Ecole d’été Maths et Cerveau
Jeudi 16 juin 2005

Probabilistic models of 3D shape and motion perception

---o---

Jacques Droulez
Laboratoire de Physiologie de la Perception et de l’Action
CNRS – Collège de France

« *Le hasard n'est que la mesure de notre ignorance* »

Henri Poincaré, La science et l'hypothèse, 1902

$\Rightarrow$ *Probability is the best way to quantify knowledge*

The problem ...

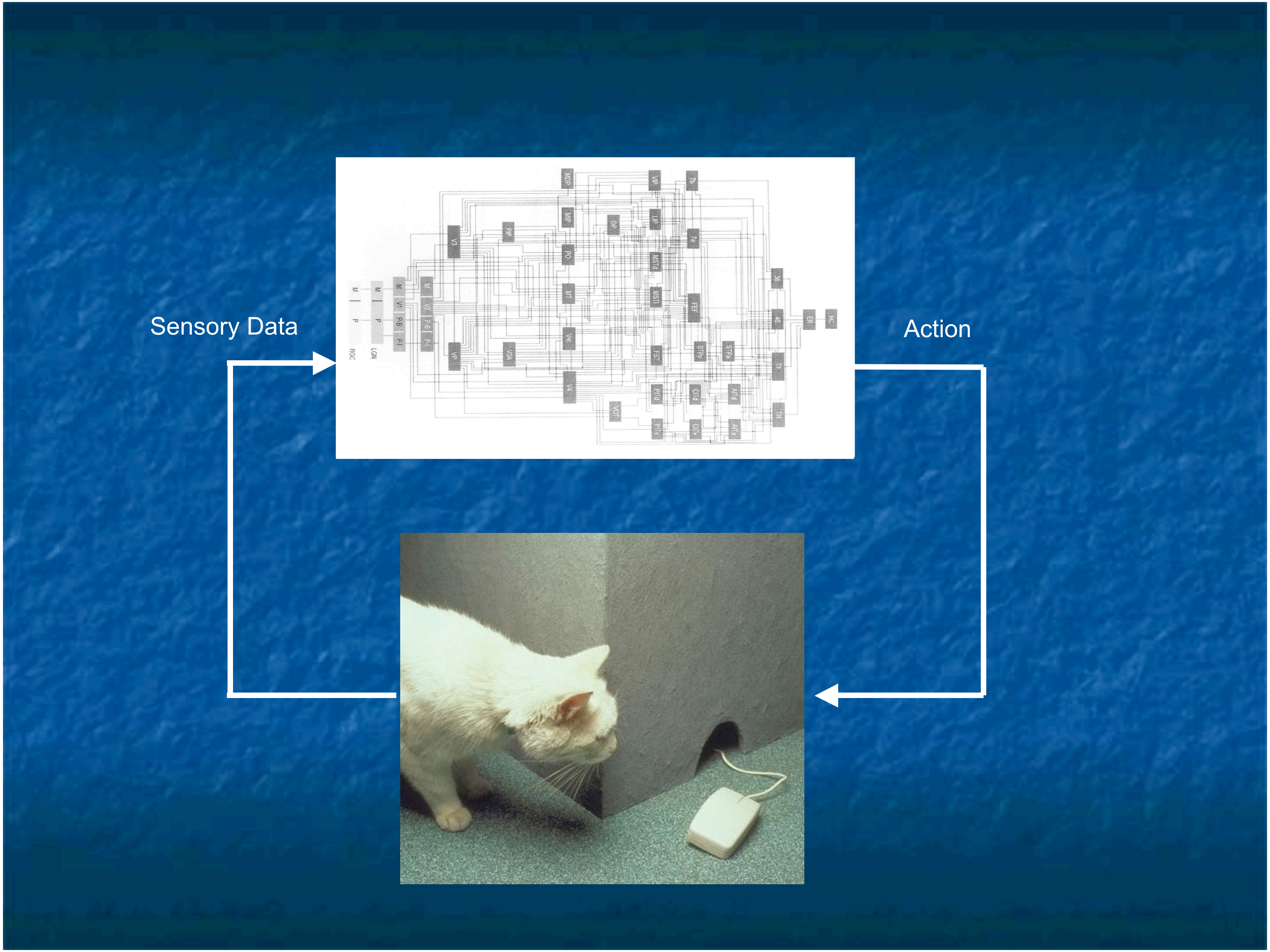
Deep Blue beats Kasparov (1997)

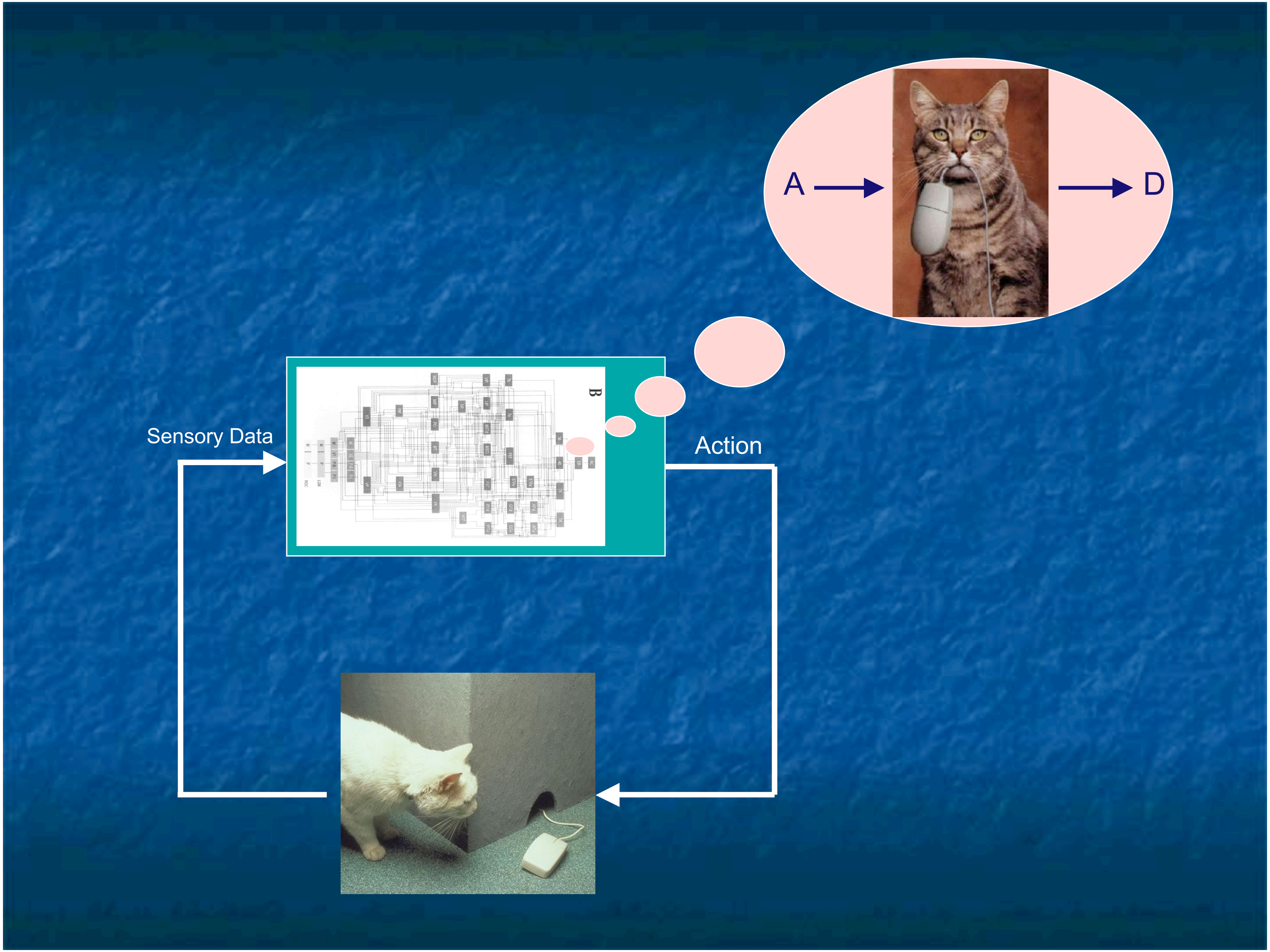
Children outmatch the most powerful machines in object perception

Imperfect knowledge about brain functions

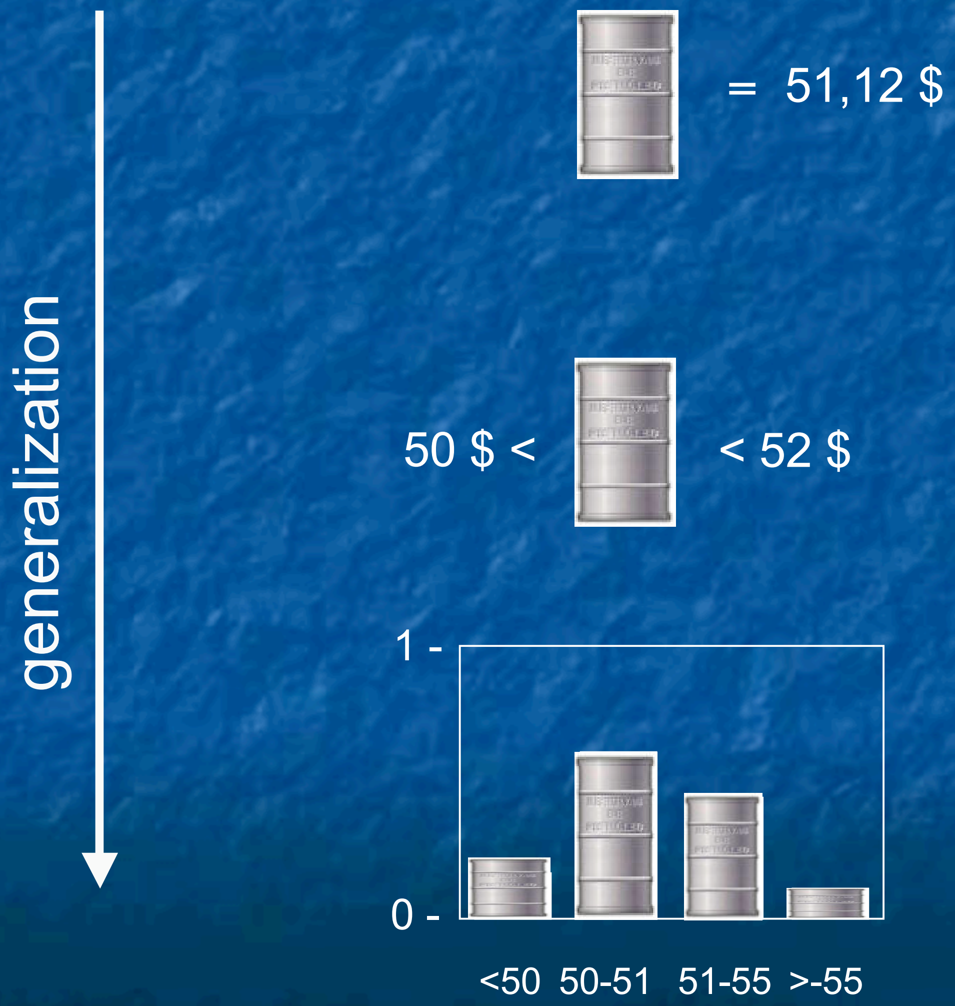
Brain’s imperfect knowledge about outside world:

- uncertainty and incompleteness of sensory data
- limited actions, temporal constraints
- internal model incompleteness


Sensory Data
Action


A
D
B
Sensory Data
Action

How to express various forms of knowledge ?

