Bayesian models for perception, action, inference and learning

Pierre Bessière

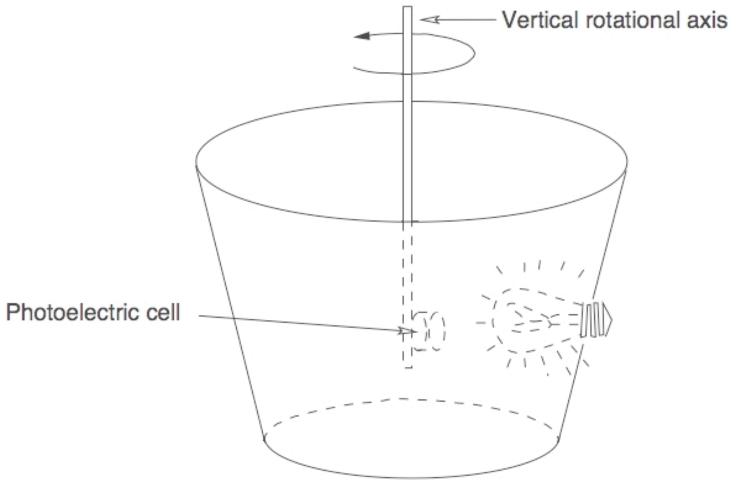
CNRS - University of Grenoble

www-laplace.imag.fr

A Bayesian brain ?

- Is Bayesian inference and learning a relevant model of cognition ?
- Is the brain performing probabilistic computation ?

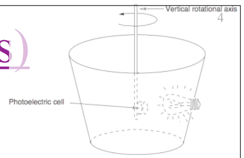
Beam-in-the-Bin experiment (Set-up)

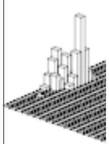


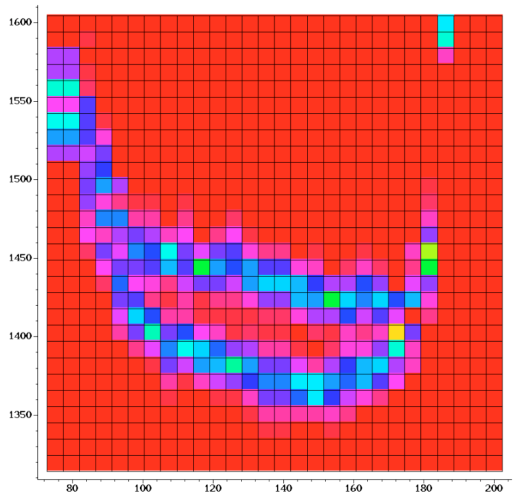
Bessière©2005

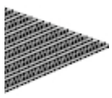
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Beam-in-the-Bin experiment (Results)





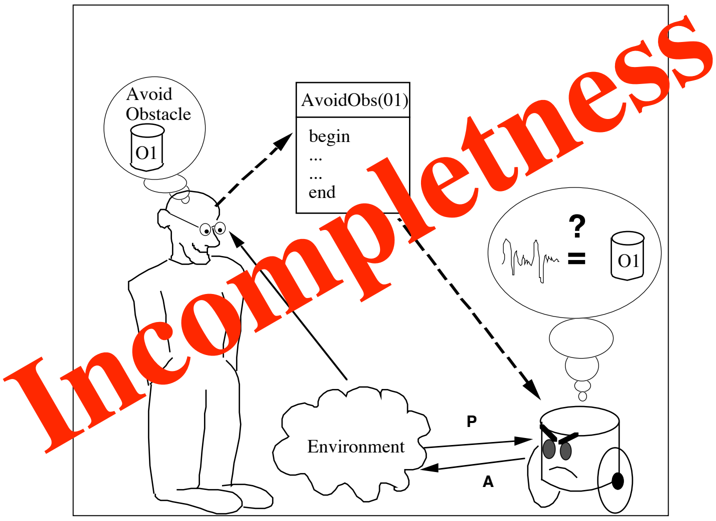




Bessière©2005

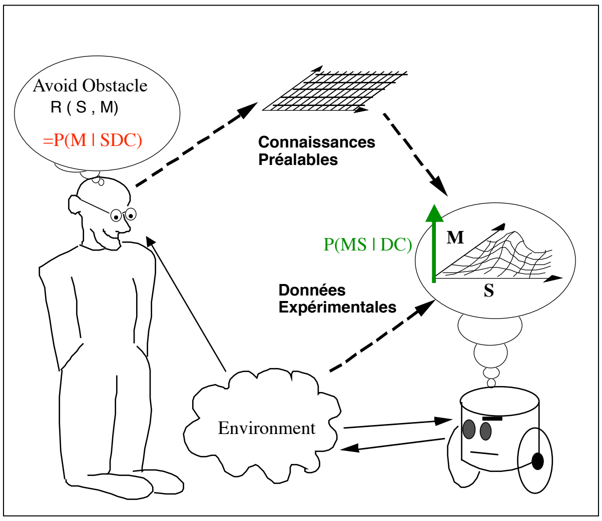
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Logical Paradigm



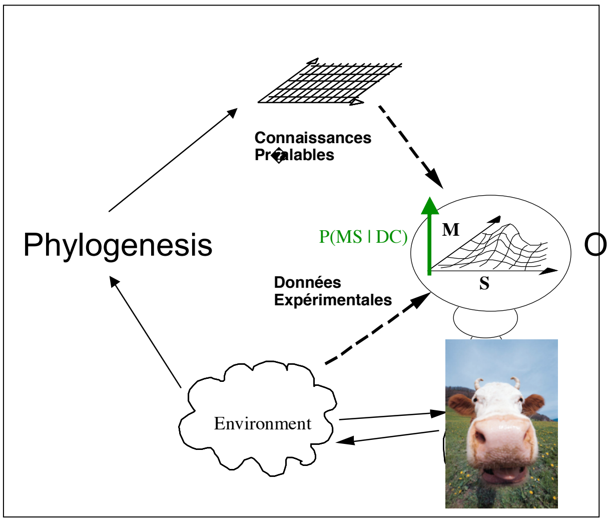
Bessière©2005

Bayesian Paradigm



Bessière©2005

Bayesian Paradigm



Ontogenesis

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Principle

Incompleteness

Preliminary Knowledge
+
Experimental Data
=
Probabilistic Representation

Maximum Entropy Principle

$$-\sum P_i \log(P_i)$$

Uncertainty

Bayesian Inference

P(AB|C)=P(A|C)P(B|AC)=P(B|C)P(A|BC)

P(A|C)+P(¬A|C) = 1

Decision

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Thesis

Probabilistic inference and learning theory, considered as a **model of reasonning**, is a new paradigm (an alternative to logic) to explain and understand perception, inference, decision, learning and action.

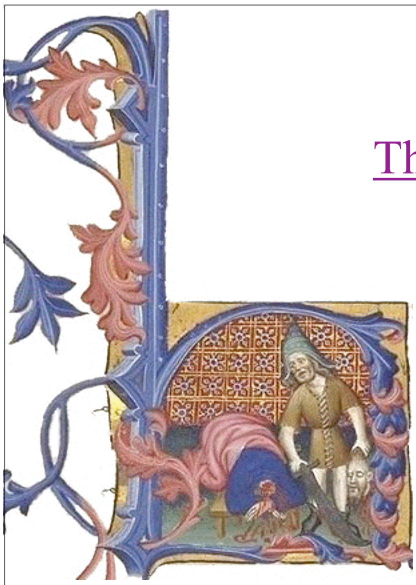
La théorie des probabilités n'est rien d'autre que le sens commun fait calcul.
Marquis Pierre-Simon de Laplace

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

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 .*
E.T. Jaynes

Subjectivist vs Objectivist epistemology of probabilities ?

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE



Theoretical Basis

Content:

- Definitions and notations
- Inference rules
- Bayesian program
- Model specification
- Model identification
- Model utilization

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Logical Proposition

Logical Proposition are denoted by lowercase name: a

Usual logical operators:

$$a \wedge b$$
$$a \vee b$$
$$\neg a$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Probability of Logical Proposition

We assume that to assign a probability to a given proposition a, it is necessary to have at least some **preliminary knowledge**, summed up by a proposition π.

$$P(a \mid \pi) \in [0,1]$$

Of course, we will be interested in reasoning on the probabilities of the conjunctions, disjunctions and negations of propositions, denoted, respectively, by:

$P(a \wedge b \mid \pi)$ $P(a \vee b \mid \pi)$ $P(\neg a \mid \pi)$

We will also be interested in the probability of proposition a conditioned by both the preliminary knowledge π and some other proposition b:

$$P(a \mid b \wedge \pi)$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Normalization and Conjunction Postulates

$$P(a \mid \pi) + P(\neg a \mid \pi) = 1$$

$$\begin{aligned} P(a \wedge b \mid \pi) &= P(a \mid \pi) \times P(b \mid a \wedge \pi) \\ &= P(b \mid \pi) \times P(a \mid b \wedge \pi) \end{aligned}$$

$$P(a \vee b \mid \pi) = P(a \mid \pi) + P(b \mid \pi) - P(a \wedge b \mid \pi)$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Syllogisms

$a \equiv$ "x divisible by 9"
$b \equiv$ "x divisible by 3"

- Logical Syllogisms:
  - Modus Ponens: $a \wedge [a \Rightarrow b] \mapsto b$
  - Modus Tollens: $\neg b \wedge [a \Rightarrow b] \mapsto \neg a$
- Probabilistic Syllogisms:
  - Modus Ponens: $P(b \mid a \wedge \pi) = 1$
  - Modus Tollens: $P(b \mid a \wedge \pi) = 1 \Leftrightarrow P(\neg a \mid \neg b \wedge \pi) = 1$
  - $P(b \mid a \wedge \pi) = 1 \Rightarrow P(a \mid b \wedge \pi) \geq P(a \mid \pi)$
  - $P(b \mid a \wedge \pi) = 1 \Rightarrow P(b \mid \neg a \wedge \pi) \leq P(b \mid \pi)$
  - $P(f \mid g \wedge \pi) \geq P(f \mid \pi) \Rightarrow P(g \mid f \wedge \pi) \geq P(g \mid \pi)$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Discrete Variable

- Variable are denoted by name starting with one uppercase letter: X
- By definition a discrete variable X is a set of propositions x_i
  - Mutually exclusive: $i \neq j \Rightarrow \left[x_i \wedge x_j\right] = false$
  - Exhaustive: at least one is true

The cardinal of X is denoted: $|X|$

Variable Conjunction

$$X \wedge Y = \{x_i \wedge y_j\}$$

$$X \vee Y = \{x_i \vee y_j\}$$

Not a variable

Conjunction rule

$$\forall x_i \in X, \forall y_j \in Y$$

$$P(x_i \wedge y_j \mid \pi) = P(x_i \mid \pi) \times P(y_j \mid x_i \wedge \pi)$$
$$= P(y_j \mid \pi) \times P(x_i \mid y_j \wedge \pi)$$

$$P(X \wedge Y \mid \pi) = P(X \mid \pi) \times P(Y \mid X \wedge \pi)$$
$$= P(Y \mid \pi) \times P(X \mid Y \wedge \pi)$$

Normalization rule

$$\sum_X P(X \mid \pi) = 1$$

Marginalization rule

$$\sum_X \mathrm{P}(X \wedge Y \mid \pi) = \mathrm{P}(Y \mid \pi)$$

Contraction/Expansion rule

$$X \wedge Y = A \Rightarrow \mathrm{P}(X \wedge Y \mid \pi) = \mathrm{P}(A \mid \pi)$$

Rules

$$P(X \wedge Y \mid \pi) = P(X \mid \pi) \times P(Y \mid X \wedge \pi)$$
$$= P(Y \mid \pi) \times P(X \mid Y \wedge \pi)$$

$$\sum_X P(X \mid \pi) = 1$$

$$\sum_X P(X \wedge Y \mid \pi) = P(Y \mid \pi)$$

$$X \wedge Y = A \Rightarrow P(X \wedge Y \mid \pi) = P(A \mid \pi)$$

Bayesian Program

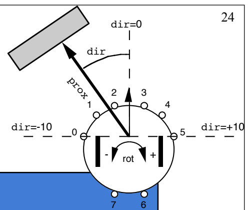
- Program
  - Description
    - *Specification*
      - Variables
      - Decomposition
      - Parametrical Forms or Recursive Question
      - Preliminary Knowledge π
    - *Identification*
      - Experimental Data δ
  - Question
    - Utilization

Learning Reactive Behaviors

Khepera

- Avoiding Obstacle
- Contour Following
- Piano mover
- Phototaxy
- etc.

Pushing Objects (model)



Program

Description

Question

Specification

- Variables
  - Dir, Prox, Vrot
- Decomposition

$$P(Dir \wedge Prox \wedge Vrot \mid \pi)$$
$$= P(Dir \mid \pi) \times P(Prox \mid \pi) \times P(Vrot \mid Dir \wedge Prox \wedge \pi)$$

- Parametrical Forms

$P(Dir \wedge Prox \mid \pi) \leftarrow$ Uniform

$P(Vrot \mid Dir \wedge Prox \wedge \pi) \leftarrow$ Gaussians

➔ Preliminary Knowledge π

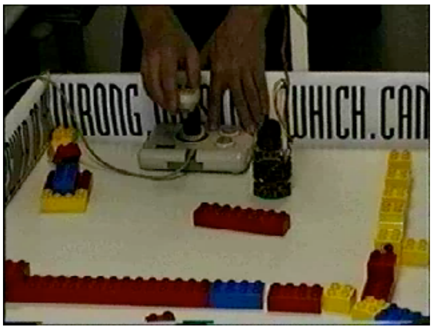
Ide

$$P(Dir \wedge Prox \wedge Vrot \mid \delta 1 \wedge \pi)$$

Utili

$$P(Vrot \mid [Dir = d] \wedge [Prox = p] \wedge \delta 1 \wedge \pi)$$

Pushing Objects (Result)



Contour Following (Model)

Program — Description — Question

Specification

- Variables

    Dir, Prox, Vrot

- Decomposition

$$P(Dir \wedge Prox \wedge Vrot \mid \pi)$$

$$= P(Dir \mid \pi) \times P(Prox \mid \pi) \times P(Vrot \mid Dir \wedge Prox \wedge \pi)$$

- Parametrical Forms

$$P(Dir \wedge Prox \mid \pi) \leftarrow \text{Uniform}$$

$$P(Vrot \mid Dir \wedge Prox \wedge \pi) \leftarrow \text{Gaussians}$$

➔ Preliminary Knowledge π

Identification

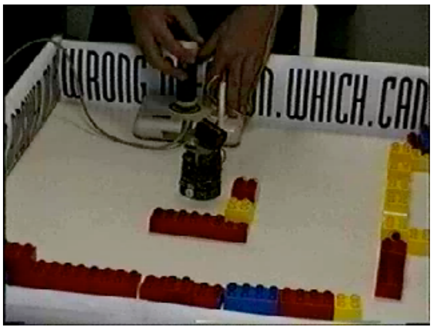
- Joystick Remote Control ➔ Experimental Data δ2

$$P(Dir \wedge Prox \wedge Vrot \mid \delta 2 \wedge \pi)$$

Utilization

$$P(Vrot \mid [Dir = d] \wedge [Prox = p] \wedge \delta 2 \wedge \pi)$$

Contour Following (Result)



