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?

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?

Probability as alternative to logic

Uncertainty is not in things but in our head: uncertainty is a lack of knowledge.

Jacob Bernouilli, Ars Conjectandi (Bernouili, 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)

Probability as alternative to logic

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.

E.T. Jaynes, Probability theory theory: the logic of science (Jaynes, 2003)

Probability
as an alternative to logic

Incompleteness

Probability as an alternative to logic

Incompleteness

Preliminary Knowledge
+
Experimental Data
=
Probabilistic Representation

Learning

Entropy Principles

Uncertainty

Probability as an alternative to logic

Incompleteness

Preliminary Knowledge
+
Experimental Data
=
Probabilistic Representation

Learning

Entropy Principles

Uncertainty

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

$$P(a \wedge b) = P(a) \times P(b \mid a)$$
$$= P(b) \times P(a \mid b)$$

Bayesian inference

Decision

Bayesian Programming & ProBT®

Bayesian Program

Description

Specification

- Variables

$S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t$

- Decomposition

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

- Parametric Forms

$$\mathbf{P}(S^0) \equiv \mathbf{G}(S^0, \mu, \sigma)$$
$$\mathbf{P}(S^i \mid S^{i-1}) \equiv \mathbf{G}(S^i, A \bullet S^{i-1}, Q)$$
$$\mathbf{P}(O^i \mid S^i) \equiv \mathbf{G}(O^i, H \bullet S^i, R)$$

Identification

- Learning from instances

Question

$$\mathbf{P}(S^t \mid O^0 \wedge \ldots\ldots \wedge O^t)$$

Bayesian Programming & ProBT®

Bayesian Program

Description

```
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 de 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);
```

Question

$$\mathbf{P}(S^t \mid O^0 \wedge \ldots\ldots \wedge O^t)$$

Bayesian Programming & ProBT®

Bayesian Program

Description

Question

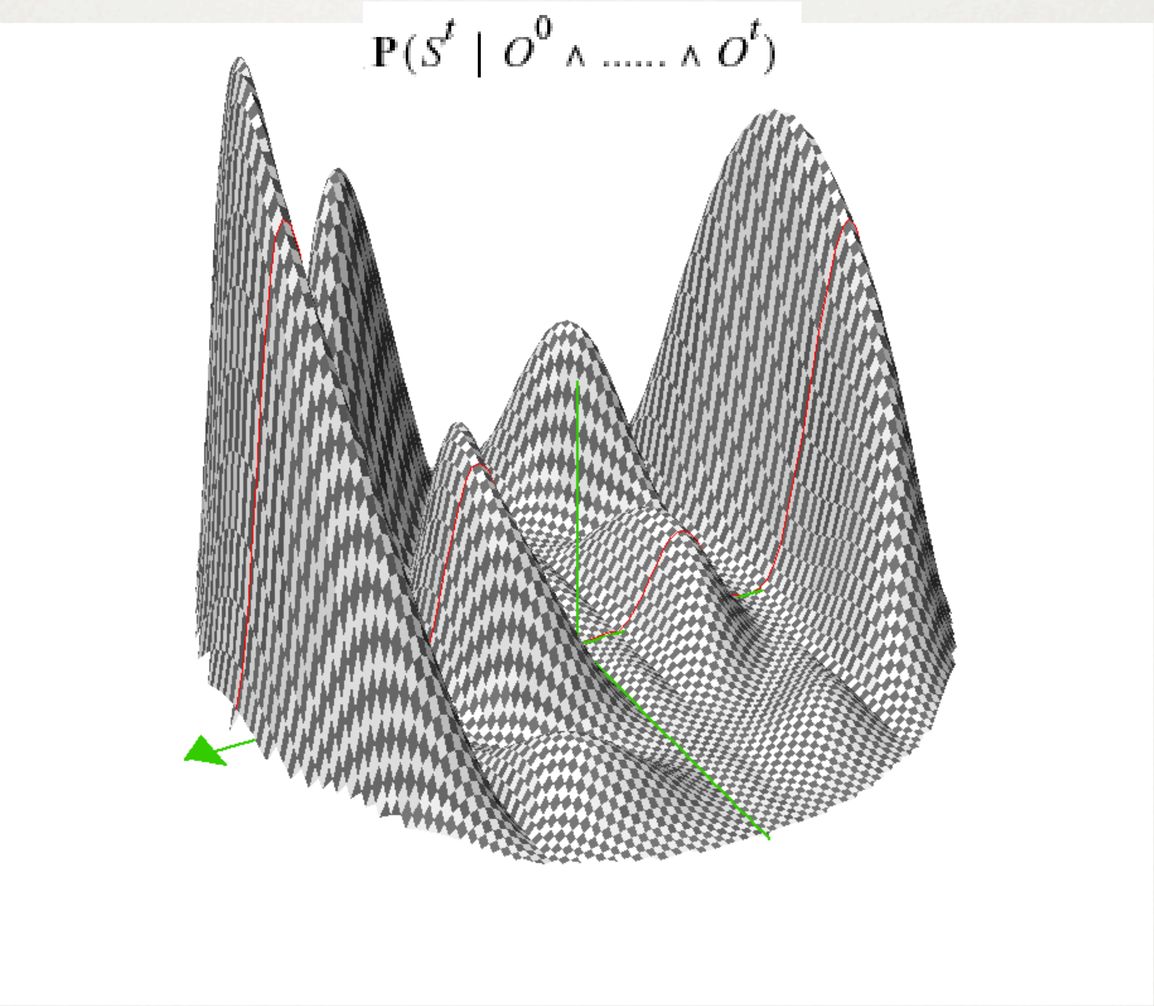
```
main ()
{
  //Variables
  plFloat read_time;
  plIntegerType id_type(0,1);
  plFloat times[5] = {1,2,3,5,10};
  plSparseType time_type(5,times);
  plSymbol id("id",id_type);
  plSymbol time("time",time_type);

  //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 de 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);

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

  //Read a time from the key board
  cout<<"P(id,time)= "<<Pid_t<<"\n";
  cout<<"Time? : ";
  cin>>read_time;

  //Getting P(id | time = read_time)
  plKernel Pid_readTime;
```

Bayesian Programming & ProBT®

Bayesian Program

Description

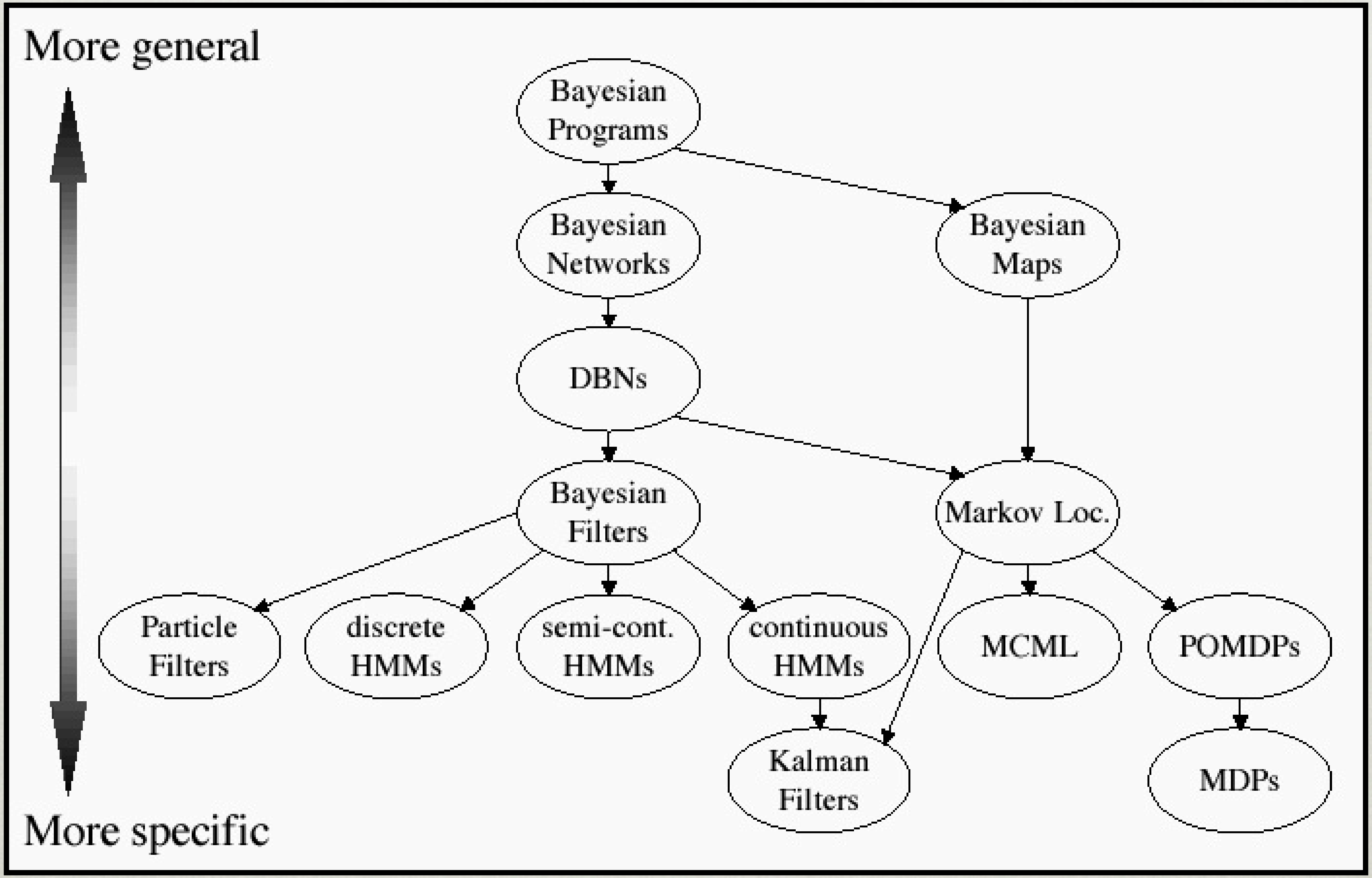
Question



ProBAYES.com

Bayesian-Programming.org

Bayesian Programming
Related formalisms

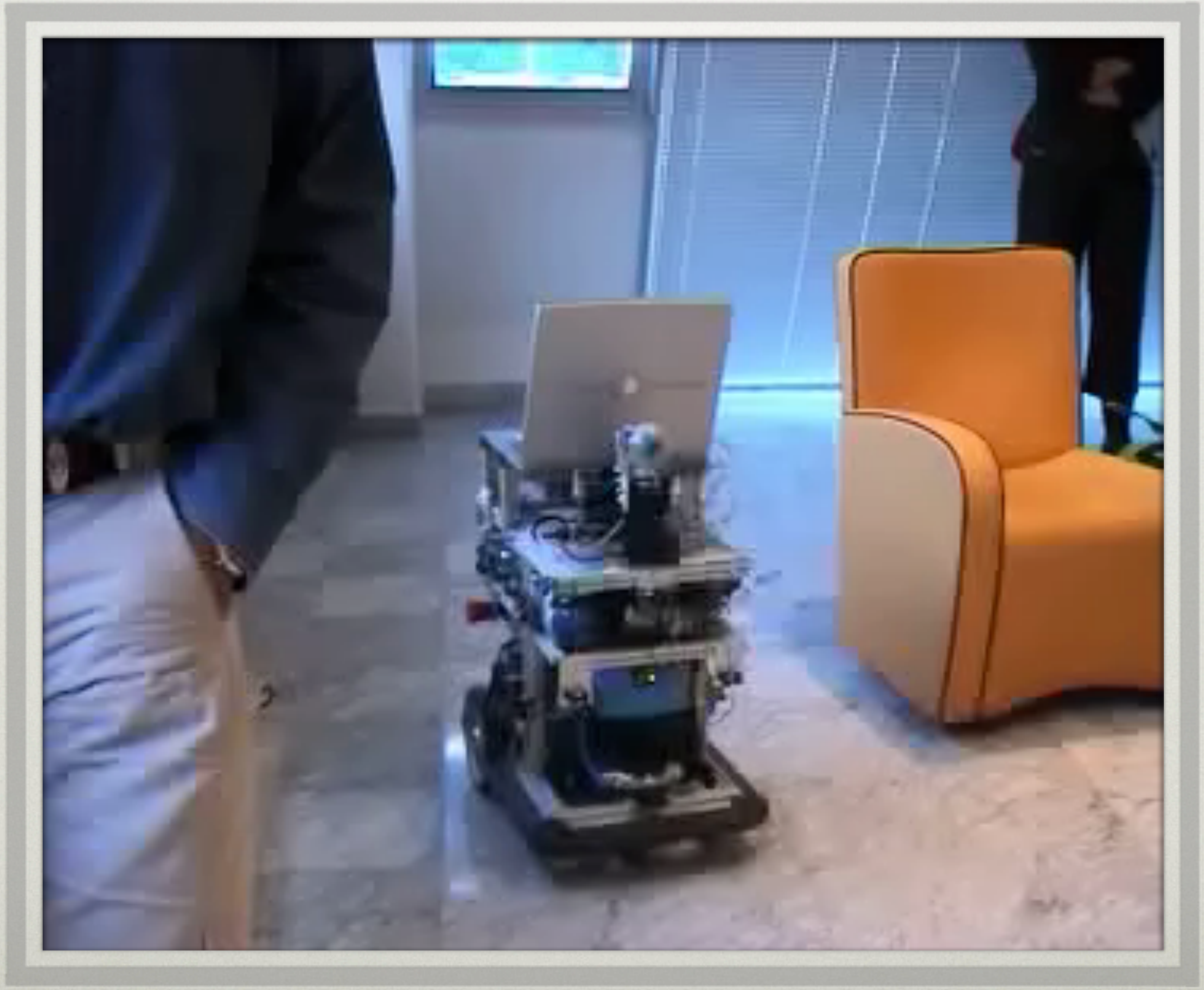


Overview

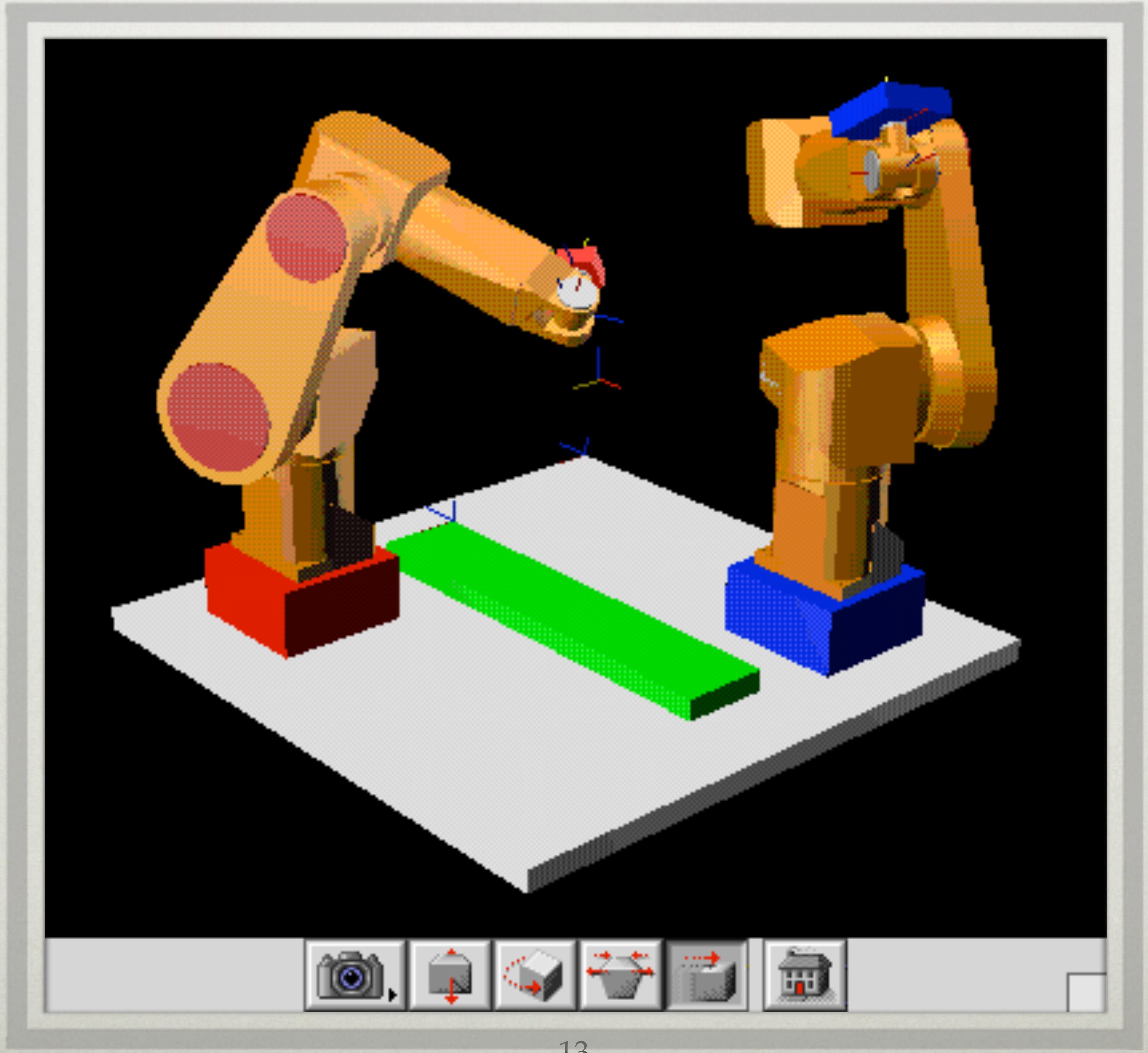
- How to survive (perceive, reason, learn, decide and act) with incomplete information ?
  - Probability as an alternative to logic
    - How to develop better artifacts using Bayesian reasoning?
    - Biological plausibility of Bayesian reasoning at a macroscopic level?
    - Biological plausibility of Bayesian reasoning at a microscopic level?