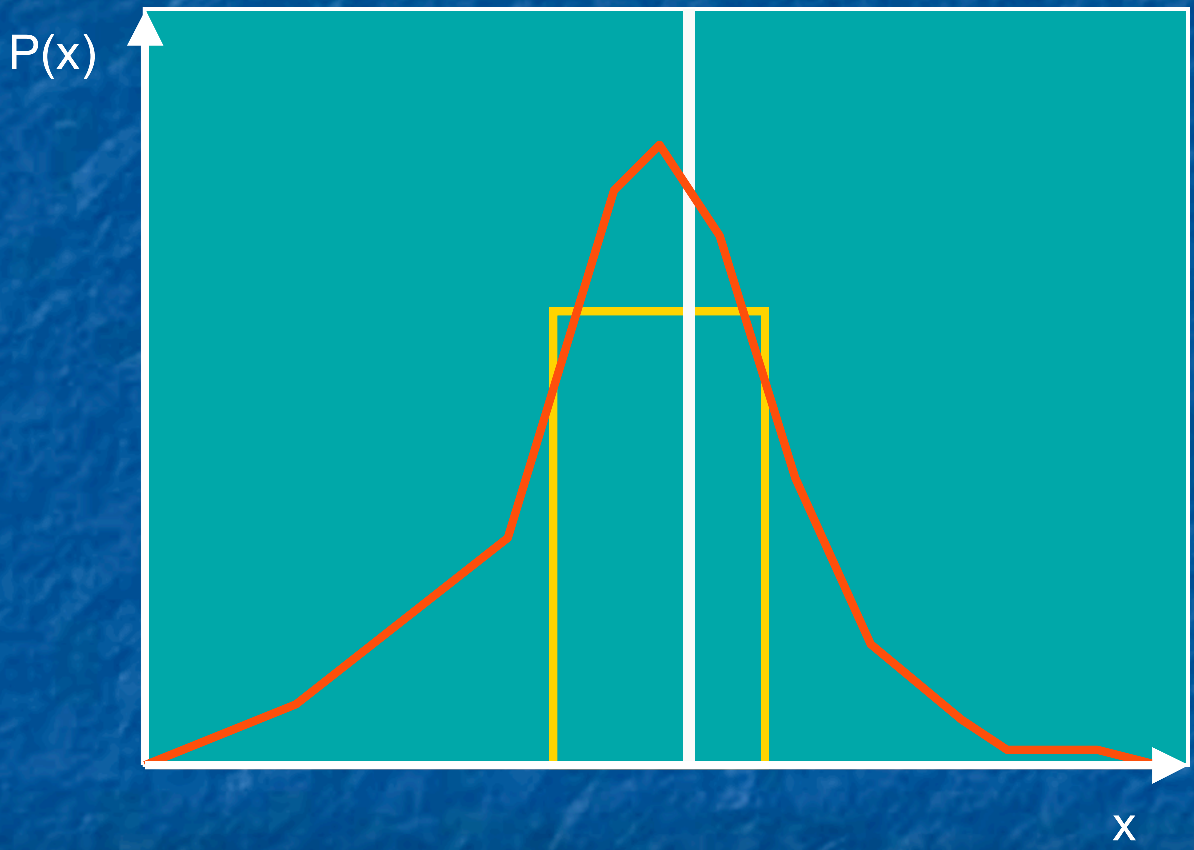


Ex.: Data, facts, ...

Ex.: thresholds, validity domains, ...

Ex.: uncertainty, variability, ...

⇒ Probabillity distribution as a common language for all knowledge forn

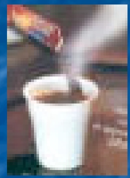


x ∈ E _ [0, 1]

$\Sigma_x P(X) = 1$

If … Then …

?

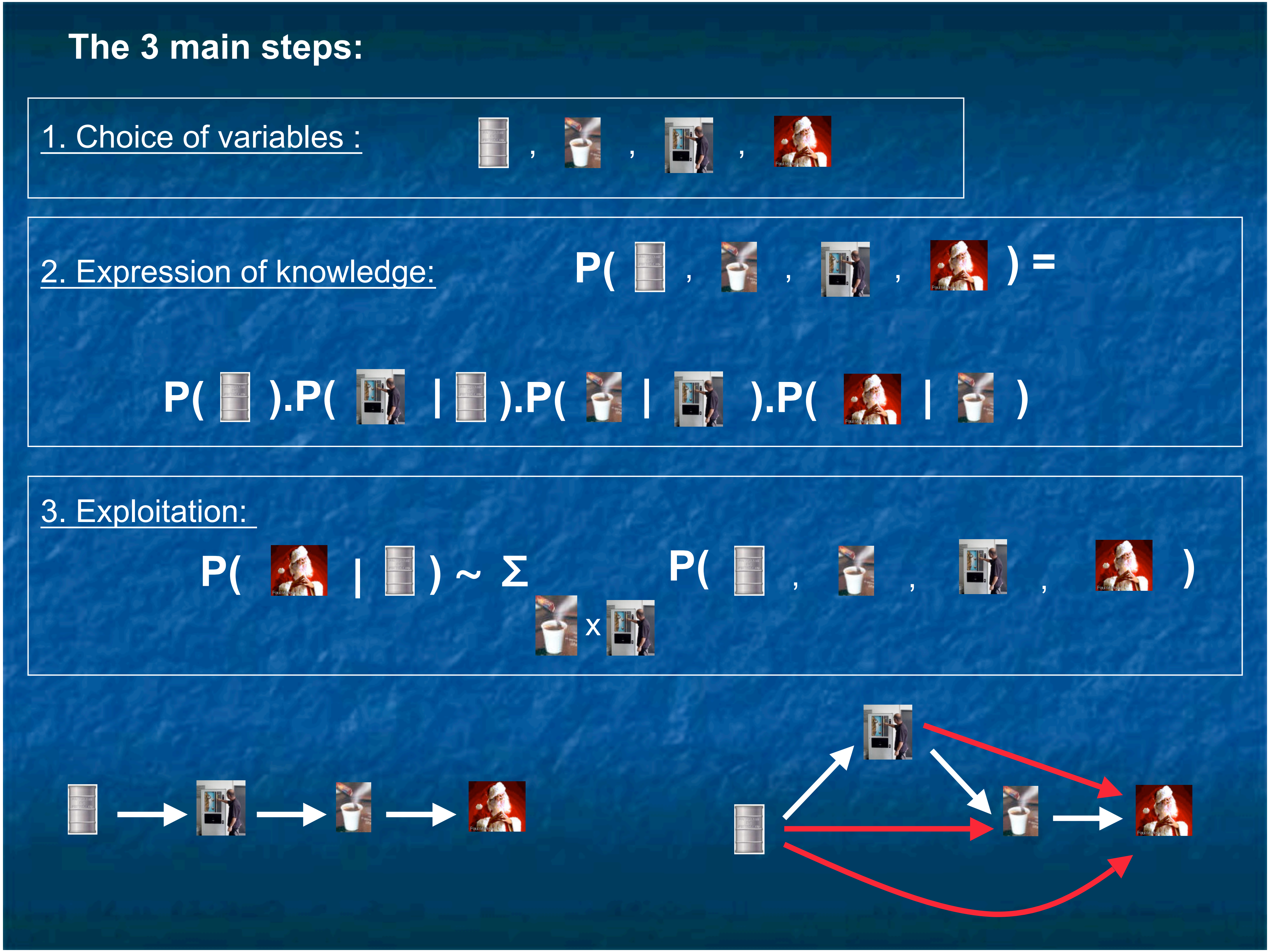


generalization

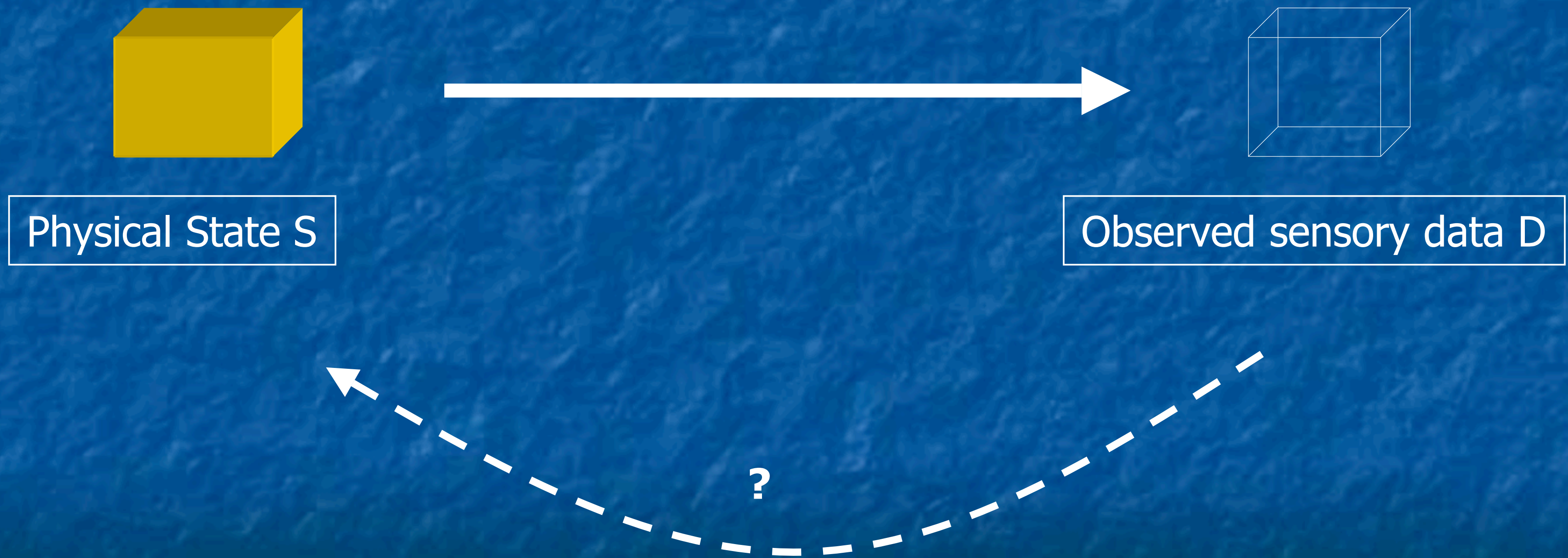
“If I know x, then y is exactly known” : function $x \rightarrow y = f(x)$

“If I know x, then y is confined in some subset” : $R(x,y)$ = true

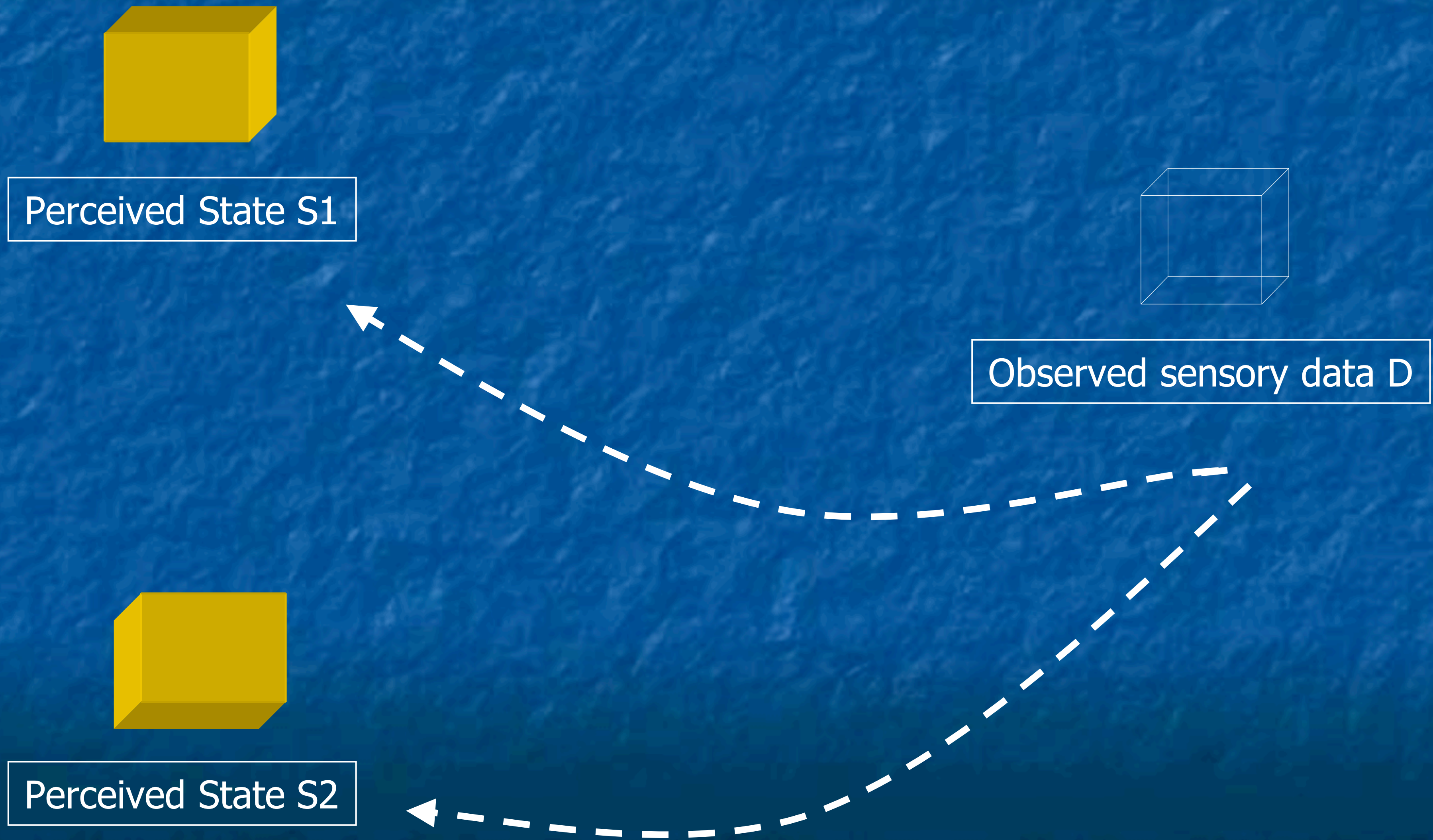
“If I know x, then I can improve my knowledge on y”: $P(y \mid x) = P(x, y) / P(x)$


The 3 main steps:
1. Choice of variables :
2. Expression of knowledge:
P(, , ,) =
P( ).P(|).P(|).P(|)
3. Exploitation:
P(|) ~ Σ P(, , ,)
x

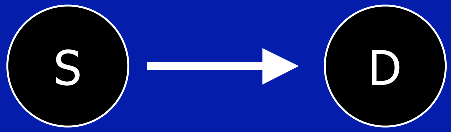
The perception is an inverse problem…



Worse, it is an ill-posed inverse problem...