Description

The purpose of a description is to specify an effective method to compute a joint distribution on a set of variables:

$\{X^1, X^2, \ldots, X^n\}$ $\{Dir, Prox, Vrot\}$

Given some preliminary knowledge π and a set of experimental data δ.
This joint distribution is denoted as:

$P(X^1 \wedge X^2 \wedge \ldots \wedge X^n \mid \pi)$ $P(Dir \wedge Prox \wedge Vrot \mid \pi)$

Decomposition

Partion in K subsets:

$L^i = X^{k_1} \wedge X^{k_2} \wedge \ldots$

$L^1 = Dir$ $L^2 = Prox$ $L^3 = Vrot$

Conjunction rule:

$$P(X^1 \wedge X^2 \wedge \ldots \wedge X^n \mid \pi) \\ = P(L^1 \mid \pi) \times P(L^2 \mid L^1 \wedge \pi) \times \ldots \times P(L^k \mid L^{k-1} \wedge \ldots \wedge L^1 \wedge \pi)$$

$$P(Dir \wedge Prox \wedge Vrot \mid \pi) \\ = P(Dir \mid \pi) \times P(Prox \mid Dir \wedge \pi) \\ \times P(Vrot \mid Dir \wedge Prox \wedge \pi)$$

Conditional independance:

$$P(L^i \mid L^{i-1} \wedge \ldots \wedge L^1 \wedge \pi) \\ = P(L^i \mid R^i \wedge \pi)$$

$$P(Prox \mid Dir \wedge \pi) = P(Prox \mid \pi)$$

Decomposition:

$$P(X^1 \wedge X^2 \wedge \ldots \wedge X^n \mid \pi) \\ = P(L^1 \mid \pi) \times P(L^2 \mid R^2 \wedge \pi) \times \ldots \times P(L^k \mid R^k \wedge \pi)$$

$$P(Dir \wedge Prox \wedge Vrot \mid \pi) \\ = P(Dir \mid \pi) \times P(Prox \mid \pi) \\ \times P(Vrot \mid Dir \wedge Prox \wedge \pi)$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Parametrical Forms or Recursive Questions

Parametrical form:

$$P(L^i \mid R^i \wedge \delta \wedge \pi) = f_{\mu(R^i, \delta)}(L^i)$$

$$P(Dir \wedge Prox \mid \delta \wedge \pi) \leftarrow \text{Uniform} \\ P(Vrot \mid Dir \wedge Prox \wedge \delta \wedge \pi) \leftarrow \text{Gaussians}$$

Recursive Question:

$$P(L^i \mid R^i \wedge \delta \wedge \pi) = P(L^i \mid R^i \wedge \delta' \wedge \pi')$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Question

Given a description, a question is obtained by partitionning the set of variables into 3 subsets: the searched variables, the known variables and the unknown variables.

We define the *Search*, *Known* and *Unknown* as the conjunctions of the variables belonging to these three sets.

We define the corresponding question as the distribution:

$$P(Search \mid Known \wedge \delta \wedge \pi)$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Inference

$$P(Search \mid Known \wedge \delta \wedge \pi) = \sum_{Unknown} P(Search \wedge Unknown \mid Known \wedge \delta \wedge \pi)$$

$$= \frac{\sum_{Unknown} P(Search \wedge Unknown \wedge Known \mid \delta \wedge \pi)}{P(Known \mid \delta \wedge \pi)}$$

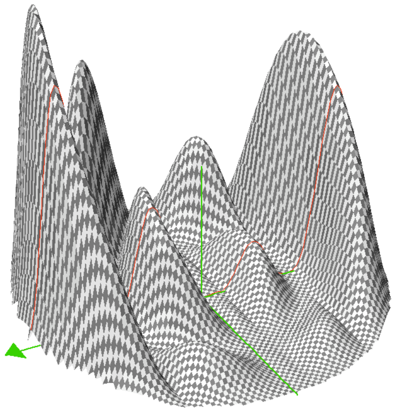
$$= \frac{\sum_{Unknown} P(Search \wedge Unknown \wedge Known \mid \delta \wedge \pi)}{\sum_{Search, Unknown} P(Search \wedge Unknown \wedge Known \mid \delta \wedge \pi)}$$

$$= \frac{1}{Z} \times \sum_{Unknown} P(Search \wedge Unknown \wedge Known \mid \delta \wedge \pi)$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

2 optimization problems

$\text{Draw}(P(Search \mid Known \wedge \delta \wedge \pi))$



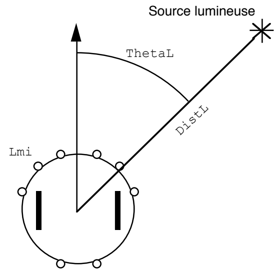
$$P(Search \mid Known \wedge \delta \wedge \pi) = \frac{1}{Z} \times \sum_{Unknwon} P(Search \wedge Known \wedge Unknown \mid \delta \wedge \pi)$$

Perception, Sensor Fusion and Multimodality

TEST 1

Daniel J. Simons
Christopher Chabris
Harvard University

Sensor Fusion



Objective: Find the position of a light source

- Difficulty:
  The robot do not have any sensor
  to find directly the position of a light source.

- Solution:
  - Model of each sensor
  - Fusion of the 8 models

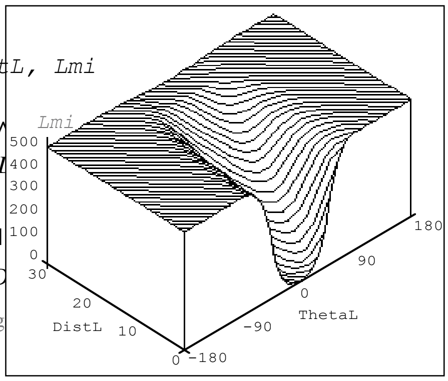
Model of a Light Sensor

Program
Description
Question

Specification

- Variables
  `ThetaL, DistL, Lmi`
- Decomposition
  $P(ThetaL \wedge DistL \wedge$
  $= P(ThetaL \wedge DistL$
- Parametrical Forms
  $P(ThetaL \wedge DistL \mid$
  $P(Lmi \mid ThetaL \wedge D$

➔ Preliminary Knowledg



Identification

- A priori specification

Utilization

$P(ThetaL \mid [Lmi = li] \wedge \delta_i \wedge \pi_{Sensor})$

$P(DistL \mid [Lmi = li] \wedge \delta_i \wedge \pi_{Sensor})$

Model of a Light Sensor (2)

Bayesian Inference: Inverse Problem

Description:

$$P(ThetaL \wedge DistL \wedge Lmi \mid \delta_i \wedge \pi_{Sensor})$$
$$= P(ThetaL \wedge DistL \mid \pi_{Sensor}) \times P(Lmi \mid ThetaL \wedge DistL \wedge \delta_i \wedge \pi_{Sensor})$$

Question 1:

$$P(ThetaL \mid lmi \wedge \delta_i \wedge \pi_{Sensor})$$
$$= \frac{1}{Z} \times \sum_{Distl} P(lmi \mid ThetaL \wedge DistL \wedge \delta_i \wedge \pi_{Sensor})$$

Proof

Question 2:

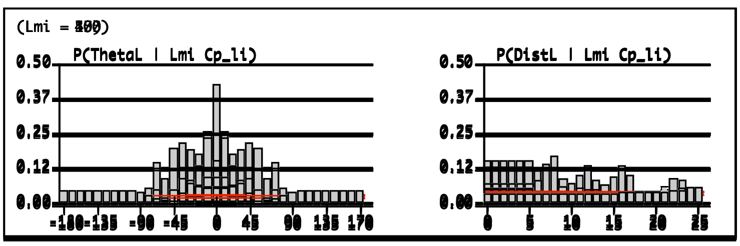
$$P(DistL \mid lmi \wedge \delta_i \wedge \pi_{Sensor})$$
$$= \frac{1}{Z} \times \sum_{ThetaL} P(lmi \mid ThetaL \wedge DistL \wedge \delta_i \wedge \pi_{Sensor})$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Model of a Light Sensor (3)

P(ThetaL | Lmi)

P(DistL | Lmi)



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Sensor Fusion (2)

Program
- Description
  - *Specification*
    - Variables

      `ThetaL, DistL, Lm0, …, Lm7`
    - Decomposition (Conditional Independance Hypothesis)

      $$P(ThetaL \wedge DistL \wedge Lm0 \wedge \ldots \wedge Lm7 \mid \wedge \pi_{Fusion})$$

      $$= P(ThetaL \wedge DistL \mid \pi_{Fusion}) \times \prod_{i=0}^{7} P(Lmi \mid ThetaL \wedge DistL \wedge \pi_{Fusion})$$
    - Parametrical Forms

      $$P(ThetaL \wedge DistL \mid \pi_{Fusion}) \leftarrow \text{Uniform}$$

      $$P(Lmi \mid ThetaL \wedge DistL \wedge \pi_{Fusion}) \leftarrow P(Lmi \mid ThetaL \wedge DistL \wedge \delta_i \wedge \pi_{Sensor})$$
  - *Identification*
    - No free parameters
- Question
  - Utilization

    $$P(ThetaL \mid lm0 \wedge \ldots \wedge lm7 \wedge \pi_{Fusion})$$

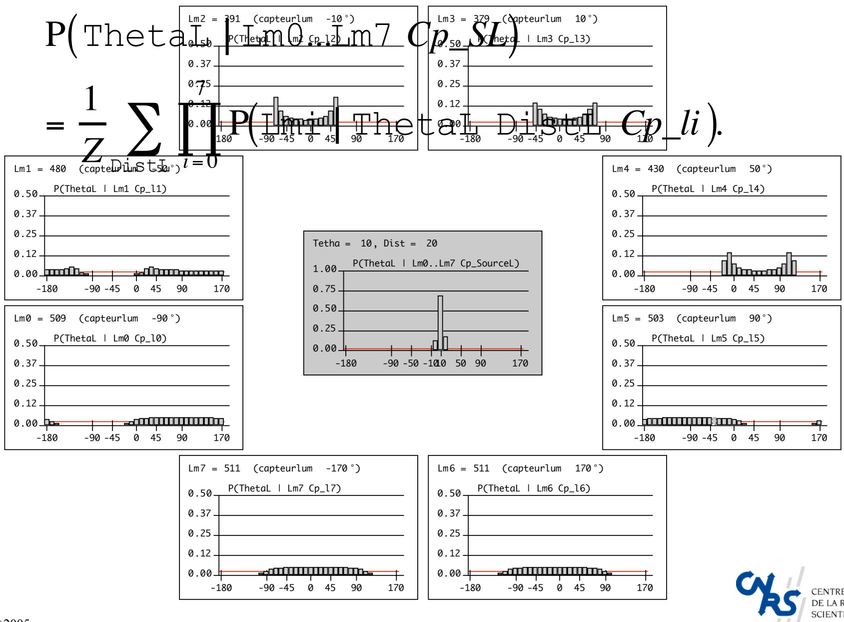
    $$P(Lm3 \mid lm2 \wedge lm4 \wedge ThetaL \wedge \pi_{Fusion})$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Sensor Fusion (3)

$$P(\texttt{ThetaL} \mid \texttt{Lm0..Lm7}\ Cp\_SL)$$

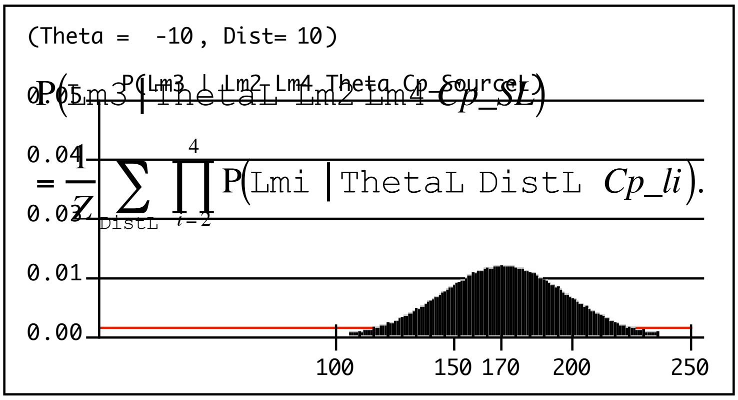
$$= \frac{1}{Z} \sum_{\texttt{DistL}} \prod_{i=0}^{7} P(\texttt{Lmi} \mid \texttt{ThetaL DistL}\ Cp\_li).$$



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

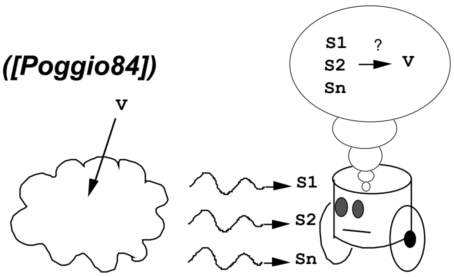
Sensor Fusion (4)

$$P(Lm3 \mid ThetaL\ Lm2\ Lm4\ Cp\_SL) = \frac{1}{Z} \sum_{DistL} \prod_{i=2}^{4} P(Lmi \mid ThetaL\ DistL\ Cp\_li).$$



Sensor Fusion (5)

Perception: an inverse problem ([Poggio84])



Bayesian answer :

Inversion + conditional independance hypothesis :

P(S1 S2 ... Sn V | C)

= P(V | C)P(S1 | V C)P (S2 | V C)...P(Sn | V C)

43

McGurk effect

Test 2

M-A. Cathiard
J-L. Schwartz
C. Abry

Cathiard, M.-A., Schwartz, J.-L. & Abry, C. (2001). Asking a naive question to the McGurk effect : why does audio [b] give more [d] percepts with usual [g] than with visual [d] ? In Proceedings of the /Auditory Visual Speech processing, AVSP'2001/, Aalborg, Copenhague, 138-142.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

44

McGurk (model)

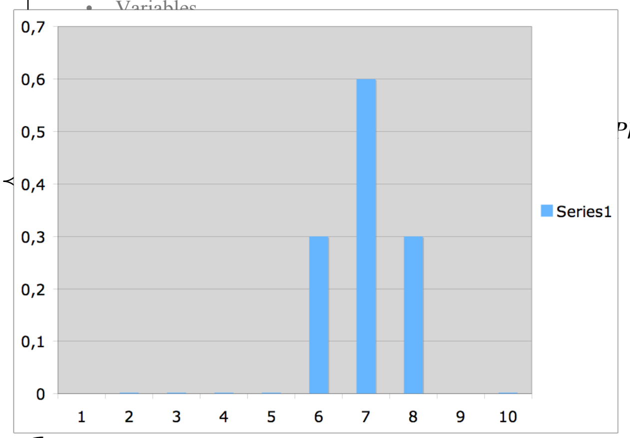
Program

Description

Specification

• Variables

Phoneme)