Olivier Lebeltel's Ph.D

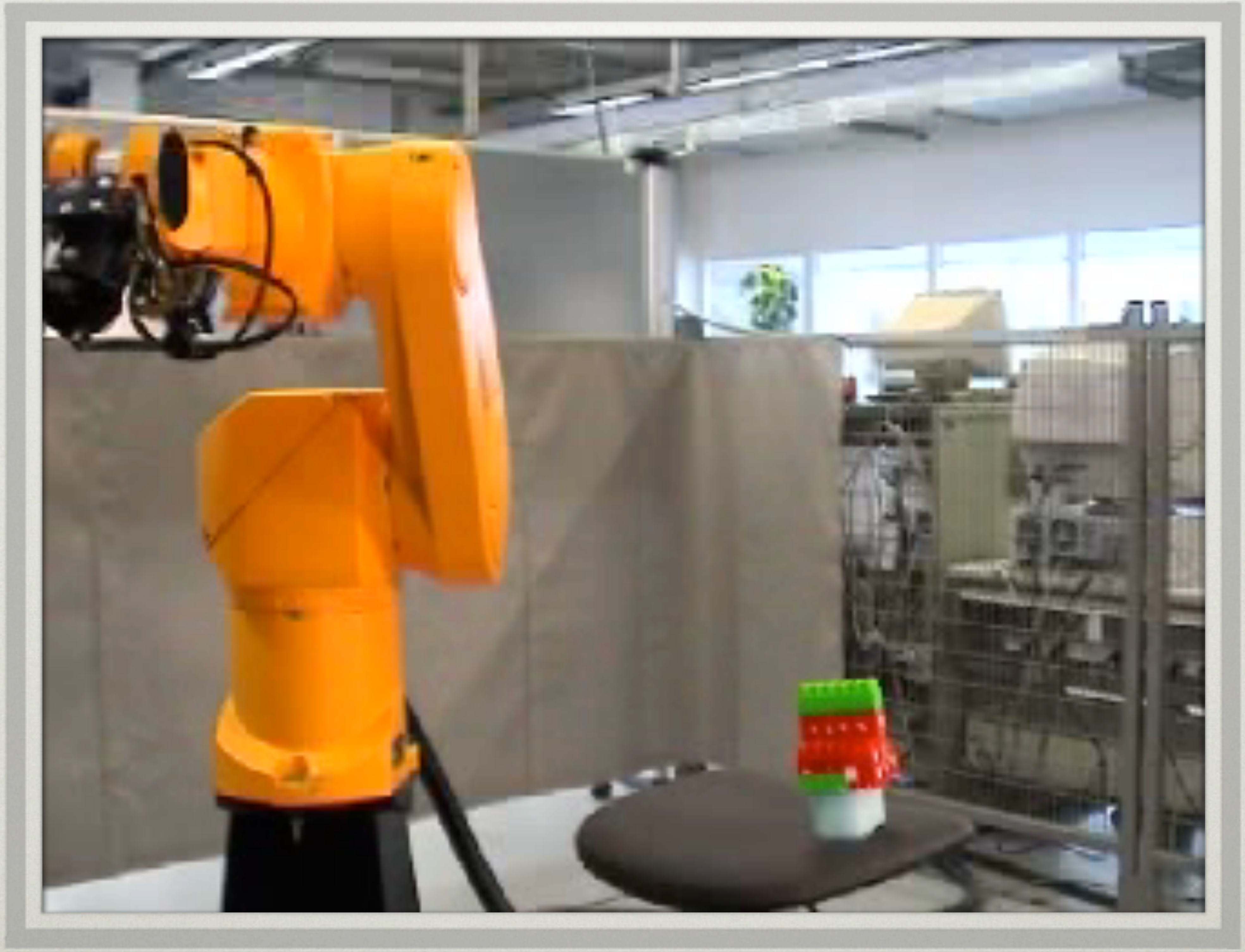
Carla Koike's Ph.D



Kamel Mekhnacha's Ph.D



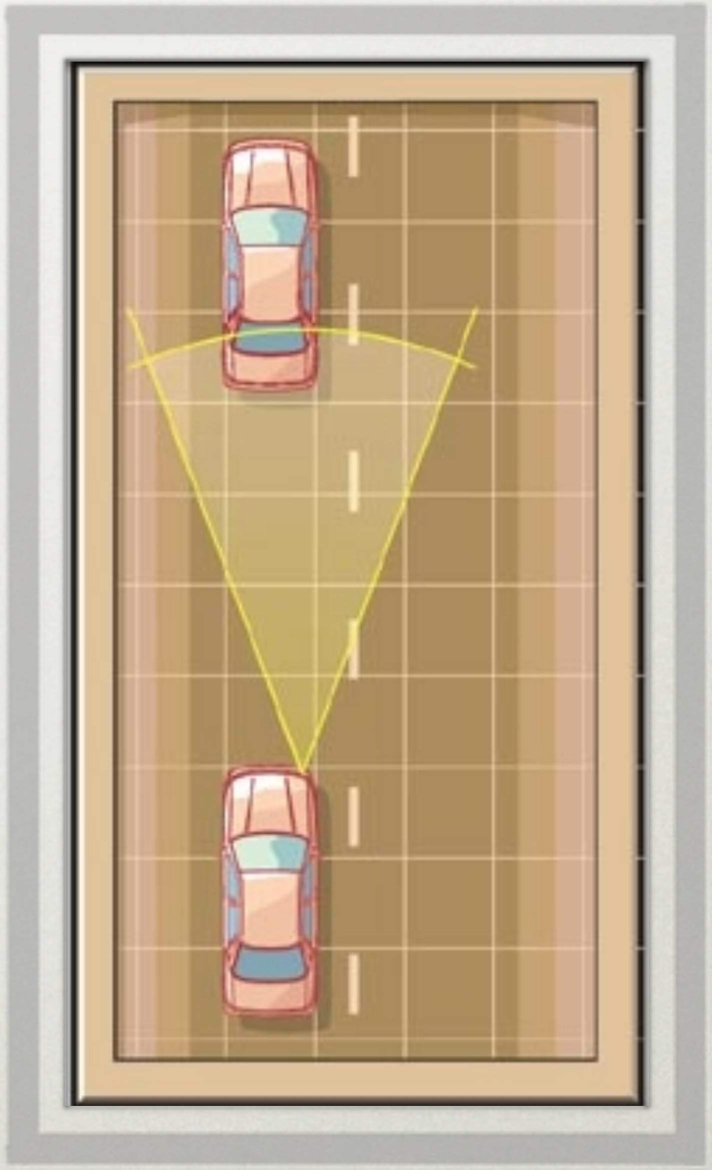
Ruben Garcia's Ph.D



Ronan Le Hy's Ph.D

Bayesian Occupancy Filter (BOF) for Avanced 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)

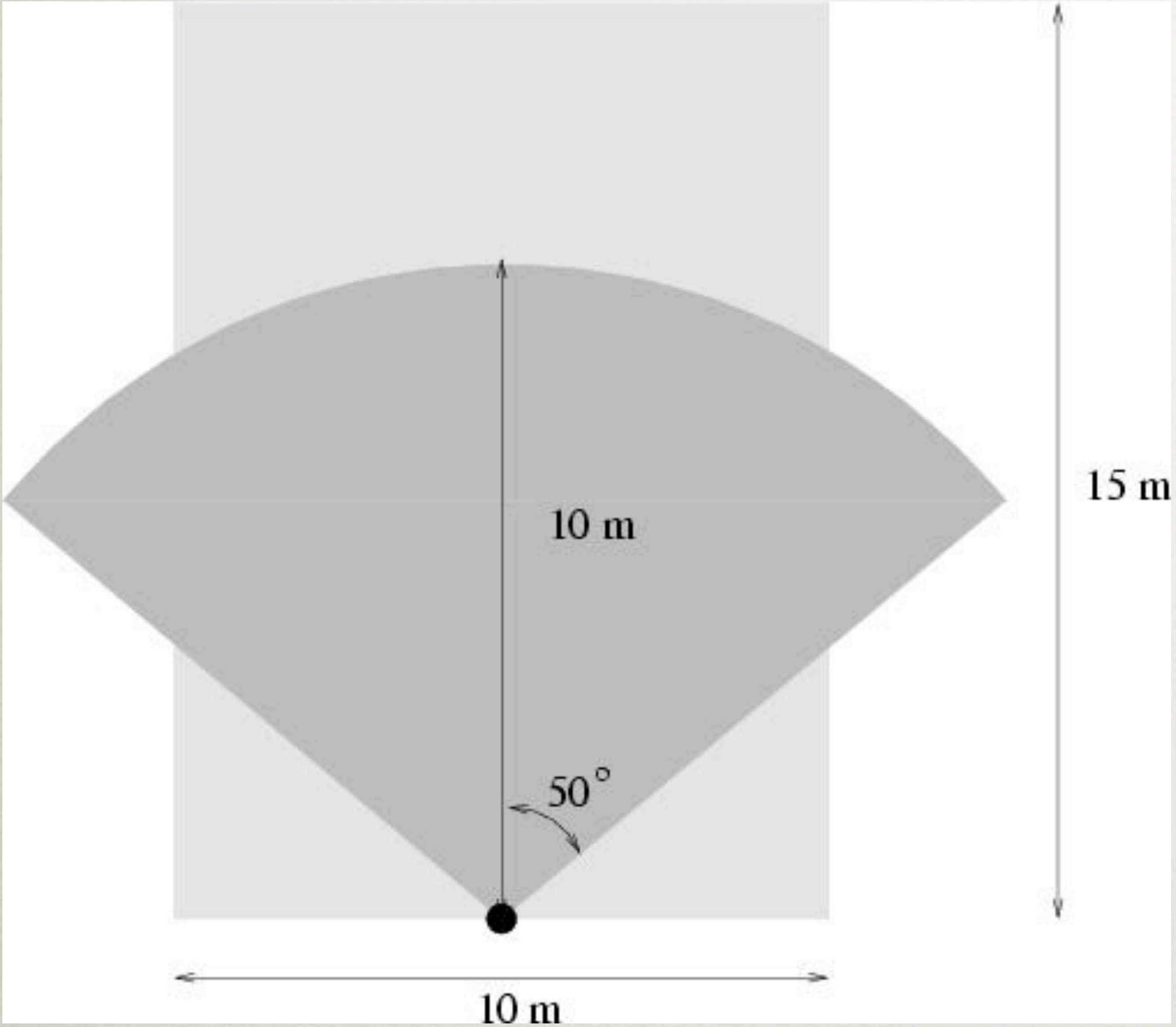


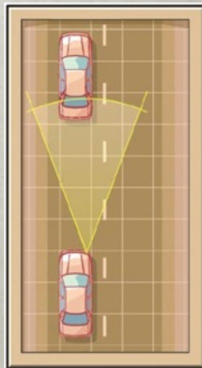
PhD thesis of Christophe Coué

Coué, C., Pradalier, C., Laugier, C., Fraichard, T. & Bessière, P. (2006) Bayesian Programming multi-target tracking: an automotive application; *IJRR (International Journal of Robotic Research);* Vol. 25, N° 1, pp. 19-30

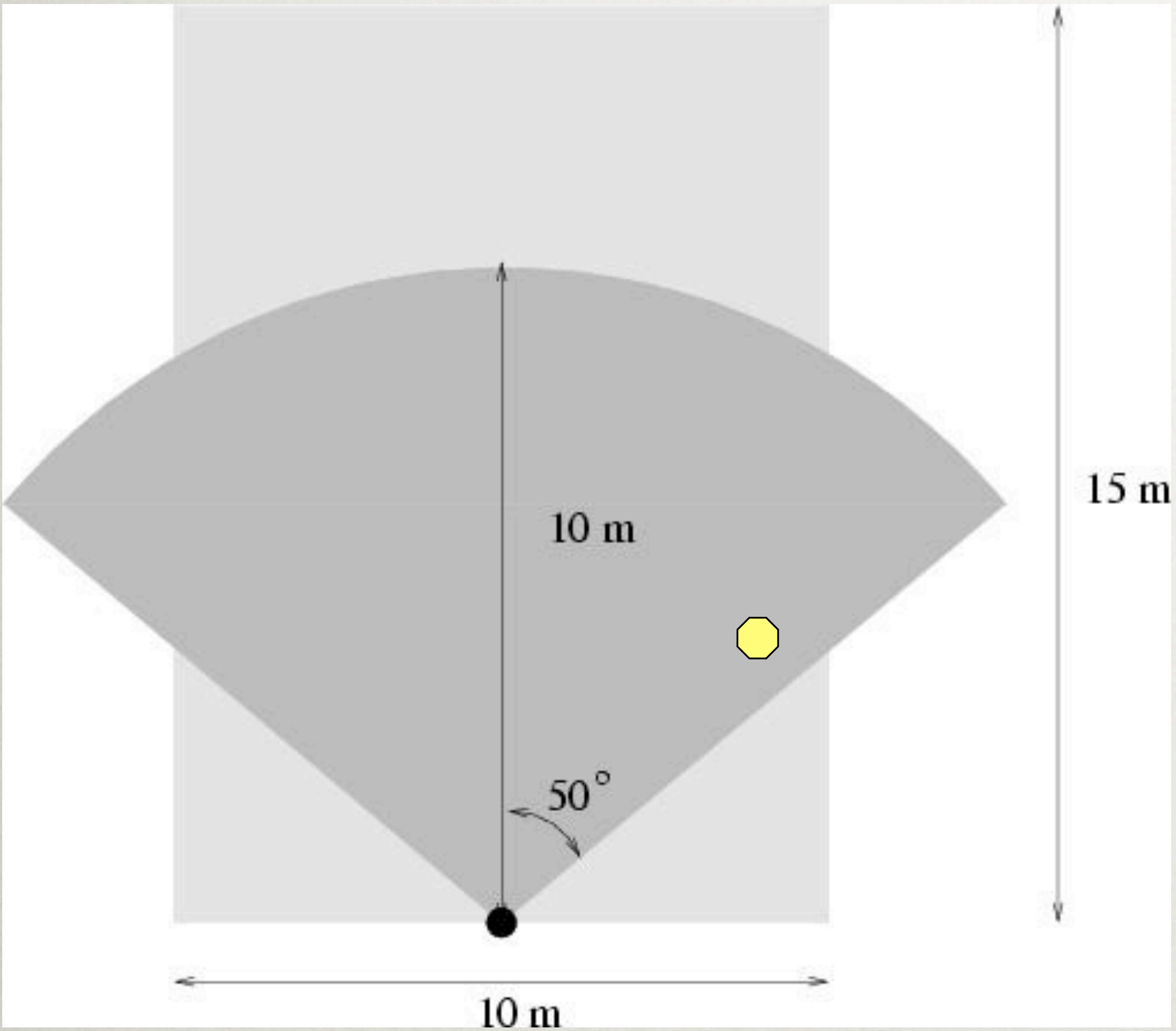
Coué, C. (2003) *Fusion d'information capteur pour l'aide à la conduite automobile*; PhD thesis, INPG

