Summary of A tutorial on the free-energy framework for modelling perception

Marco Casari

Introducti

Single variable model

Multiple variable model

Conclusi

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

Marco Casari

University of Turin

Complex system in neuroscience, 12 December 2023



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

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

Marco

Complex system in neuroscience, 12 December 2023

• In this presentation the main topics of the paper are presented, in the same order they appear in the paper.

_ . .

Introduction

• Predictive coding model of Rao and Ballard.

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 framework for modelling perception and learning by Rafal

Marco Casari

Introduction

Single variable model

Multiple variables model

Canalusia

Introduction

2023-12-1

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.

- 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.

Introd

Predictive coding model of Rao and Ballard.

• Free-energy model of Friston.

Hebbian plasticity.

free-energy framework for modelling perception and learning by Rafal Bogacz

Marco Casari

Introduction

Single variable model

Multiple variables model

Conclusion

Introduction

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

· Hebbian plasticity

- 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.

Introduction

Marco Casari

Introduction

Predictive coding model of Rao and Ballard.

- Free-energy model of Friston.
- Hebbian plasticity.
- Free energy minimization.

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

Introduction

2023-12

-Introduction

. Predictive coding model of Rao and Ballard

- 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.

Introduction

Working hypotheses

Local computation.

Summary of A tutorial on the free-energy framework 2023-12-10 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.

Working hypotheses

Local computation



Introduction

Local computation.

Local plasticity.

Working hypotheses

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

· Local plasticity.

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.



Introduction

Local computation.

Basic neuronal computation.

Local plasticity.

Working hypotheses

2023-12

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

Introduction

Working hypotheses

Local elasticity

-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.
- 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

Marco Casari

• Feature is a scalar variable $v \in \Omega_v$.

• Stimulus is a scalar variable $u \in \Omega_u$.

Single variable model

2023-12-

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

-Single variable model

-Single variable model

Single variable model

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

the free-energy framework or modellin perception and learning

Marco Casari

Introductio

Single variable model

Multiple variables model

C l

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 $g: \Omega_V \to \Omega_U$.

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.

Single variable model

Relation between feature and stimulus is a differentiable

2. In general inferred variable and sensory input are related by some smooth function.



Single variable model

4 D > 4 A > 4 B > 4 B > B 9 Q C

free-energy framework for modellin perception and learning by Rafal

Marco Casari

Introductio

Single variable model

Multiple variables model

Conclusion

- 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 $g: \Omega_V \to \Omega_{II}$.
- Sensory input p(u|v) is affected by gaussian noise and it has mean g(v) and variance Σ_u .

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.

Stimulus is a scalar variable u ∈ Ω_u.
 Relation between feature and stimulus is a differentiable.

function $g : \Omega_v \rightarrow \Omega_u$. Sensory input g(u|v) is affected by gaussian noise and it

- 2. In general inferred variable and sensory input are related by some smooth function.
- 3. Sensory input and stimulus are drafted from the same space.

immary of

free-energy framework for modellin perception and learning by Rafal

Marco Casari

Introductio

Single variable model

Multiple variables model

Conclusio

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 $g: \Omega_V \to \Omega_U$.
- Sensory input p(u|v) is affected by gaussian noise and it has mean g(v) and variance Σ_u .
- Prior knowledge of the feature p(v) follows a gaussian distribution with mean v_p and variance Σ_p .

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.

Stimulus is a scalar variable u ∈ Ω_u.
 Relation between feature and stimulus is a differentiable.

function $g:\Omega_v \to \Omega_o$.

• Sensory input p(u|v) is affected by gaussian noise and it has mean g(v) and variance Σ_o .

• Prior knowledge of the feature p(v) follows a gaussian distribution with mean v_o and variance Σ_o .

- 2. In general inferred variable and sensory input are related by some smooth function.
- 3. Sensory input and stimulus are drafted from the same space.
- 4. Information gained and constantly updated from previous experience.



free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introductio

Single variable model

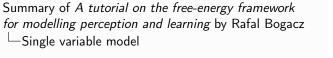
Multiple variables

_ . .

Exact solution to the inference problem

• Bayes theorem:

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





Exact solution to the inference problem

2023-12-10

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.



free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introduction

Single variable model

Multiple variables model

Conclusion

Exact solution to the inference problem

• Bayes theorem:

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

Marginal likelihood of stimuli:

$$p(u) = \int_{\Omega_V} p(v)p(u|v) \, \mathrm{d}v \quad . \tag{2}$$

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

2023-12-10

- 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.
- 2. In general, marginal likelihood is difficult to evaluate.



Single variable model

Exact solution to the inference problem

Bayes theorem:

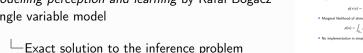
$$p(v|u) = \frac{p(v)p(u|v)}{p(u)} \quad . \tag{2}$$

Marginal likelihood of stimuli:

$$p(u) = \int_{\Omega_v} p(v)p(u|v) \, \mathrm{d}v \quad . \tag{2}$$

• No implementation in simple biological systems.