A simplified perception scheme:



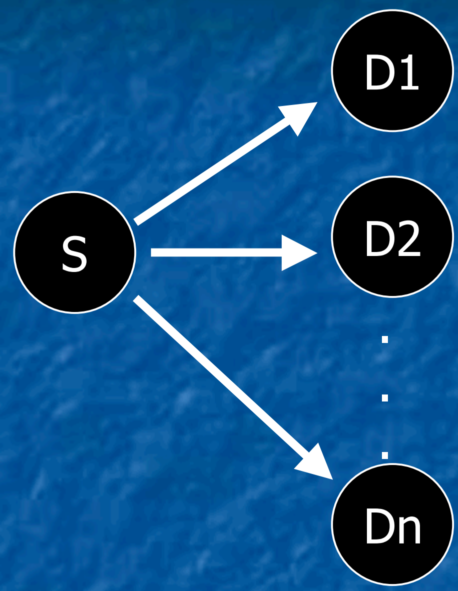
S: perceived state
D: sensory data
Prior: P(S)
Sensor model: P(D | S)

Perception:
P(S | D) ~ P(S).P(D | S)

Sensory Data

Action

Perception viewed as an internal model explaining sensory data



$P(S \mid D1\ D2 \ldots Dn) \sim P(S).\_{}_{i=1,n}\ P(Di \mid S)$

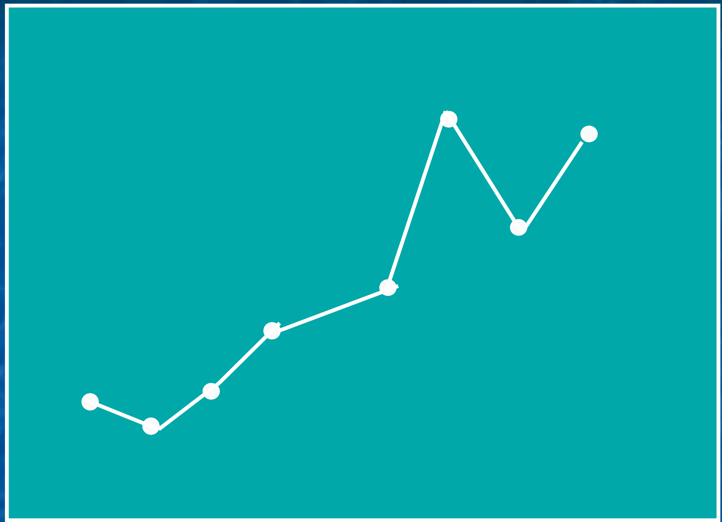
$\log P(S \mid D1\ D2 \ldots Dn) = \alpha + \log P(S) + \Sigma_{i=1,n}\ \log P(Di \mid S)$

Ex.: $\log P(S \mid D1\ D2 \ldots Dn) = \alpha - \_\ |S|^2 - \_\ \Sigma_{i=1,n}\ |Di - Si|^2/\sigma_m^2$

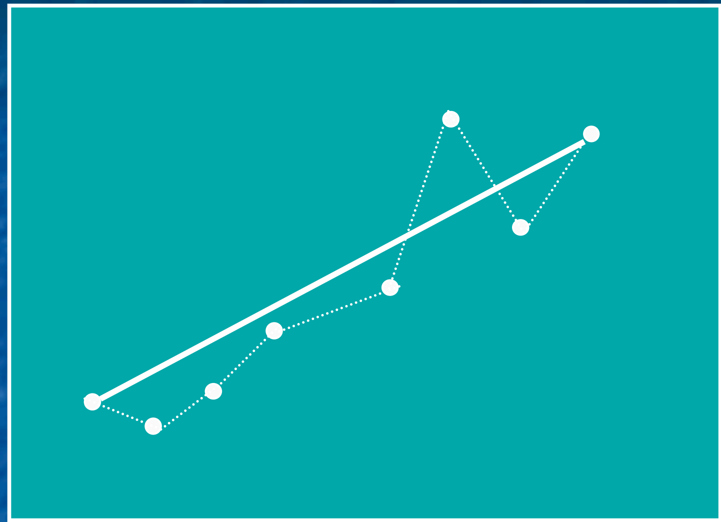
Ex. : polygonal segmentation from laser proximeter data

$P(S) = (1/Z).\ \mathrm{Exp}(-\beta.\ \text{Nb of } 1)$

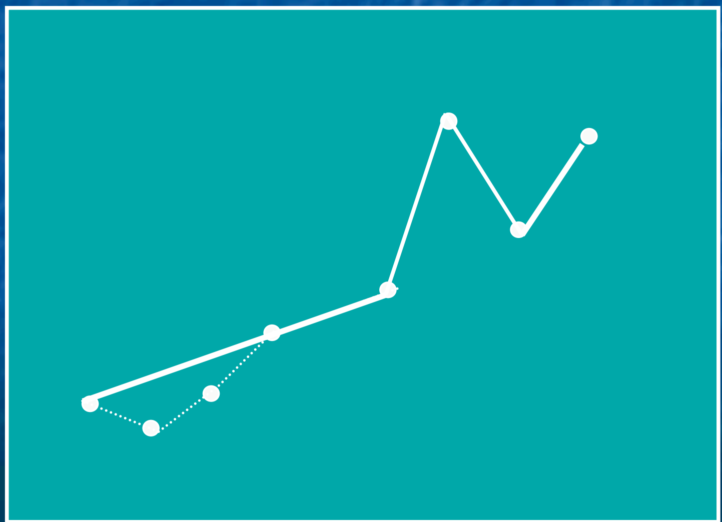
β = regularization constant



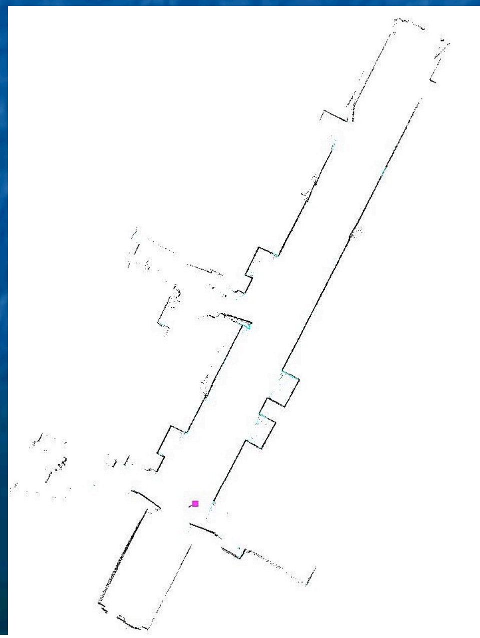
S=11111111



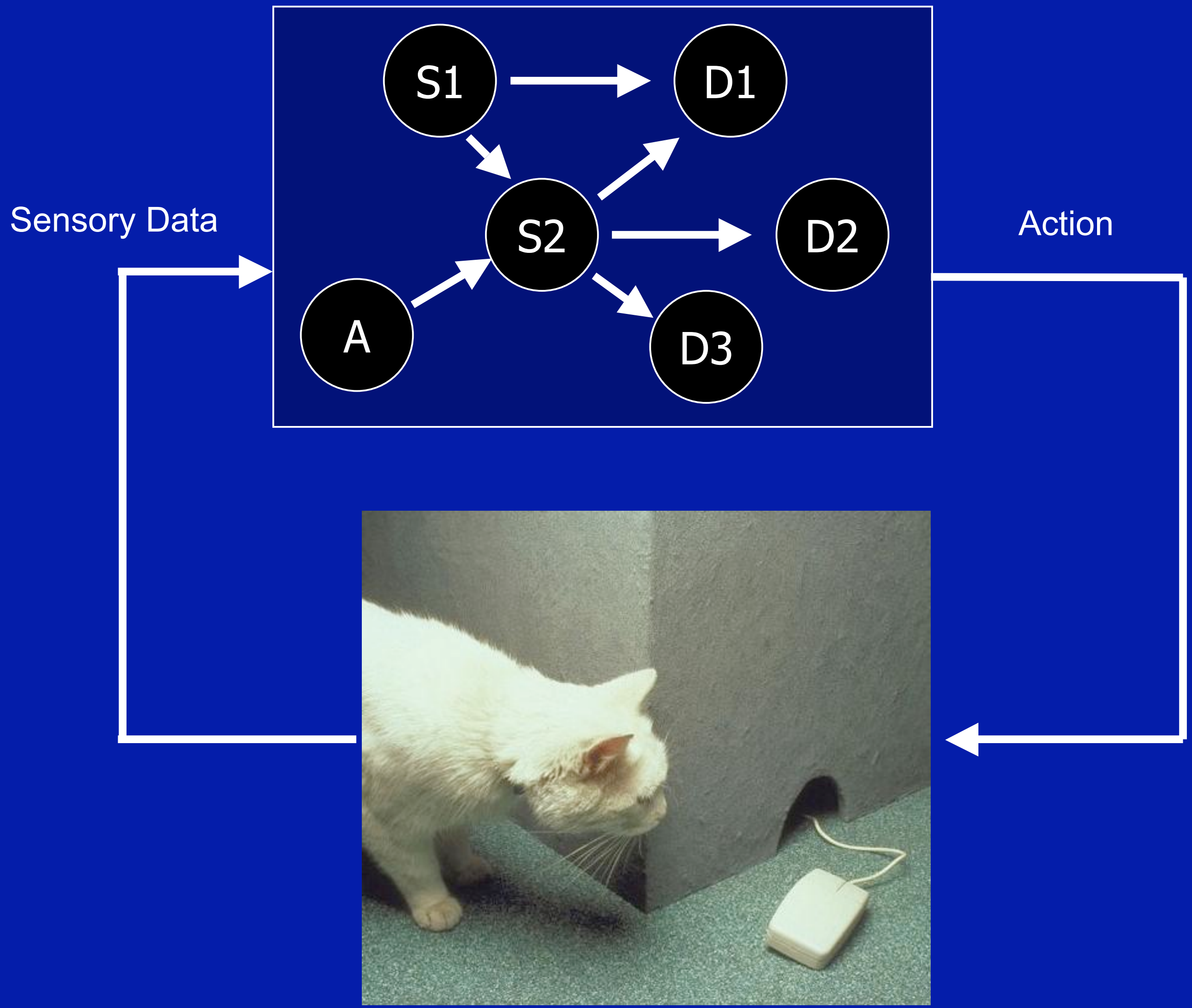
S=10000001



S=10001111



More complex structures can be captured by Bayesian networks

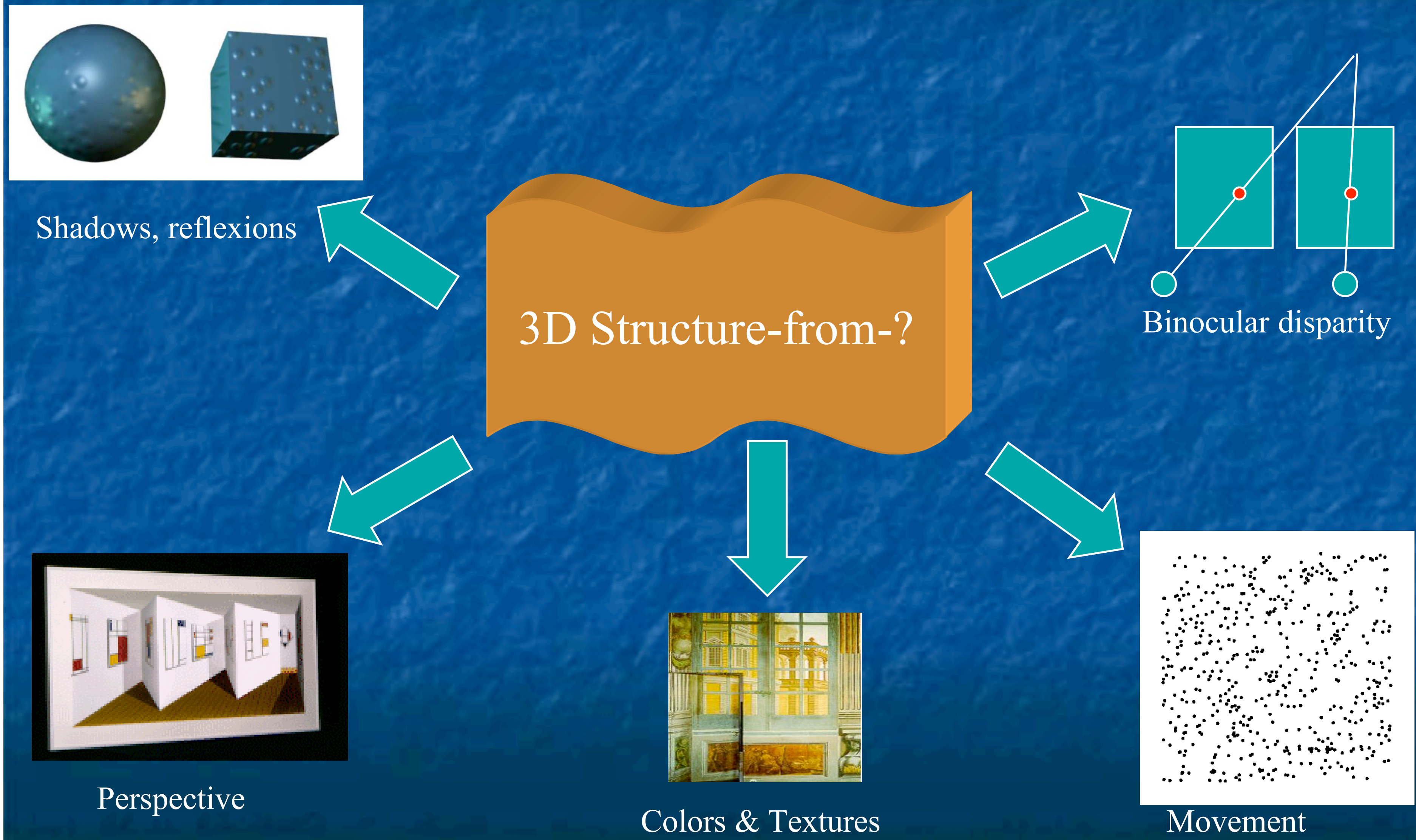


Two examples:

1. **3D shape from motion**
2. **3D self-motion perception**

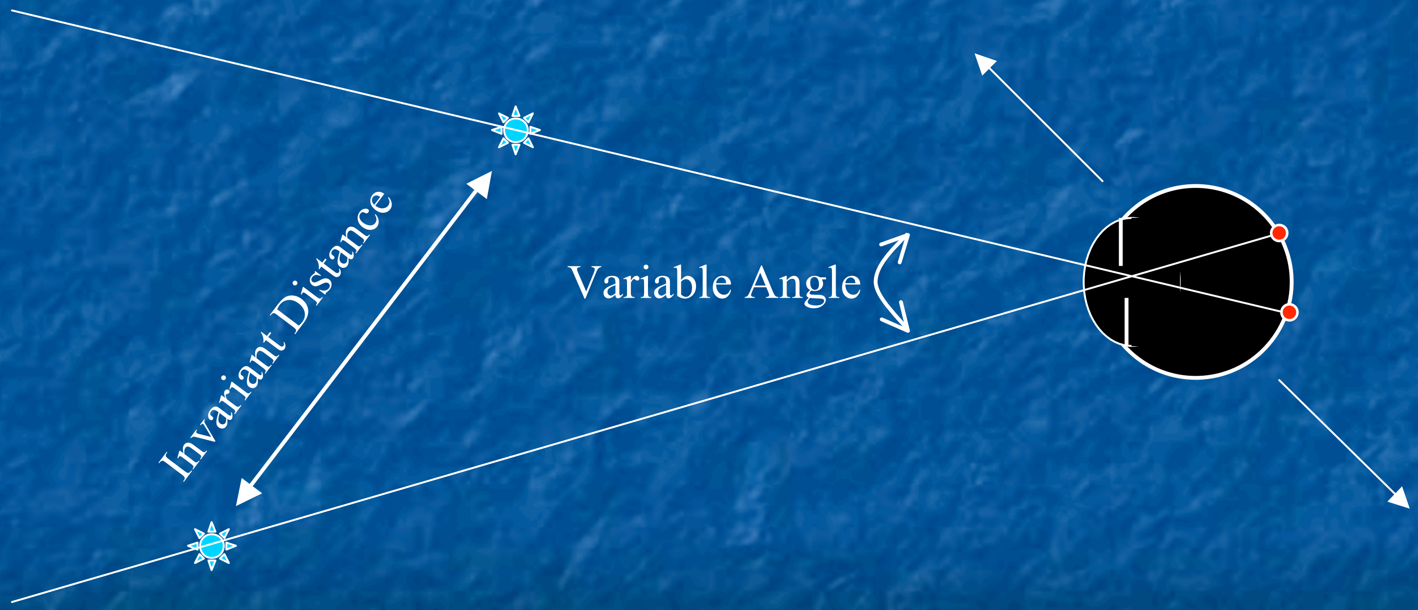
Object Perception

Numerous sources of information
Various characteristics (uncertainty, ambiguity)



Motion cues:

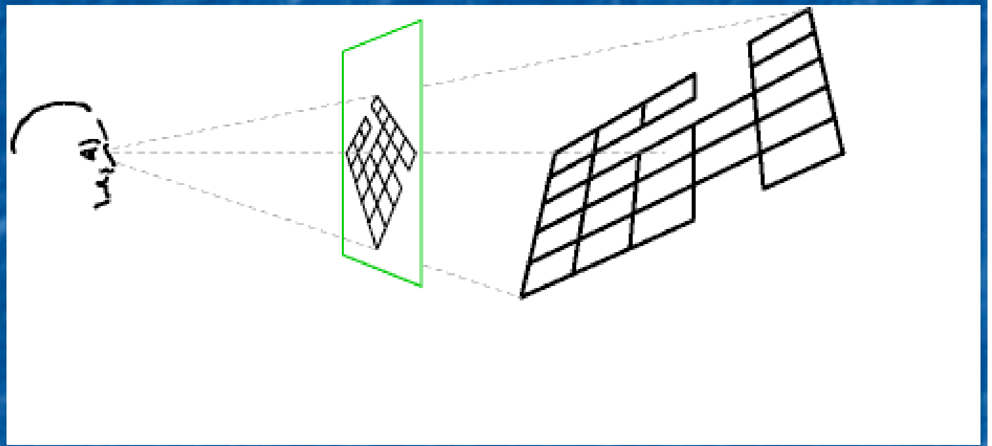
The rigidity assumption : the relative movement is a 3D isometric transformation



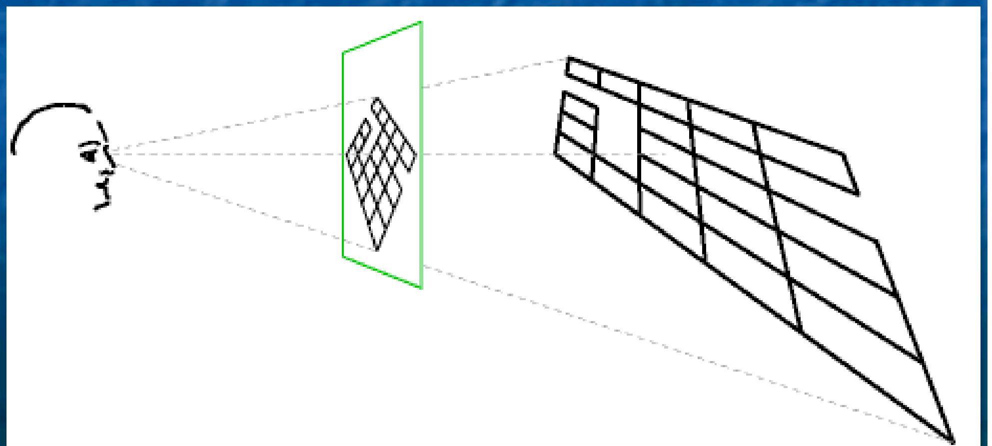
Perspective cues:

prior knowledge favoring regular texture on 3D surface

The image of a regular plane ...

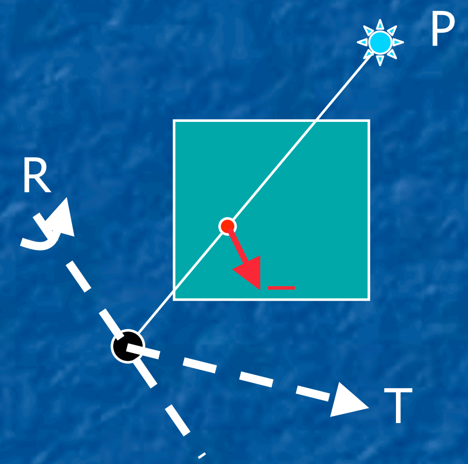


…back projected on another plane
⇒ « trompe-l’œil »



According to the rigidity assumption:
Optic Flow = rotation + translation velocity fields
_ = R + p.T **p = proximity map (1/distance)**

Object geometry (p) and relative 3D motion (R,T) determine the optic flow (_)
Knowing _, how to compute shape and movement parameters (p, R, T) ?



The direct functional model is quite simple:

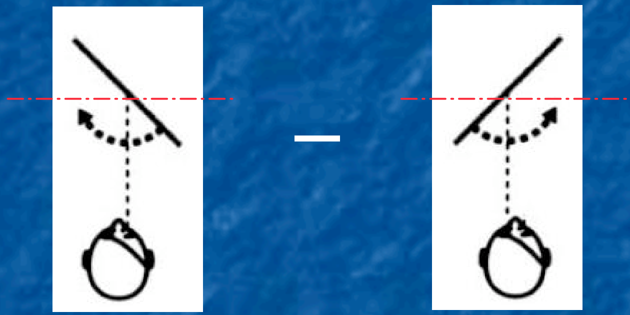
$$\__{x}(x,y) = p(x,y).(T_x - x.T_z) + (1 + x^2).R_y - y.R_z - xy.R_x$$
$$\__{y}(x,y) = p(x,y).(T_y - y.T_z) - (1 + y^2).R_x + x.R_z + xy.R_y$$

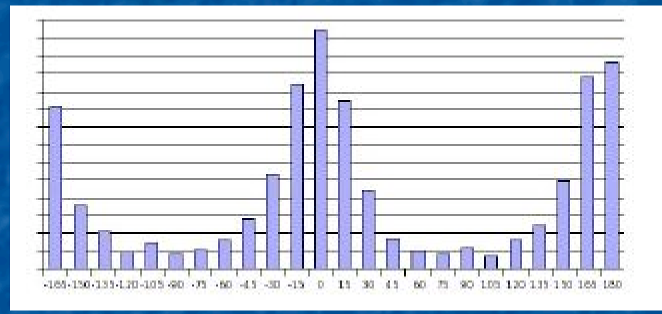
But the inverse problem is quite difficult ...

Non linear equations + “noise” + high dimension (~ 12)
⇒ General Algorithms are typically not robust

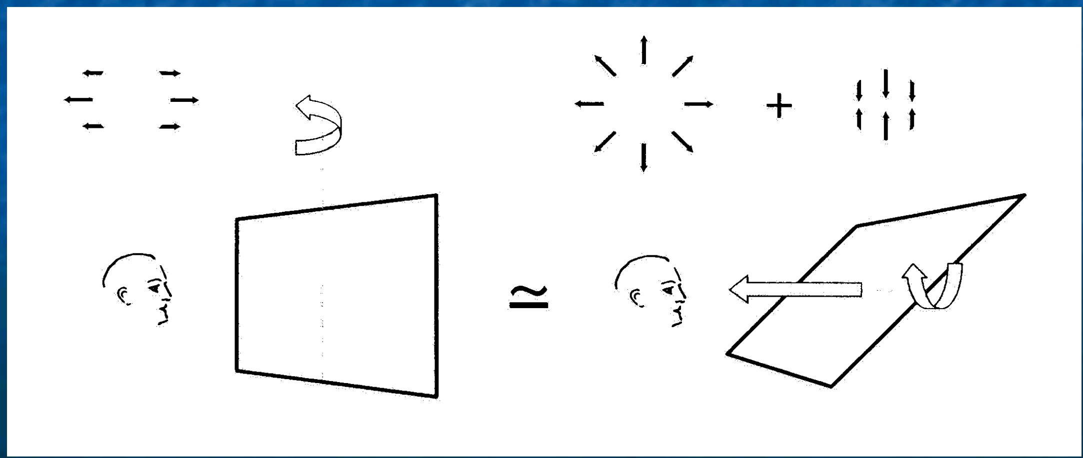
Several examples of optic flow ambiguities