1 sensor - 1 object

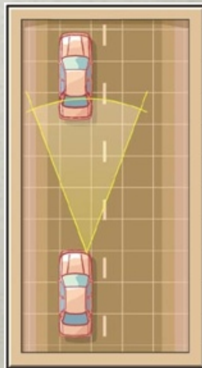




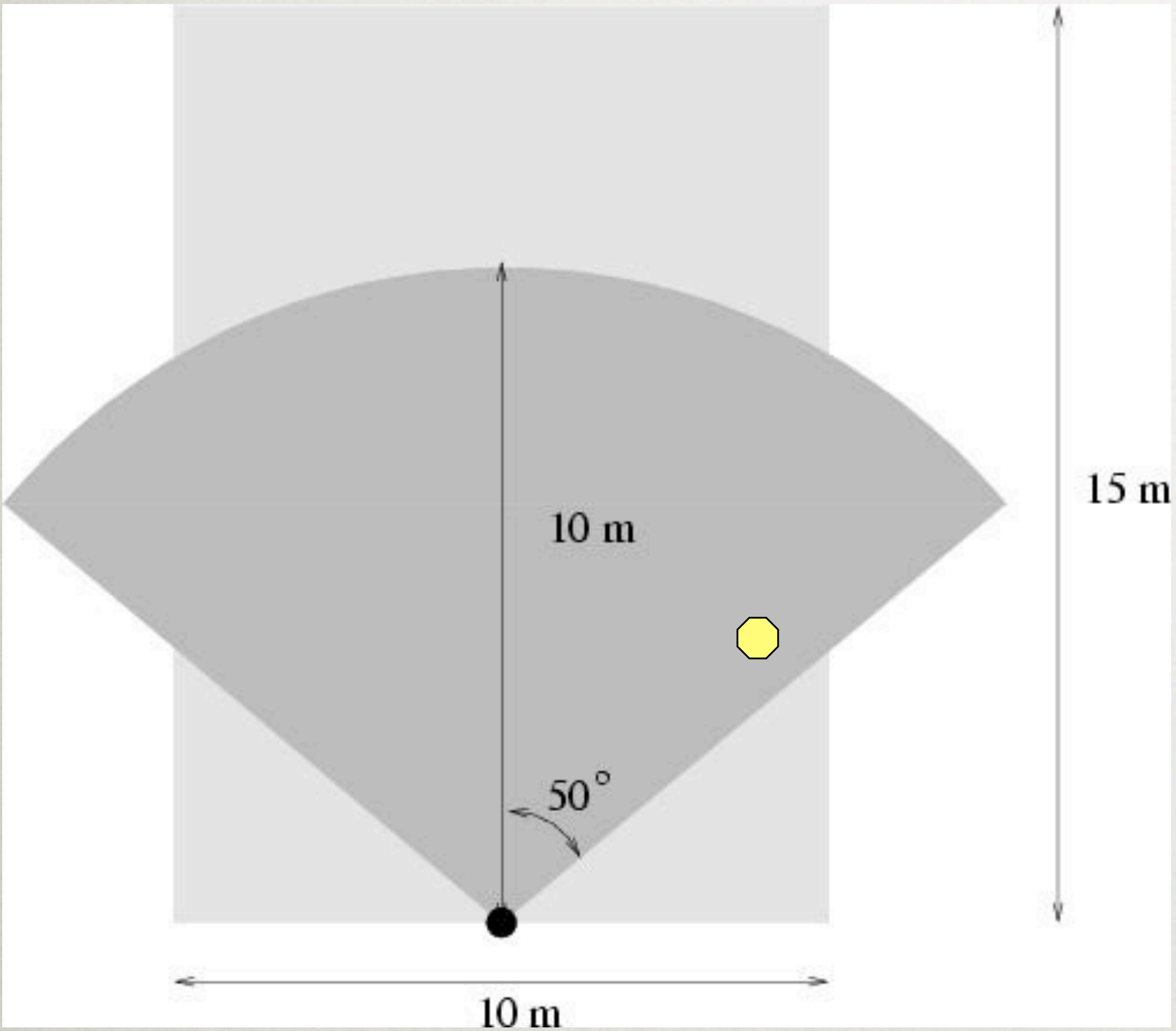
1 SENSOR - 1 OBJECT

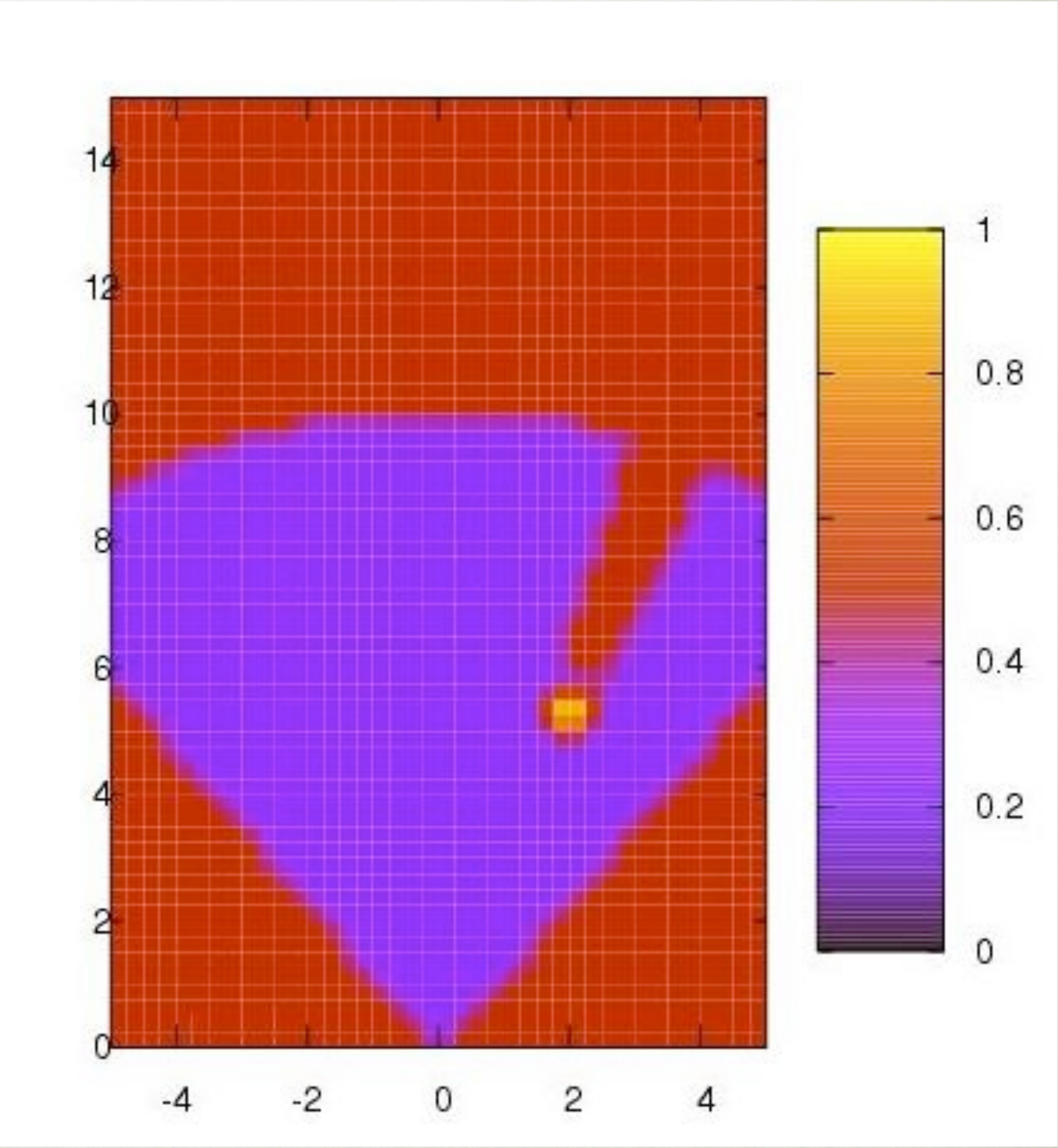


$$z = (5, 2, 0, 0)$$



1 sensor - 1 object

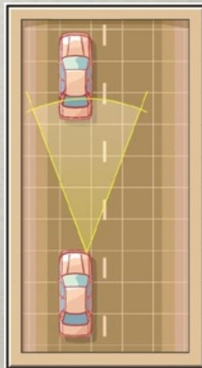




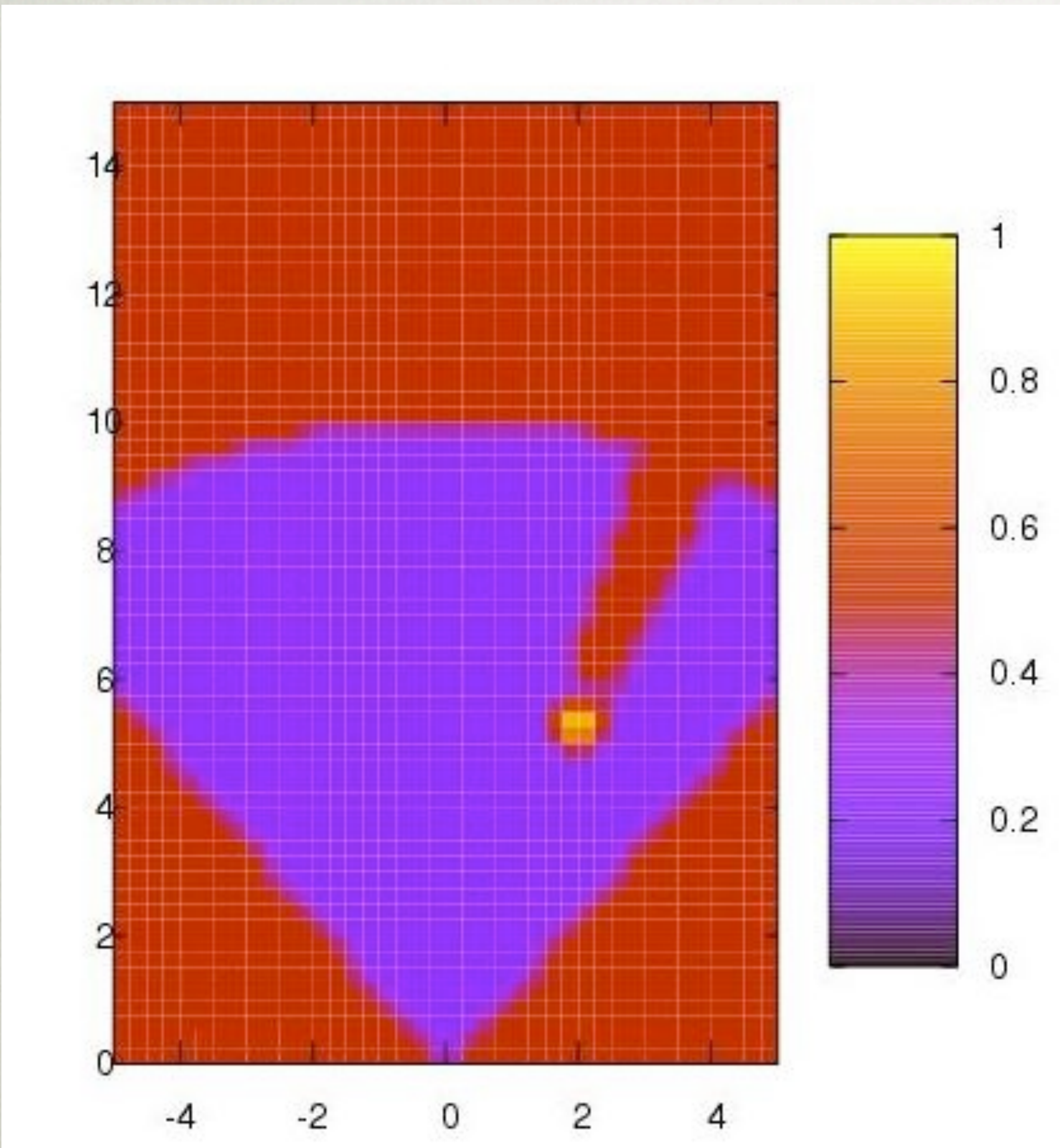
$z = (5, 2, 0, 0)$

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

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

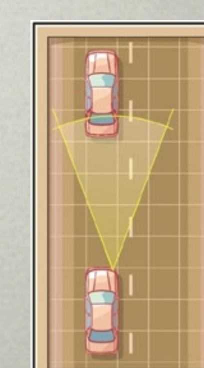


1 sensor - 1 object

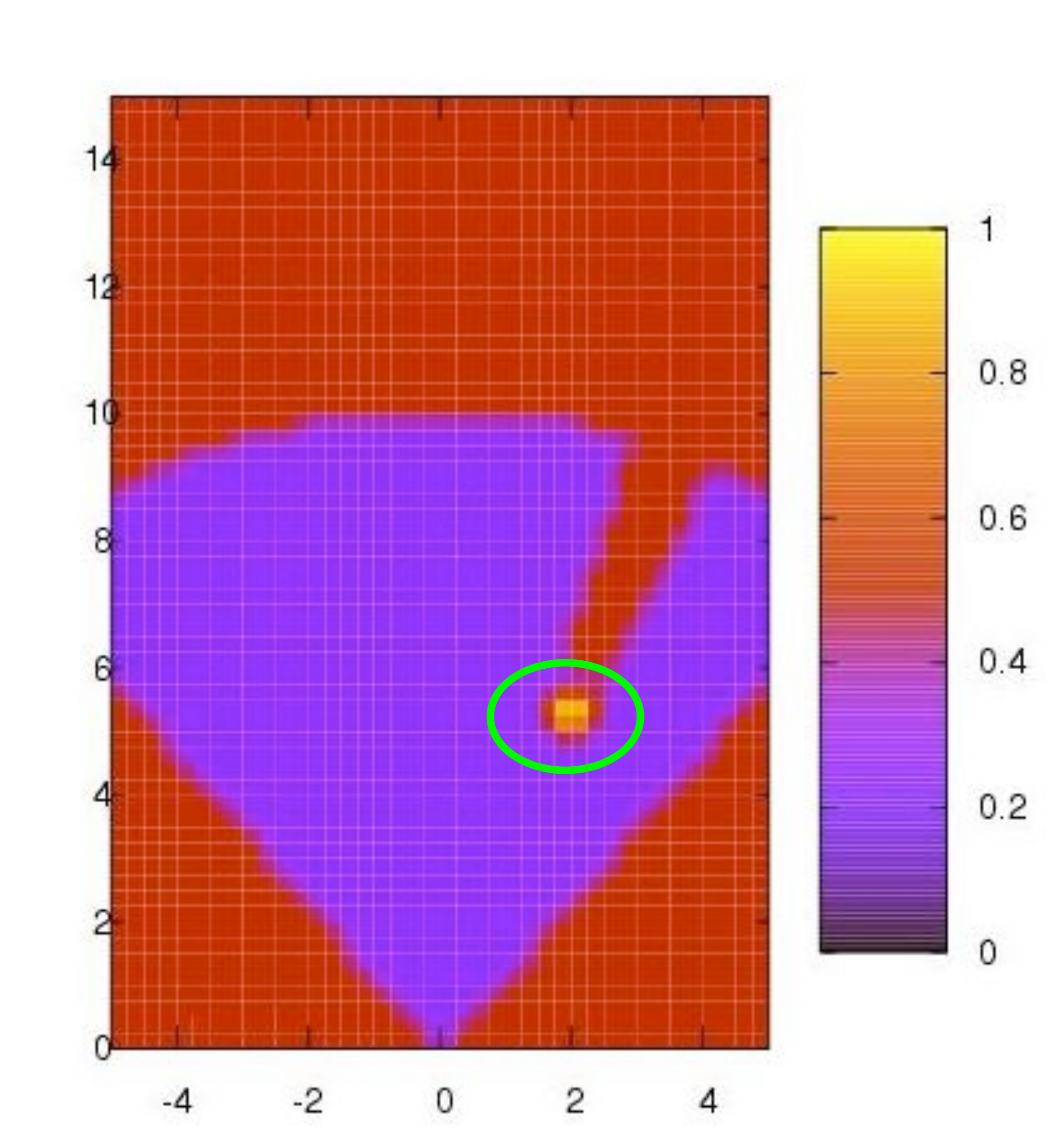


$$P(\,[E_c=1] \mid z\ c)$$

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

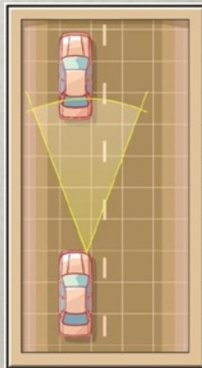


1 sensor - 1 object

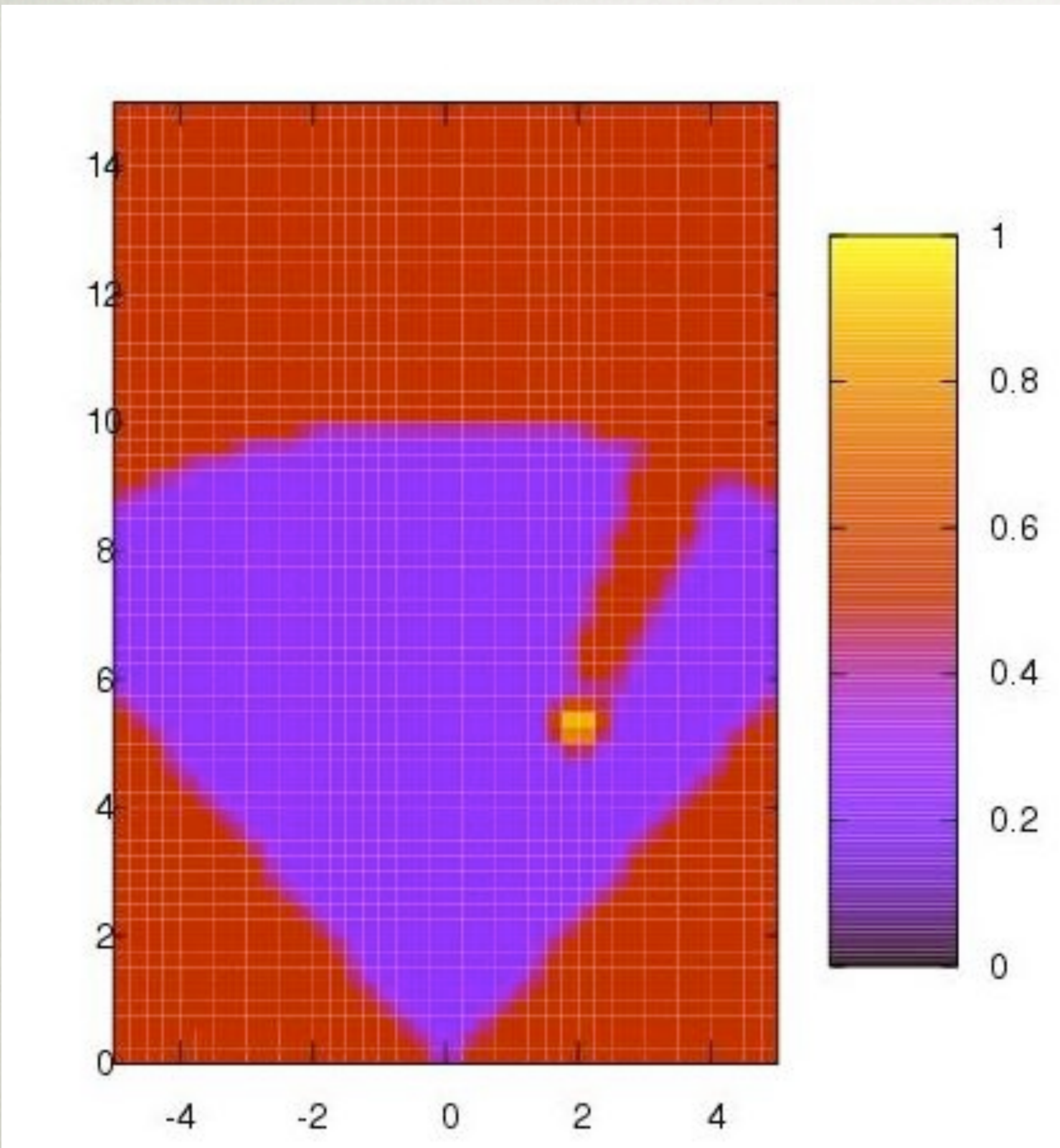


- Occupied space

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

