Probabilistic Models
Of
Sensory-motor Systems

Pierre Bessière, Gabriel Synnaeve
E-Motion team at INRIA Grenoble
CNRS - LPPA - Collège de France

Bayesian-Programming.org
Intelligence

Who is the most clever?

Baron Wolfgang von Kempelen (1769)
**Overview**

- **How to survive (perceive, reason, learn, decide and act) with incomplete information?**

- **Probability as an alternative to logic**

- **How to develop better artifacts using Bayesian reasoning?**

- **Biological plausibility of Bayesian reasoning at a macroscopic level?**

- **Biological plausibility of Bayesian reasoning at a microscopic level?**
How to survive (perceive, reason, learn, decide and act) with incomplete information?

Probability as an alternative to logic

How to develop better artifacts using Bayesian reasoning?

Biological plausibility of Bayesian reasoning at a macroscopic level?

Biological plausibility of Bayesian reasoning at a microscopic level?
Uncertainty is not in things but in our head: **uncertainty is a lack of knowledge.**

*Jacob Bernouilli, Ars Conjectandi (Bernouilli, 1713)*

Probability theory is nothing else than **common sense made calculus.**

*Marquis Pierre-Simon de Laplace, Théorie analytique des probabilités (Laplace 1812)*

The actual science of logic is conversant at present only with things either certain, impossible, or entirely doubtful, none of which (fortunately) we have to reason on. Therefore **the true logic for this world is the calculus of Probabilities**, which takes account of the magnitude of the probability which is, or ought to be, in a reasonable man's mind.

*James Clerk Maxwell (1850)*
Randomness is just the measure of our ignorance. To undertake any probability calculation, and even for this calculation to have a meaning, we have to admit, as a starting point, an hypothesis or a convention, that always comprises a certain amount of arbitrariness. In the choice of this convention, we can be guided only by the principle of sufficient reason. From this point of view, everything in science would just be unconscious applications of the calculus of probabilities. Condemning this calculus would be condemning the whole science.

Henri Poincaré, La science et l’hypothèse (Poincaré, 1902)

By inference we mean simply: deductive reasoning whenever enough information is at hand to permit it; inductive or probabilistic reasoning when - as is almost invariably the case in real problems - all the necessary information is not available. Thus the topic of « Probability as Logic » is the optimal processing of uncertain and incomplete knowledge.

Probability
as an alternative to logic

Incompleteness
Probability as an alternative to logic

Incompleteness

Preliminary Knowledge + Experimental Data = Probabilistic Representation

Uncertainty

Learning

Entropy Principles
Probability as an alternative to logic

Incompleteness

Preliminary Knowledge + Experimental Data = Probabilistic Representation

Uncertainty

Learning

Entropy Principles

Bayesian inference

Decision

\[ P(a) + P(\neg a) = 1 \]

\[ P(a \land b) = P(a) \times P(b \mid a) \]

\[ = P(b) \times P(a \mid b) \]
Bayesian Programming & ProBT®

Bayesian Program

**Description**
- **Specification**
  - *Variables*
    - $S^0, ..., S^t, O^0, ..., O^t$
  - *Decomposition*
    - $P(S^0 \land ... \land S^t \land O^0 \land ... \land O^t) = P(S^0) \times P(O^0 | S^0) \times \prod_{i=2}^{t} [P(S^i | S^{i-1}) \times P(O^i | S^i)]$
  - *Parametric Forms*
    - $P(S^0) = G(S^0, \mu, \sigma)$
    - $P(S^i | S^{i-1}) = G(S^i, A \cdot S^{i-1}, Q)$
    - $P(O^i | S^i) = G(O^i, H \cdot S^i, R)$

**Identification**
- *Learning from instances*
  - $P(S^t | O^0 \land ... \land O^t)$

**Question**
Bayesian Programming

main ()
{
// Variables
plFloat read_time;
plIntegerType id_type(0,1);
plFloat times[5] = {1,2,3,5,10};
plSparseType time_type(5,times);
plSymbol id("id",id_type);
plSymbol time("time",time_type);

// Parametrical forms
// Construction of P(id)
plProbValue id_dist[2] = {0.75,0.25};
plProbTable P_id(id,id_dist);