Question

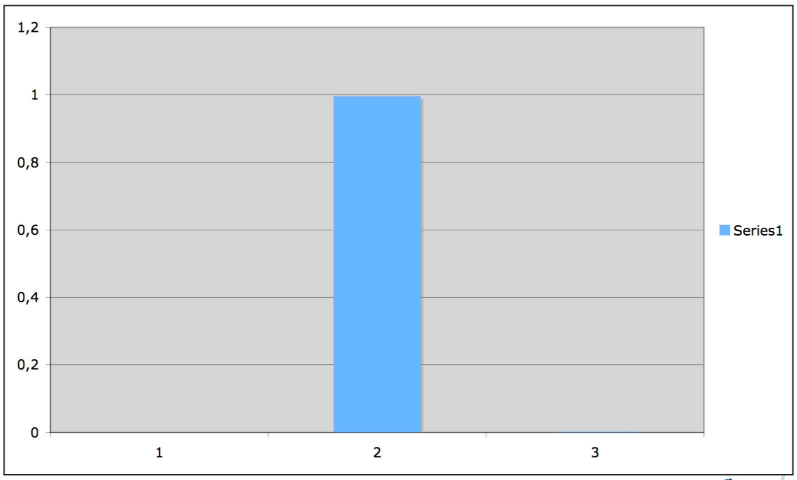
Utilization

$P(Phoneme \mid Occlusion \wedge Lips)$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

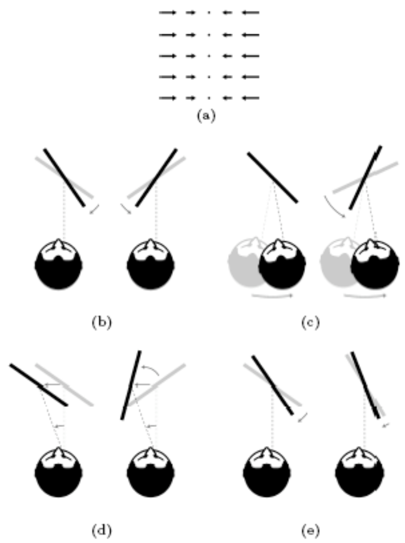
McGurk (result)

$$P(Phoneme \mid [Occlusion=2] \wedge [Lips=open])$$
$$= \frac{1}{Z} \times P([Occlusion=2] \mid Phoneme) \times P([Lips=open] \mid Phoneme)$$



Active perception

« Shape from Motion » (F. Colas)



- Inverse problem
- Ill posed problem

- Stationnarity
- Rigidity

Shape from motion (model)

Program

Description

Specification

- Variables

  $X\ Y\ \vec{\Sigma}\ \Lambda\ \vec{\Omega}\ \vec{T}\ \vec{\Phi}$

- Decomposition

$$
\begin{aligned}
& P(X\ Y\ \vec{\Sigma}\ \Lambda\ \vec{\Omega}\ \vec{T}\ \vec{\Phi}) \\
= \ & P(X\ Y) P(\vec{\Sigma}) P(\Lambda) \\
\times \ & P(\vec{\Omega}\ \vec{T} \mid \vec{\Sigma}) \\
\times \ & P(\Phi^0 \mid T_x\ T_y) \\
\times \ & P(\Phi^1 \mid X\ Y\ \vec{\Omega}\ \vec{T}) \\
\times \ & P(\Phi^2 \mid X\ Y\ \vec{\Omega}\ \vec{T}\ \Lambda)
\end{aligned}
$$

Stationarity ($P(\vec{\Omega}\ \vec{T} \mid \vec{\Sigma})$)

Rigidity ($P(\Phi^1 \mid \ldots)$, $P(\Phi^2 \mid \ldots)$)

- Parametrical Forms

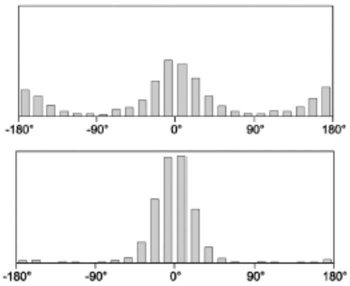
  Physical and physiological law → distributions

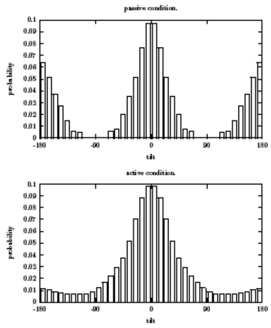
Identification

Question

$$
\begin{aligned}
& P(X\ Y \mid \vec{\varphi}\ \vec{\sigma}\ \lambda) \\
\propto \ & P(X\ Y) \sum_{\vec{\Omega}, \vec{T}} \left\{ \begin{array}{ll} & P(\vec{\Omega}\ \vec{T} \mid \vec{\sigma}) \\ \times & P(\varphi^0 \mid T_x\ T_y) \\ \times & P(\varphi^1 \mid X\ Y\ \vec{\Omega}\ \vec{T}) \\ \times & P(\varphi^2 \mid X\ Y\ \vec{\Omega}\ \vec{T}\ \lambda) \end{array} \right\}
\end{aligned}
$$

Orientation and stationarity

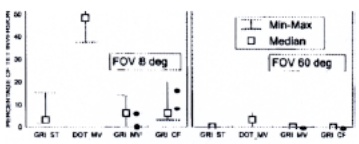


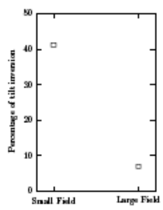


J.J.A. van Boxtel, M. Wexler, and J. Droulez. Perception of plane orientation from self-generated and passively observed optic flow. *Journal of Vision*, 3(5):318–332, 2003.

CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

Orientation and field of vision

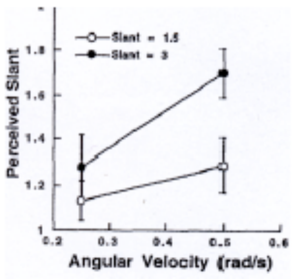


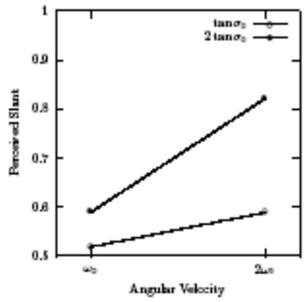


V. Cornilleau-Pérès, M. Wexler, J. Droulez, E. Marin, C. Miège, and B. Bourdoncle. Visual perception of planar orientation: dominance of static depth cues over motion cues. *Vision Research*, 42:1403–1412, 2002.

CENTRE NATIONAL
DE LA RECHERCHE
SCIENTIFIQUE

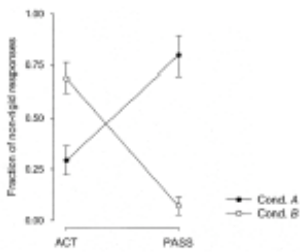
Inclination and angular speed

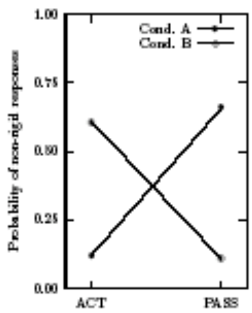




F. Domini and C. Caudex. Perceiving surface slant from deformation of optic flow. *J Exp Psychol Hum Percept Perform.*, 25(2):426–44, 1999.

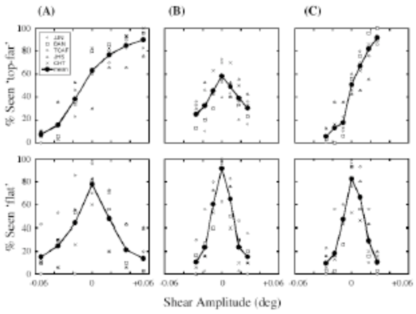
Stationarity prevails rigdity



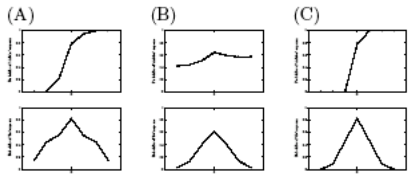


M. Wexler, I. Lamouret, and J. Droulez. The stationarity hypothesis: an allocentric criterion in visual perception. *Vision Research*, 41:3023–3037, 2001.

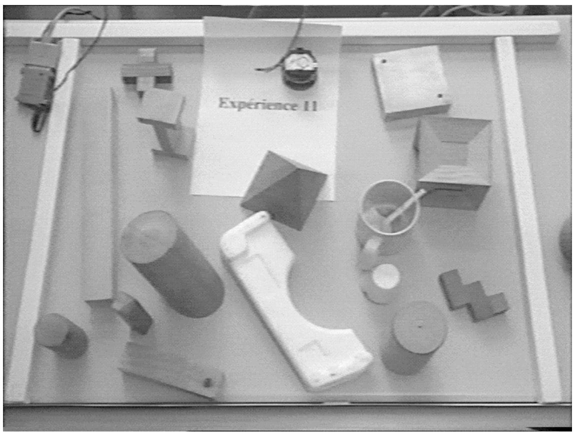
Pursuit



J.J. Naji and T.C.A. Freeman. Perceiving depth order during pursuit eye movement. *Vision Research*, 44:3025–34, 2004.



Object recognition and discovering of novelty



Object Recognition (Model)

Program

Description

Specification

- Variables

  `Nlt, Nlr, Per, Llsl, O`

- Decomposition (Conditional Independance Hypothesis)

$$\mathrm{P}(O \wedge Nrt \wedge Nlt \wedge Per \wedge Llsl \mid \delta \wedge \pi)$$
$$= \mathrm{P}(O \mid \delta \wedge \pi) \times \mathrm{P}(Nrt \mid O \wedge \delta \wedge \pi) \times \mathrm{P}(Nlt \mid O \wedge \delta \wedge \pi)$$
$$\times \mathrm{P}(Per \mid O \wedge \delta \wedge \pi) \times \mathrm{P}(Llsl \mid O \wedge \delta \wedge \pi)$$

- Parametrical Forms

$$\mathrm{P}(O \mid \delta \wedge C) = \text{Uniform}$$

$\mathrm{P}(Nlt \mid O \wedge \delta \wedge \pi) = \text{Laplace}$	$\mathrm{P}([Nlt = 0] \mid O \wedge \delta \wedge \pi) = \text{Uniform}$
$\mathrm{P}(Nrt \mid O \wedge \delta \wedge \pi) = \text{Laplace}$	$\mathrm{P}([Nrt = 0] \mid O \wedge \delta \wedge \pi) = \text{Uniform}$
$\mathrm{P}(Per \mid O \wedge \delta \wedge \pi) = \text{Gaussian}$	$\mathrm{P}([Per = 0] \mid O \wedge \delta \wedge \pi) = \text{Uniform}$
$\mathrm{P}(Llsl \mid O \wedge \delta \wedge \pi) = \text{Gaussian}$	$\mathrm{P}([Llsl = 0] \mid O \wedge \delta \wedge \pi) = \text{Uniform}$

Identification

- No free parameters

Question

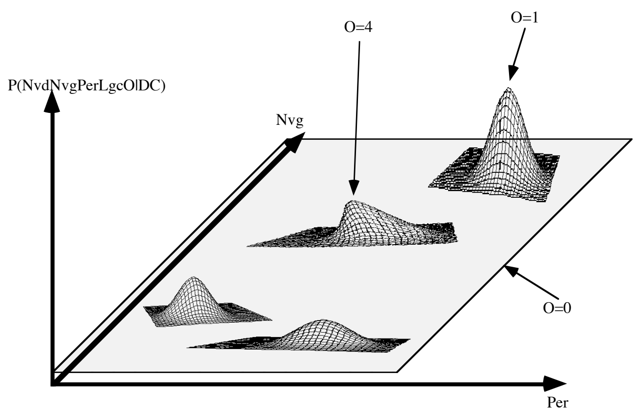
Utilization

$$\mathrm{P}(O \mid nlt \wedge nrt \wedge per \wedge llsl \wedge \delta \otimes \pi)$$

CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

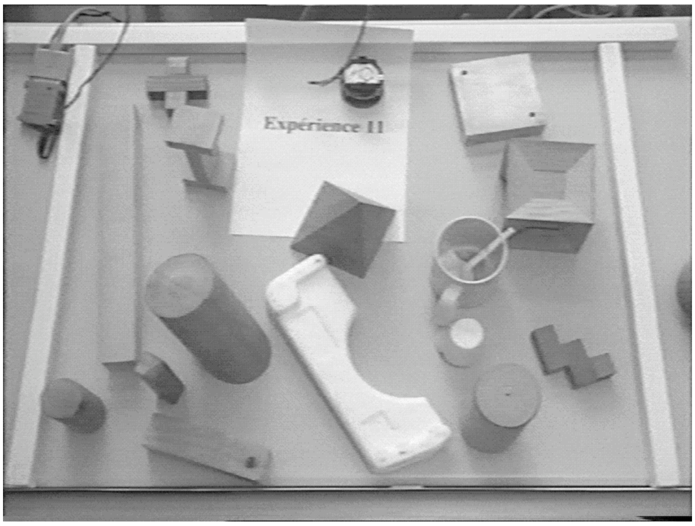
Object Recognition (Question)

$$\mathrm{P}(O \mid nlt \wedge nrt \wedge per \wedge llsl \wedge \delta \otimes \pi)$$



CNRS CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Object Recognition (Result)

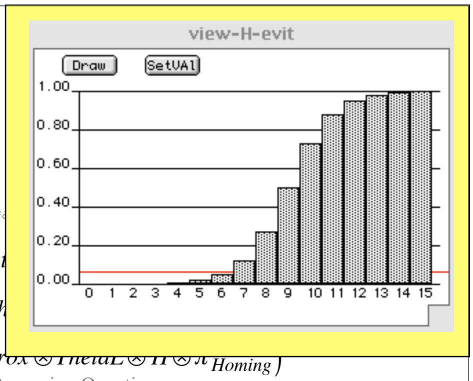


CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Behaviour Combination

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Homing Behavior (1)



Program

Description

Specification

- Variables

  Dir, Prox, ThethaL, …

  H : {a, p}   [H=…

- Decomposition

  $P(Dir \otimes Prox \otimes The\ldots$

  $= P(Dir \otimes Prox \otimes Th\ldots$

  $\times P(Vrot \mid Dir \otimes Prox \otimes ThetaL \otimes H \otimes \pi_{Homing})$

- Parametrical Forms and Recursive Questions

  $P(Dir \otimes Prox \otimes ThetaL \mid \pi_{Homing}) = Uniform$

  $P(H \mid Prox \otimes \pi_{Homing}) = f(Prox)$

  $P(Vrot \mid Dir \otimes Prox \otimes ThetaL \otimes [H = a] \otimes \pi_{Homing})$

  $= P(Vrot \mid Dir \otimes Prox \otimes \pi_{Avoidance})$

  $P(Vrot \mid Dir \otimes Prox \otimes ThetaL \otimes [H = p] \otimes \pi_{Homing})$