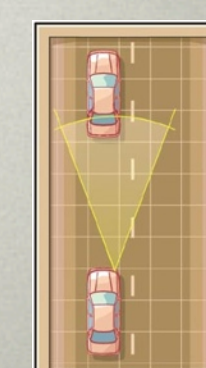


1 sensor - 1 object

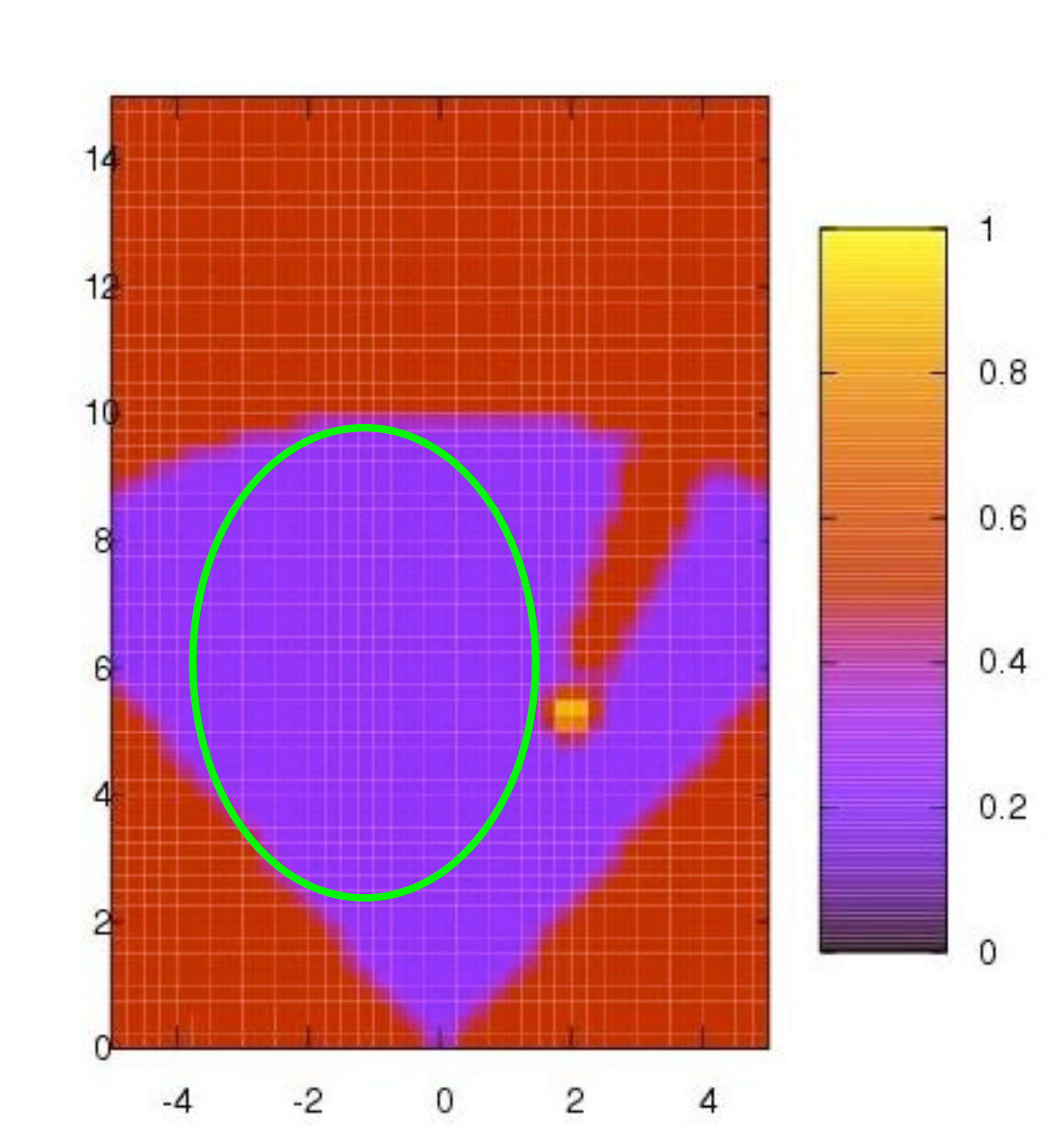


- Occupied space

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

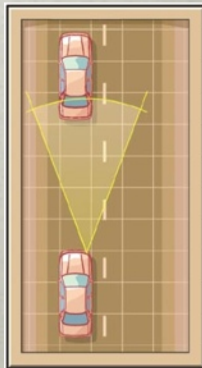


1 sensor - 1 object

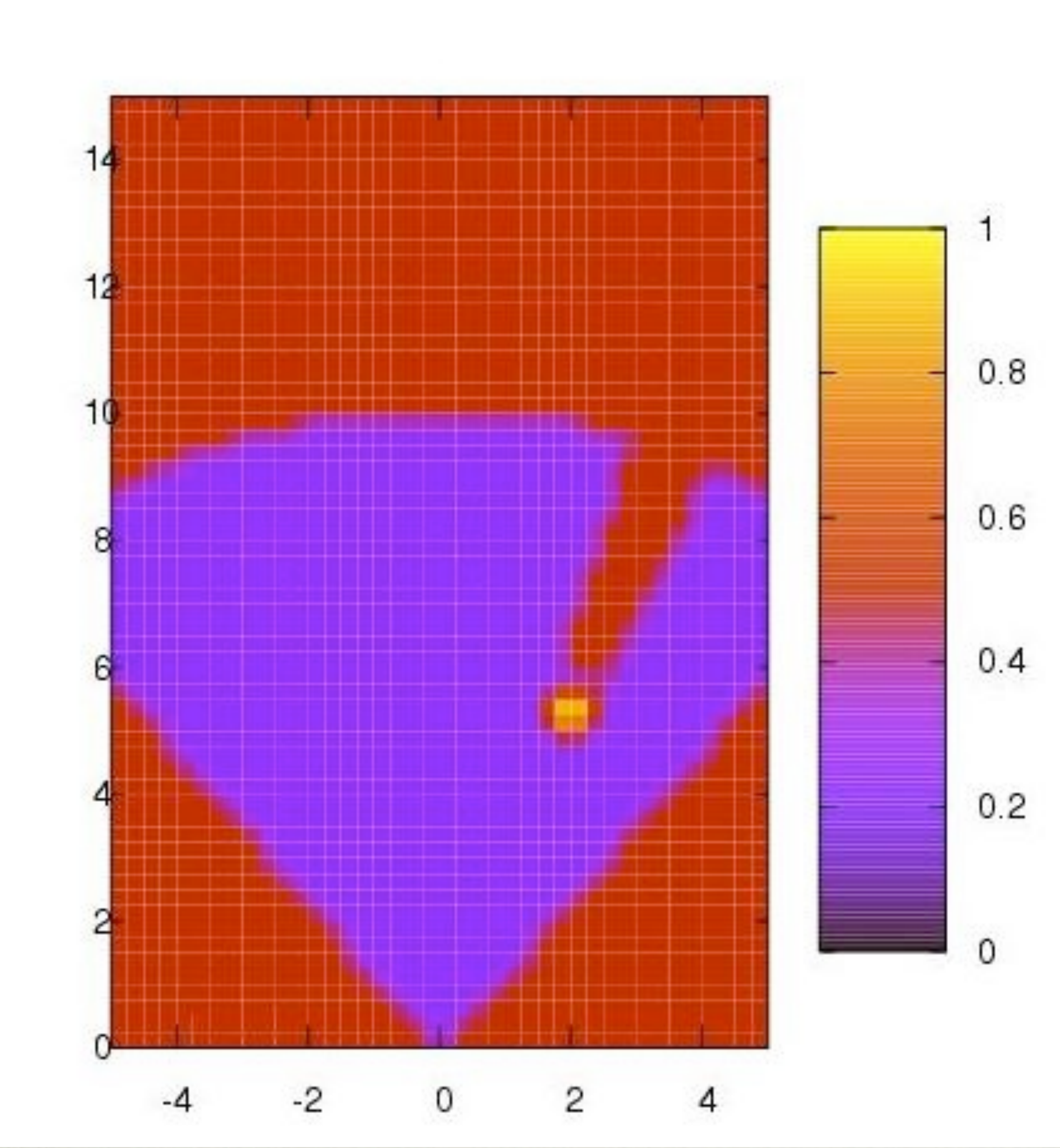


- Occupied space
- Free space

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



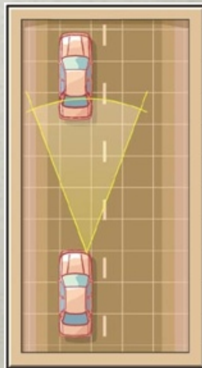
1 sensor - 1 object



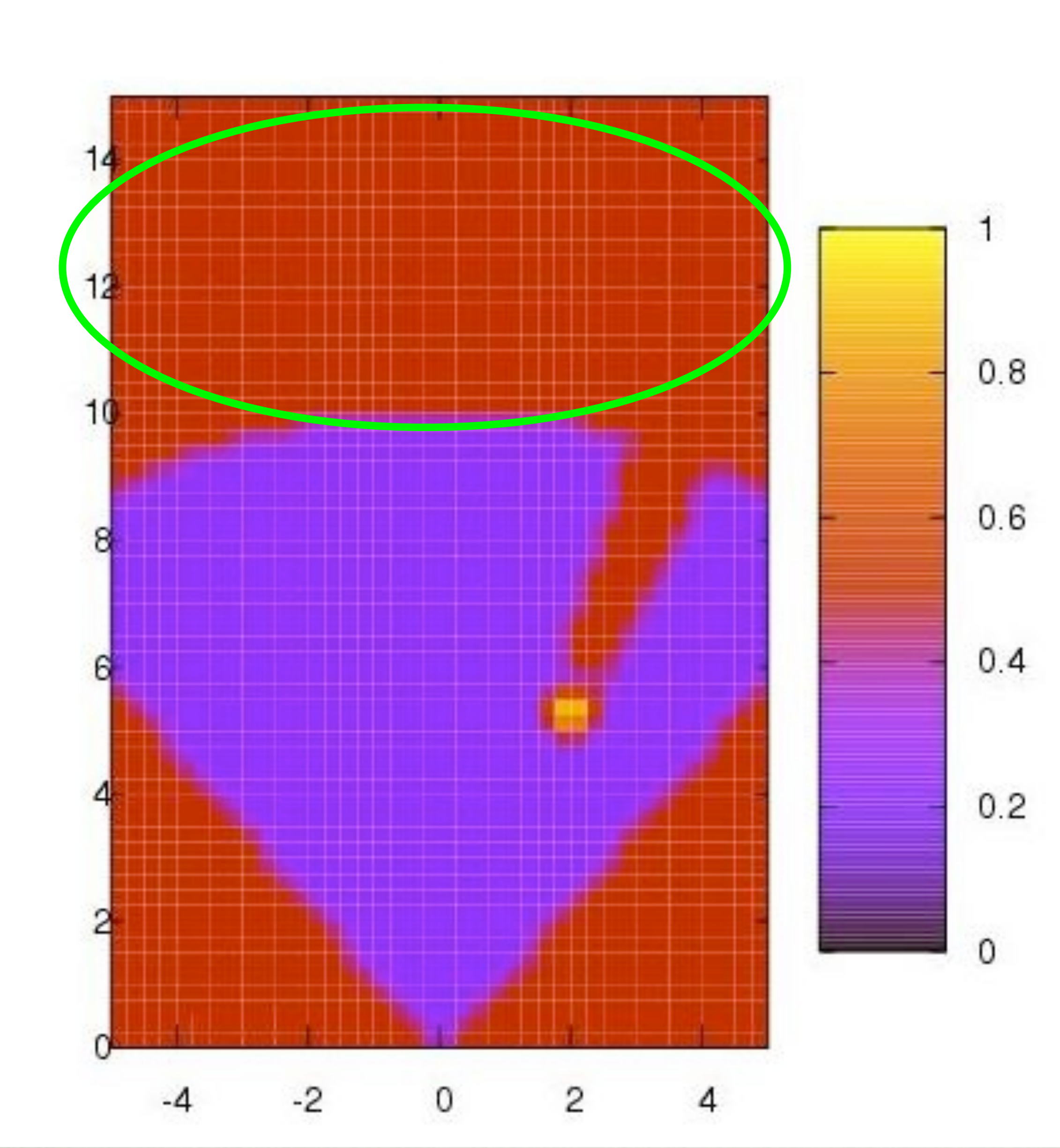
$P([E_c=1] \mid z\ c)$

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

- Occupied space
- Free space



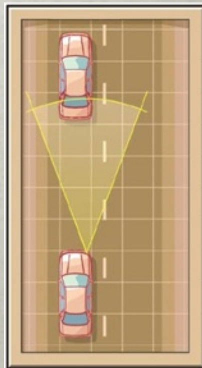
1 sensor - 1 object



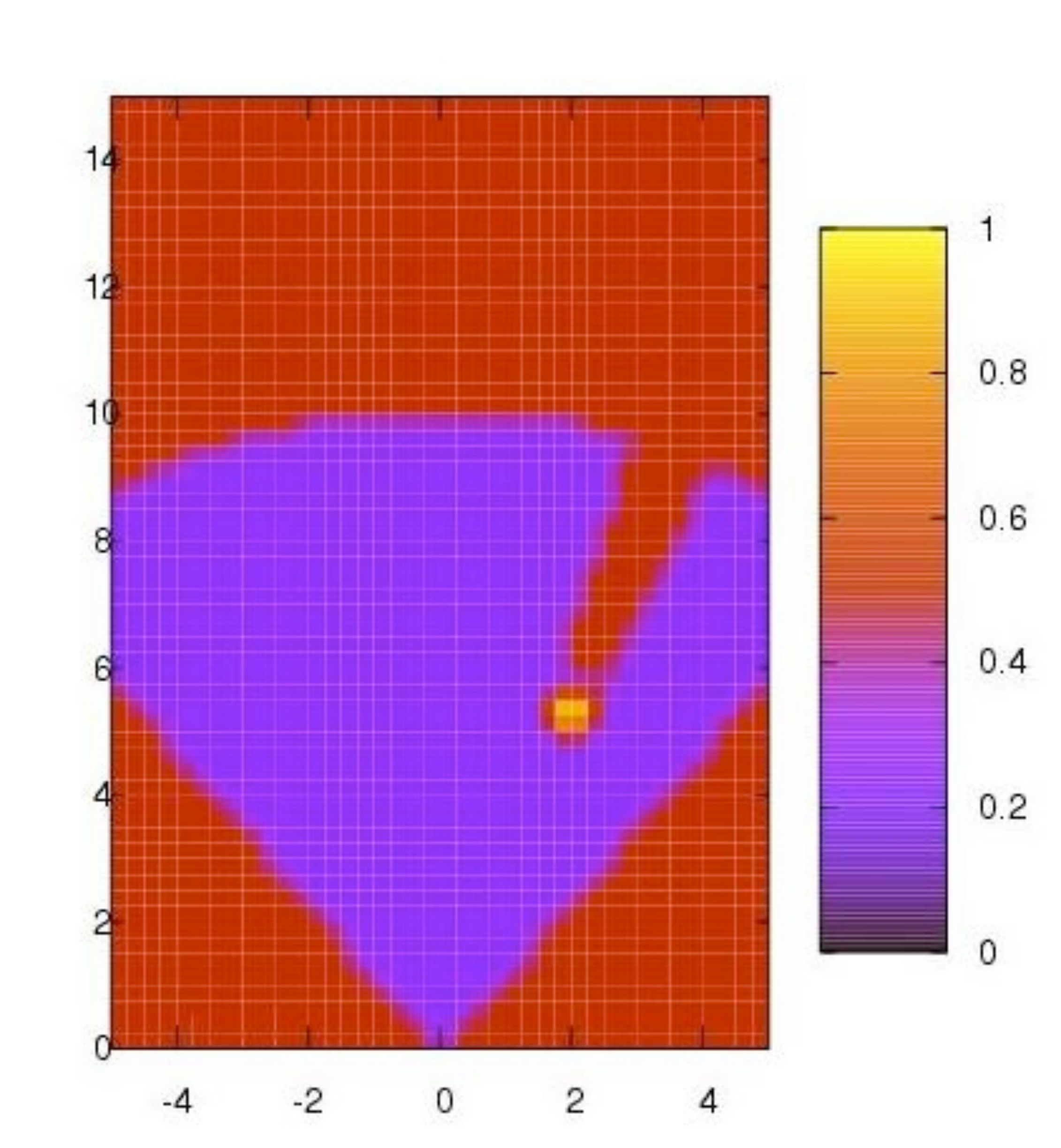
- Occupied space
- Free space
- Nonobservable space

P(*[E$_c$=1] | z c*)