1. Perceptive inversion (Fronto-parallel plane symmetry for both object & motion)

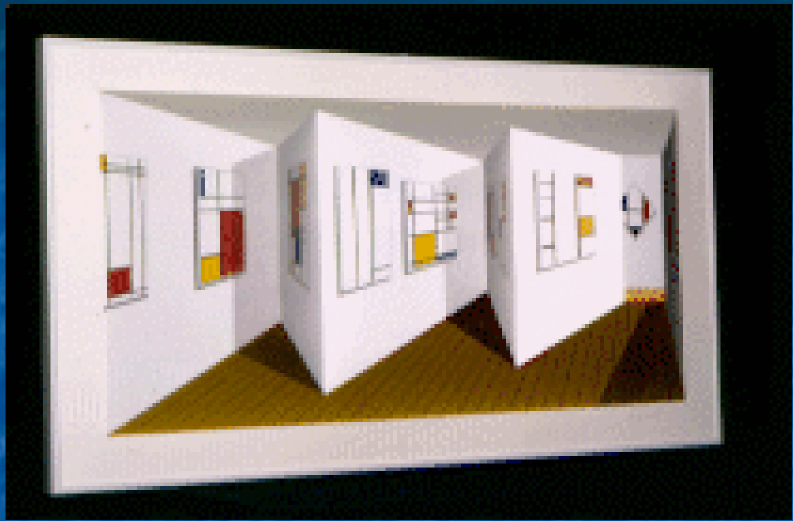


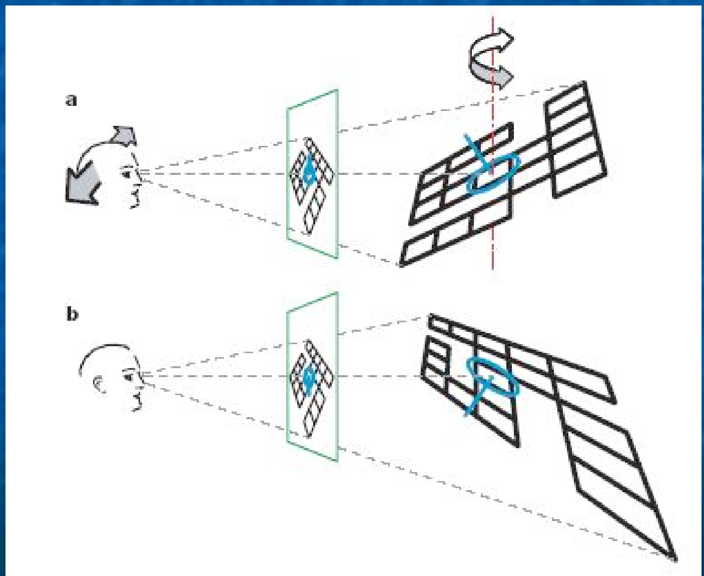


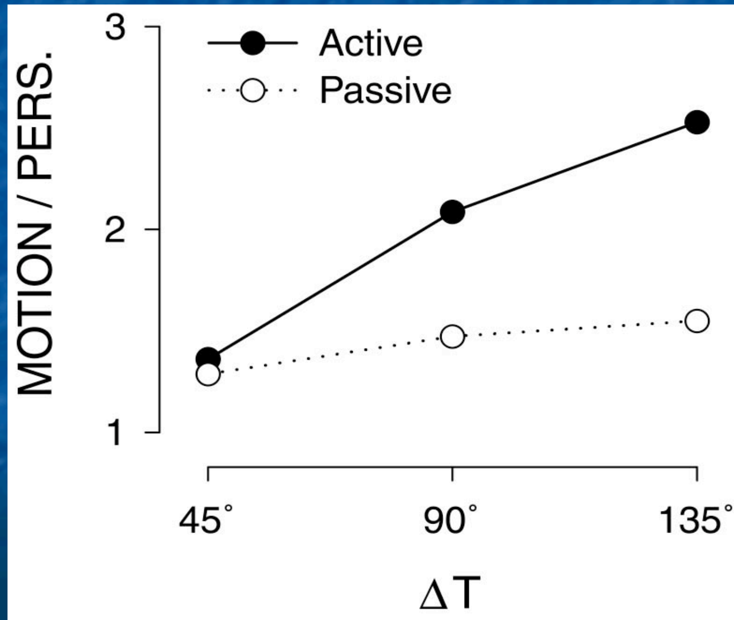
2. Similar optic flows result from different combinations of rotation and translation



3. Conflict between motion & pictorial cues



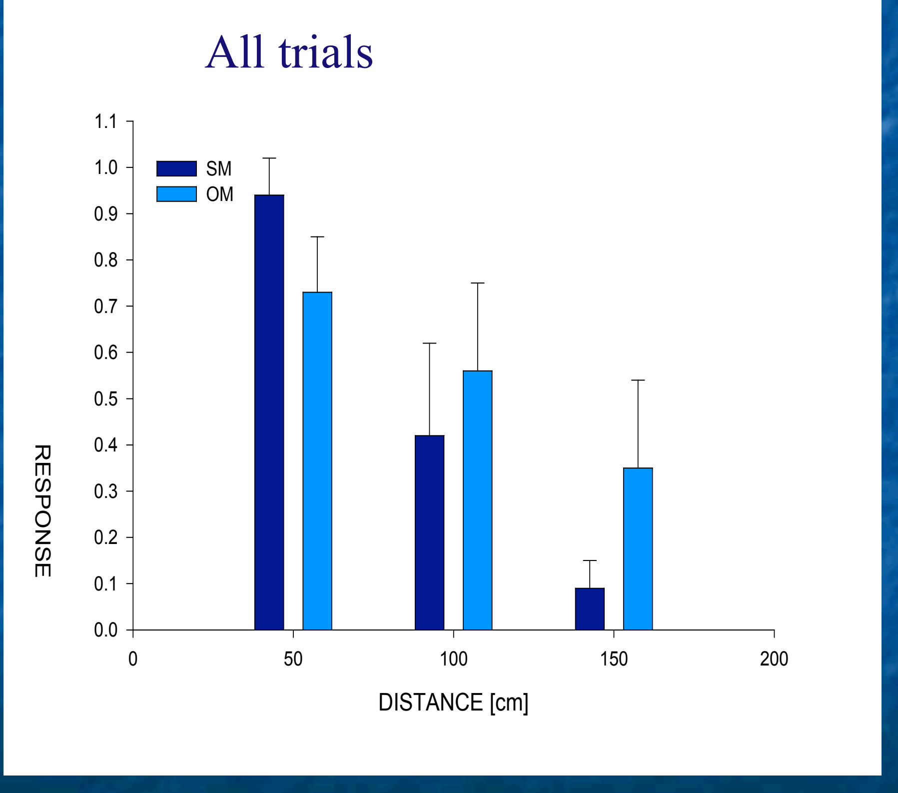


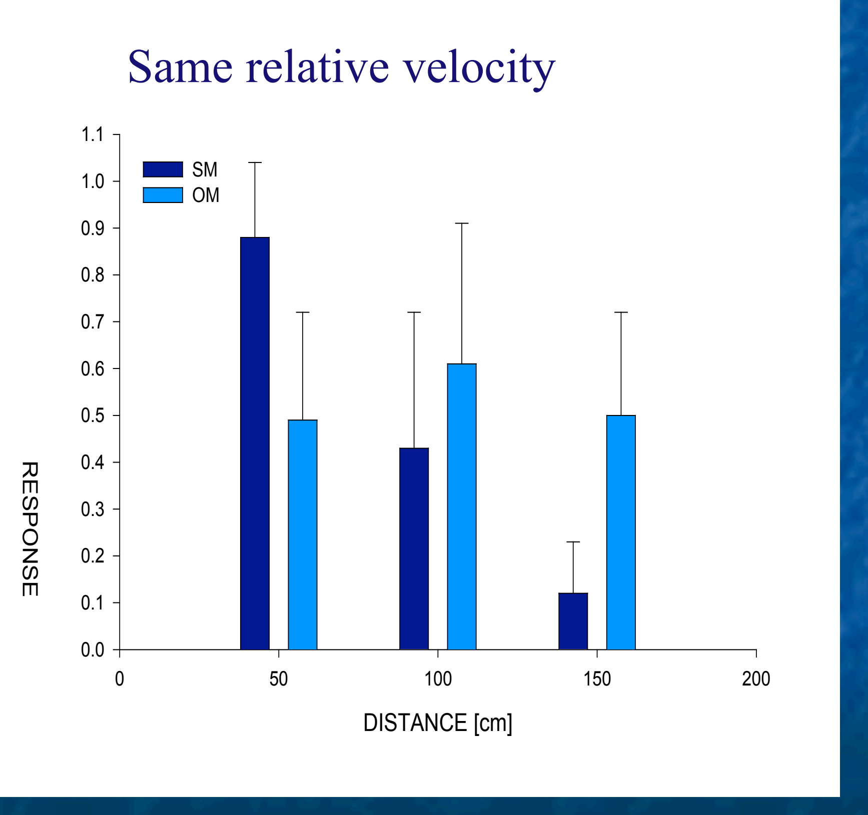


M. Wexler, F. Panerai, I. Lamouret & J. Droulez, Nature, 409, 85-88 (2001)

4. Contribution of self motion to depth perception (scale ambiguity)

Subject Motion (SM) versus Object Motion (OM)
Task: report whether or not object distance is less than one meter

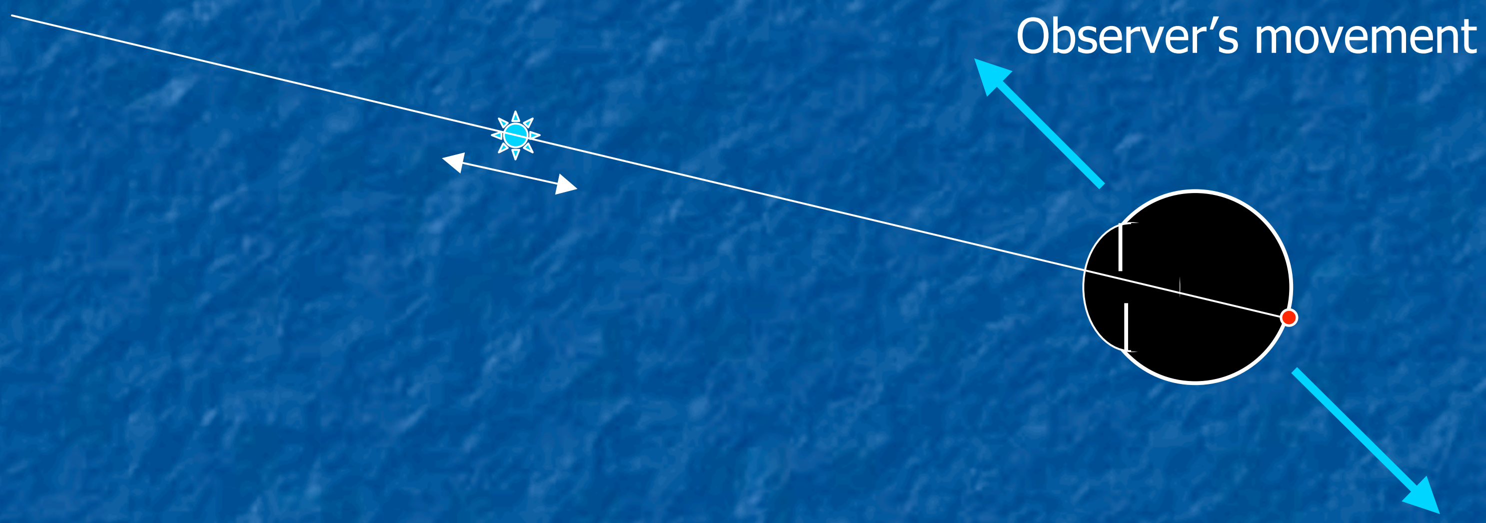




Panerai, Cornilleau-Pérès & Droulez, Perception & Psychophysics, 64: 717-731 (2002)

The stationarity assumption :

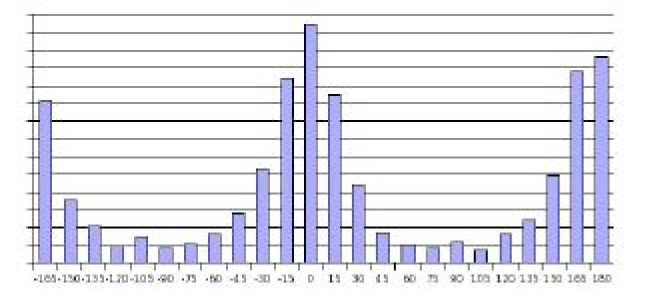
Preference for small allocentric object movement



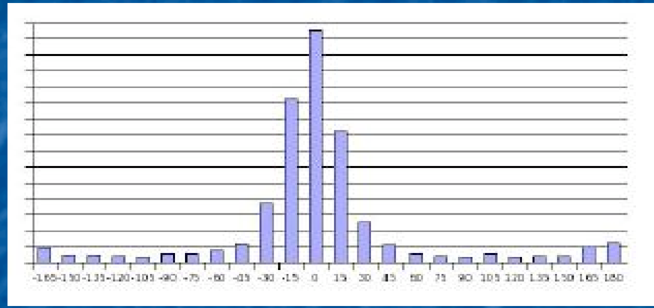
Knowledge about self motion (observer's displacement)

Can be used to remove optic flow ambiguities ⇒ Stationarity Assumption

Suppression of perceptive inversion



Same movement but produced by the observer



Preference for the most stationary solution (even if it is less rigid)

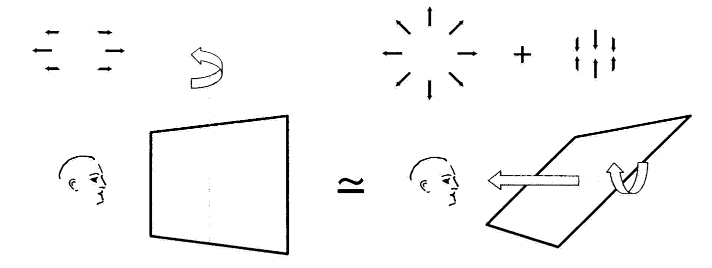


Fig. 4. Geometry of Experiment 2. On the left is a plane with horizontal tilt, rotating about a vertical axis. On the right is a plane with vertical tilt, approaching the observer while rotating about a horizontal axis. Both motions results in the same first-order optic flow, whose components are shown above the corresponding human figure.