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.
- 2. In general, marginal likelihood is difficult to evaluate.
- 3. Complex calculations and infinite nodes are needed to represent each value of the posterior.



free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introduction

Single

variable model

variable: model

Canalusia

Approximate solution to the inference problem

• Most likely value of the feature is a scalar variable $\phi \in \Omega_{\nu}$.



2023-12-

Summary of A tutorial on the free-energy framework for modelling perception and learning by Rafal Bogacz —Single variable model

Approximate solution to the inference problem

1. Evaluating the mode of the posterior instead of the whole function is more biologically plausible.

Approximate solution to the inference problem

immary of

free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introductio

Single variable model

Multiple variable

Canalusia

Approximate solution to the inference problem

- Most likely value of the feature is a scalar variable $\phi \in \Omega_{\nu}$.
- Equivalent to maximize negative free energy with respect to the feature:

$$F(v,u) = \ln (p(v)p(u|v)) \quad . \tag{3}$$



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

Approximate solution to the inference problem

2023-12

1. Evaluating the mode of the posterior instead of the whole function is more biologically plausible.

Approximate solution to the inference problem

2. The most likely feature value is the fixed point of the gradient descent method applied to the negative free energy.

tutorial the ree-energ ramewoo

free-energy framework for modelling perception and learning by Rafal Bogacz

Marco Casa

Introduction

Single variable model

variables model

Conclusion

Approximate solution to the inference problem

- Most likely value of the feature is a scalar variable $\phi \in \Omega_{\nu}$.
- Equivalent to maximize negative free energy with respect to the feature:

$$F(v,u) = \ln (p(v)p(u|v)) \quad . \tag{3}$$

Prediction errors:

$$\varepsilon_p = \frac{v - v_p}{\Sigma_p} \quad , \tag{4}$$

$$\varepsilon_u = \frac{u - g(v)}{\Sigma_u} \quad . \tag{5}$$



Summary of A tutorial on the free-energy framework for modelling perception and learning by Rafal Bogacz \sqsubseteq Single variable model

2023-12

Approximate solution to the inference problem

Approximate solution to the inference problem Most Boby value of the feature is a scalar variable $\phi \in \Omega_r$. Equivalent to maintain negative free energy with respect to the feature: $F(v, \omega) = \ln \left(\mu(v) p(u|v) \right) . \tag{3}$ Prediction errors: $\varepsilon_F = \frac{v - \mu_r}{\Sigma_r} . \tag{4}$

- 1. Evaluating the mode of the posterior instead of the whole function is more biologically plausible.
- 2. The most likely feature value is the fixed point of the gradient descent method applied to the negative free energy.
- 3. Prediction errors are introduced as new variables to extend the dynamical system and satisfy Hebbian plasticity.

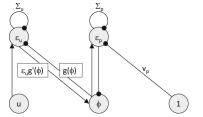


Fig. 3 from article: network implementation of the dynamical system

$$\begin{cases} \dot{\phi} = \varepsilon_{u} g'(\phi) - \varepsilon_{p} \\ \dot{\varepsilon_{p}} = \phi - v_{p} - \Sigma_{p} \varepsilon_{p} \\ \dot{\varepsilon_{u}} = u - g(\phi) - \Sigma_{u} \varepsilon_{u} \end{cases}$$
(6)

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

• Note that all three hypotheses are satisfied.

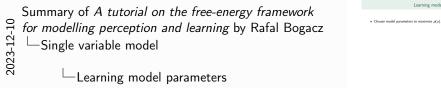
☐ Neural implementation

Neural implementation

Single variable model

Learning model parameters

• Choose model parameters to maximize p(u).



Learning model parameters

1. Model parameters are mean and variance of variables.



free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introduction

Single variable model

> Multiple variables model

Conclusion

Learning model parameters

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

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

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

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

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.

* Equivalent to maintain engation free energy with respect to parameters: $\frac{\partial F}{\partial z} = \frac{\partial - v_y}{\Sigma_p}, \qquad (7)$ $\frac{\partial F}{\partial z_p} - \frac{1}{2} \left(\frac{\partial - v_y}{2} \right)^2 - \frac{1}{\Sigma_p} \right), \qquad (8)$ $\frac{\partial F}{\partial z_p} - \frac{1}{2} \left(\frac{\partial - v_y}{2} \right)^2 - \frac{1}{\Sigma_p} \right), \qquad (9)$

Learning model parameters

- 2. The fixed point of this dynamical system exists only as sample mean over the occured events of perception.



free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introduction

Single variable model

Multiple variables model

Conclusion

Learning model parameters

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

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

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

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

Hebbian plasticity is satisfied using prediction errors.

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

Learning model parameters

2023-12

• Choose modal parameters to maximize p(x).
• Equivalent to maximize negative free energy with respect to parameters. $\frac{\partial G}{\partial y} = \frac{\partial - y_x}{\partial x_y}.$ $\frac{\partial G}{\partial x_y} = \frac{1}{2} \left(\frac{(x-y_y)^2}{x_y} - \frac{1}{x_y} \right).$ (8) $\frac{\partial G}{\partial x_y} = \frac{1}{2} \left(\frac{(x-y)^2}{x_y} - \frac{1}{x_y} \right).$ (9)

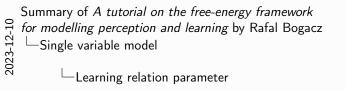
Learning model parameters

- 1. Model parameters are mean and variance of variables.
- 2. The fixed point of this dynamical system exists only as sample mean over the occured events of perception.
- 3. Without prediction errors, the computation is still local.