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

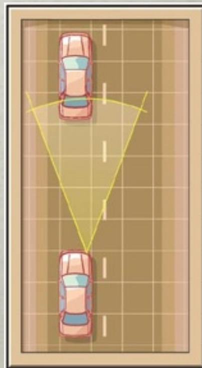


1 sensor - 1 object

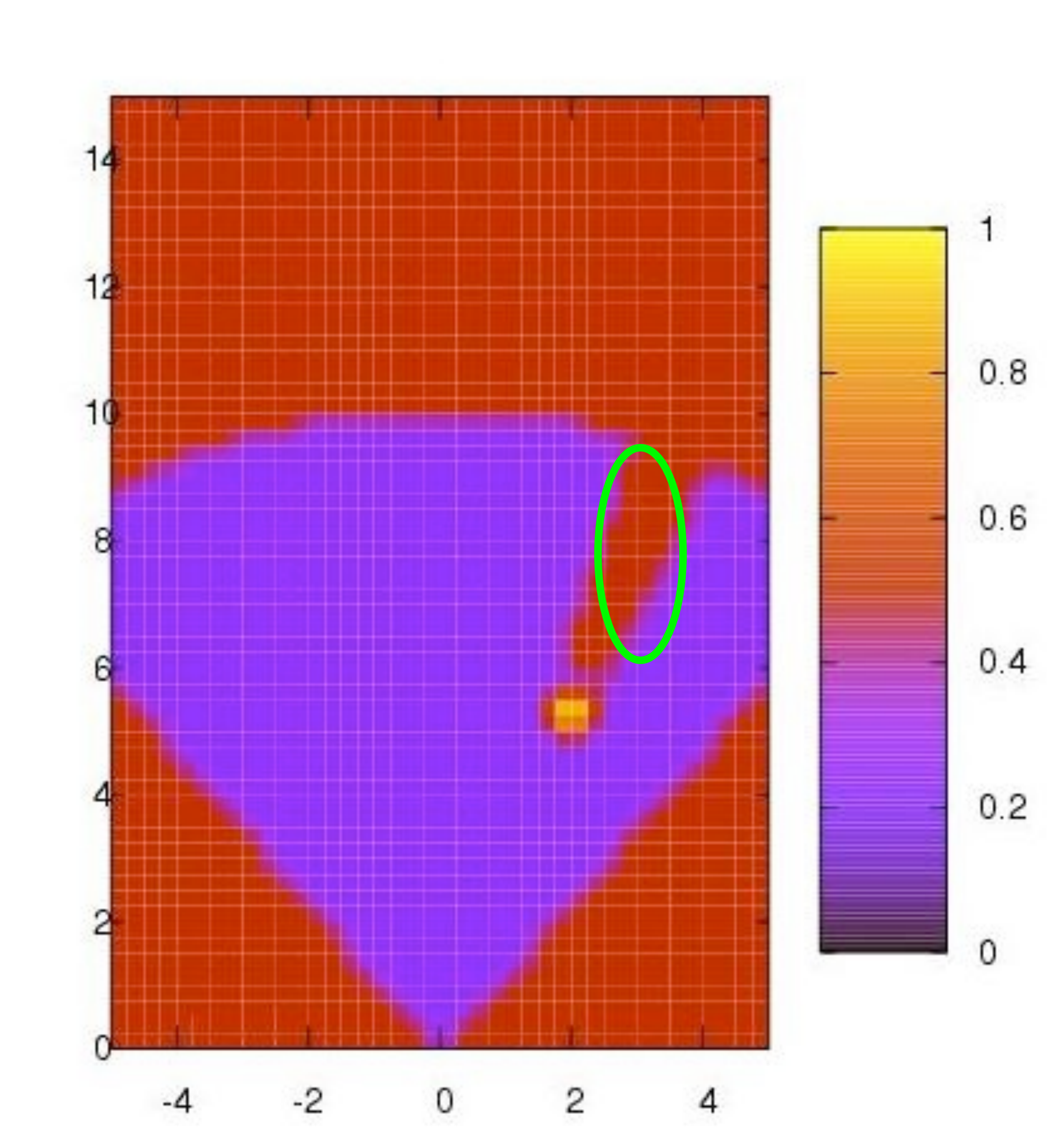


$P([E_c=1] \mid z\ c)$
$c = [x, y, 0, 0]$

- Occupied space
- Free space
- Nonobservable space

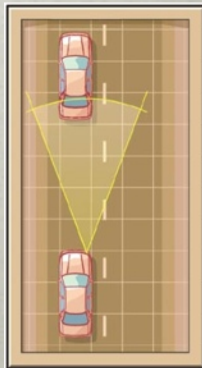


1 SENSOR - 1 OBJECT

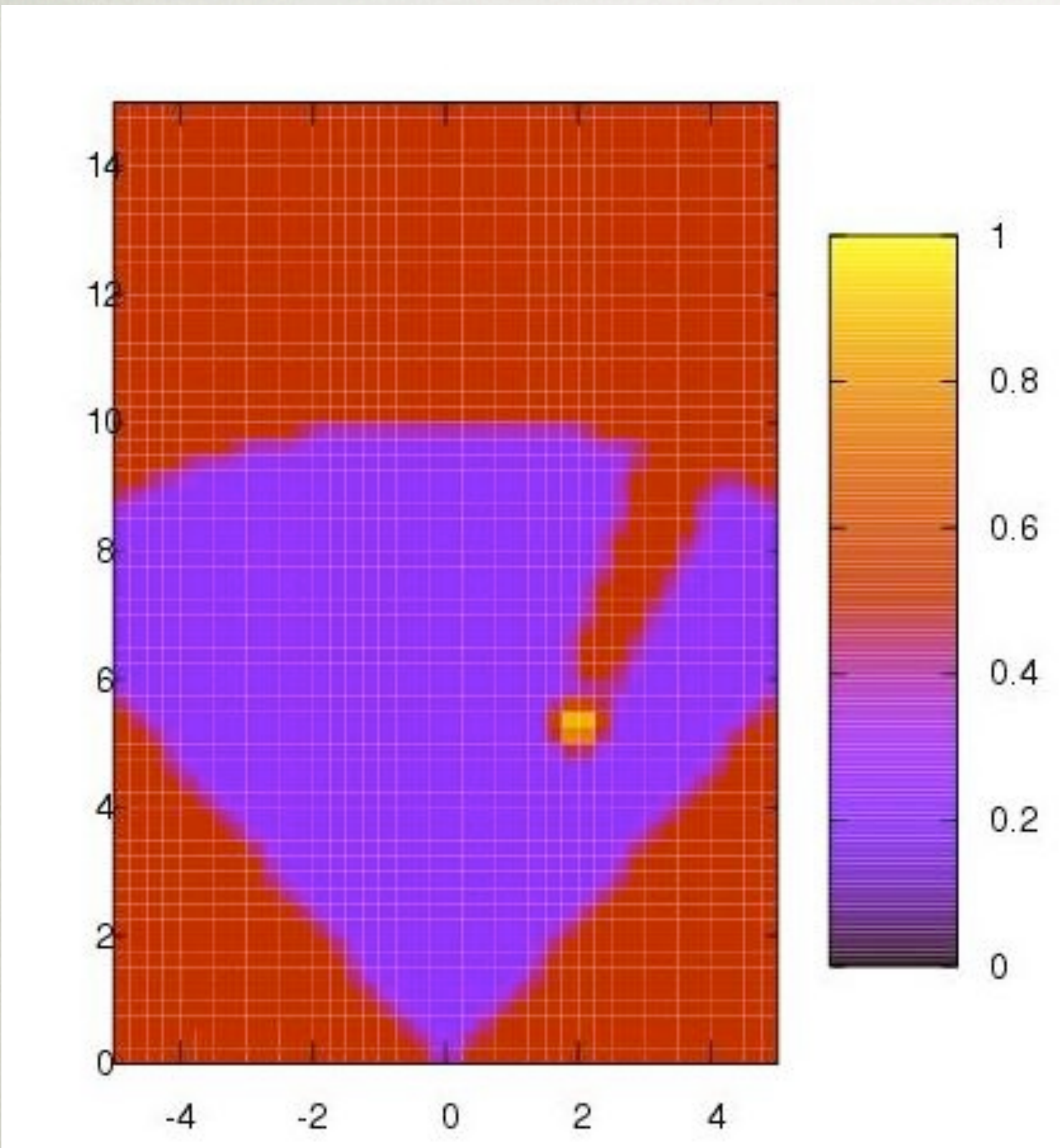


- Occupied space
- Free space
- Nonobservable space
- Occultated space

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

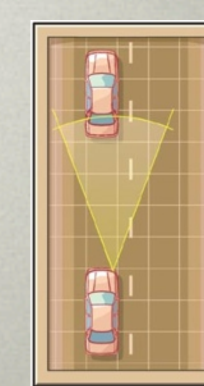


1 sensor - 1 object

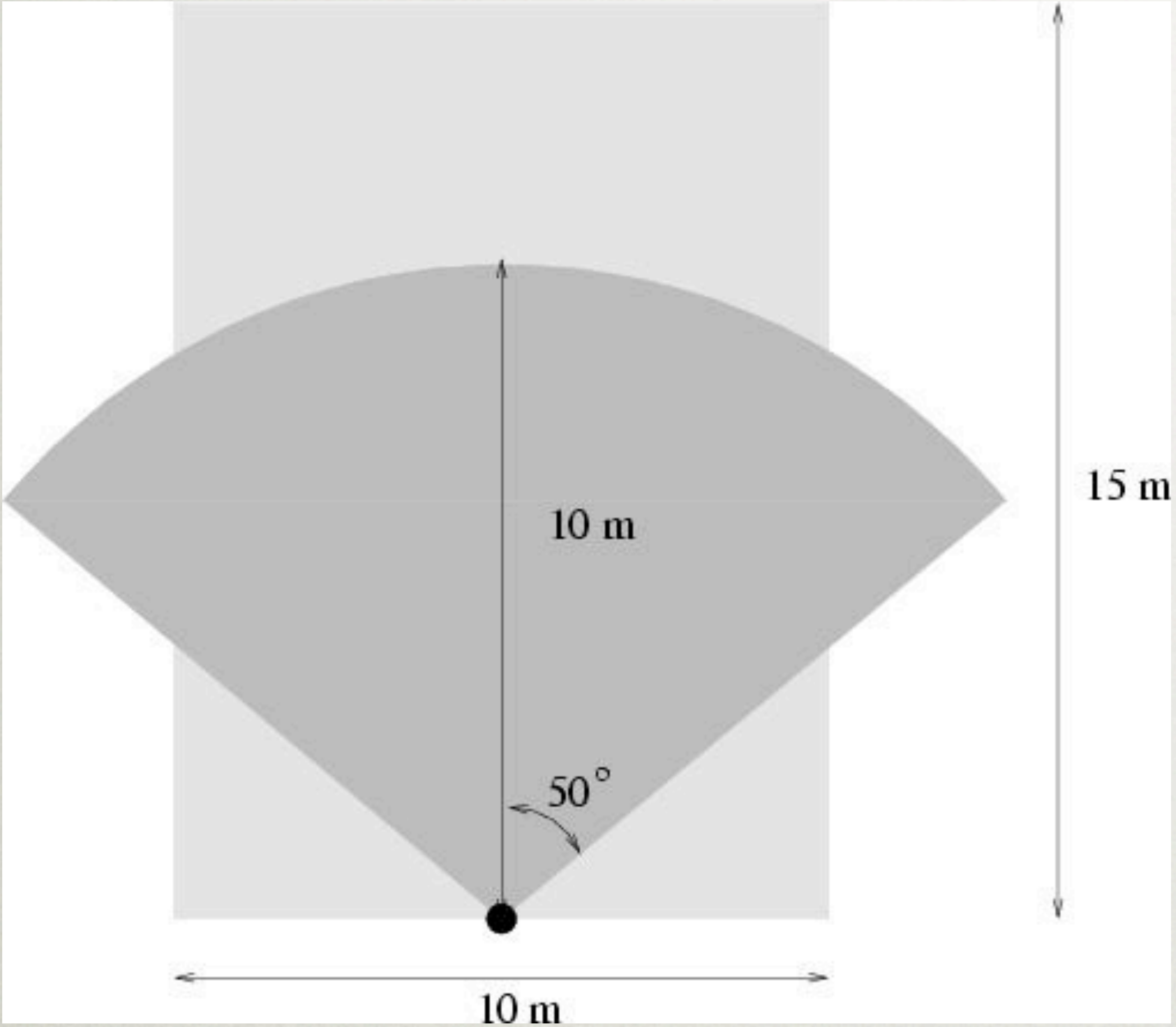


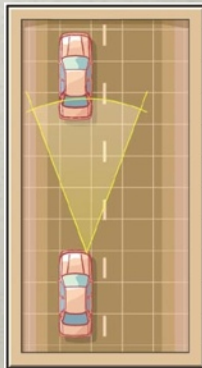
$$P(\ [E_c=1]\ |\ z\ c)$$
$$c = [x, y, 0, 0]$$