  $= P(Vrot \mid ThetaL \otimes \pi_{Phototaxy})$

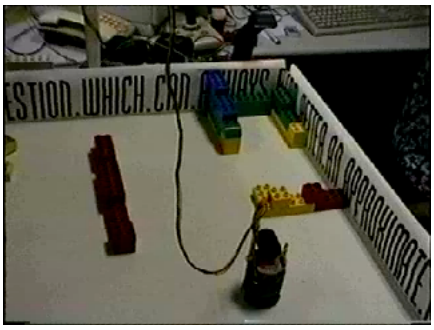
Question

Utilization

$P(Vrot \mid dir_t \otimes prox_t \otimes lum_t \otimes \pi_{Homing})$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Homing Behavior (2)



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Nightwatchman Task (2)

Question : $P(M \mid s \otimes \pi_{Watchman})$

$$P\left(\texttt{Vrot Vtrans} \;\middle|\; \begin{array}{l}\texttt{px0..px7 lm0..lm7 veille feu obj?}\\ \texttt{eng tach\_t-1 td\_t-1 tempo tour}\\ \texttt{dir prox dirG proxG vtrans\_c}\\ \texttt{dnv mnv mld per}\end{array} \; \pi_{Watchman}\right).$$

Solution :

$P(\texttt{Vrot Vtrans} \mid \texttt{px0 px1 ... lm7 veille feu ... per})$

$\propto \sum_{\texttt{ThetaL DistL Td Tach H Base}} P(\ldots)$

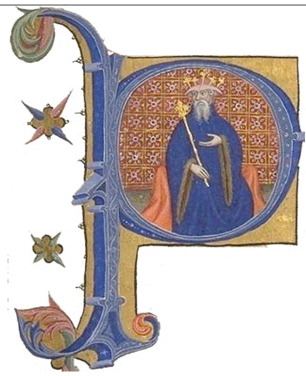
Nightwatchman Task (3)

Answer !

$$P\left(\texttt{Vrot Vtrans} \;\middle|\; \begin{array}{l}\texttt{px0..px7 lm0..lm7 veille feu obj?}\\ \texttt{eng tach\_t-1 td\_t-1 tempo tour}\\ \texttt{dir prox dirG proxG vtrans\_c}\\ \texttt{dnv mnv mld per}\end{array} \; \pi_{Watchman}\right)$$

$$= \frac{1}{Z} \sum_{\substack{\texttt{Td}\\ \texttt{ThetaL}\\ \texttt{H}}} \left(\begin{array}{l} \sum_{\texttt{Tach}} \left(\begin{array}{l} P\left(\texttt{Td} \;\middle|\; \begin{array}{l}\texttt{Tach}\\ \texttt{td\_t-1 tempo tour}\end{array} \; \pi_{Moove}\right) \\ \sum_{\texttt{Base}} \left(\begin{array}{l} P\left(\texttt{Tach} \;\middle|\; \begin{array}{l}\texttt{Base}\\ \texttt{veille feu obj?}\\ \texttt{eng tach\_t-1}\end{array} \; \pi_{Task}\right) \\ P\left(\texttt{Base} \;\middle|\; \begin{array}{l}\texttt{px0...px7}\\ \texttt{lm0...lm7}\end{array} \; \pi_{Base}\right) \end{array} \right) \end{array} \right) \\ \sum_{\texttt{DistL}} P(\texttt{ThetaL DistL} \mid \texttt{lm0..lm7} \; \pi_{Fusion}) \\ P(\texttt{H} \mid \texttt{prox} \; \pi_{Homing}) \\ P\left(\texttt{Vrot Vtrans} \;\middle|\; \begin{array}{l}\texttt{H Td ThetaL}\\ \texttt{dir prox dirG proxG vtrans\_c}\end{array} \; \pi_{Watchman}\right) \end{array} \right).$$

Nightwatchman Task: Result



Prediction short term memory and time

Recursive Bayesian Filter (Model)

$$
\mathbf{Pr}\begin{cases}
\mathbf{Ds}\begin{cases}
\mathbf{Sp}\begin{cases}
\mathbf{Va} \\
S^0, \ldots\ldots, S^T, O^0, \ldots\ldots, O^T \\
\mathbf{Dc} \\
\mathbf{P}(S^0 \wedge \ldots\ldots \wedge S^T \wedge O^0 \wedge \ldots\ldots \wedge O^T) = \mathbf{P}(S^0) \times \mathbf{P}(O^0 \mid S^0) \times \prod_{i=1}^{T} [\mathbf{P}(S^i \mid S^{i-1}) \times \mathbf{P}(O^i \mid S^i)] \\
\mathbf{Fo} \\
\mathbf{P}(S^0) \\
\mathbf{P}(S^i \mid S^{i-1}) \\
\mathbf{P}(O^i \mid S^i)
\end{cases} \\
\mathbf{Id}
\end{cases} \\
\mathbf{Qu} \\
\mathbf{P}(S^{t+k} \mid O^0 \wedge \ldots\ldots \wedge O^t) \qquad (k=0) \equiv Filtering \qquad (k>0) \equiv Prediction \qquad (k<0) \equiv Smoothing
\end{cases}
$$

Bayesian Filter (Recursive Calculation)

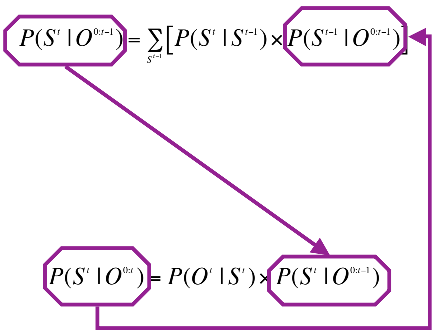
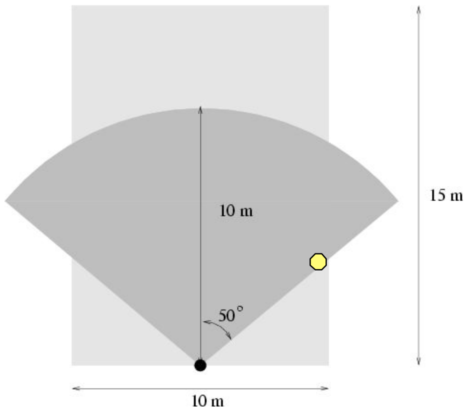
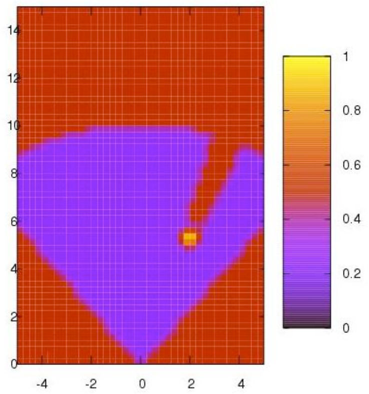


Illustration (1 sensor - 1 object)





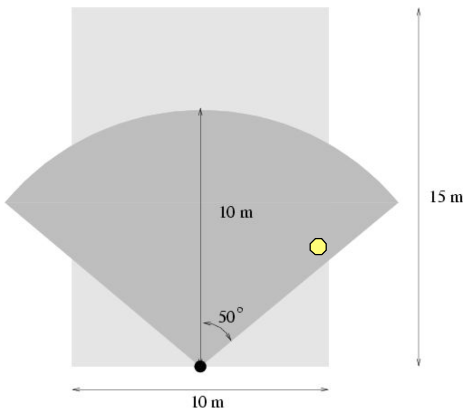
$z = (\ 5, 2, 0, 0)$

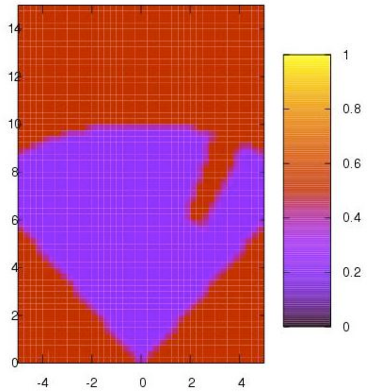
$P(\ [E_c=1]\ |\ z\ c)$

$c = [x, y, 0, 0]$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Illustration (1 sensor - 1 object)





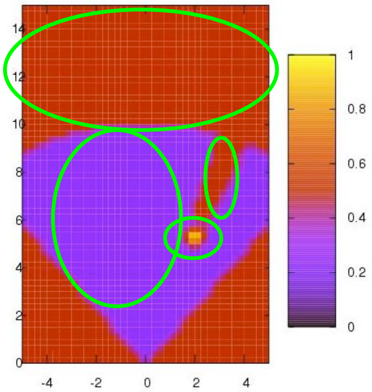
$z = (\ 5, 2, 0, 0)$

$P(\ [E_c=1]\ |\ z\ c)$

$c = [x, y, 0.8, -1.0]$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Illustration (1 sensor - 1 object)

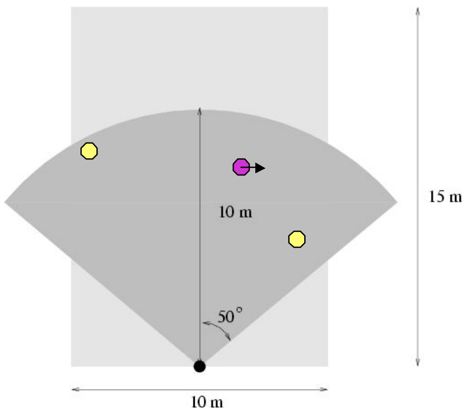


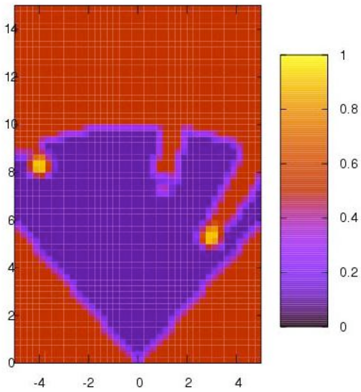
- espace occupé
- espace inoccupé
- espace inobservable
- espace occulté

P($[E_c=1]$ / z c)
c = [x, y, 0, 0]

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Illustration (1 sensor - 3 Objects)



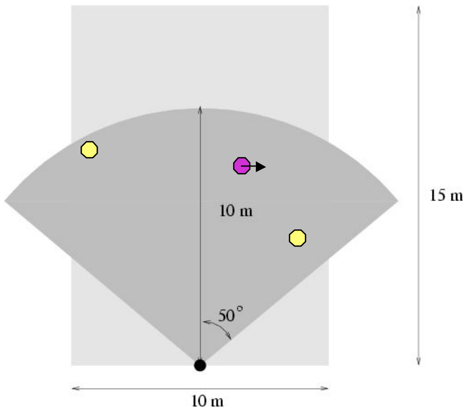


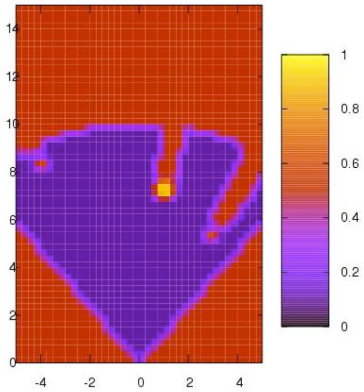
z_1 = (8.3, -4, 0, 0)
z_2 = (7.3, 1.9, 0, 0.8)
z_3 = (5, 3, 0, 0)

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

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Illustration (1 sensor - 3 Objects)



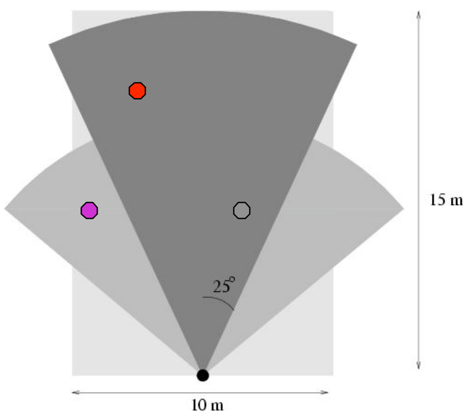


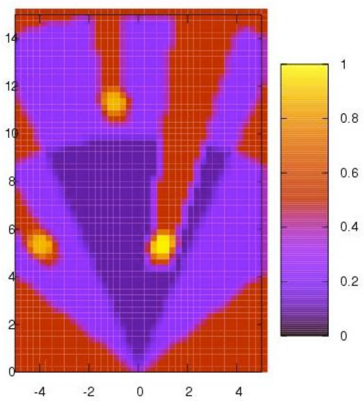
z_1 = (8.3, -4, 0, 0)
z_2 = (7.3, 1.9, 0, 0.8)
z_3 = (5, 3, 0, 0)

$P([E_c=1] \mid z_1\ z_2\ z_3\ c)$
$c = [x, y, 0, 0.8]$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Illustration (2 sensors - 3 objects)



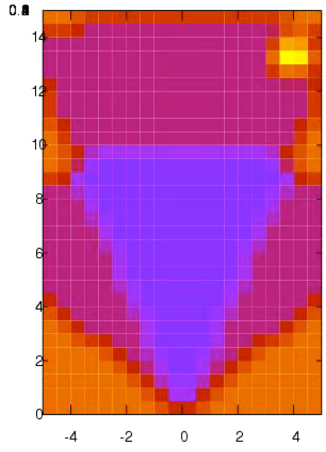


$z_{1,1}$ = (5.5, -4, 0, 0) $z_{1,2}$ = (5.5, 1, 0, 0)
$z_{2,1}$ = (11, -1, 0, 0) $z_{2,2}$ = (5.4, 1.1, 0,0)

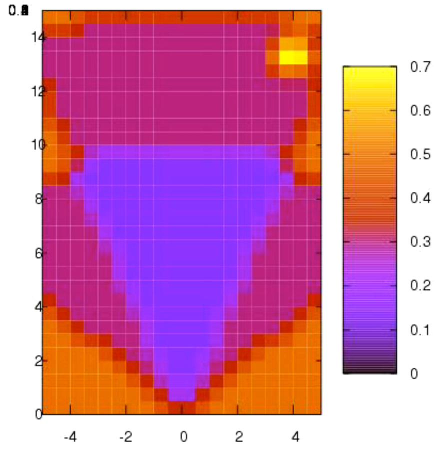
$P([E_c=1] \mid z_{1,1}\ z_{1,2}\ z_{2,1}\ z_{2,2}\ c)$
$c = [x, y, 0, 0]$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Static vs Dynamic