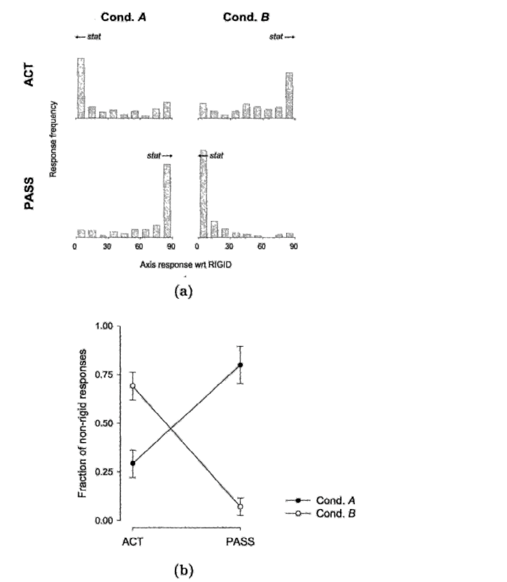
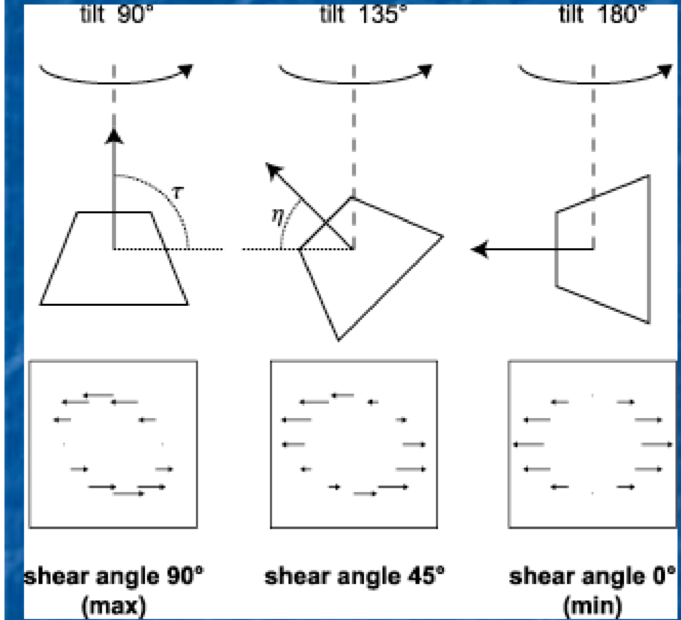
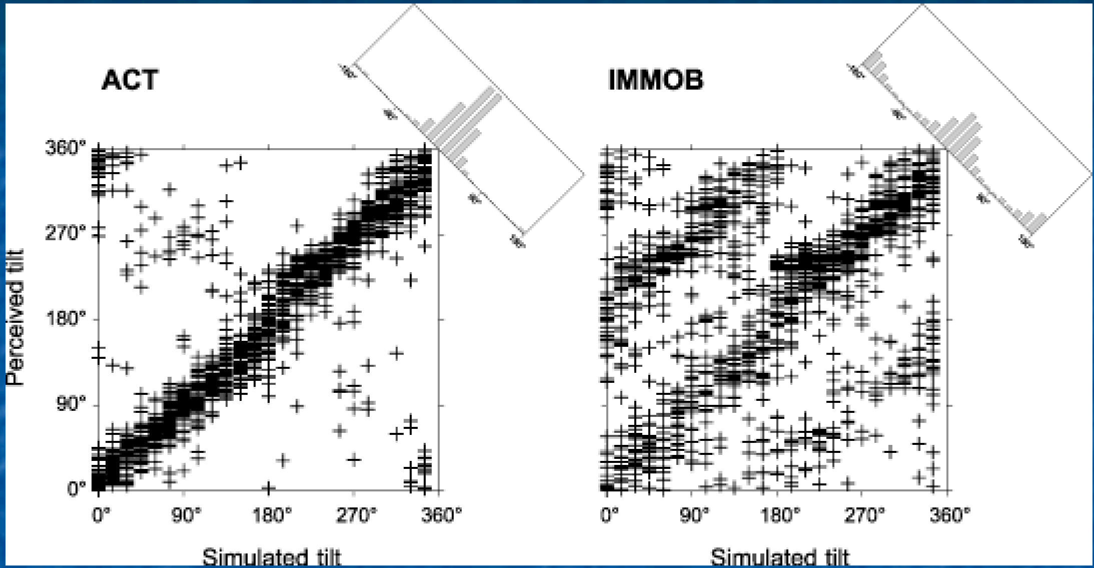


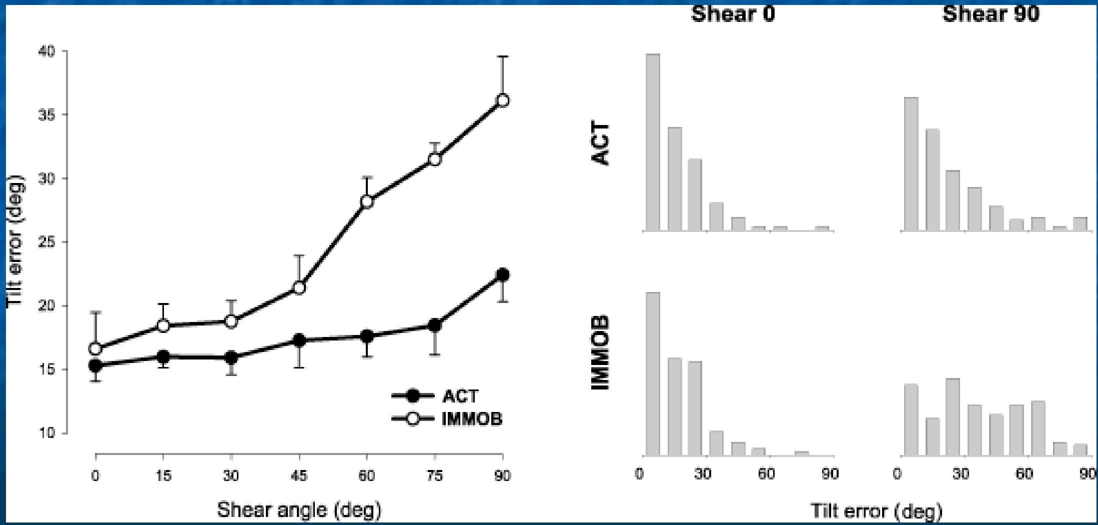
Fig. 6. Results of Experiment 2. (a) Distribution of axis responses for all subjects, in the ACT and PASS conditions, in conditions *A* and *B*. Response 0° indicates the rigid solution, 90° the deforming solution. Predictions of the stationary hypothesis are marked by 'stat'. (b) Fraction of responses closer to the non-rigid solution than to the rigid solution. Error bars represent between-subject standard errors.

Wexler, Lamouret & Droulez, Vision Research, 41, 3023-3037 (2001)

Variability of perceptive responses (« shear effect »)







J. Van Boxtel, M. Wexler & J. Droulez, Journal of Vision 3(5) : 318-332. (2003)

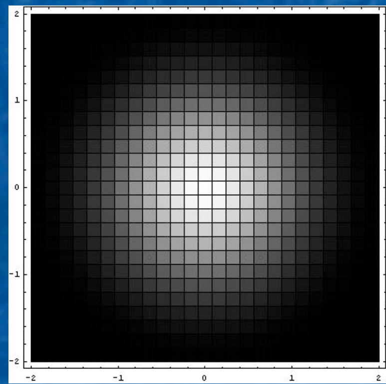
Structure of the probabilistic model

Variables : Object structure, Observer motion, Relative Motion, Optic Flow

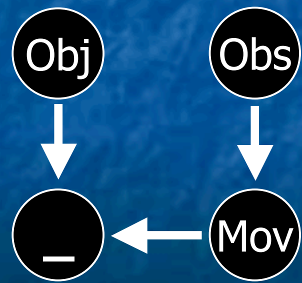
Knowledge Expression:

P(Obj, Obs, Move, Flow) = P(Obj).P(Obs).P(Move | Obs).P(Flow | Move, Obj)

P(Obj) = “Fronto-parallel plane prior”
P(Obs) = “Self-motion information”
P(Move | Obs) = “Stationarity assumption”
P(Flow | Move, Object) = “Rigidity assumption”



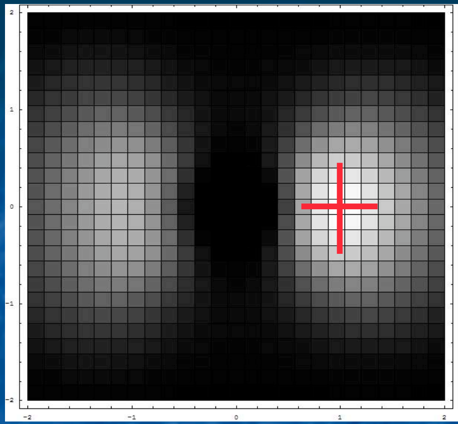
Question: P(Obj | Obs, Flow) ?



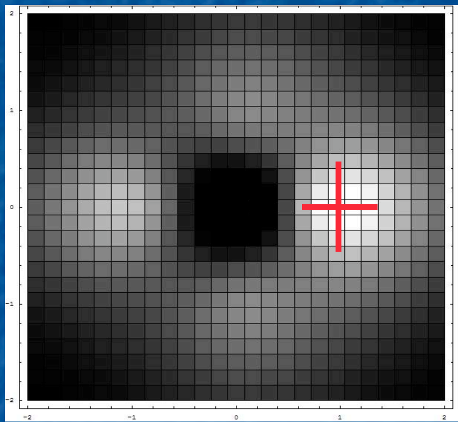
Experimental results to be explained:
- Perceptive Inversion (suppressed in active condition)
- Perceptive variability due to shear (reduced in active condition)
- 90° Rotation of perceived orientation with added depth translation

F. Colas, J. Droulez, M. Wexler & P. Bessière (2005)

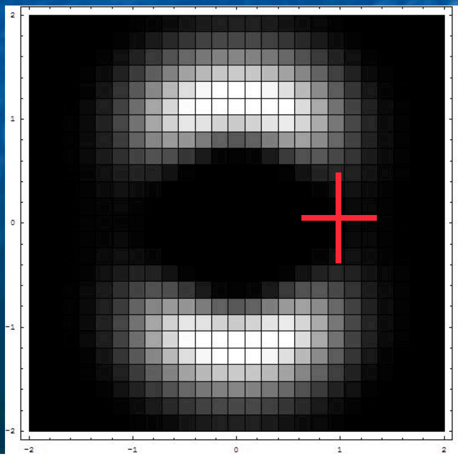
Probabilistic model (results)



