inferred variable and sensory input are related by some smooth function.

- Feature is a scalar variable  $v \in \Omega_v$ .
- Stimulus is a scalar variable  $u \in \Omega_u$ .
- Relation between feature and stimulus is a differentiable function  $g: \Omega_v \rightarrow \Omega_u$ .





## Exact solution to the inference problem

- Bayes theorem:

$$p(v|u) = \frac{p(v)p(u|v)}{p(u)} \quad . \quad (1)$$

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

- └ Single variable model
  - └ Exact solution to the inference problem

1. Knowledge of feature depending on a given stimulus is the posterior. Prior knowledge on the feature is the prior, distribution of stimulus is the likelihood.













# Neural implementation

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

└ Single variable model

└ Neural implementation

- Note that hypotheses and Hebbian plasticity are satisfied.

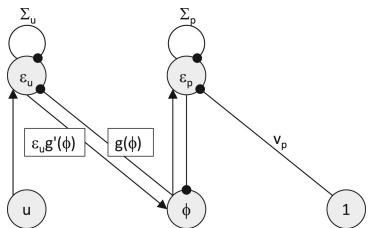


Fig. 3 from article: network implementation of the dynamical system

$$\begin{cases} \dot{\phi} = \varepsilon_u g'(\phi) - \varepsilon_p \\ \dot{\varepsilon}_p = \phi - v_p - \Sigma_p \varepsilon_p \\ \dot{\varepsilon}_u = u - g(\phi) - \Sigma_u \varepsilon_u \end{cases} \quad (6)$$

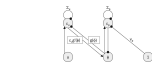


Fig. 3 from article: network implementation of the dynamical system

$$\begin{cases} \dot{\phi} = \varepsilon_u g'(\phi) - \varepsilon_p \\ \dot{\varepsilon}_p = \phi - v_p - \Sigma_p \varepsilon_p \\ \dot{\varepsilon}_u = u - g(\phi) - \Sigma_u \varepsilon_u \end{cases} \quad (6)$$

## Learning model parameters

- Choose model parameters to maximize  $p(u)$ .

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

- Single variable model

- Learning model parameters

1. Model parameters are mean and variance of variables.

## Learning model parameters

- Choose model parameters to maximize  $p(u)$ .
- Equivalent to maximize negative free energy with respect to parameters:

$$\frac{\partial F}{\partial v_p} = \frac{\phi - v_p}{\Sigma_p} \quad , \quad (7)$$

$$\frac{\partial F}{\partial \Sigma_p} = \frac{1}{2} \left( \frac{(\phi - v_p)^2}{\Sigma_p^2} - \frac{1}{\Sigma_p} \right) \quad , \quad (8)$$

$$\frac{\partial F}{\partial \Sigma_u} = \frac{1}{2} \left( \frac{(u - g(\phi))^2}{\Sigma_u^2} - \frac{1}{\Sigma_u} \right) \quad . \quad (9)$$

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

- Single variable model

- Learning model parameters

1. Model parameters are mean and variance of variables.
2. The fixed point of this dynamical system exists only as sample mean over the occurred events of perception, where most likely feature value and stimulus are known.

$$\frac{\partial F}{\partial v_0} = \frac{\phi - v_p}{\Sigma_0}, \quad (7)$$

$$\frac{\partial F}{\partial \Sigma_-} = \frac{1}{2} \left( \frac{(\phi - v_\phi)^2}{\Sigma_-^2} - \frac{1}{\Sigma_-} \right), \quad (8)$$

$$\frac{\partial F}{\partial x} = \frac{1}{2} \left( \frac{(u - g(\phi))^2}{x^2} - \frac{1}{x} \right) \quad (9)$$

# Learning model parameters

- Choose model parameters to maximize  $p(u)$ .
- Equivalent to maximize negative free energy with respect to parameters:

$$\frac{\partial F}{\partial v_p} = \frac{\phi - v_p}{\Sigma_p} \quad , \quad (7)$$

$$\frac{\partial F}{\partial \Sigma_p} = \frac{1}{2} \left( \frac{(\phi - v_p)^2}{\Sigma_p^2} - \frac{1}{\Sigma_p} \right) \quad , \quad (8)$$

$$\frac{\partial F}{\partial \Sigma_u} = \frac{1}{2} \left( \frac{(u - g(\phi))^2}{\Sigma_u^2} - \frac{1}{\Sigma_u} \right) \quad . \quad (9)$$

- Hebbian plasticity is satisfied using prediction errors.

2023-12-11

## Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

### Single variable model

#### Learning model parameters

- Model parameters are mean and variance of variables.
- The fixed point of this dynamical system exists only as sample mean over the occurred events of perception, where most likely feature value and stimulus are known.
- Without prediction errors, the computation is still local.