// Construction of P(time | id = john)
plProbValue t_john_dist[5] = {20,30,10,5,2};
plProbTable P_t_john(time,t_john_dist);

// Construction of P(time | id = bill)
plProbValue t_bill_dist[5] = {2,6,10,40,20};
plProbTable P_t_bill(time,t_bill_dist);

// Construction of P(time | id)
plKernelTable Pt_id(time,id);
plValues t_and_id(time^id);
  t_and_id[id] = 0;
  Pt_id.push(P_t_john,t_and_id);
  t_and_id[id] = 1;
  Pt_id.push(P_t_bill,t_and_id);

// Decomposition
// P(time, id) = P(id) P(time | id)
plJointDistribution jd(time^id,P_id*Pt_id);
}

Bayesian Program Description

QUESTION

P(\theta | \alpha_0 \wedge \ldots \wedge \alpha_d)

// Decomposition
// P(time, id) = P(id) P(time | id)
// Construction of P(id)
plProbValue id_dist[2] = {0.75,0.25};
plProbTable P_id(id,id_dist);
// Construction of P(time | id = john)
plProbValue t_john_dist[5] = {20,30,10,5,2};
plProbTable P_t_john(time,t_john_dist);
// Construction of P(time | id = bill)
plProbValue t_bill_dist[5] = {2,6,10,40,20};
plProbTable P_t_bill(time,t_bill_dist);
// Construction of P(time | id)
plKernelTable Pt_id(time,id);
plValues t_and_id(time^id);
  t_and_id[id] = 0;
  Pt_id.push(P_t_john,t_and_id);
  t_and_id[id] = 1;
  Pt_id.push(P_t_bill,t_and_id);
// Decomposition
// P(time, id) = P(id) P(time | id)
plJointDistribution jd(time^id,P_id*Pt_id);
main () {
    // Variables
    plFloat read_time;
    plIntegerType id_type(0, 1);
    plFloat times[5] = {1, 2, 3, 5, 10};
    plSparseType time_type(5, times);
    plSymbol id("id", id_type);
    plSymbol time("time", time_type);

    // Parametrical forms
    // Construction of \( P(id) \)
    plProbValue id_dist[2] = {0.75, 0.25};
    plProbTable P_id(id, id_dist);

    // Construction of \( P(time | id = john) \)
    plProbValue t_john_dist[5] = {20, 30, 10, 5, 2};
    plProbTable P_t_john(time, t_john_dist);

    // Construction of \( P(time | id = bill) \)
    plProbValue t_bill_dist[5] = {2, 6, 10, 40, 20};
    plProbTable P_t_bill(time, t_bill_dist);

    // Construction of \( P(time | id) \)
    plKernelTable P_t_id(time, id);
    plValues t_and_id(time^id);
    t_and_id[id] = 0;
    P_t_id.push(P_t_john, t_and_id);
    t_and_id[id] = 1;
    P_t_id.push(P_t_bill, t_and_id);

    // Decomposition
    // \( P(time, id) = P(id) \cdot P(time | id) \)
    plJointDistribution jd(time^id, P_id*P_t_id);

    // Question
    // Getting the question \( P(id | time) \)
    plCndKernel Pid_t;
    jd.ask(Pid_t, id, time);

    // Question
    // Getting the question \( P(id | time) \)
    plKernel Pid_readTime;
}

Bayesian Programming & ProBT®
Bayesian Programming & ProBT®

\[ P(S^t | O^0 \land \ldots \land O^t) \]

ProBAYES.com
Bayesian-Programming.org
Bayesian Programming
Related formalisms

More general

Bayesian Programs
Bayesian Networks
DBNs
Bayesian Filters
Particle Filters
discrete HMMs
semi-cont. HMMs
continuous HMMs
Kalman Filters

Bayesian Maps
Markov Loc.
MCML
POMDPs
MDPs

More specific
HOW TO SURVIVE (PERCEIVE, REASON, LEARN, DECIDE AND ACT) WITH INCOMPLETE INFORMATION?

**Probability as an alternative to logic**

HOW TO DEVELOP BETTER ARTIFACTS USING BAYESIAN REASONING?

