the History of the Future of the Bayesian Brain

Jawook Gu
Dept. of Bio and Brain Engineering, KAIST

2015/05/20

Author

Karl J. Friston

- British neuroscientist
- Major in brain imaging
- Famous for *Statistical parametric mapping* (SPM)

Contents of paper

- Introduction
- The Bayesian brain
- Prehistory: functional integration club
- History: optimality, natural selection and value
- The Bayesian paradigm
- Bayesian brain & optimization
- Epilogue

Contents of paper

- Introduction
- The Bayesian brain (1967)
- Prehistory: functional integration club (1990)
- History: optimality, natural selection and value (1991)
- The Bayesian paradigm (1994)
- Bayesian brain & optimization
- Epilogue

Characters

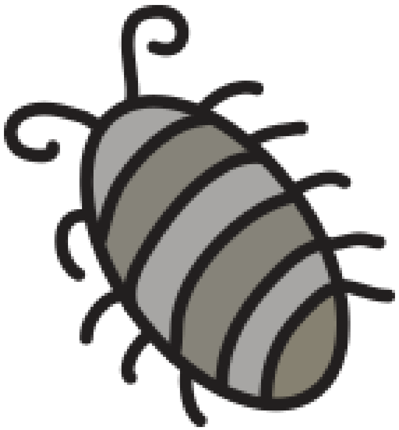
Samir Zeki

Horace Barlow

Gerry Edelman

Graeme Mitchison

Peter Foldiak
Peter Dayan
Giulio Tononi
Olaf Sporns
Tim Shallice
Gerffrey Hinton
….



Wooldice

Contents of presentation

- About paper
  - Author
  - Contents of paper
  - Character
- Background knowledge
  - Bayesian statistics
  - Bayesian optimal classifier
- Part I: The rise of Bayesian thinking
  - Functional segregation and integration of brain
  - The notion of optimality
- Part II: The Bayesian Brain
  - Optimal decision
  - Value learning

Background knowledge:

Bayesian statistics

Bayesian statistics

- Classical definition of probability
  - Frequency
  - $P(event) = \lim_{n \to \infty} \frac{number\ of\ the\ event}{number\ of\ trial}$
  - How to get P(event) when n =1?

- Bayesian probability
  - Calculate the probability using data, logic, and hypothesis

- Bayes' theorem

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$

Bayesian optimal classifier

$$y = \mathrm{argmax}_{c_j \in C} \sum_{h_i \in H} P(c_j|h_i)P(T|h_i)P(h_i)$$

- C_j = class, h_i = Hypothesis, T: training set
- No other classifier can overcome Bayesian optimal classifier.

- http://en.wikipedia.org/wiki/Ensemble_learning

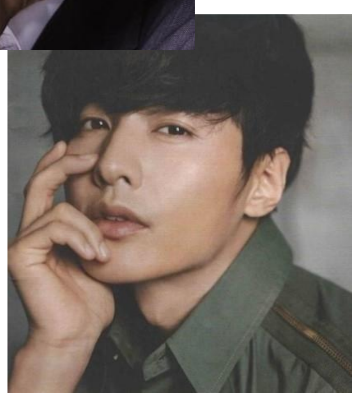
Bayesian optimal classifier

$$y = \mathrm{argmax}_{c_j \in C} \sum_{h_i \in H} P(c_j|h_i)P(T|h_i)P(h_i)$$

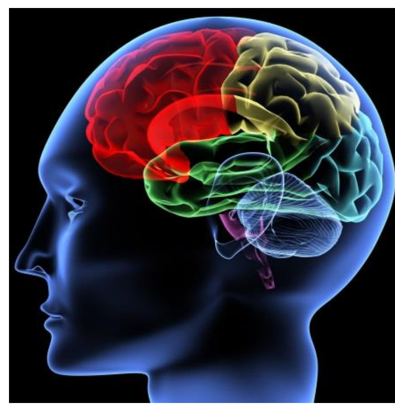
- C_j = class, h_i = Hypothesis, T: training set

Prior probability: P(h)

y = arg max P(.)



Brain



Bayesian optimal classifier

B. Won

JS. Jung

Part I:

The rise of Bayesian thinking

Functional segregation / integration

- Regionally specific activations
  - Statistical parametric mapping

- Interactions mediated by effective connectivity
  - Dynamic causal modeling

Notion of optimality

- Brain is optimal in some sense.
  - What is optimized?
  - Information theory: Bayes optimal

- Bayes brain
  - Optimal decision theory
  - Value learning

Part II:

The Bayesian brain

Bayesian brain

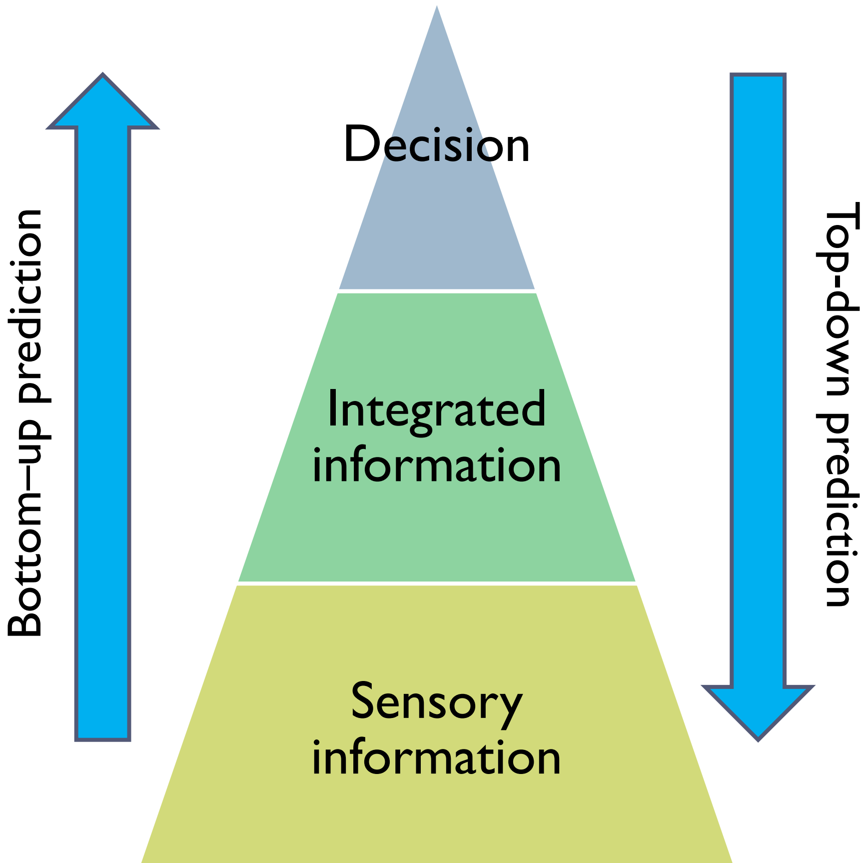
- Information theory
  - maximize the mutual information between sensory input & internal representations
- Value learning /selection
  - value or adaptive fitness
- Free energy minimization
  - marginal likelihood of a model

- All 3 processes are same thing

- Self organizing system (brain) minimize entropy

Epilogue

- Further research of Bayesian brain



- *Top-down predictions suppress errors of bottom-up prediction* :
  example of minimizing free energy

Research of neuronal infrastructures :
functional integration
effective connectivity
dynamic causal modeling
Bayesian evidence based modeling