- Choose model parameters to maximize  $p(u)$ .
- Equivalent to maximize negative free energy with respect to parameters:
 
$$\frac{\partial F}{\partial v_p} = \frac{\phi - v_p}{\Sigma_p} \quad , \quad (7)$$

$$\frac{\partial F}{\partial \Sigma_p} = \frac{1}{2} \left( \frac{(\phi - v_p)^2}{\Sigma_p^2} - \frac{1}{\Sigma_p} \right) \quad , \quad (8)$$

$$\frac{\partial F}{\partial \Sigma_u} = \frac{1}{2} \left( \frac{(u - g(\phi))^2}{\Sigma_u^2} - \frac{1}{\Sigma_u} \right) \quad . \quad (9)$$
- Hebbian plasticity is satisfied using prediction errors.







- ◀ ◻ ▶ ◀ ◻ ▶ ◀ ≡ ▶ ◀ ≡ ▶ ≡

- Minimization of Kullback-Leibler divergence:

$$KL(q(v)||p(v|u)) = \int_{\Omega_v} q(v) \ln \left( \frac{q(v)}{p(v|u)} \right) dv \quad . \quad (13)$$

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

- └ Single variable model
  - └ Free energy framework

1. In general, the posterior is approximated by a simpler probability distribution and the divergence between the two is minimized.

- Minimization of Kullback-Leibler divergence:

$$KL(q(v)||p(v|u)) = \int_{\Omega_v} q(v) \ln \left( \frac{q(v)}{p(v|u)} \right) dv \quad (13)$$





## Multiple variables model

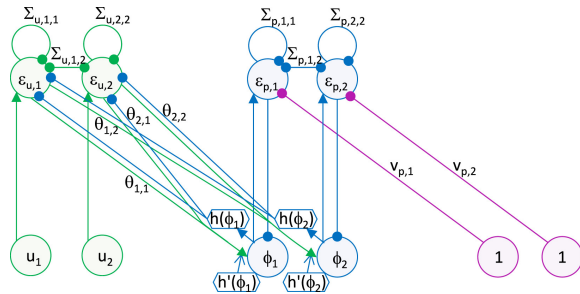


Fig. 5 from article: example of a model with 2 features and 2 stimuli. Equations are rewritten using matrix notation, but local plasticity is not satisfied.

2023-12-11

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

└ Multiple variables model

└ Multiple variables model

- Calculus rules are extended to work elementwise on vectors and matrices, multivariate gaussian distribution and nonlinear relation between variables and stimuli are used.
- The inverse of covariance matrix depends on non-adjacent neurons, Hebbian plasticity is again partially satisfied.

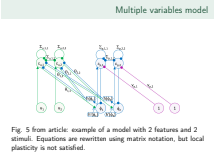


Fig. 5 from article: example of a model with 2 features and 2 stimuli. Equations are rewritten using matrix notation, but local plasticity is not satisfied.





# Recover local plasticity

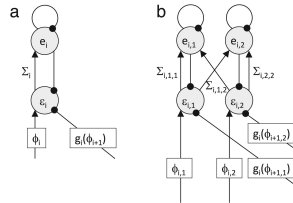


Fig. 7 from article: networks satisfying local plasticity for (a) single variable model and (b) multiple variables model. They implement the generalized dynamical system

$$\begin{cases} \dot{\vec{e}}_i = \vec{\phi}_i - g_i(\vec{\phi}_{i+1}) - \vec{e}_i \\ \dot{\vec{e}}_i = \mathbf{\Sigma}_i \vec{e}_i - \vec{e}_i \end{cases} \quad (19)$$

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

└ Multiple variables model

└ Recover local plasticity

- The update rule for the model parameters is Hebbian and contains the learning rate as hyperparameter of the model.
- Convergence of prediction errors to the sample variances is guaranteed if the most likely feature values change at slower time scales.


Fig. 7 from article: networks satisfying local plasticity for (a) single variable model and (b) multiple variables model. They implement the generalized dynamical system

$$\begin{cases} \dot{\vec{e}}_i = \vec{\phi}_i - g_i(\vec{\phi}_{i+1}) - \vec{e}_i \\ \dot{\vec{e}}_i = \mathbf{\Sigma}_i \vec{e}_i - \vec{e}_i \end{cases} \quad (19)$$



# Conclusion

- Stimuli weighted by noise.
- Learn covariance of stimuli.
- Attentional modulation.

A series of small, faint navigation icons typically found in Beamer presentations, including symbols for back, forward, search, and other navigation functions.

2023-12-11 Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

- └ Conclusion
- └ Conclusion

Summary of *A tutorial on the free-energy framework for modelling perception and learning* by Rafal Bogacz

## Conclusion

## └ Conclusion

- Stimuli weighted by noise.
- Learn covariance of stimuli.
- Attentional modulation.

- Stimuli weighted by noise.
- Learn covariance of stimuli.
- Attentional modulation.