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

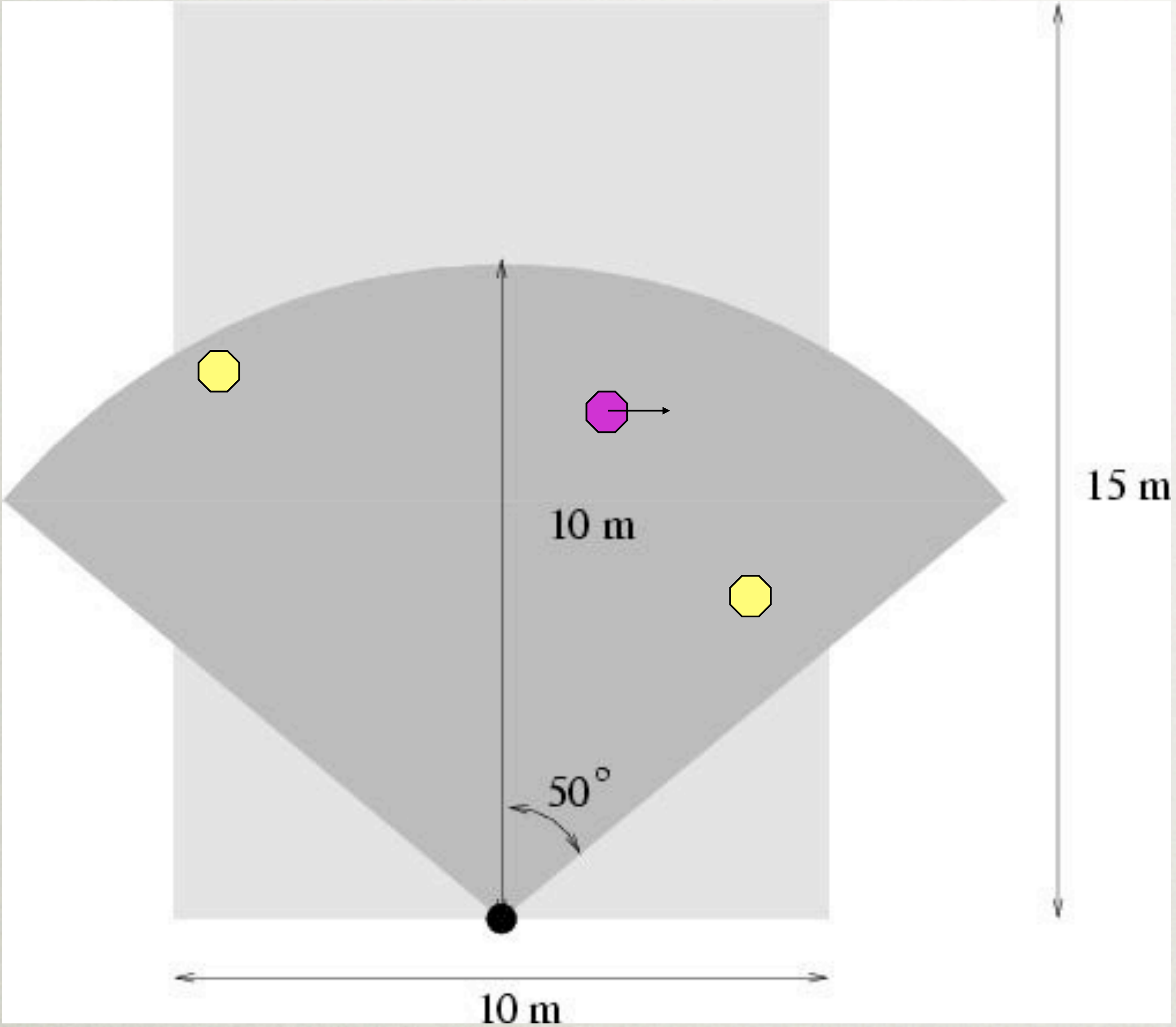


1 sensor - multiple target





1 SENSOR - MULTIPLE TARGET

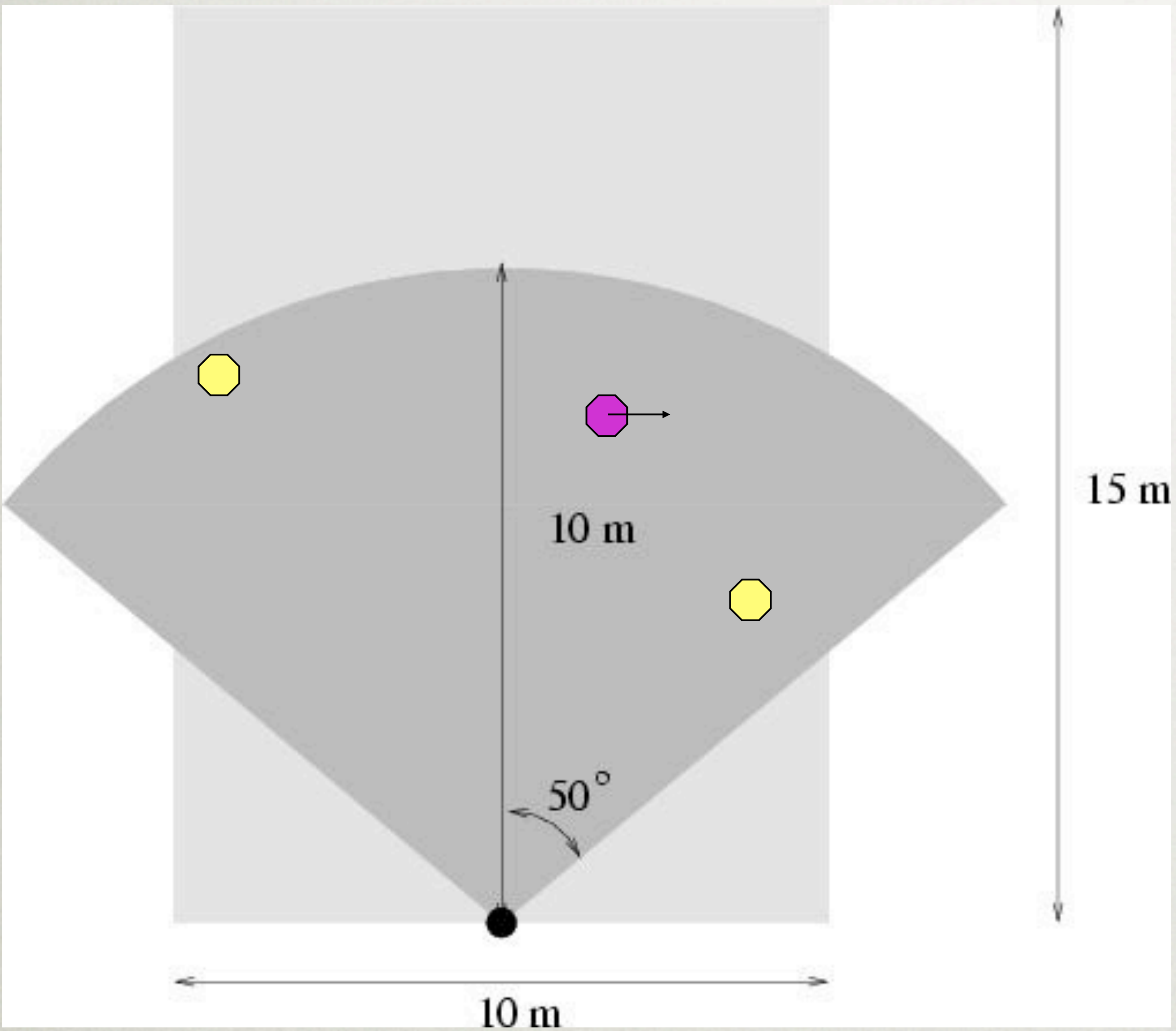


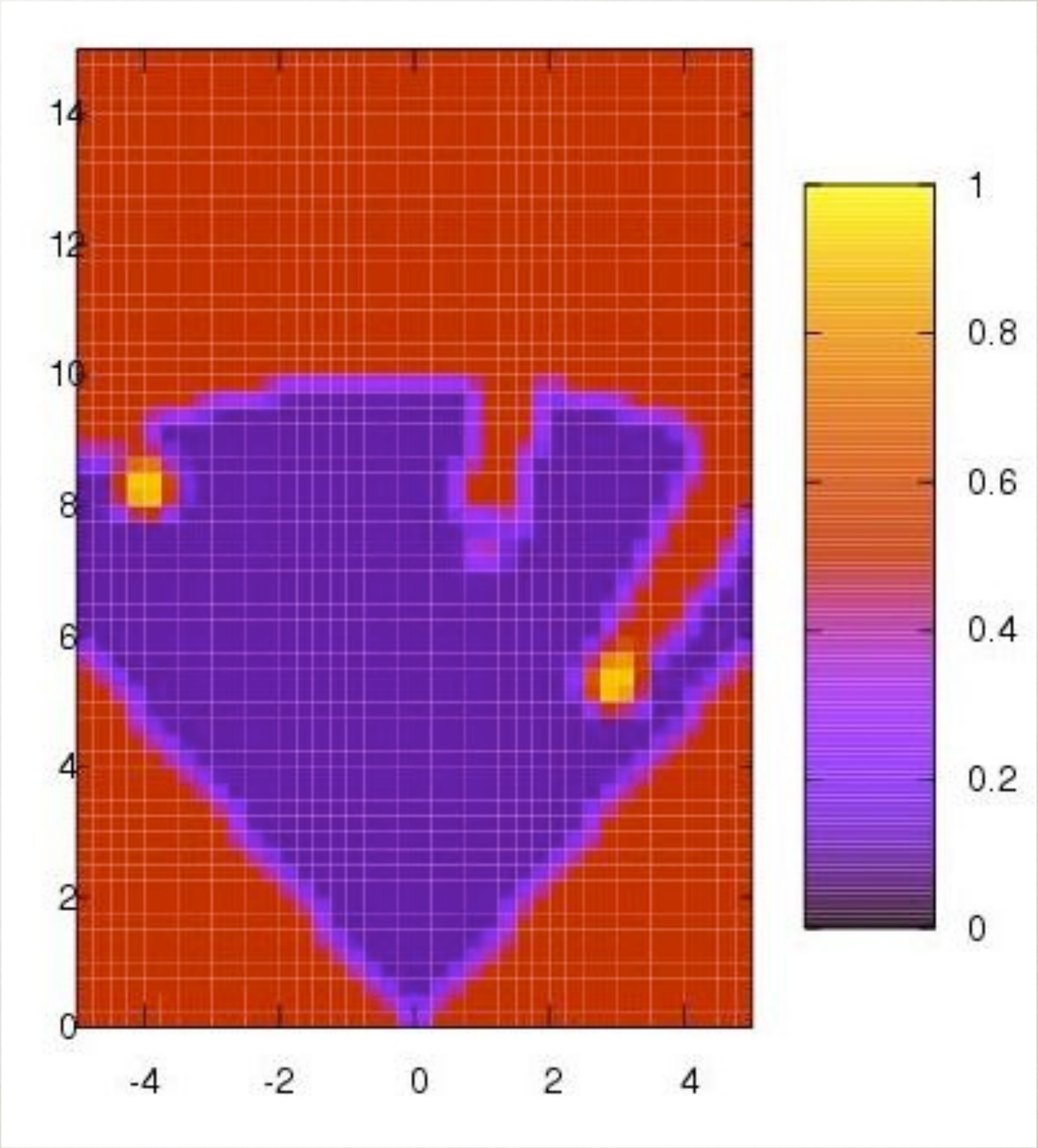
$$z_1 = (\ 8.3, -4, 0, 0)$$

$$z_2 = (\ 7.3, 1.9, 0, 0.8)$$

$$z_3 = (\ 5, 3, 0, 0)$$

1 sensor - multiple target





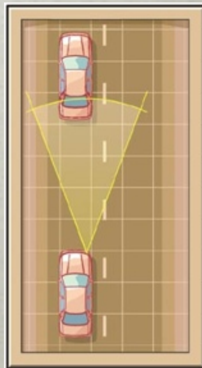
$z_1 = (\ 8.3, -4, 0, 0)$

$z_2 = (\ 7.3, 1.9, 0, 0.8)$

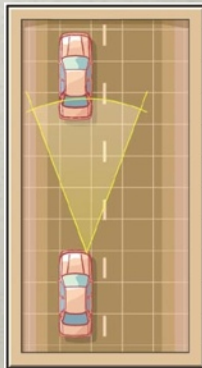