**Biological plausibility of Bayesian reasoning at a macroscopic level?**

**Biological plausibility of Bayesian reasoning at a microscopic level?**
Olivier Lebeltel’s Ph.D

Object recognition
Carla Koike’s Ph.D
Ruben Garcia’s Ph.D
Ronan Le Hy’s Ph.D
Bayesian Occupancy Filter (BOF) for Advanced Driver Assist. Syst.

- Take uncertainty into account explicitly
- No “data association problem”
- Robustness to object occlusions/disappearances
- Can be implemented on dedicated hardware (GPU or even DSP)

PhD thesis of Christophe Coué


1 SENSOR - 1 OBJECT
$z = (5, 2, 0, 0)$
1 SENSOR - 1 OBJECT

\[ z = (5, 2, 0, 0) \]

\[ P(\ [E_c=1] / z \ c) \]

\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

\[ P( [E_c=1] \mid z_c) \]
\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

• Occupied space

\[ P( [E_c=1] \mid z_c) \]
\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

\[ P([E_c=1] \mid z_c) \]
\[ c = [x, y, 0, 0] \]

• Occupied space
1 SENSOR - 1 OBJECT

- Occupied space
- Free space

P( [E_c=1] | z c)  
\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

- Occupied space
- Free space

\[ P( [E_c=1] | z_c) \]

\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

- Occupied space
- Free space
- Nonobservable space

\[ P( [E_c=1] \mid z \, c) \]
\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

- Occupied space
- Free space
- Nonobservable space

\[ P( [E_c=1] \mid z, c) \]
\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

- Occupied space
- Free space
- Nonobservable space
- Occultated space

\[ P( [E_c=1] \mid z_c) \]
\[ c = [x, y, 0, 0] \]
1 SENSOR - 1 OBJECT

- Occupied space
- Free space
- Nonobservable space
- Occultated space

$$P(\ [E_c=1] \mid z \ c)$$
$$c = [x, y, 0, 0]$$
1 SENSOR - MULTIPLE TARGET

10 m

15 m

50°
1 SENSOR - MULTIPLE TARGET

\[ z_1 = (8.3, -4, 0, 0) \]
\[ z_2 = (7.3, 1.9, 0, 0.8) \]
\[ z_3 = (5, 3, 0, 0) \]
1 SENSOR - MULTIPLE TARGET

\[ P([E_c=1] \mid z_1, z_2, z_3, c) \]
\[ c = [x, y, 0, 0] \]

\[ z_1 = (8.3, -4, 0, 0) \]
\[ z_2 = (7.3, 1.9, 0, 0.8) \]
\[ z_3 = (5, 3, 0, 0) \]
2 SENSOR - 3 TARGETS
2 SENSOR - 3 TARGETS
2 SENSOR - 3 TARGETS

\[ z_{1,1} = (5.5, -4, 0, 0) \]
\[ z_{1,2} = (5.5, 1, 0, 0) \]
\[ z_{2,1} = (11, -1, 0, 0) \]
\[ z_{2,2} = (5.4, 1.1, 0, 0) \]
2 SENSOR - 3 TARGETS

$z_{1,1} = (5.5, -4, 0, 0)$  $z_{1,2} = (5.5, 1, 0, 0)$  $z_{2,1} = (11, -1, 0, 0)$  $z_{2,2} = (5.4, 1.1, 0,0)$

$P([E_{c}=1] \mid z_{1,1} z_{1,2} z_{2,1} z_{2,2} c) \quad c = [x, y, 0, 0]$
Bayesian Filter
Bayesian Filter

Program

Description

Question
Bayesian Filter

Program

Question

Description

Specification

• Variables

Identification
Bayesian Filter

Specification
- Variables
  \[ S^0, \ldots, S^t, O^0, \ldots, O^t \]
Bayesian Filter

Program

Description

Specification
- Variables
  \[ S^0, ..., S^t, O^0, ..., O^t \]
- Decomposition (Conditional Independance Hypothesis)

Identification

Question
Bayesian Filter

Specification

- Variables

\[ S^0, \ldots, S^t, O^0, \ldots, O^t \]

- Decomposition (Conditional Independence Hypothesis)

\[
P(S^0 \land \ldots \land S^t \land O^0 \land \ldots \land O^t) = P(S^0) \times P(O^0 | S^0) \times \prod_{i=2}^{t} [P(S^i | S^{i-1}) \times P(O^i | S^i)]
\]
Bayesian Filter