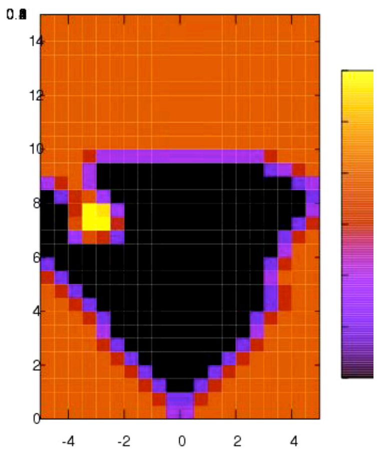


Static Estimation

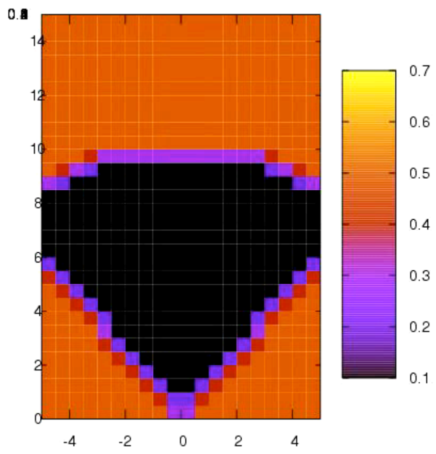


Dynamic Estimation

Static vs Dynamic

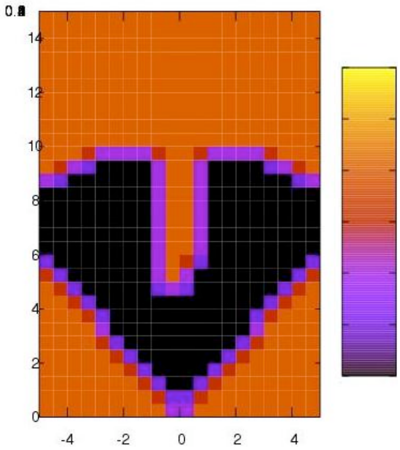


Static Estimation

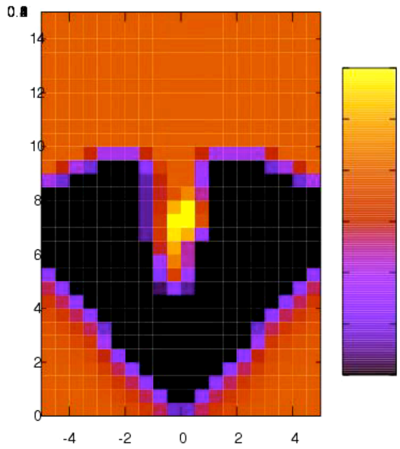


Dynamic Estimation

Static vs Dynamic



Static Estimation



Dynamic Estimation

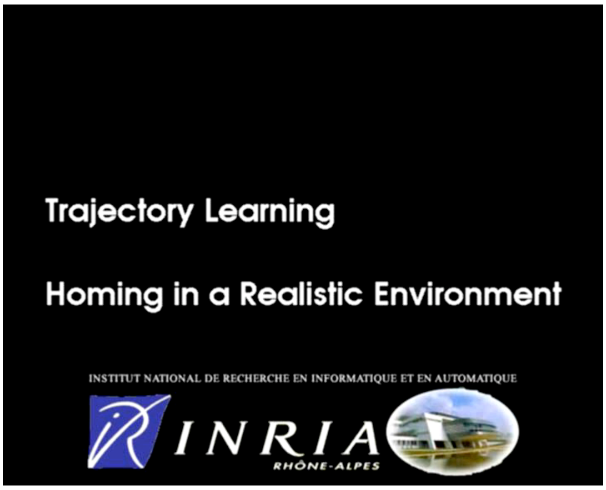
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Application

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

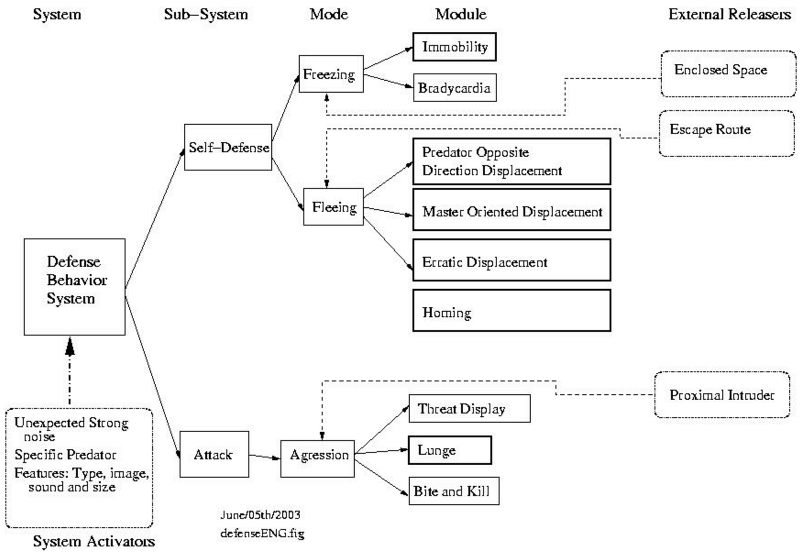
Application

Application



Action selection and attention focalization

Ethologic behaviours

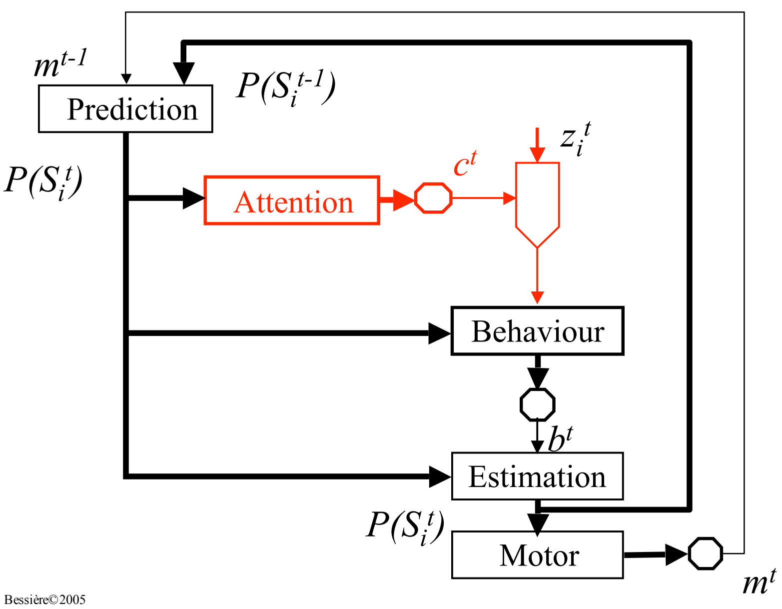


Decomposition

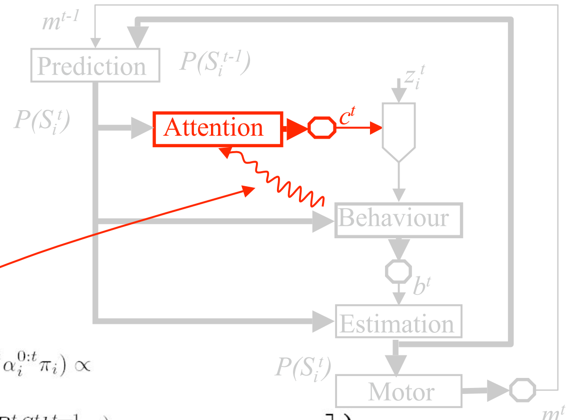
$$P(M^{0:t} S_i^{0:t} Z_i^{0:t} B^{0:t} C^{0:t} \lambda_i^{0:t} \beta_i^{0:t} \alpha_i^{0:t} | \pi_i) =$$

$$= \left| \begin{array}{l} \prod_{j=1}^{t} \left[\begin{array}{ll} P(S_i^j | S_i^{j-1} M^{j-1} \pi_i) \times & \textbf{Dynamic} \\ P(Z_i^j | S_i^j C^j \pi_i) \times & \textbf{Observation} \\ P(B^j | \pi_i) \times P(\beta_i^j | B^j S_i^j B^{j-1} \pi_i) \times & \textbf{Behaviour Selection} \\ P(C^j | \pi_i) \times P(\alpha_i^j | C^j S_i^j B^j \pi_i) \times & \textbf{Focus of Attention} \\ P(M^j | \pi_i) \times P(\lambda_i^j | M^j B^j S_i^j M^{j-1} \pi_i) & \textbf{Motor} \end{array} \right] \\ P(M^0 | \pi_i) \times P(S_i^0 | \pi_i) \times P(C^0 | \pi_i) \times \\ P(B^0 | \pi_i) \times P(Z_i^0 | S_i^0 C^0 \pi_i) \times \\ P(\lambda_i^0 | M^0 S^0 B^0 \pi_i) \times P(\beta_i^0 | B^0 S^0 \pi_i) \times P(\alpha_i^0 | C^0 B^0 S^0 \pi_i) \end{array} \right.$$

Decision process



Selective Perception Question
Attention



$$P(C^t | z_i^{0:t-1} c^{0:t-1} b^{0:t-1} m^{0:t-1} \lambda_i^{0:t-1} \beta_i^{0:t} \alpha_i^{0:t} \pi_i) \propto$$

$$\sum_{B^t} \left[\begin{array}{l} P(B^t|\pi_i) \times P(\beta_i^t | B^t S_i^t b^{t-1} \pi_i) \times \\ P(C^t|\pi_i) \times P(\alpha_i^t | C^t S_i^t B^t \pi_i) \times \\ P(S_i^t | z_i^{0:t-1} c^{0:t-1} b^{0:t-1} m^{0:t-1} \lambda_i^{0:t-1} \beta_i^{0:t-1} \alpha_i^{0:t-1} \pi_i) \end{array} \right]$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Representation of space

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Bayesian maps

Programme
- Description
  - Question

Spécification
- Variables :
  - **P** : variable de perception
  - $\mathbf{L_t}$: lieu au temps t
  - $\mathbf{L_{t'}}$: lieu au temps t' (t' > t)
  - **A** : variable d'action
- Décomposition : quelconque
- Formes paramétriques : quelconque

Identification
- Quelconque

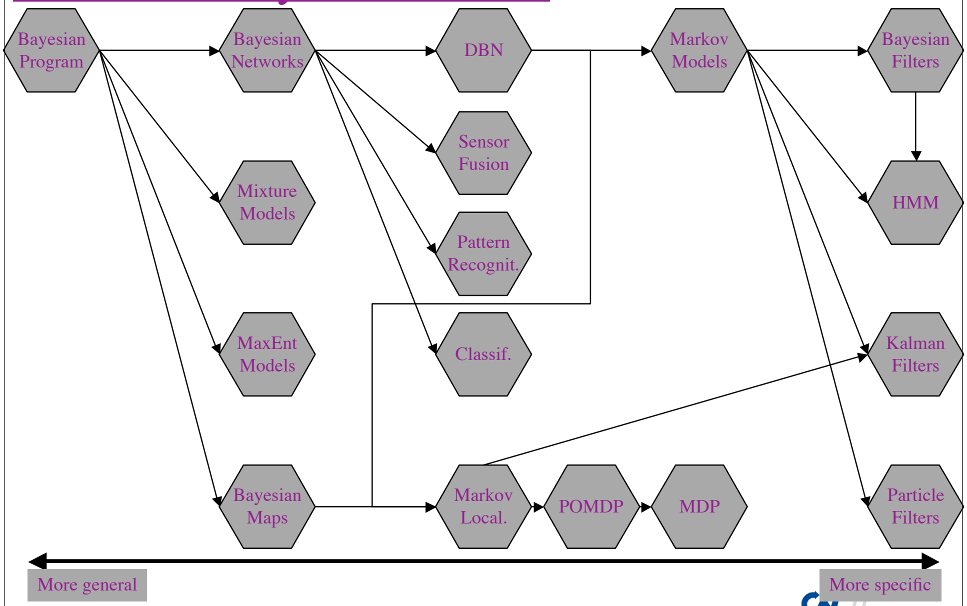
Utilisation
- Localisation $P(L_t \mid P)$
- Prédiction $P(L_{t'} \mid A\ L_t)$
- Contrôle $P(A \mid L_t\ L_{t'})$
- Définition de comportements : $P(A_i \mid \ldots)$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

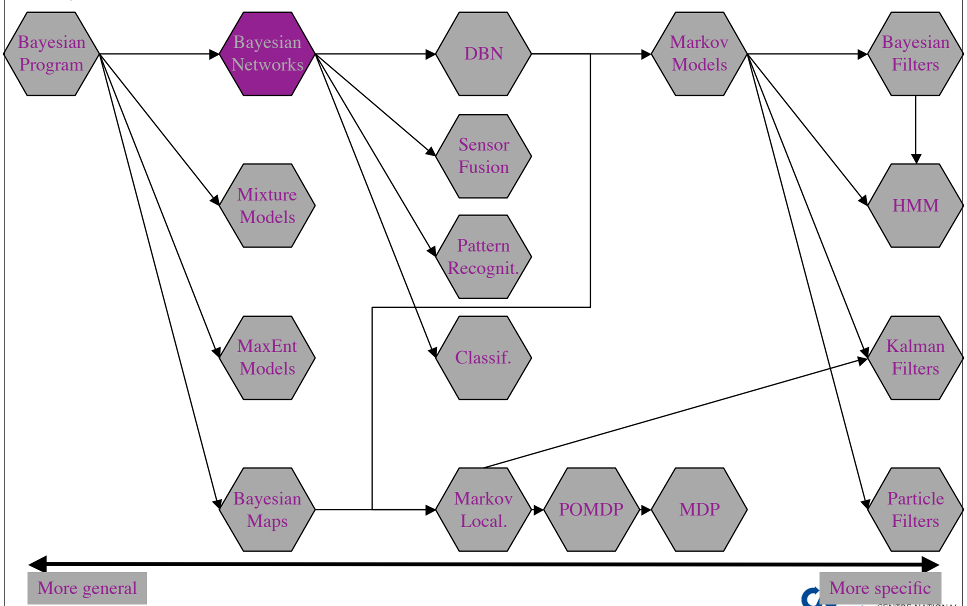
Review of Bayesian models

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Review of Bayesian models



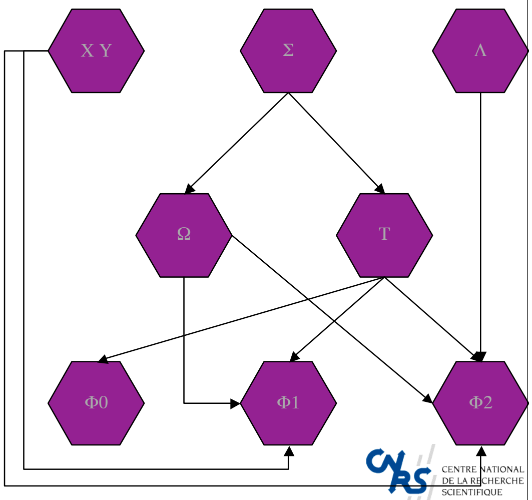
Bayesian Networks



Bayesian Network

$$
\begin{aligned}
& P(X\ Y\ \vec{\Sigma}\ \Lambda\ \vec{\Omega}\ \vec{T}\ \vec{\Phi}) \\
= \; & P(X\ Y)P(\vec{\Sigma})P(\Lambda) \\
\times \; & P(\vec{\Omega}\ \vec{T} \mid \vec{\Sigma}) \\
\times \; & P(\Phi^0 \mid T_x\ T_y) \\
\times \; & P(\Phi^1 \mid X\ Y\ \vec{\Omega}\ \vec{T}) \\
\times \; & P(\Phi^2 \mid X\ Y\ \vec{\Omega}\ \vec{T}\ \Lambda)
\end{aligned}
$$