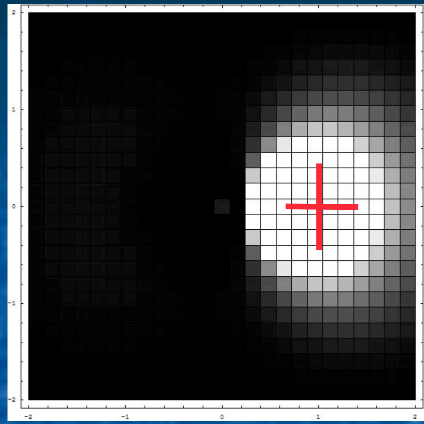
Immobile
Subject
Shear 0°



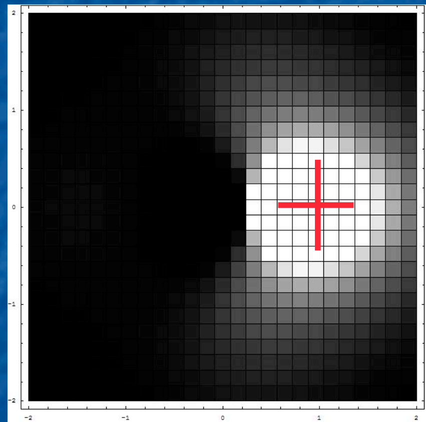
Immobile
Subject
Shear 90°



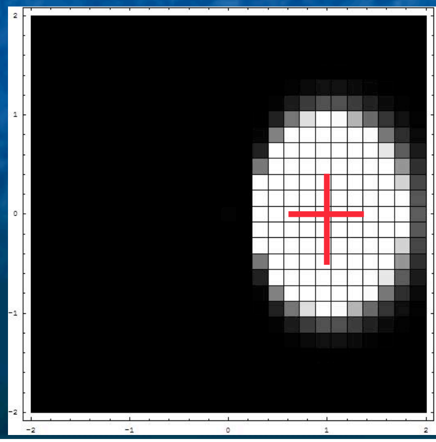
Immobile
Subject
+TZ



Moving
Subject
Shear 0°



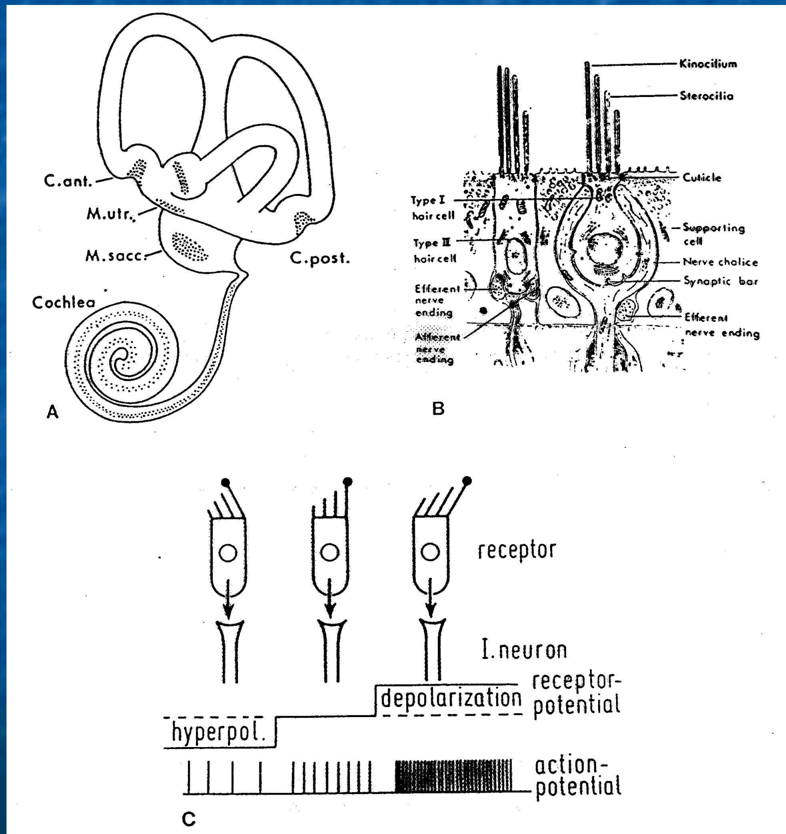
Moving
Subject
Shear 90°



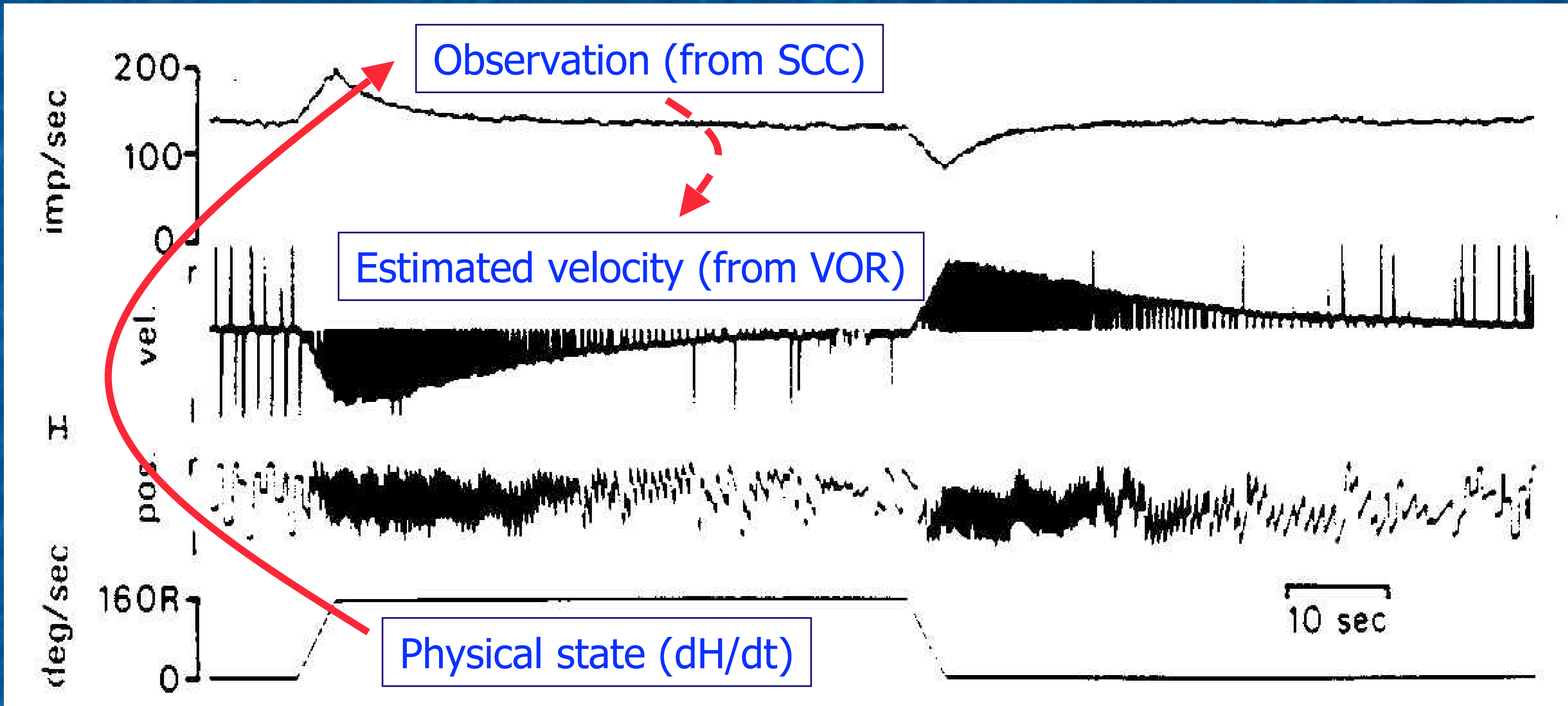
Moving
Subject
+TZ

2nd example: Self-motion perception

The vestibular sensor : 3 semi-circular canals (head angular acceleration)
+ 2 otolithic organs (head linear acceleration)



A first example of ambiguity: how to estimate the sustained angular velocity ?



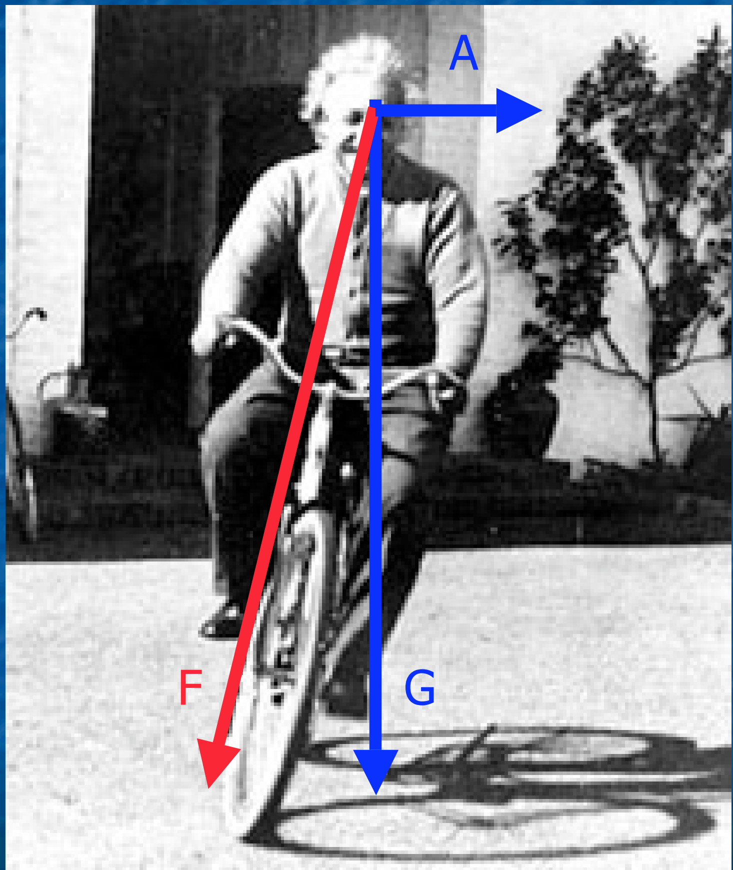
Data from Büttner & Waespe (81)

While an exact integration (from filtered acceleration to velocity) is mathematically straightforward, it would yield error accumulation with noisy sensory data !
⇒ The brain favors low estimated velocity

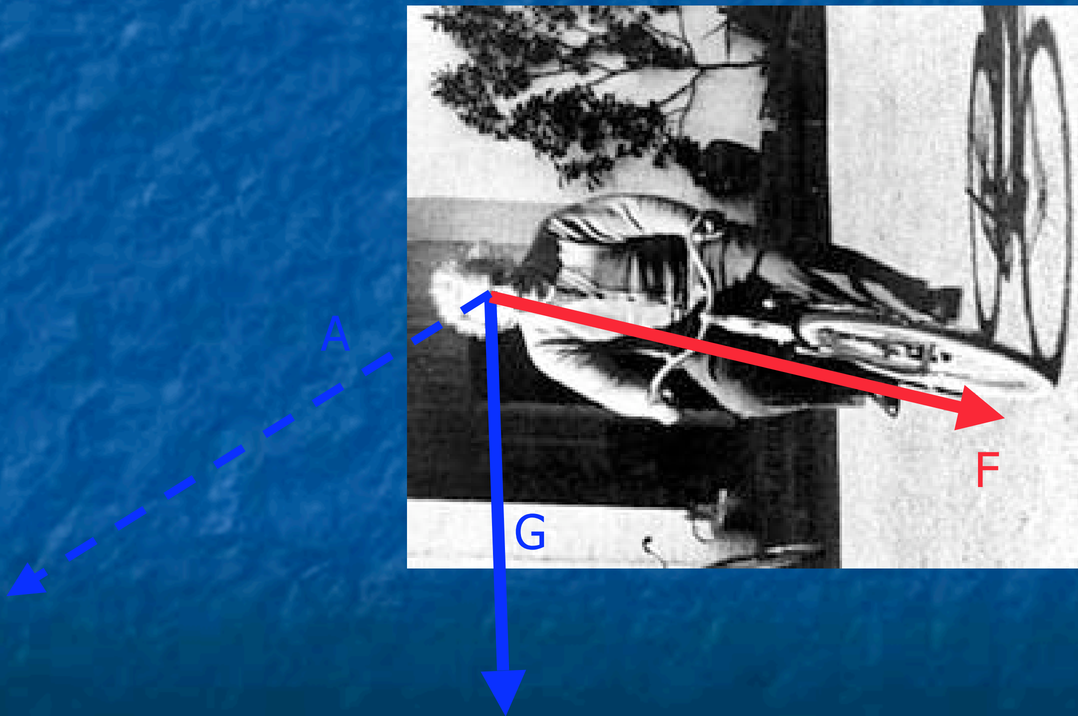
Another well-known example of ambiguity:
how to distinguish the inertial linear acceleration from gravity ?

F = G - A

The physical state (A,G) cannot be inferred from the observed otolithic signal (F)



The actual solution

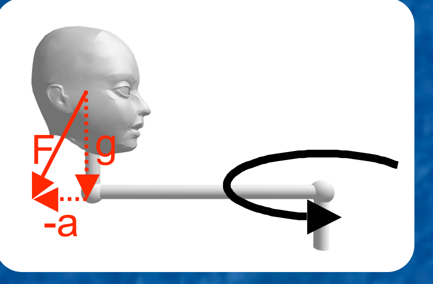


Another solution to the inverse problem

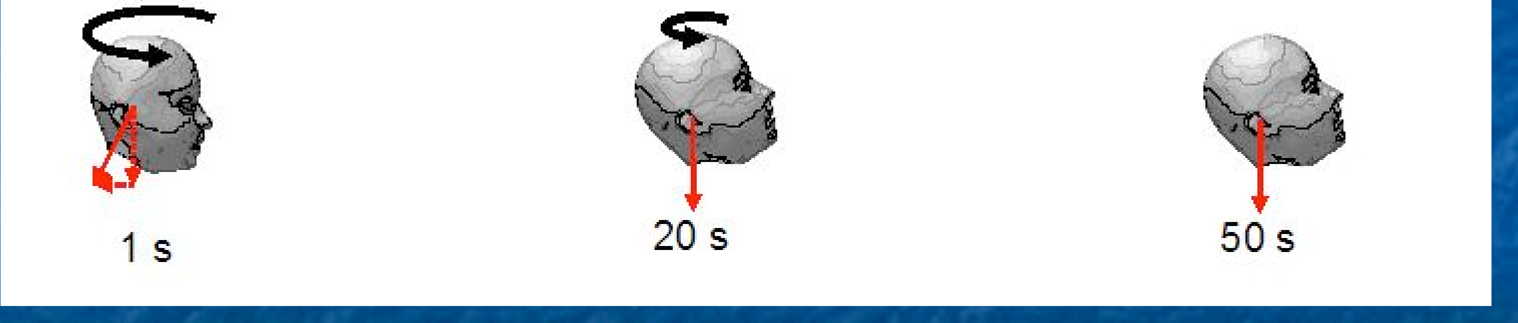
Both ambiguities combine each other !