Specification

- Variables
  
  \[ S^0, \ldots, S^t, O^0, \ldots, O^t \]

- Decomposition (Conditional Independence Hypothesis)

\[
\mathbb{P}(S^0 \land \ldots \land S^t \land O^0 \land \ldots \land O^t) = \mathbb{P}(S^0) \times \mathbb{P}(O^0 | S^0) \times \prod_{i = 2}^{t} \left[ \mathbb{P}(S^i | S^{i-1}) \times \mathbb{P}(O^i | S^i) \right]
\]

Identification

- Parametric Forms or Bayesian Subroutines

Program

Description

Question
Bayesian Filter

Specification

- Variables
  \[S^0, \ldots, S^t, O^0, \ldots, O^t\]
- Decomposition (Conditional Independance Hypothesis)

\[
P(S^0 \land \ldots \land S^t \land O^0 \land \ldots \land O^t) = P(S^0) \times P(O^0 \mid S^0) \times \prod_{i=2}^{t} [P(S^i \mid S^{i-1}) \times P(O^i \mid S^i)]
\]

Identification

- Parametric Forms or Bayesian Subroutines

Program

- \(P(S^0) \equiv G(S^0, \mu, \sigma)\)
- \(P(S^i \mid S^{i-1}) \equiv G(S^i, A \cdot S^{i-1}, Q)\)
- \(P(O^i \mid S^i) \equiv G(O^i, H \cdot S^i, R)\)
Bayesian Filter

Specification

- Variables
  \[ S^0, \ldots, S^t, O^0, \ldots, O^t \]
- Decomposition (Conditional Independance Hypothesis)

\[
P(S^0 \wedge \ldots \wedge S^t \wedge O^0 \wedge \ldots \wedge O^t) = P(S^0) \times P(O^0 | S^0) \times \prod_{i=2}^{t} [P(S^i | S^{i-1}_i) \times P(O^i | S^i)]
\]

Program

- Parametric Forms
  or Bayesian Subroutines

Identification

- Learning from instances

\[
P(S^0) = G(S^0, \mu, \sigma)
\]
\[
P(S^i | S^{i-1}) = G(S^i, A \cdot S^{i-1}, Q)
\]
\[
P(O^i | S^i) = G(O^i, H \cdot S^i, R)
\]
Bayesian Filter

Specification
- Variables
  \[ S^0, \ldots, S^t, O^0, \ldots, O^t \]
- Decomposition (Conditional Independance Hypothesis)

\[
P(S^0 \land \ldots \land S^t \land O^0 \land \ldots \land O^t) = P(S^0) \times P(O^0 | S^0) \times \prod_{i=2}^{t} [P(S^i | S^{i-1}) \times P(O^i | S^i)]
\]

Program
- Parametric Forms
  or Bayesian Subroutines

Identification
- Learning from instances

\[
P(S^t | O^0 \land \ldots \land O^t)
\]
**Bayesian Filter**

\[
P(S' | O^{t-1}) = \sum_{S^{t-1}} [P(S' | S^{t-1}) \times P(S^{t-1} | O^{t-1})]
\]

\[
P(S' | O^{t}) = P(O' | S') \times P(S' | O^{t-1})
\]
**Bayesian Filter**

\[
P(S' \mid O^{0:t-1}) = \sum_{S^{t-1}} \left[ P(S' \mid S^{t-1}) \times P(S^{t-1} \mid O^{0:t-1}) \right]
\]

\[
P(S' \mid O^{0:t}) = P(O' \mid S') \times P(S' \mid O^{0:t-1})
\]
Bayesian Filter

\[
P(S' \mid O^{0:t-1}) = \sum_{S^{t-1}} \left[ P(S' \mid S^{t-1}) \times P(S^{t-1} \mid O^{0:t-1}) \right]
\]

\[
P(S' \mid O^{0:t}) = P(O' \mid S') \times P(S' \mid O^{0:t-1})
\]
Without vs with filtering

(Videos)
Real time filtering
Real time filtering
How to survive (perceive, reason, learn, decide and act) with incomplete information?