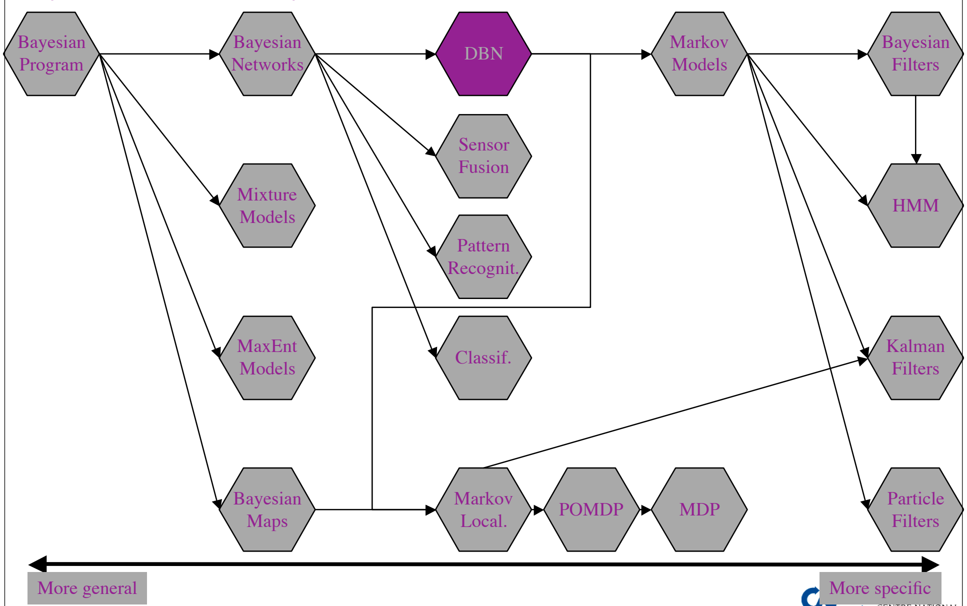
- Pr
  - Ds
    - Sp
      - Va
        - $X_1, \ldots, X_N$
      - Dc
        - $P(X_1 \wedge \ldots \wedge X_N) = \prod_{i=1}^{N} P(X_i \mid R_i)$
      - Fo
        - *Any*
    - Id
  - Qu
    - $P(X_i \mid Known)$



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Bessière©2005

Dynamical Bayesian Networks



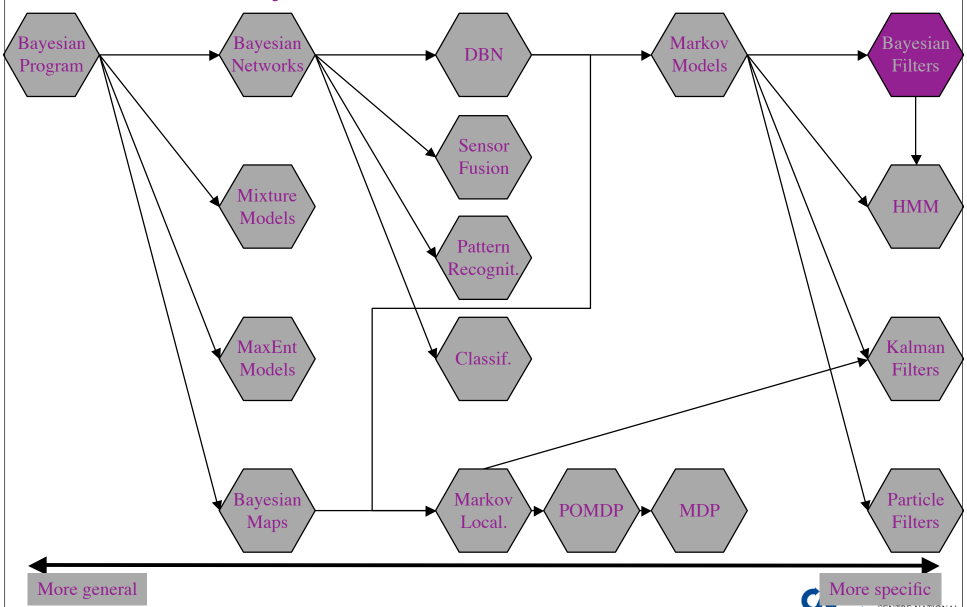
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Bessière©2005

Dynamical Bayesian Network

$$
\mathbf{Pr}\begin{cases}
\mathbf{Ds}\begin{cases}
\mathbf{Sp}\begin{cases}
\mathbf{Va} \\
X_1^0, \ldots\ldots, X_N^0, \ldots\ldots, X_1^T, \ldots\ldots, X_N^T \\
\mathbf{Dc} \\
\mathbf{P}(X_1^0 \wedge \ldots\ldots \wedge X_N^T) = \mathbf{P}(X_1^0 \wedge \ldots\ldots \wedge X_N^0) \times \prod_{t=0}^{T}\prod_{i=1}^{N} \mathbf{P}(X_i^t \mid R_i^t) \\
\mathbf{Fo} \\
Any
\end{cases} \\
\mathbf{Id}
\end{cases} \\
\mathbf{Qu} \\
\mathbf{P}(X_i^T \mid known)
\end{cases}
$$

Recursive Bayesian Filters

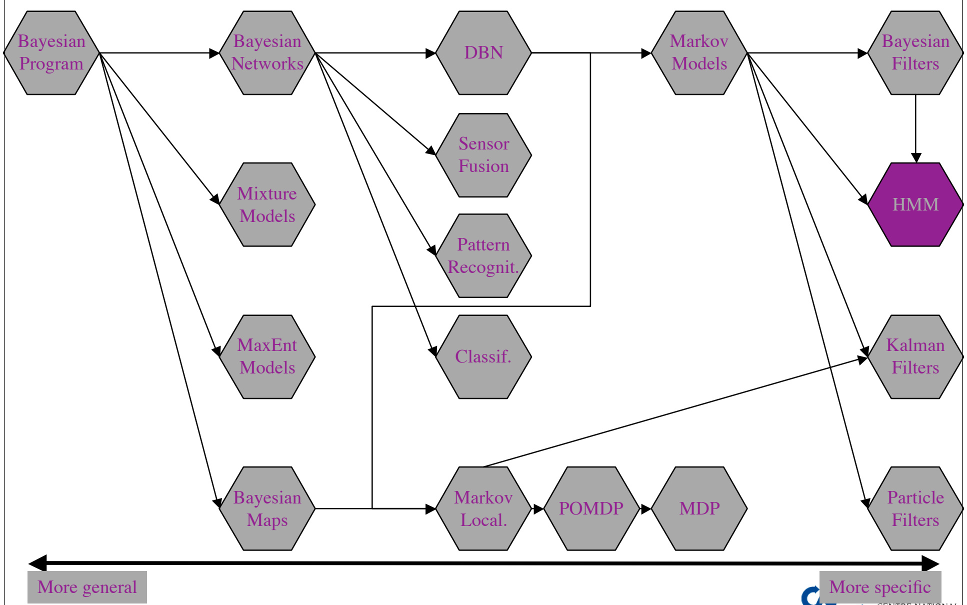


Recursive Bayesian Filters

$$
\mathbf{Pr}\begin{cases}
\mathbf{Ds}\begin{cases}
\mathbf{Sp}\begin{cases}
\mathbf{Va} \\
S^0, \ldots\ldots, S^T, O^0, \ldots\ldots, O^T \\
\mathbf{Dc} \\
\mathbf{P}(S^0 \wedge \ldots\ldots \wedge S^T \wedge O^0 \wedge \ldots\ldots \wedge O^T) = \mathbf{P}(S^0) \times \mathbf{P}(O^0 \mid S^0) \times \prod_{i=1}^{T} [\mathbf{P}(S^i \mid S^{i-1}) \times \mathbf{P}(O^i \mid S^i)] \\
\mathbf{Fo} \\
\mathbf{P}(S^0) \\
\mathbf{P}(S^i \mid S^{i-1}) \\
\mathbf{P}(O^i \mid S^i)
\end{cases} \\
\mathbf{Id}
\end{cases} \\
\mathbf{Qu} \\
\mathbf{P}(S^{t+k} \mid O^0 \wedge \ldots\ldots \wedge O^t) \qquad (k = 0) \equiv Filtering \qquad (k > 0) \equiv Prediction \qquad (k < 0) \equiv Smoothing
\end{cases}
$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Hidden Markov Models



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

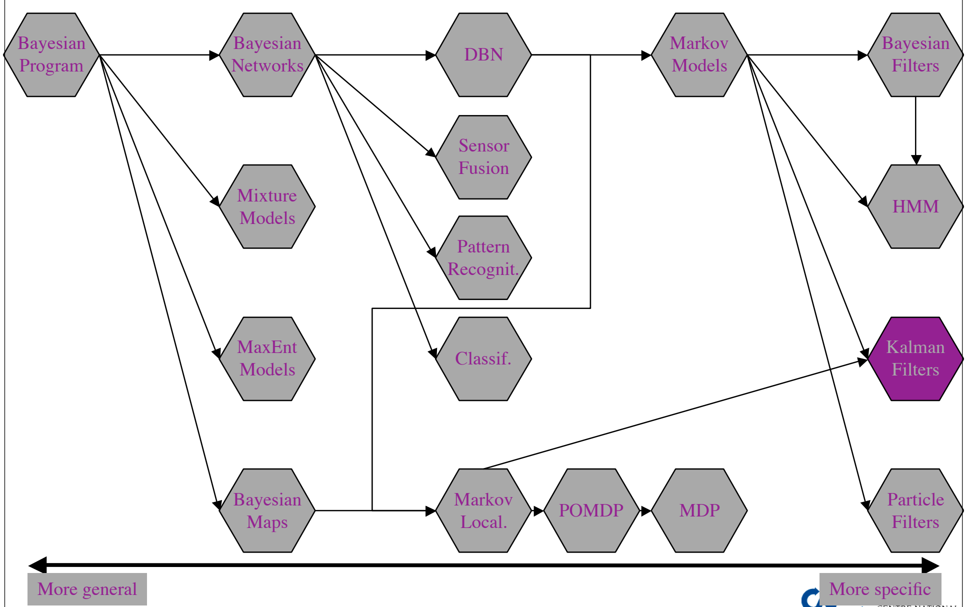
Hidden Markov Models

$$
\text{Pr}\begin{cases}
\text{Ds}\begin{cases}
\text{Sp}\begin{cases}
\text{Va} \\
S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t \\
\text{Dc} \\
\mathbf{P}(S^0 \wedge \ldots\ldots \wedge S^t \wedge O^0 \wedge \ldots\ldots \wedge O^t) = \mathbf{P}(S^0) \times \mathbf{P}(O^0 \mid S^0) \times \prod_{i=1}^{t} [\mathbf{P}(S^i \mid S^{i-1}) \times \mathbf{P}(O^i \mid S^i)] \\
\text{Fo} \\
\mathbf{P}(S^0) \equiv Matrix \\
\mathbf{P}(S^i \mid S^{i-1}) \equiv Matrix \\
\mathbf{P}(O^i \mid S^i) \equiv Matrix
\end{cases} \\
\text{Id}
\end{cases} \\
\text{Qu} \\
\text{MAX}_{S^1 \wedge S^2 \wedge \ldots\ldots \wedge S^{t-1}}[\mathbf{P}(S^1 \wedge S^2 \wedge \ldots\ldots \wedge S^{t-1} \mid S^t \wedge O^0 \wedge \ldots\ldots \wedge O^t)]
\end{cases}
$$

Bessière© 2005

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Kalman Filters



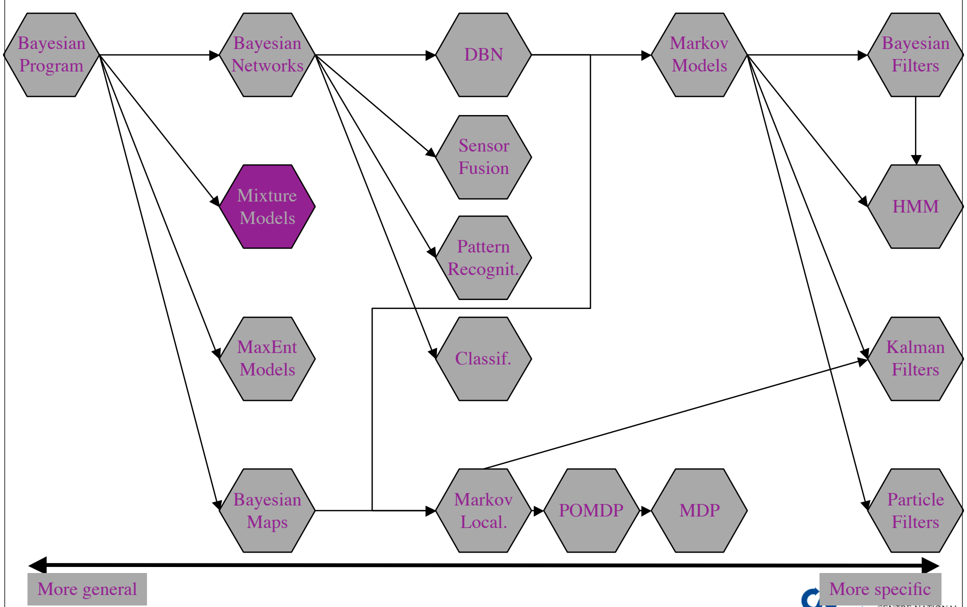
CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Bessière© 2005

Kalman Filters

$$
\text{Pr}\left\{\begin{array}{l}
\text{Ds}\left\{\begin{array}{l}
\text{Sp}\left\{\begin{array}{l}
\textbf{Va} \\
S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t \\
\textbf{Dc} \\
\mathbf{P}(S^0 \wedge \ldots\ldots \wedge S^t \wedge O^0 \wedge \ldots\ldots \wedge O^t) = \mathbf{P}(S^0) \times \mathbf{P}(O^0 \mid S^0) \times \prod_{i=1}^{t} [\mathbf{P}(S^i \mid S^{i-1}) \times \mathbf{P}(O^i \mid S^i)] \\
\textbf{Fo} \\
\mathbf{P}(S^0) \equiv \mathbf{G}(S^0, \mu, \sigma) \\
\mathbf{P}(S^i \mid S^{i-1}) \equiv \mathbf{G}(S^i, A \bullet S^{i-1}, Q) \\
\mathbf{P}(O^i \mid S^i) \equiv \mathbf{G}(O^i, H \bullet S^i, R)
\end{array}\right. \\
\text{Id}
\end{array}\right. \\
\textbf{Qu} \\
\mathbf{P}(S^t \mid O^0 \wedge \ldots\ldots \wedge O^t)
\end{array}\right.
$$