Ex.: during off-axis rotation (centrifugation)

Physical state

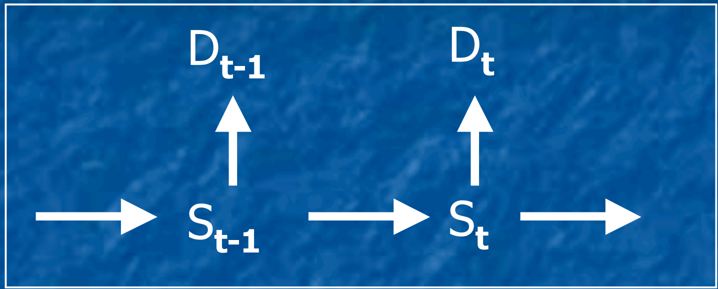


Perceived states



Decreasing the estimated angular velocity
⇒ alignment of estimated gravity with F
⇒ decreasing estimated linear acceleration to 0

Dealing with temporal series of variables: Bayesian Filters



Variables: $D_1, \ldots, D_t, S_0, S_1, \ldots, S_t$

Knowledge expression :

$P(D_1, \ldots, D_t, S_0, S_1, \ldots, S_t) = P(S_0).P(S_1 \mid S_0).P(D_1 \mid S_1)\ldots.P(S_t \mid S_{t-1}).P(D_t \mid S_t)$

Observation : $P(D_t \mid S_t)$ « sensor models »

Transition: $P(S_t \mid S_{t-1})$ « dynamic models »

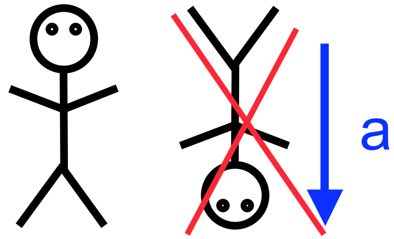
Exploitation:

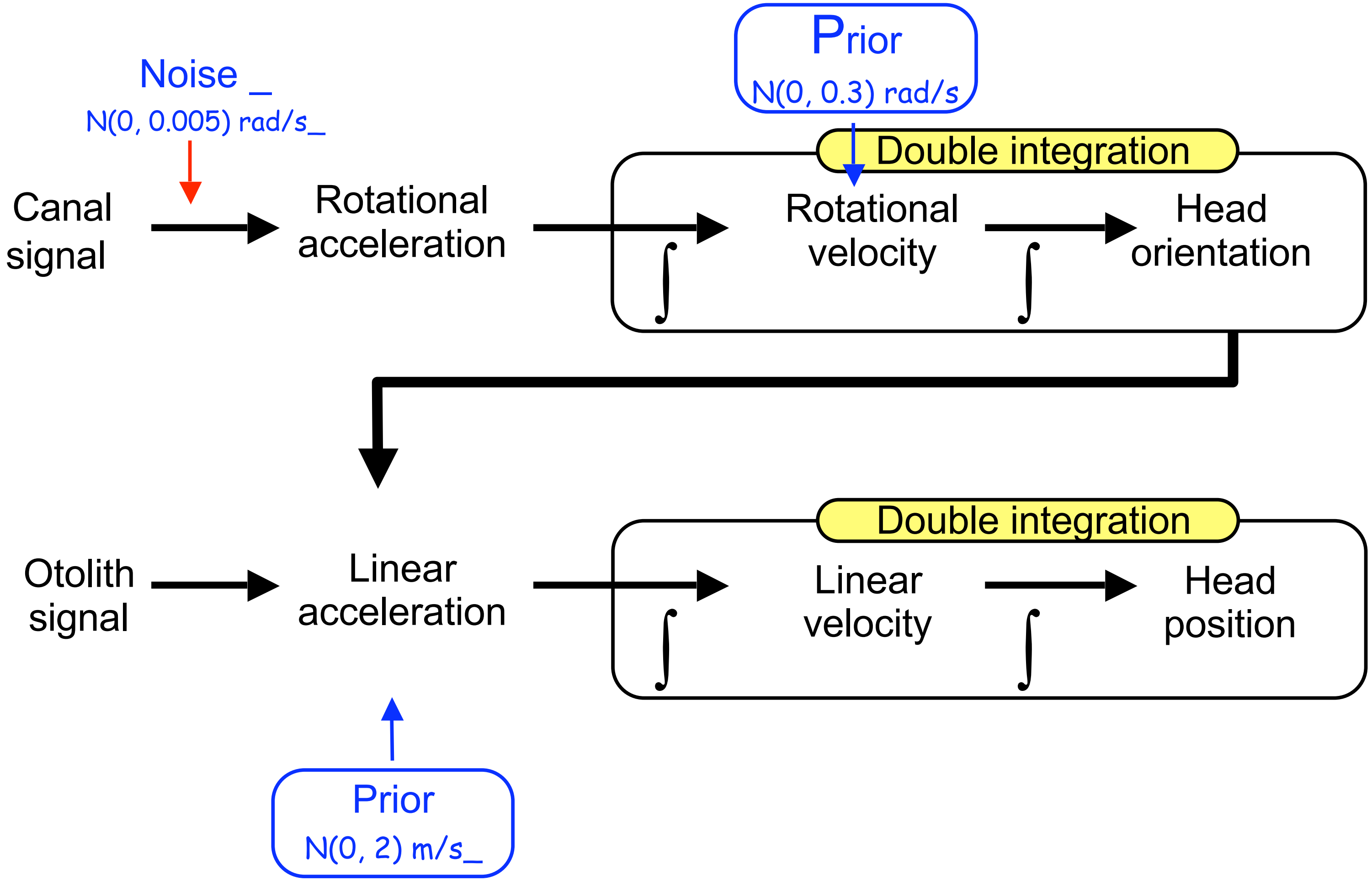
$P(S_t \mid d_1, \ldots, d_t) \sim P(d_t \mid S_t). \Sigma_{St-1} P(S_t \mid S_{t-1}).P(S_{t-1} \mid d_1, \cdots, d_{t-1})$

Particular cases: HMM, Kalman

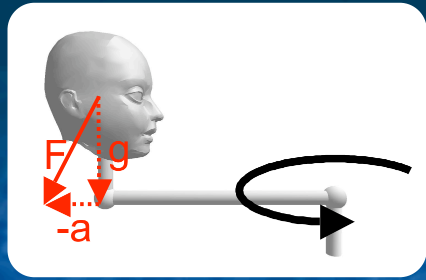
Dynamics + priors

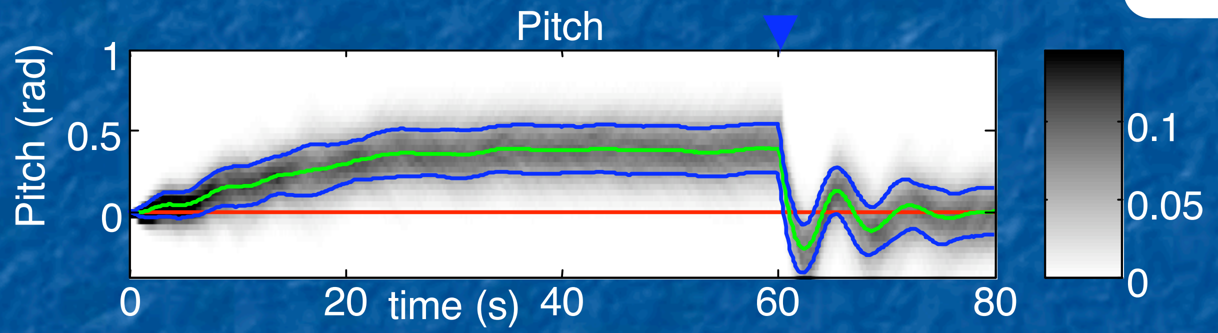
(Low angular velocity & linear acceleration)

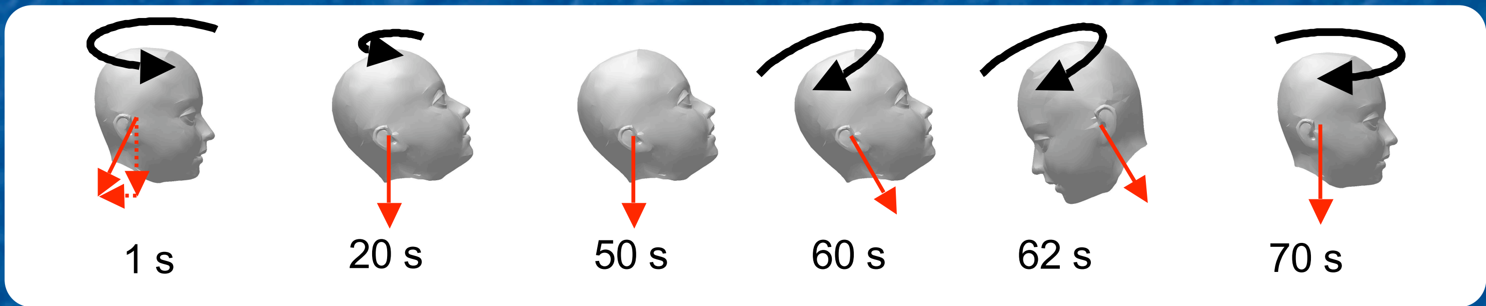


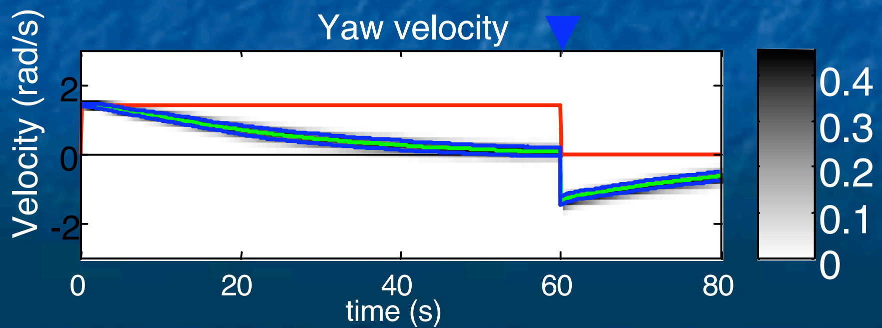


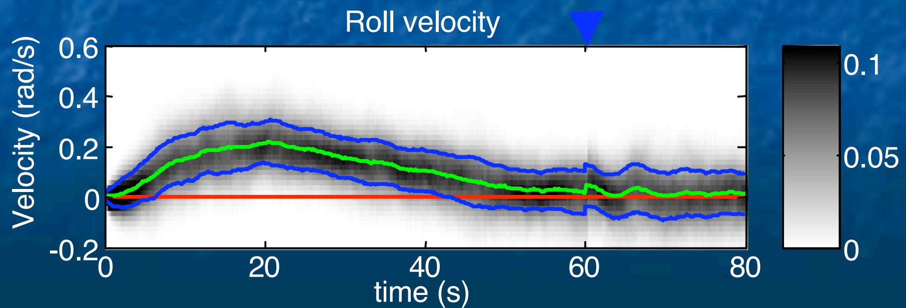
Centrifugation (Off Axis Rotation)









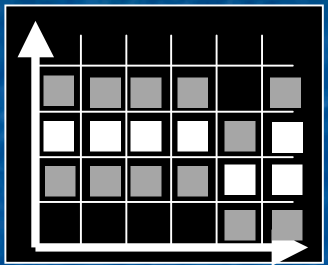


J. Laurens & J. Droulez, 2004

Neural Implementation of probabilistic computations:
(the 3rd person viewpoint)

Main issues:

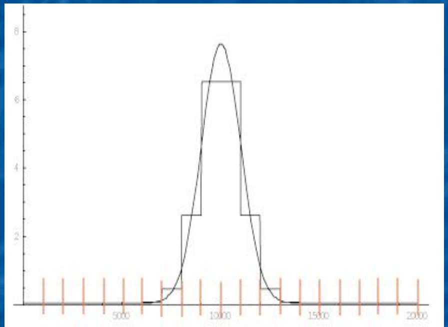
1. Relevant variables ?



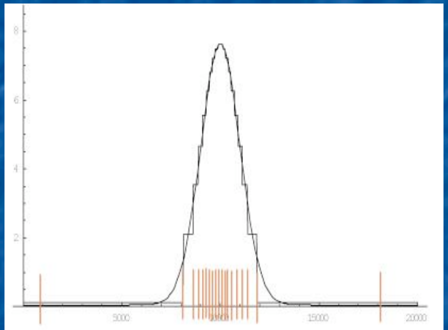
Coordinates: x1, x2...

Space-time subsets: S1, S2, ...

2. Neural code for P(x) ?

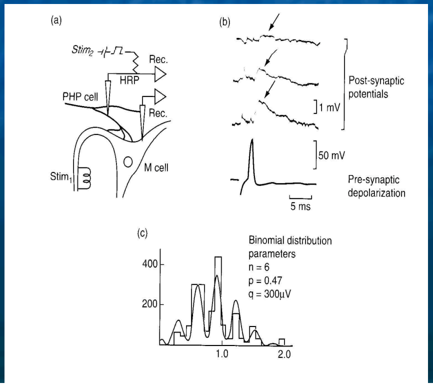


$Ri(t) \sim \int_{\Delta t} p(xit).dt$



$\int_{t1}^{t2} p(xit).dt = \Delta$

3. Reduction of computational costs ?



Sampling through random neuroT release

(H. Korn & DH Faber, 87)

Acknowledgements

Pierre Bessière
Jeroen van Boxtel
Simon Capern
Francis Colas
Valérie Cornilleau-Pérès
Frédéric Davesne
Ivan Lamouret
Jean Laurens
Francesco Panerai
Mark Wexler

Thank you for your attention,
and for your patience …