$z_3 = (\ 5, 3, 0, 0)$

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

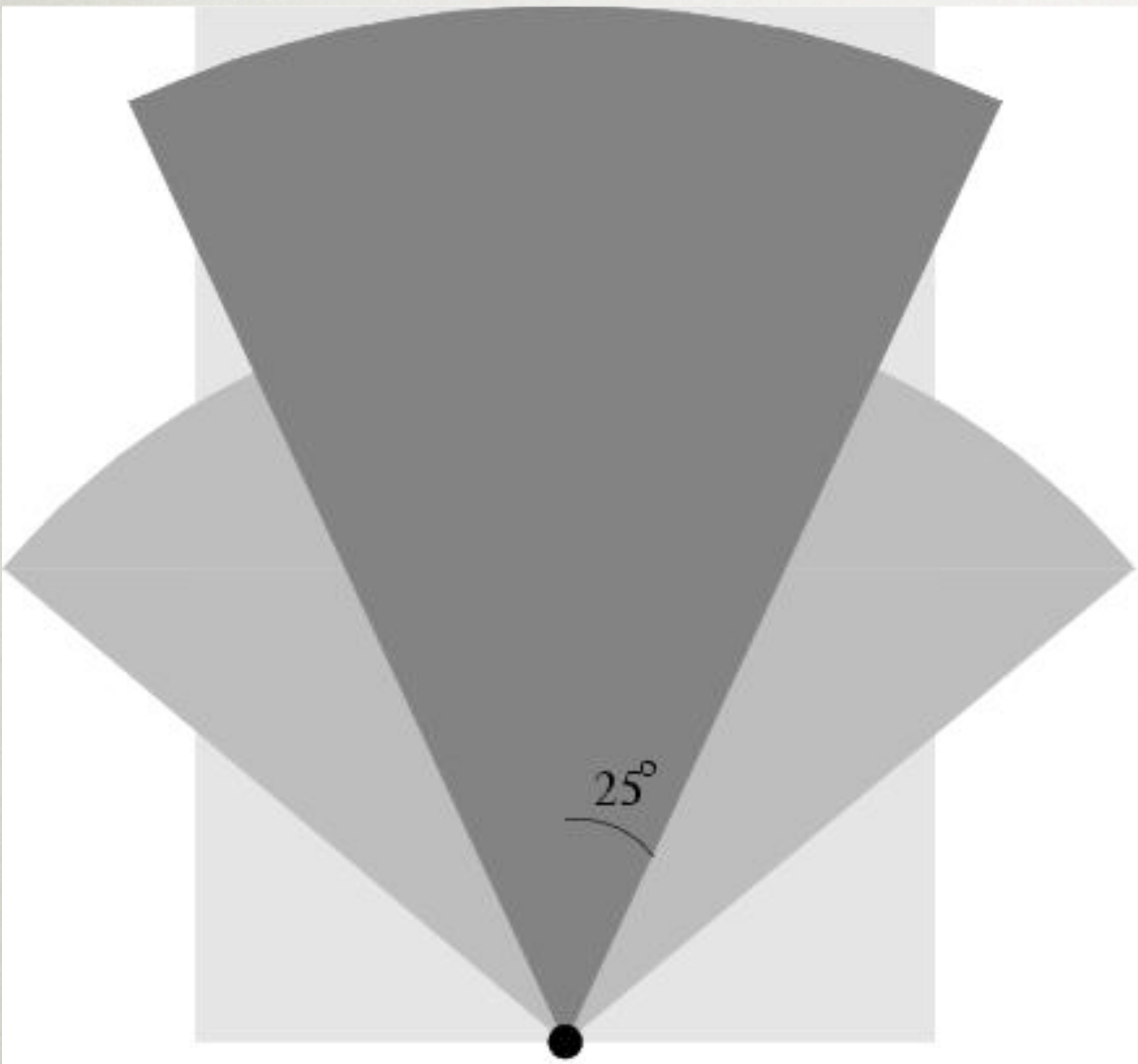
$c = [x, y, 0, 0]$

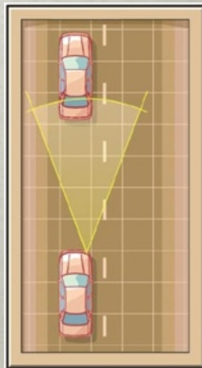


2 sensor - 3 targets

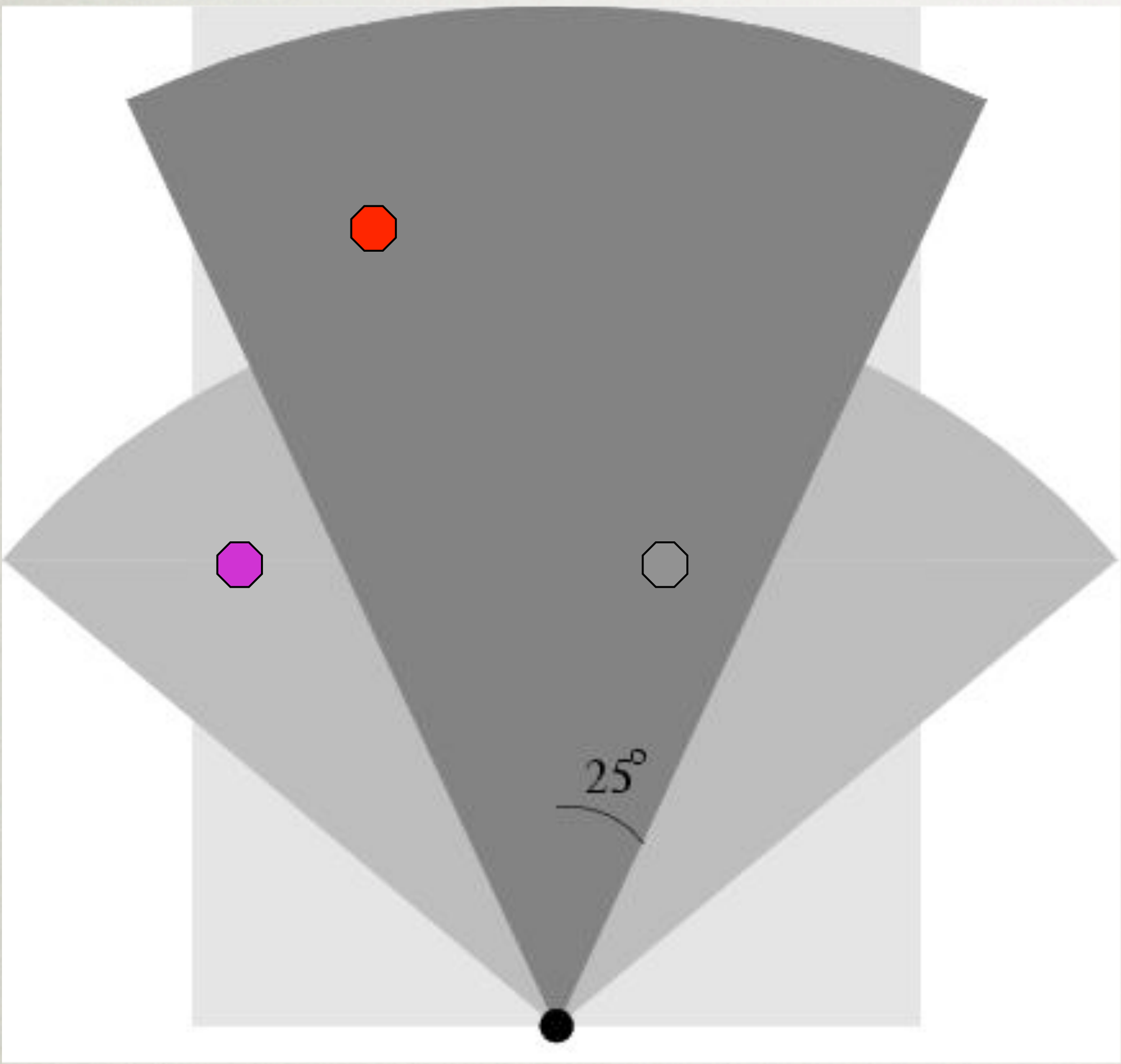


2 sensor - 3 targets



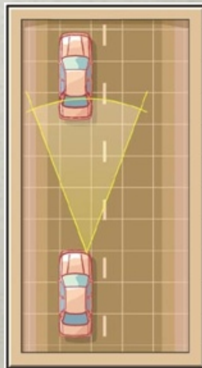


2 sensor - 3 targets

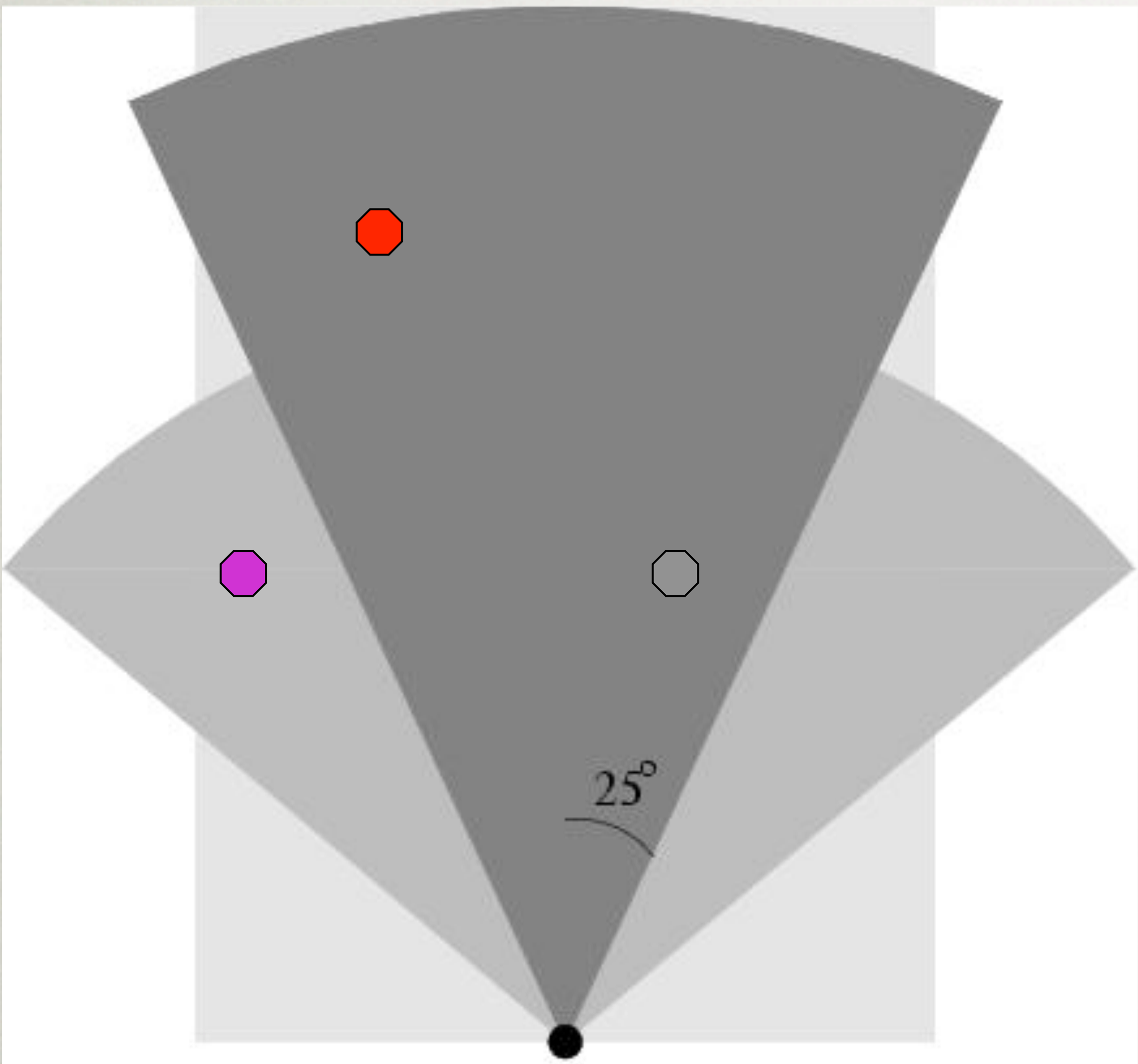


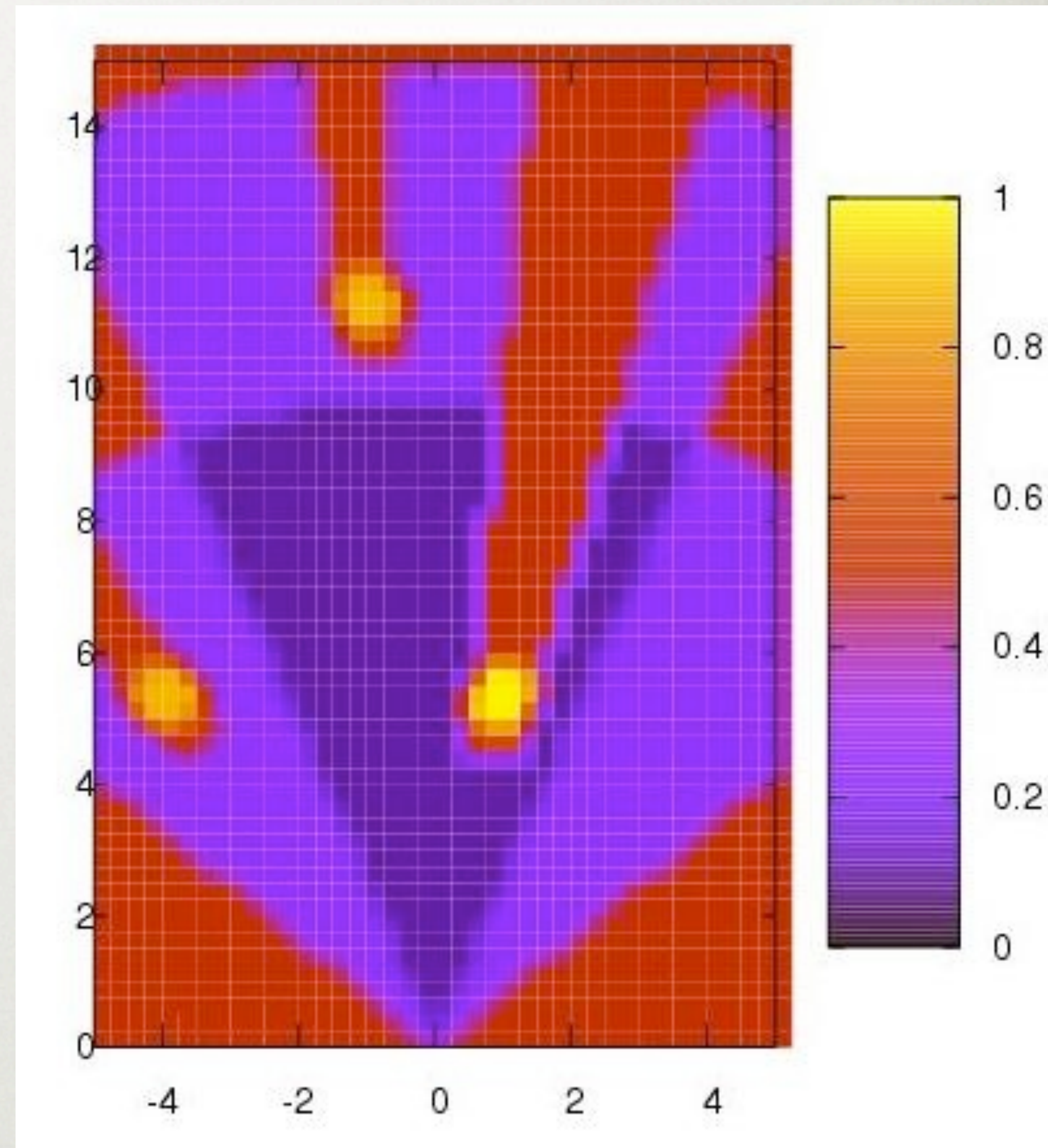
$z_{1,1} = (5.5, -4, 0, 0)$ $z_{1,2} = (5.5, 1, 0, 0)$