Mixture Models

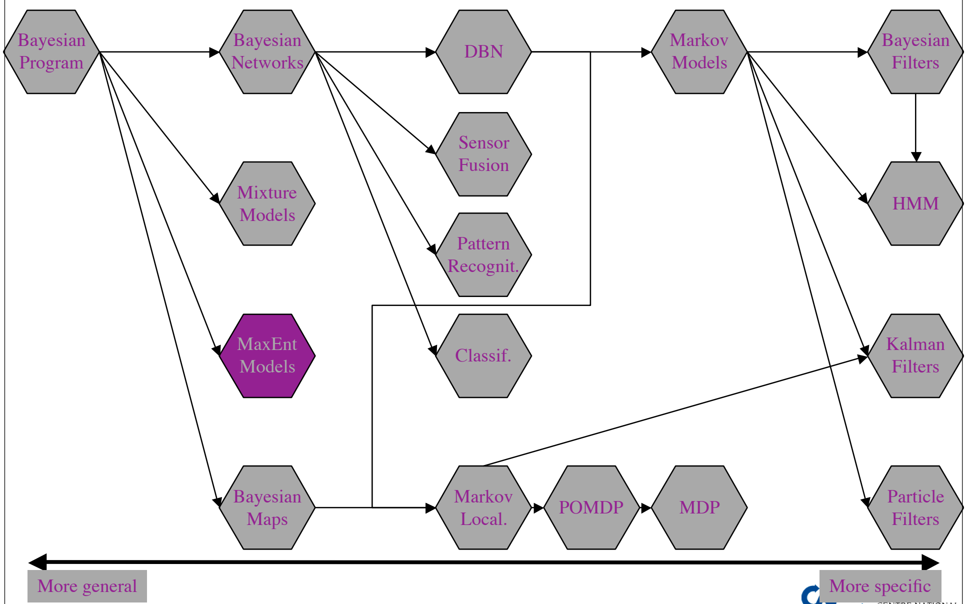


Mixture Models

$$
\text{Pr}\begin{cases}
\text{Ds}\begin{cases}
\text{Sp}\begin{cases}
\text{Va} \\
X_1, \ldots\ldots, X_N \\
\text{Dc} \\
\mathbf{P}(X_1 \wedge \ldots\ldots \wedge X_N) = \sum_{i=1}^{M} [\alpha_i \times \mathbf{P}(X_1 \wedge \ldots\ldots \wedge X_N \mid \pi_i)] \\
\text{Fo} \\
\mathbf{P}(X_1 \wedge \ldots\ldots \wedge X_N \mid \pi_i) \equiv \mathbf{G}(X_1 \wedge \ldots\ldots \wedge X_N, \mu_i, \sigma_i)
\end{cases} \\
\text{Id}
\end{cases} \\
\text{Qu}
\end{cases}
$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Maximum Entropy Models



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Maximum Entropy Models

$$\mathbf{Pr}\begin{cases} \mathbf{Ds}\begin{cases} \mathbf{Sp}\begin{cases} \mathbf{Va} \\ X_1, \ldots\ldots, X_N \\ \mathbf{Dc} \\ \mathbf{P}(X_1 \wedge \ldots\ldots \wedge X_N) \\ = \prod_{i=0}^{M}[e^{-[\lambda_i \times f_i(X_1 \wedge \ldots\ldots \wedge X_M)]}] = e^{-\sum_{i=0}^{M}[\lambda_i \times f_i(X_1 \wedge \ldots\ldots \wedge X_M)]} \\ \mathbf{Fo} \\ f_0 = 1 \qquad f_1, \ldots\ldots, f_M \text{ M observable functions} \end{cases} \\ \mathbf{Id} \\ \{\lambda_1, \ldots\ldots, \lambda_M\} \end{cases} \\ \mathbf{Qu} \\ \mathbf{P}(Searched \mid Known) \end{cases}$$

Snapshot on the inference algorithms

Question

$$P(X^1 \wedge X^2 \wedge \ldots \wedge X^n) = P(L^1) \times P(L^2 \mid R^2) \times \ldots \times P(L^k \mid R^k)$$

$$P(Search \mid Known)$$

$$= \sum_{Free} P(Search \wedge Free \mid Known)$$

$$= \sum_{Free} \frac{P(Search \wedge Free \wedge Known)}{P(Known)}$$

$$= \frac{1}{Z} \times \sum_{Free} P(Search \wedge Free \wedge Known)$$

$$= \frac{1}{Z} \times \sum_{Free} P(L^1) \times P(L^2 \mid R^2) \times \ldots \times P(L^k \mid R^k)$$

- Marginalisation (integration) in high-dimensional space
- Searching the modes in high-dimensional space

2 modes: exact or approximated

- Symbolic simplification
- Numeric computation

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Symbolic simplification

$$P(Search \mid Known) = \frac{1}{Z} \times \sum_{Free} \left[\prod_{i=0}^{k} P(L^i \mid R^i) \right]$$

Factorization:

$$P(Search \mid Known) = \frac{1}{Z} \times \prod_{j \in J} P(L^j \mid R^j) \times \sum_{Free1} \left[\prod_{k \in K} P(L^k \mid R^k) \right]$$

Sum to 1:
[Bessière et al. 2003]

$$P(Search \mid Known) = \frac{1}{Z} \times \prod_{j \in J} P(L^j \mid R^j) \times \sum_{Free2} \left[\prod_{l \in L} P(L^l \mid R^l) \right]$$

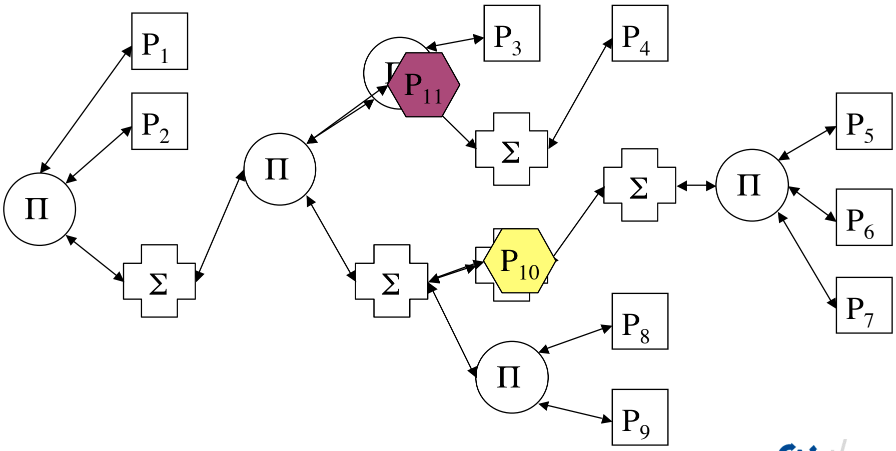
Distributivity:
Generalized distributive law [Aji & McEliece2000]
Restrictions successives [Raoult & Smail2003]

$$P(Search \mid Known) = \frac{1}{Z} \times \prod_{j \in J} P(L^j \mid R^j) \times \sum_{Free3} \left[\prod_{m \in M} P(L^m \mid R^m) \times \sum_{Free4} \left[\prod_{n \in N} P(L^n \mid R^n) \times \ldots \times \sum_{FreeX} \left[\prod_{o \in O} P(L^o \mid R^o) \right] \right] \right]$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Numeric computation (exact)

$$P(Search \mid Known) = \frac{1}{Z} \times \prod_{j \in J} P(L^j \mid R^j) \times \sum_{Free3} \left[\prod_{m \in M} P(L^m \mid R^m) \times \sum_{Free4} \left[\prod_{n \in N} P(L^n \mid R^n) \times \ldots \times \sum_{FreeX} \left[\prod_{o \in O} P(L^o \mid R^o) \right] \right] \right]$$



Arithmetic of very small numbers

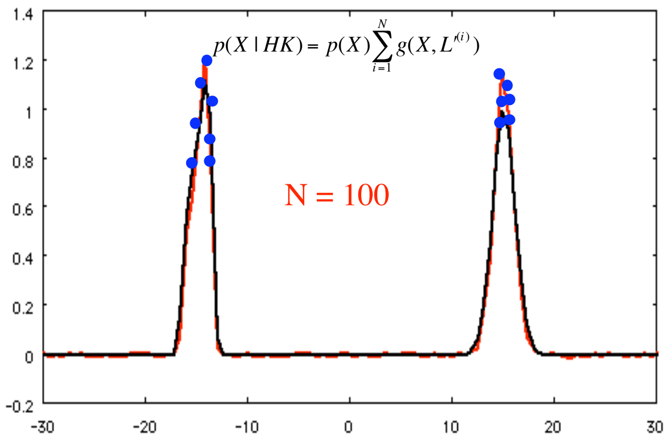
- PlFloat a new type of number specific to encode probabilities
- 64 bits representation with 32 bits mantissa
- Very efficient implementation of the corresponding arithmetic

Numeric computation (approximated)

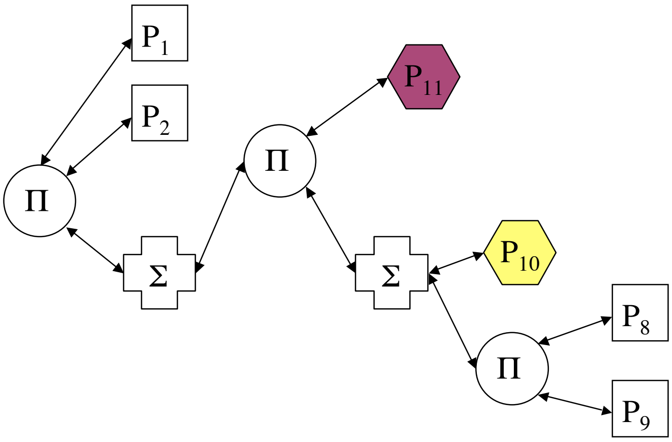
- MCMC for integration

- Genetic Algorithm to search the modes

MCMC + Annealing



Multi-Resolution Binary Tree (MRBT)

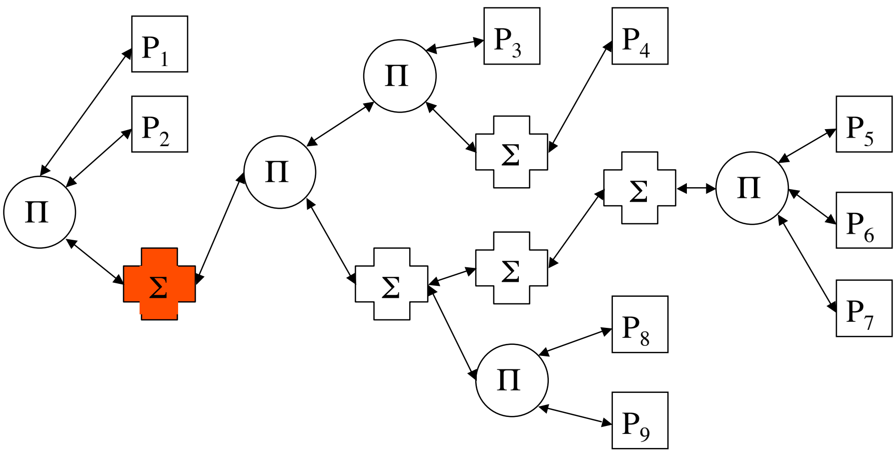


Multi-Resolution Binary Tree (MRBT)

- Very compact representation
- Multi-resolution: more precision in high probability area
- Very efficient draw
- Incremental and anytime representation

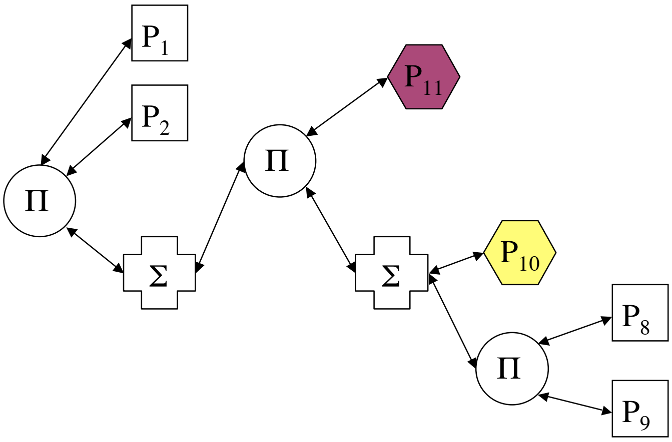
Bayesian Hardware (2)

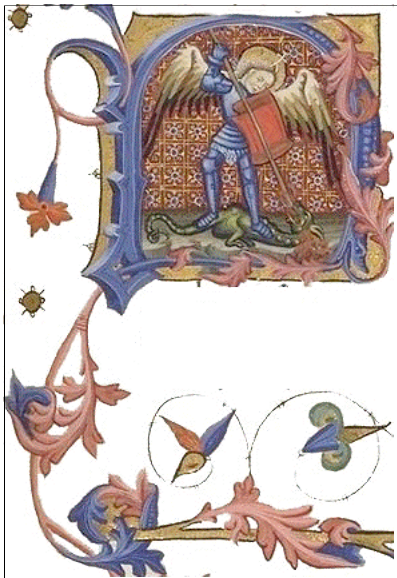
Parallelization (1)



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Parallelization (2)





Data and Preliminary knowledge

How to Deal with Data?