- Probability as an alternative to logic
- How to develop better artifacts using Bayesian reasoning?
- Biological plausibility of Bayesian reasoning at a macroscopic level?
- Biological plausibility of Bayesian reasoning at a microscopic level?
MODELING BEHAVIORS

- PhD Olivier Lebeltel
- PhD Jean Laurens
- PhD C. Moulin-Frier
- PhD Francis Colas
- PostDoc Francis Colas

Diagram showing anatomical parts including:
- canals
- otoliths
- vestibular nerve
- collea

Image with labeled points:
- TD
- LP
- LH
- TT
- J
- Lx
- TB

Images of models and anatomical structures.
Bayesian Action Perception:
Handwriting experiments

Ph.D Estelle Gilet

Motor Equivalence?
Motor Equivalence?

- Writer “style”
  ➢ [Wright90]
- Common activated motor areas
  ➢ [Wing00]

[Serratrice93]
Simulation of action during perception?

[Longcamp03]

Writing

Pseudo letter reading

Letter reading
Reading

- OCR
  - [Meulenbroek96]
  - [Flash95]
- Human models
  - [Crettez98]
  - [Vuori02]
  - [Dehaene07]
Writing

- [Hinton05]
- [Meulenbroek96]
- [Flash95]
BAP Model

 Perception

 Read trajectory

 Action

 Generated trajectory

 Sensations

 Motor gestures
From motor perception theory to Bayesian Action Perception
From motor perception theory to Bayesian Action Perception

Motor perception theory
[Liberman57]

Perception for action control theory
[Schwartz01]
A common space for motor and perception

Internal representation
Common features for both representations

\[
\frac{dx}{dt}(t) = 0 \lor \frac{dy}{dt}(t) = 0
\]
**Here come the probabilities**

\[ P(C_L^{0:N} | L W) = P(C_L^0 | L W) P(L) P(W) \]

\[ = \prod_{n=1}^{N} \left( \begin{array}{c} P(C_{Lx}^n | C_{Lx}^{n-1}, L W) P(C_{Ly}^n | C_{Ly}^{n-1}, L W) \\ P(C_{Lx}^n | C_{Lx}^{n-1}, L W) P(C_{Ly}^n | C_{Ly}^{n-1}, L W) \end{array} \right) \]

\[ \times P(L) P(W) \]
Learning succession of control points

\[
P(C_{Lx}^3 \mid [C_{Lx}^2 = 15] [L = l] [W = Julienne]) = \frac{p_i + \alpha}{N + k\alpha}.
\]
Learning succession of control points

\[ P(C_{Lx}^3 \mid C_{Lx}^2 [L = I] [W = Julienne]) \]
Learning succession of control points

\[ P(C^3_{Lx} | C^2_{Lx} [L = I] [W = Julienne]) \]
Learning succession of control points

\[ P(C^3_{Lx} \mid C^2_{Lx} [L = l] [W = Julienne]) \]
Learning succession of control points

\[ P(C_{Lx}^3 \mid C_{Lx}^2 \mid L = l \mid W = Julienne) \]
Learning succession of control points

\[ P(C^3_{Lx} \mid C^2_{Lx} [L = l] [W = Julienne]) \]
BAP Model

- Letter
  - \( \lambda_L \)
  - Perceptive internal representation
    - \( \lambda_V \)
      - Perceptive internal representation
        - \( V \)
          - Read trajectory
        - \( C_V \)
      - Motor internal representation
        - \( \lambda_P \)
          - Generated trajectory
            - \( P \)
              - Effector
            - \( E \)
          - Simulated internal representation
            - \( \lambda_S \)
              - Simulated generated trajectory
                - \( S \)
  - Writer
    - \( W \)
      - Simulated internal representation
        - \( C_L \)

BAP MODEL

\[
P \begin{pmatrix}
C_{LV} & C_{LP} & C_{LS} & C_{V} & \lambda_{P} & \lambda_{V} & \lambda_{S} & \lambda_{I} & E \\
C_{P} & C_{S} & V & P & L & W & S
\end{pmatrix}
\]

\[
\begin{align*}
P(C_{LV} | L \ W) & \cdot P(C_{LP} | L \ W) \cdot P(C_{LS} | L \ W) \\
P(L) & \cdot P(W) \cdot P(\lambda_{I} | C_{LV} \ C_{LP}) \\
P(\lambda_{V} | C_{LV} \ C_{V}) & \cdot P(C_{V} | V) \cdot P(V) \\
P(\lambda_{P} | C_{LP} \ C_{P}) & \cdot P(P | C_{P}) \cdot P(E | P) \\
P(\lambda_{S} | C_{LS} \ C_{S}) & \cdot P(C_{S} | S) \cdot P(S | P)
\end{align*}
\]
BAP Model
BAP model
BAP Model
BAP Model
Letter recognition
knowing the scripter

\[ P(L \mid [V_X^{0:M} = v_X^{0:M}] [V_Y^{0:M} = v_Y^{0:M}] [W = w] [\lambda_V = 1]) \]
**Letter recognition**

