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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
model

Conclusion

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

- Predictive coding model of Rao and Ballard.

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

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

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

Marco Casari

- 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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
model

Conclusion

1. Local computation.

1. Local computation.
2. Local plasticity.

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

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

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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
model

Conclusion

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 $g : \Omega_v \to \Omega_u$.

- 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 $g : \Omega_v \rightarrow \Omega_u$.
- Sensory input $p(u|v)$ is affected by gaussian noise and it has mean $g(v)$ and variance Σ_u.

- 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 $g : \Omega_v \rightarrow \Omega_u$.
- 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.

- Bayes theorem:

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

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

Marco Casari

Introduction

**Single
variable
model**

Multiple
variables
model

Conclusion

Exact solution to the inference problem

- Bayes theorem:

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

- Marginal likelihood of stimuli:

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

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

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:

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

- No implementation in simple biological systems.

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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
model

Conclusion

Approximate solution to the inference problem

- Most likely value of the feature is a scalar variable $\phi \in \Omega_v$.

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

Marco Casari

Introduction

Single variable model

Multiple variables model

Conclusion

Approximate solution to the inference problem

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

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

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

Marco Casari

Introduction

**Single
variable
model**

Multiple
variables
model

Conclusion

Approximate solution to the inference problem

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

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

- 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

Marco Casari

Introduction

Single
variable
model

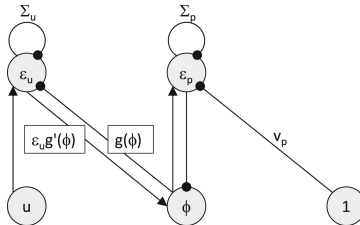Multiple
variables
model

Conclusion

Neural implementation

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

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

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

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

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

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

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

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

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:

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

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

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

- Hebbian plasticity is satisfied using prediction errors.

- Linear relation:

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

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

Marco Casari

Introduction

**Single
variable
model**

Multiple
variables
model

Conclusion

Learning relation parameter

- Linear relation:

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

- Nonlinear relation:

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

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

Marco Casari

Introduction

Single
variable
model

Multiple
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:

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

- Minimization of Kullback-Leibler divergence:

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

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

Marco Casari

Introduction

Single variable model

Multiple variables model

Conclusion

Free energy framework

- Minimization of Kullback-Leibler divergence:

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

- Definition of negative free energy:

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

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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
model

Conclusion

Free energy framework

- Minimization of Kullback-Leibler divergence:

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

- Definition of negative free energy:

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

- 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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
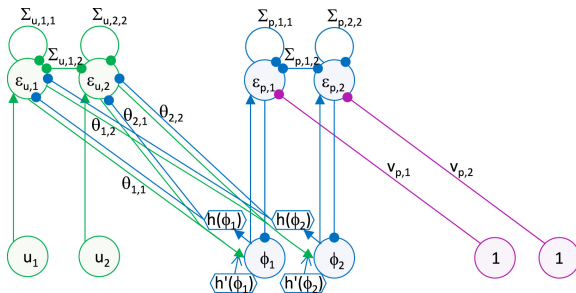model

Conclusion

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.

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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
model

Conclusion

Hierarchical structure implementation

- Parallel to structure of cortical areas.

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

Marco Casari

Introduction

Single
variable
model

Multiple
variables
model

Conclusion

Hierarchical structure implementation

- Parallel to structure of cortical areas.
- Generalized equations for the inference task:

$$\dot{\vec{\phi}}_i = -\vec{\varepsilon}_i + h'(\vec{\phi}_i) \times \mathbf{\Theta}_{i-1}^{\mathsf{T}} \vec{\varepsilon}_{i-1} \quad , \qquad (15)$$

$$\dot{\vec{\varepsilon}}_i = \vec{\phi}_i - \mathbf{\Theta}_i h(\vec{\phi}_{i+1}) - \mathbf{\Sigma}_i \vec{\varepsilon}_i \quad . \qquad (16)$$

- Generalized equations for the learning task:

$$\frac{\partial F}{\partial \mathbf{\Sigma}_i} = \frac{1}{2}(\vec{\varepsilon}_i \vec{\varepsilon}_i^{\mathsf{T}} - \mathbf{\Sigma}_i^{-1}) \quad , \qquad (17)$$

$$\frac{\partial F}{\partial \mathbf{\Theta}_i} = \vec{\varepsilon}_i h(\vec{\phi}_{i+1})^{\mathsf{T}} \quad . \qquad (18)$$

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

Marco Casari

Introduction

Single
variable
model

Multiple
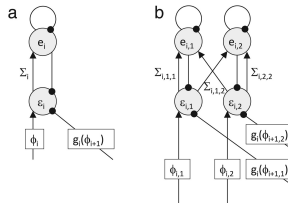variables
model

Conclusion

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

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

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