Using Preliminary Knowledge

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

$$P\left(X^1 \otimes X^2 \otimes ... \otimes X^n \mid \delta \otimes \pi\right)$$

$$P(\Delta \otimes \Pi) = P(\Pi) \times P(\Delta \mid \Pi) = P(\Delta) \times P(\Pi \mid \Delta)$$

- Direct problem:

$$P(\Delta \mid \Pi) = \frac{P(\Delta) \times P(\Pi \mid \Delta)}{P(\Pi)}$$

- Inverse problem:

$$P(\Pi \mid \Delta) = \frac{P(\Pi) \times P(\Delta \mid \Pi)}{P(\Delta)}$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Bernoulli's Urn (1)

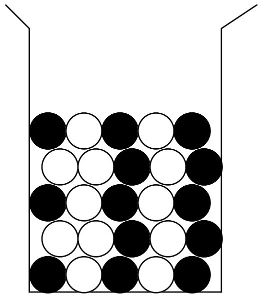
- Variables

  $Draw$
- Decomposition

  $P(Draw \mid \pi)$
- Parametrical Form

  $$P([Draw = black] \mid \pi) = \frac{b}{w+b}$$

  $$P([Draw = white] \mid \pi) = \frac{w}{w+b}$$



Preliminary Knowledge π: "We draw from an urn containing w white balls and b black balls"

Bernoulli's Urn (2)

- Variables:

  $\Delta = Draw_1 \otimes Draw_2 \otimes ... \otimes Draw_m$

  W

  B

  $\Pi = \pi' \otimes W \otimes B$
- Decomposition:

  $$P(\Delta \mid \Pi) = P(Draw_1 \otimes ... \otimes Draw_m \mid \pi' \otimes W \otimes B)$$
- Parametrical Form:

  $$P([Draw_i = white] \mid \pi' \otimes w \otimes b) = \frac{w}{w+b}$$

$$P(\Delta \mid \Pi) = \frac{\binom{\omega}{w} \times \binom{\beta}{b}}{\binom{m}{w+b}}$$

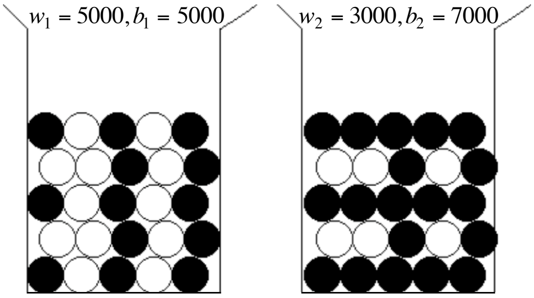
Note:

$$P(Draw_2 \mid [Draw_1 = white] \otimes \pi' \otimes w \otimes b) = \frac{w-1}{w+b-1}$$

$$= P(Draw_1 \mid [Draw_2 = white] \otimes \pi' \otimes w \otimes b)$$

Bernoulli's Urn (3)

$$P(\Pi \mid \Delta) = \frac{P(\Pi) \times P(\Delta \mid \Pi)}{P(\Delta)}$$



$$\frac{P(\pi_1 \mid \delta)}{P(\pi_2 \mid \delta)} = \frac{P(\pi_1) \times P(\delta \mid \pi_1)}{P(\pi_2) \times P(\delta \mid \pi_2)}$$

$$= \frac{P(\pi_1)}{P(\pi_2)} \times \frac{\binom{\omega}{w_1} \times \binom{\beta}{b_1} \times \binom{m}{w_2 + b_2}}{\binom{\omega}{w_2} \times \binom{\beta}{b_2} \times \binom{m}{w_1 + b_1}}$$

m,ω	$\frac{P(\pi_1 \mid \delta)}{P(\pi_2 \mid \delta)}$
2,1	1.2
4,2	1.4
10,5	2.4
20,10	5.7
50,25	80
100,50	6728

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Parameters Identification

Variables:

$\Delta \, , \, \Pi = \Pi' \otimes \Psi$

Decomposition:

$$P(\Delta \otimes \Pi) = P(\Delta \otimes \Pi' \otimes \Psi) = P(\Pi') \times P(\Psi \mid \Pi') \times P(\Delta \mid \Psi \otimes \Pi')$$

$$P(\Psi \mid \delta \otimes \pi') = \frac{P(\delta \otimes \pi' \otimes \Psi)}{P(\delta \otimes \pi')}$$

$$= \frac{1}{\omega} \times P(\Psi \mid \pi') \times P(\delta \mid \Psi \otimes \pi')$$

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Model Selection

Variables:

$\Delta\ ,\ \Pi = \Pi' \otimes \Psi$

Decomposition:

$$P(\Delta \otimes \Pi) = P(\Delta \otimes \Pi' \otimes \Psi) = P(\Pi') \times P(\Psi \mid \Pi') \times P(\Delta \mid \Psi \otimes \Pi')$$

$$P(\Pi' \mid \delta) = \sum_{\Psi} \frac{P(\delta \otimes \Psi \otimes \Pi')}{P(\delta)}$$
$$= \frac{1}{\omega} \times P(\Pi') \times \sum_{\Psi} \left[P(\Psi \mid \Pi') \times P(\delta \mid \Psi \otimes \Pi') \right]$$

Summary

Preliminary Knowledge	Do we need it? Where does it come from? How to specify it?
Variables	Pre-treatments Post-treatments
Decomposition	Model Selection
Parametrical Forms	
Parameters values	Learning

Entropy Principles

Content:

- Entropy Principle Statement
- Frequencies and Laplace succession law
- Observables and Exponential laws
- Wolf's dice

Entropy Principle Statement

20 000 flip of a coins: 9553 heads
Probability distribution of the coin?

$$H(P) = -\sum_{i=1}^{q}\left[P_i \times \log(P_i)\right]$$

Observables and Exponential Laws

Observable:

$$\mathrm{f}_i(V) \rightarrow \Re$$

Constraint levels:

$$\forall j, j \in \{1, ..., m\}, \quad \sum_V \left[\mathrm{P}(V) \times \mathrm{f}_j(V) \right] = F_j$$

Maximum Entropy Distribution:

$$\mathrm{P}^*(V) = \frac{1}{Z(\lambda_1, ..., \lambda_m)} \times e^{-\sum_{j=1}^{m} \left[\lambda_j \times \mathrm{f}_j(V) \right]}$$

Partition Function:

$$Z(\lambda_1, ..., \lambda_m) = \sum_V e^{-\lambda_1 \mathrm{f}_1(V) - ... - \lambda_m \mathrm{f}_m(V)}$$

Constraints differential equation:

$$\frac{\partial}{\partial \lambda_j} \log\left(Z(\lambda_1, ..., \lambda_m) \right) + F_j = 0$$

20 000 Flips

Observable:

$$\mathrm{f}_1(V) = 1, \text{ if V is "Head"; else } \mathrm{f}_1(V) = 0$$

Constraint levels:

$$\sum_V \left[\mathrm{P}(V) \times \mathrm{f}_1(V) \right] = F_1 = \frac{9553}{20000}$$

Maximum Entropy Distribution:

$$\mathrm{P}^*(V) = \frac{1}{Z(\lambda_1)} \times e^{-\lambda_1 \mathrm{f}_1(V)}$$

Partition Function:

$$Z(\lambda_1) = \sum_V e^{-\lambda_1 \mathrm{f}_1(V)} = 1 + e^{-\lambda_1}$$

Constraints differential equation:

$$\frac{\partial}{\partial \lambda_1} \log\left(Z(\lambda_1) \right) + F_1 = 0$$

$$\frac{\partial}{\partial \lambda_1} \log\left(1 + e^{-\lambda_1} \right) + F_1 = 0$$

$$\Leftrightarrow -\frac{e^{-\lambda_1}}{1 + e^{-\lambda_1}} + F_1 = 0$$

$$\Leftrightarrow \left(\frac{1}{F_1} - 1 \right) e^{-\lambda_1} = 1$$

$$\Leftrightarrow \lambda_1 = \log\left(\frac{1 - F_1}{F_1} \right)$$

$$\mathrm{P}^*(V) = \frac{1}{1 + e^{-\log\left(\frac{1-F_1}{F_1} \right)}} \times e^{-\log\left(\frac{1-F_1}{F_1} \right) \mathrm{f}_1(V)}$$

$$\mathrm{P}^*([V = Head]) = F_1$$

132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646134646446444464
611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611113
144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164611645641146416411311166464146161413435116163164161144611646
411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164611645641146416411311166464146
161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164
611645641146416411311166464146161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414
641231464646343436011614641164611645641146416411311166464146161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316
3416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646461346
464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111
161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416
114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131
116646414616141343511616316416114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601
161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131
461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111161616111314413161414646461346464464444646113131641
614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111161616111314413161 41
464646134646446444464611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552
161161411116161611131441316141464646134646446444464611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351
161631641611446116464114113552161161411116161611131441316141464646134646446444464611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411
464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646134646446444464611313164161441643163416414164132415413146164654146412314646
463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646134646446444464611313164161441643163416414164
132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646134646446444464
611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611113
144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164611645641146416411311166464146161413435116163164161144611646
411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164611645641146416411311166464146
161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164
611645641146416411311166464146161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414
641231464646343436011614641164611645641146416411311166464146161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316
3416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646461346
464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111
161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416
114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131
116646414616141343511616316416114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601
161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131
461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111161616111314413161414646461346464464444646113131641
614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111161616111314413161 41
464646134646446444464611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552
161161411116161611131441316141464646134646446444464611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351
161631641611446116464114113552161161411116161611131441316141464646134646446444464611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411
464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646134646446444464611313164161441643163416414164132415413146164654146412314646
463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646134646446444464611313164161441643163416414164
132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646134646446444464
611313164161441643163416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611113
144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164611645641146416411311166464146161413435116163164161144611646
411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164611645641146416411311166464146
161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414641231464646343436011614641164
611645641146416411311166464146161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416413241541314616465414
641231464646343436011614641164611645641146416411311166464146161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316
3416414164132415413146164654146412314646463434360116146411646116456411464164113111664641461614134351161631641611446116464114113552161161411116161611131441316141464646461346
464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416114461164641141135521611614111
161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131116646414616141343511616316416
114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601161464116461164564114641641131
116646414616141343511616316416114461164641141135521611614111161616111314413161414646461346464464444646113131641614416431634164141641324154131461646541464123146464634343601
1614641164611645641146416411311166464146161413435116163164161144611646411411355216116141111616161113144131614146464613464644644446461131316416144164316341641416455511 12666

Frequencies and Laplace Succession Law

Preliminary Knowledge:

1- Each of the 20 000 digit is a number
2- The 20 000 data come from the same phenomenon
3- A single variable *V* has been observed 20 000 times
4- The order of these observations is not relevant

$$P^*([V = i]) = \frac{n_i}{n}$$

5- The variable V may take 6 different values

$$P^*([V = i]) = \frac{n_i + 1}{n + \lfloor V \rfloor}$$

	V=1	V=2	V=3	V=4	V=5	V=6	Total
ni	3246	3449	2897	2841	3635	3932	20000
Uniform	0,16666	0,16666	0,16666	0,16666	0,16666	0,16666	1
Laplace	0,16230	0,17245	0,14486	0,14206	0,18174	0,19659	1

Wolf's dice (1)

H1 Hypothesis: excavations shifted the gravity center

$\mathrm{f}_1(V) = V \qquad F_1 = 3.5983$

$$P^*(V) = \frac{1}{Z} \times e^{0.03372 \times V}$$

	V=1	V=2	V=3	V=4	V=5	V=6	Total
ni	3246	3449	2897	2841	3635	3932	20000
Uniform	0.16666	0.16666	0.16666	0.16666	0.16666	0.16666	1
Laplace	0.16230	0.17245	0.14486	0.14206	0.18174	0.19659	1
H1	0.15294	0.15818	0.16361	0.16922	0.17502	0.18103	1

Wolf's dice (2)

H2 Hypothesis: The dice is oblong along the 1-6 direction and the excavations shifted the gravity center

$\mathrm{f}_1(V) = V \qquad F_1 = 3.5983$

$\mathrm{f}_2(V) = -1 \text{ if } V = 1 \text{ or if } V = 0; \text{ else } \mathrm{f}_2(V) = 0 \qquad F_2 = \frac{1}{20000} \times \sum_{i=1}^{6} n_i \times \mathrm{f}_2(i) = 0.3589$

$$\mathrm{P}^*(V) = \frac{1}{Z} \times e^{0.03234 \times V - 0.1104 \times \mathrm{f}_2(V)}$$

	V=1	V=2	V=3	V=4	V=5	V=6	Total
ni	3246	3449	2897	2841	3635	3932	20000
Uniform	0.16666	0.16666	0.16666	0.16666	0.16666	0.16666	1
Laplace	0.16230	0.17245	0.14486	0.14206	0.18174	0.19659	1
H1	0.15294	0.15818	0.16361	0.16922	0.17502	0.18103	1
H2 (1-6)	0,16497	0,15259	0,15760	0,16278	0,16813	0,19393	1
H3 (2-5)	0,14803	0,16808	0,15843	0,16390	0,18612	0,17543	1
H4 (3-4)	0,16433	0,16963	0,14117	0,14573	0,18656	0,19258	1

Wolf's dice (3)

Inverse Problem:

$$P(h_i \mid \delta) = \frac{P(h_i) \times P(\delta \mid h_i)}{P(\delta)}$$

$$P(\delta \mid h_i) = W \times p_{i1}{}^{3246} \times p_{i2}{}^{3449} \times p_{i3}{}^{2897} \times p_{i4}{}^{2841} \times p_{i5}{}^{3635} \times p_{i6}{}^{3932}$$

	Laplace	h4	h2	h3	h1	Uniform
P(hi)	0.99	0.01	10^{-32}	10^{-35}	10^{-45}	10^{-59}

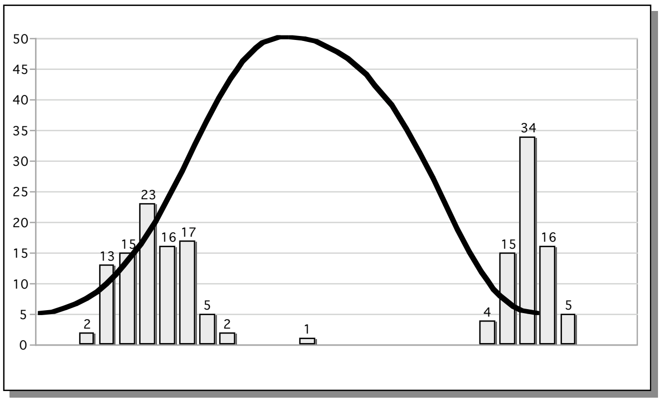
Theoretical Basis

Objective:
Justify the use of the entropy function H

Content:

- What is a good representation?
- Combinatorial justification
- Information theory justification
- Bayesian justification
- Axiomatic justification
- Entropy concentration theorems justifications

What is a Good Representation?



Combinatorial Justification

Statistical Mechanic	Probabilistic Inference
q microscopic states	q propositions
Macroscopic state $v_k = \{n_1, ..., n_q\}$	Distribution $\delta_k = \{p_1, ..., p_q\}$
$\sum_{i=1}^{q} n_i = n$	$\sum_{i=1}^{q} p_i = 1$
$\sum_{i=1}^{q} n_i \times e_i = e$	$\forall j, j \in \{1, ..., m\}; \quad \sum_{i=1}^{q} p_i \times f_j(v_i) = F_j$
$W(v_k) = \frac{n!}{n_1! \times ... \times n_q!}$	$W(\delta_k) = \frac{n!}{(n \times p_1)! \times ... \times (n \times p_q)!}$
$\log(W(v_k)) \approx -n \times \sum_{i=1}^{q} \frac{n_i}{n} \log\left(\frac{n_i}{n}\right)$	$\log(W(\delta_k)) \approx -n \times \sum_{i=1}^{q} p_i \log(p_i)$

v	w
10000,10000	10^{1386}
9553,10447	10^{1384}
0,20000	1

v	w
10000,10000	10^{1386}
9553,10447	10^{1384}
0,20000	1

Shannon's justification

Shannon C. E. (1948) ; "A Mathematical Theory of Communication" ; *Bell Systems Technical Journal* ; 27

Reprinted as Shannon C.E. & Weaver (1949) "The Mathematical Theory of Communication" ; University of Illinois Press, Urbana

Shore's Axiomatic Justification

Shore, J.E. & Johnson, R.W. (1980) ; "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy" ; *IEEE Transactions on Information Theory* ; IT-26 26-37

Entropy Concentration Theorem

Jaynes E.T. (1982) ; “On the rationale of Maximum Entropy Methods” ; Proceedings of the IEEE

Robert Claudine (1990) ; “An Entropy Concentration Theorem: Applications in Artificial Intelligence and Descriptive Statistics” ; *Journal of Applied Probabilities*