**knowing the scripter**
**Letter recognition**

**Knowing the scripter**

![Letter recognition diagram]

93.36%
Scripter recognition
knowing the letter

\[ P(W \mid [V_X^{0:M} = v_X^{0:M}] [V_Y^{0:M} = v_Y^{0:M}] [L = l] [\lambda_V = 1]) \]

<table>
<thead>
<tr>
<th></th>
<th>Estelle</th>
<th>Julienne</th>
<th>Jean-Louis</th>
<th>Christophe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estelle</td>
<td>0.76</td>
<td>0.03</td>
<td>0.07</td>
<td>0.14</td>
</tr>
<tr>
<td>Julienne</td>
<td>0.02</td>
<td>0.80</td>
<td>0.07</td>
<td>0.11</td>
</tr>
<tr>
<td>Jean-Louis</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Christophe</td>
<td>0.10</td>
<td>0.14</td>
<td>0.13</td>
<td>0.62</td>
</tr>
</tbody>
</table>

79,5%
Motor control

\[ P(\dot{\theta}_1^{0:T}, \dot{\theta}_2^{0:T} \mid [L = l] [W = w] [\lambda_p = 1]) \]
Motor control
Inter scripter variability

\[ P(\dot{\theta}_1^0:T, \dot{\theta}_2^0:T \mid [L = l] [W = w] [\lambda_P = 1]) \]
Motor equivalence
Motor equivalence

$W = Estelle \quad W = Christophe \quad W = Julienne$

Bras simulé

Bras robotique

Robot holonome
Copy

Trace copy

Letter copy
Letter recognition with motor simulation

\[ P(L \mid [V^0:M_x = v^0:M_x] [V^0:M_y = v^0:M_y] [W = w] [\lambda V = 1] [\lambda_L = 1] [\lambda_P = 1] [\lambda_S = 1]) \]
Letter recognition with motor simulation
Letter recognition with motor simulation
**Results**

<table>
<thead>
<tr>
<th></th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>k</th>
<th>l</th>
<th>m</th>
<th>n</th>
<th>o</th>
<th>p</th>
<th>q</th>
<th>r</th>
</tr>
</thead>
<tbody>
<tr>
<td>With motor simulation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Without motor simulation</td>
<td>0</td>
<td>0.9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Extracts of the probability distributions over letters, computed as solutions to the reading task with (top row) and without (bottom row) motor simulation, when presented with the truncated $g$ shown Fig. 23.

doi:10.1371/journal.pone.0020387.t003
Perspectives
Speech? (Ph.D in progress)
How to survive (perceive, reason, learn, decide and act) with incomplete information?

Probability as an alternative to logic

How to develop better artifacts using Bayesian reasoning?

Biological plausibility of Bayesian reasoning at a macroscopic level?

Biological plausibility of Bayesian reasoning at a microscopic level?
Amoeba

How is it performing probabilistic inference?
Amoeba

How is it performing probabilistic inference?

Cell signaling
8 ALLOSTERIC STATES

2 MESSENGERS
8 ALLOSTERIC STATES
2 MESSENGERS

\[
P([\Omega_1 = 0] \land [\Omega_2 = 0] \land [\Omega_3 = 0]) = \frac{1}{D}
\]

\[
P([\Omega_1 = 0] \land [\Omega_2 = 1] \land [\Omega_3 = 0]) = \frac{k_{000 \rightarrow 010} \times m_1}{D}
\]

\[
P([\Omega_1 = 1] \land [\Omega_2 = 0] \land [\Omega_3 = 0]) = \frac{k_{000 \rightarrow 100} \times m_1}{D}
\]