Learning relation parameter



$$g(v,\theta) = \theta v \quad . \tag{10}$$



1. Only one parameter is considered without loss of generality.

Learning relation parameter



free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introduction

Single variable model

Multiple variables model

Conclusion

Learning relation parameter

2023-12-10

Linear relation:

$$g(v,\theta) = \theta v \quad . \tag{10}$$

Nonlinear relation:

$$g(v,\theta) = \theta h(v) \quad . \tag{11}$$

Summary of A tutorial on the free-energy framework for modelling perception and learning by Rafal Bogacz —Single variable model

Learning relation parameter



Learning relation parameter

- 1. Only one parameter is considered without loss of generality.
- 2. The nonlinearity increases the complexity of the network and partially changes Hebbian plasticity, still keeping it local.



ummary of tutorial on

free-energy framework for modelling perception and learning by Rafal Bogacz

Marco Casari

Introductio

Single variable model

variables model

Conclusion

Learning relation parameter

• Linear relation:

$$g(v,\theta) = \theta v \quad . \tag{10}$$

Nonlinear relation:

$$g(v,\theta) = \theta h(v) \quad . \tag{11}$$

Gradient of negative free energy for learning:

$$\frac{\partial F}{\partial \theta} = \frac{u - \theta h(\phi)}{\Sigma_u} h(\phi) = \varepsilon_u h(\phi) \quad . \tag{12}$$

Summary of *A tutorial on the free-energy framework*of for modelling perception and learning by Rafal Bogacz
Single variable model
Learning relation parameter

* Linear relations $g(r,\theta) = \theta r$. (12) * Nonlinear relations $g(r,\theta) = \theta \theta r$. (12) * Gradient of suggestive few learning: $\frac{\partial r}{\partial \theta} = \frac{-\cos(\phi)}{\Sigma_{\alpha}} h(\phi) = \epsilon_{\alpha} h(\phi) \quad . \quad (12)$

Learning relation parameter

- 1. Only one parameter is considered without loss of generality.
- 2. The nonlinearity increases the complexity of the network and partially changes Hebbian plasticity, still keeping it local.
- 3. Same consideration of model parameters apply to the relation parameter.



free-energy framework for modelling perception and learning by Rafal

Marco Casari

Introductio

Single variable model

Multiple variables

. . .

Free energy framework

• Minimization of Kullback-Leibler divergence:

$$\mathit{KL}(q(v)||p(u|v)) = \int_{\Omega_V} q(v) \ln \left(\frac{q(v)}{p(v|u)} \right) \mathrm{d}v$$
 . (13)

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

Single variable model

Free energy framework

* Minimization of Kullback-Leibler divergence: $\mathcal{KL}(q(v)||g(v)v) = \int_{\Omega_v} q(v) \ln \left(\frac{q(v)}{|x|^2 |v|} \right) dv \quad . \quad (1:$

Free energy framework

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



Free energy framework

Minimization of Kullback-Leibler divergence:

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

Definition of negative free energy:

$$F(v,u) = \int_{\Omega_v} q(v) \ln \left(\frac{p(v,u)}{q(v)} \right) dv \quad . \tag{14}$$

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. 2. Minimize KL divergence or maximize negative free energy to learn
- most likely model value, maximize marginal likelihood or maximize negative free energy to learn model parameters.

Free energy framework

Minimization of Kullback-Leibler divergence:

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

Definition of negative free energy:

$$F(v,u) = \int_{\Omega_v} q(v) \ln \left(\frac{p(v,u)}{q(v)} \right) dv$$
 . (14)

• For the models discussed in the paper: $q(v) = \delta(v - \phi)$.

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.
- 2. Minimize KL divergence or maximize negative free energy to learn most likely model value, maximize marginal likelihood or maximize negative free energy to learn model parameters.
- 3. Equation (3) is recovered using delta function centered in the most likely feature value as probability distribution.



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

2023-12-10

-Multiple variables model └─Multiple variables model

Marco Casari

Multiple variables model

Learning parameters

2023-12-10

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

Learning parameters

Learning parameters

Marco Casari

Multiple variables model

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

Hierarchical structure implementation

-Multiple variables model

2023-12-10

Hierarchical structure implementation

Marco Casari

Introductio

Single variable model

Multiple variables model

Conclusion

2023-12-10 for modelling perception and learning by Rafal Bogacz

-Multiple variables model

Recover local plasticity

Marco Casari

Multiple variables model

Conclusion

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

Conclusion

for modelling p
Conclusion

└─ Conclusion

framework for modelling perception and learning by Rafal

Marco Casari

Introduction

Single variable model

Multiple variables model

Conclusion