\[
P([\Omega_1 = 1] \land [\Omega_2 = 1] \land [\Omega_3 = 0]) = \frac{k_{000 \rightarrow 010} \times k_{010 \rightarrow 110} \times m_1 \times m_2}{D}
\]

\[
P([\Omega_1 = 0] \land [\Omega_2 = 1] \land [\Omega_3 = 1]) = \frac{k_{000 \rightarrow 001}}{D}
\]

\[
P([\Omega_1 = 0] \land [\Omega_2 = 1] \land [\Omega_3 = 1]) = \frac{k_{000 \rightarrow 010} \times k_{010 \rightarrow 011} \times m_2}{D}
\]

\[
P([\Omega_1 = 1] \land [\Omega_2 = 0] \land [\Omega_3 = 1]) = \frac{k_{000 \rightarrow 100} \times k_{100 \rightarrow 101} \times m_1}{D}
\]

\[
P([\Omega_1 = 1] \land [\Omega_2 = 1] \land [\Omega_3 = 1]) = \frac{k_{000 \rightarrow 010} \times k_{010 \rightarrow 110} \times k_{110 \rightarrow 111} \times m_1 \times m_2}{D}
\]
8 allosteric states
2 messengers

\[ O([\Omega_3]) = \frac{k_{000} \rightarrow 001 + k_{000} \rightarrow 100 \times k_{100} \rightarrow 101 \times m_1}{1 + k_{000} \rightarrow 100 \times m_1 + k_{000} \rightarrow 010 \times m_2 + k_{000} \rightarrow 010 \times k_{010} \rightarrow 110 \times m_1 \times m_2} \]
Bayesian gate

\[ \Sigma = \frac{P(S = 1 | \phi_1 \phi_2 [\lambda_1 = 1][\lambda_2 = 1])}{P(S = 0 | \phi_1 \phi_2 [\lambda_1 = 1][\lambda_2 = 1])} \]

\[ = \frac{P(f_1 = 0 [f_2 = 0 [S = 1]]) + P(011)\phi_2 + P(101)\phi_1 + P(111)\phi_1 \phi_2}{P(000) + P(010)\phi_2 + P(100)\phi_1 + P(110)\phi_1 \phi_2} \]
Bayesian gate

\[ O([\Omega_3]) = \frac{k_{000} \rightarrow 001 + k_{000} \rightarrow 100 \times k_{100} \rightarrow 101 \times m_1 + k_{000} \rightarrow 010 \times k_{010} \rightarrow 011 \times m_2 + k_{000} \rightarrow 010 \times k_{000} \rightarrow 010 \times k_{010} \rightarrow 110 \times k_{110} \rightarrow 111 \times m_1 \times m_2}{1 + k_{000} \rightarrow 100 \times m_1 + k_{000} \rightarrow 010 \times m_2 + k_{000} \rightarrow 010 \times m_1 + m_1 \times m_2} \]

\[ \Sigma = \frac{P([S = 1] \mid \phi_1 \phi_2 \lambda_1 = 1 \lambda_2 = 1)}{P([S = 0] \mid \phi_1 \phi_2 \lambda_1 = 1 \lambda_2 = 1)} \]

\[ = \frac{P([f_1 = 0] [f_2 = 0] [S = 1]) + P(011) \phi_2 + P(101) \phi_1 + P(111) \phi_1 \phi_2}{P(000) + P(010) \phi_2 + P(100) \phi_1 + P(110) \phi_1 \phi_2} \]
Bayesian biochemistry: Basic ideas

- **Bayesian values** -> Concentration of messengers, membrane potential & spike frequency
- **Bayesian gates** -> Equilibrium between allosteric macromolecules & messengers
- **Bayesian inference** -> Signal transduction.
- The interplay between local biochemical mechanisms and distant electrical propagation in neurons is the key level to understand brain computation.
Bayesian biochemistry: Open questions

- How is information encoded at the different scales (molecular, intra-cellular, cellular, inter-cellular, population, system)?

- How is information processed at these different scales?

- How is information memorized at these different scales?

- What is meant by learning and adaptation at these different scales?

- Do sensory-motor systems perceive values or probabilities of values?

- How do they make decisions on the actions to perform?
Want to know more?

Bayesian-Programming.org

Pierre.Bessiere@College-de-france.fr