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



2 SENSOR - 3 TARGETS



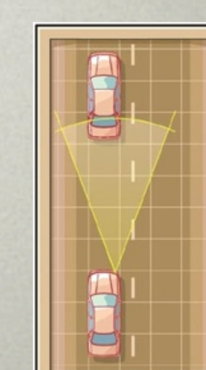


$z_{1,1} = (5.5, -4, 0, 0)$ $z_{1,2} = (5.5, 1, 0, 0)$

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

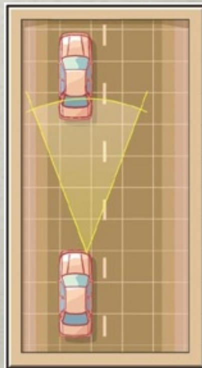
$P([E_c=1] \mid z_{1,1}\ z_{1,2}\ z_{2,1}\ z_{2,2}\ c)$

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



Bayesian Filter

Program

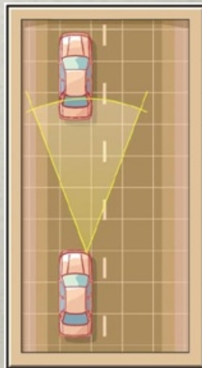


Bayesian Filter

Program

Description

Question



Bayesian Filter

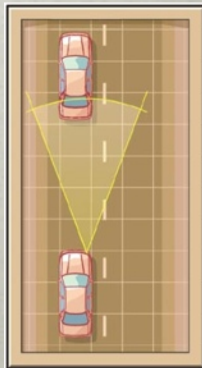
Program

Description

Question

Specification

Identification



Bayesian Filter

Program

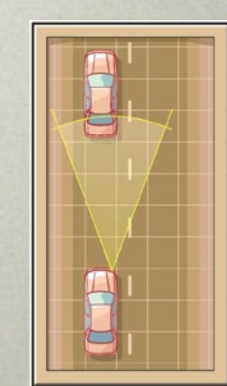
Description

Specification

- Variables

Identification

Question



Bayesian Filter

Program

Description

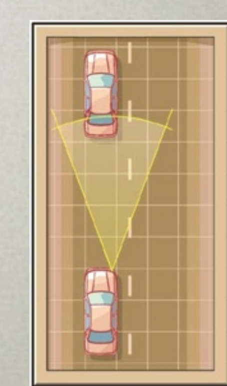
Specification

- Variables

$$S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t$$

Identification

Question



Bayesian Filter

Program

Description

Specification

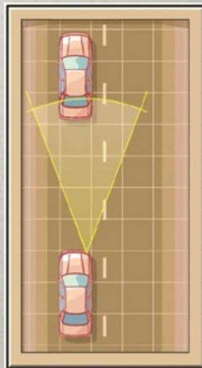
- Variables

$$S^0, ..., S^t, O^0, ..., O^t$$

- Decomposition (Conditional Independance Hypothesis)

Identification

Question



Bayesian Filter

Program

Description

Question

Specification

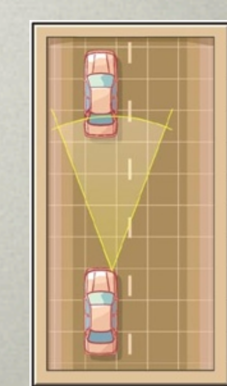
- Variables

$$S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t$$

- Decomposition (Conditional Independance Hypothesis)

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

Identification



Bayesian Filter

Program

Question

Specification

- Variables

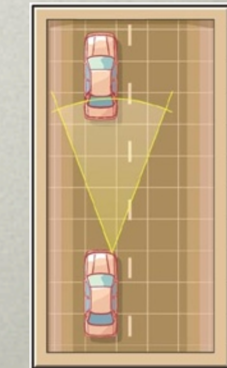
$$S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t$$

- Decomposition (Conditional Independance Hypothesis)

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

- Parametric Forms
  or Bayesian Subroutines

Identification



Bayesian Filter

Program

De ion

Specification

- Variables

$$S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t$$

- Decomposition (Conditional Independance Hypothesis)

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

- Parametric Forms
  or Bayesian Subroutines

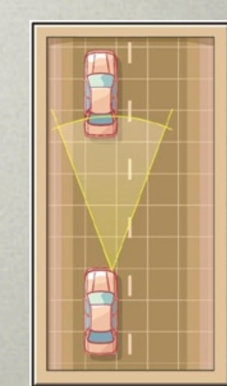
$$\mathbf{P}(S^0) \equiv \mathbf{G}(S^0, \mu, \sigma)$$

$$\mathbf{P}(S^i \mid S^{i-1}) \equiv \mathbf{G}(S^i, A \bullet S^{i-1}, Q)$$

$$\mathbf{P}(O^i \mid S^i) \equiv \mathbf{G}(O^i, H \bullet S^i, R)$$

Identification

Question



Bayesian Filter

Program

Description

Question

Specification

- Variables

$$S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t$$

- Decomposition (Conditional Independance Hypothesis)

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

- Parametric Forms
  or Bayesian Subroutines

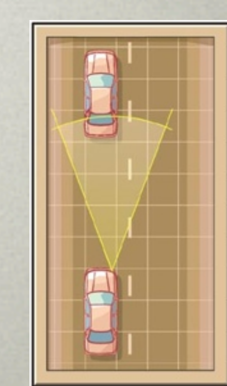
$$\mathbf{P}(S^0) \equiv \mathbf{G}(S^0, \mu, \sigma)$$

$$\mathbf{P}(S^i \mid S^{i-1}) \equiv \mathbf{G}(S^i, A \bullet S^{i-1}, Q)$$

$$\mathbf{P}(O^i \mid S^i) \equiv \mathbf{G}(O^i, H \bullet S^i, R)$$

Identification

- Learning from instances



Bayesian Filter

Program

Description

Specification

- Variables

$$S^0, \ldots\ldots, S^t, O^0, \ldots\ldots, O^t$$

- Decomposition (Conditional Independance Hypothesis)

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

- Parametric Forms
  or Bayesian Subroutines

$$\mathbf{P}(S^0) \equiv \mathbf{G}(S^0, \mu, \sigma)$$

$$\mathbf{P}(S^i \mid S^{i-1}) \equiv \mathbf{G}(S^i, A \bullet S^{i-1}, Q)$$

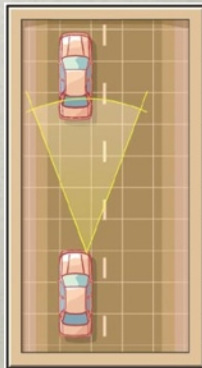
$$\mathbf{P}(O^i \mid S^i) \equiv \mathbf{G}(O^i, H \bullet S^i, R)$$

Identification

- Learning from instances

Question

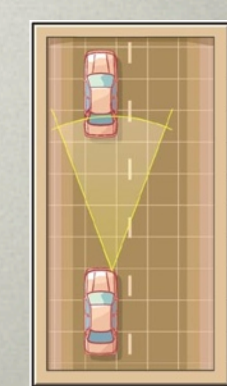
$$\mathbf{P}(S^t \mid O^0 \wedge \ldots\ldots \wedge O^t)$$



Bayesian Filter

$$P(S^t \mid O^{0:t-1}) = \sum_{S^{t-1}} \left[P(S^t \mid S^{t-1}) \times P(S^{t-1} \mid O^{0:t-1}) \right]$$

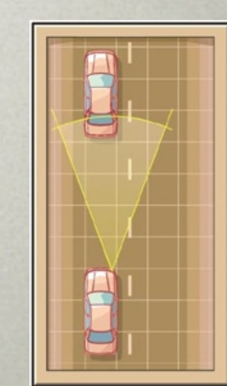
$$P(S^t \mid O^{0:t}) = P(O^t \mid S^t) \times P(S^t \mid O^{0:t-1})$$



Bayesian Filter

$$\boxed{P(S^t \mid O^{0:t-1})} = \sum_{S^{t-1}} \left[P(S^t \mid S^{t-1}) \times P(S^{t-1} \mid O^{0:t-1}) \right]$$

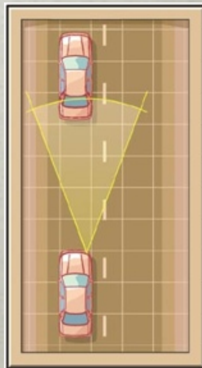
$$P(S^t \mid O^{0:t}) = P(O^t \mid S^t) \times \boxed{P(S^t \mid O^{0:t-1})}$$



Bayesian Filter

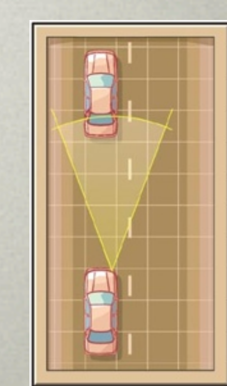
$$\boxed{P(S^t \mid O^{0:t-1})} = \sum_{S^{t-1}} \left[P(S^t \mid S^{t-1}) \times \boxed{P(S^{t-1} \mid O^{0:t-1})} \right]$$

$$\boxed{P(S^t \mid O^{0:t})} = P(O^t \mid S^t) \times \boxed{P(S^t \mid O^{0:t-1})}$$

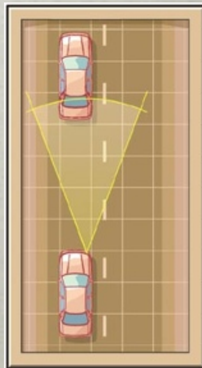


Without vs with filtering

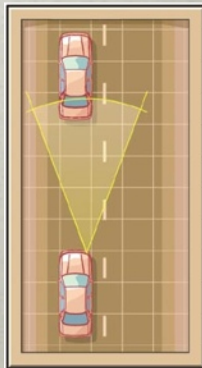
(Videos)



Real time filtering



Real time filtering

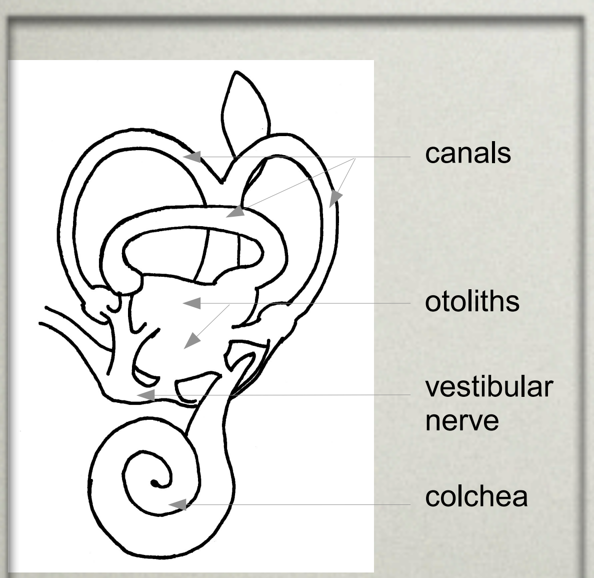


Overview

- 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?

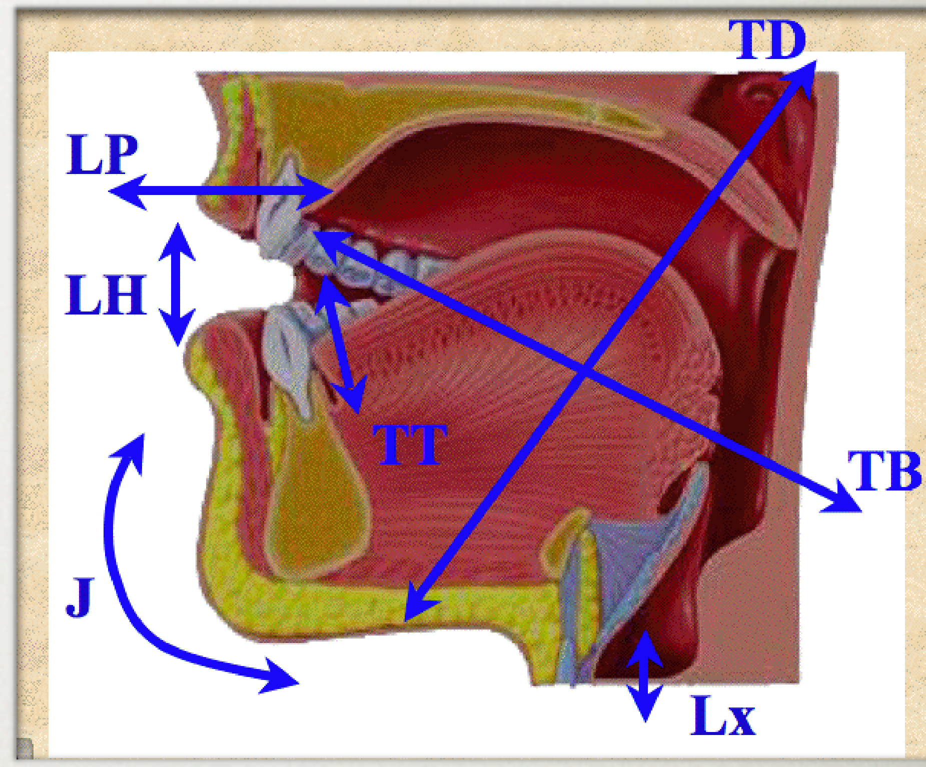
Modeling behaviors

PhD Jihene Serkhane

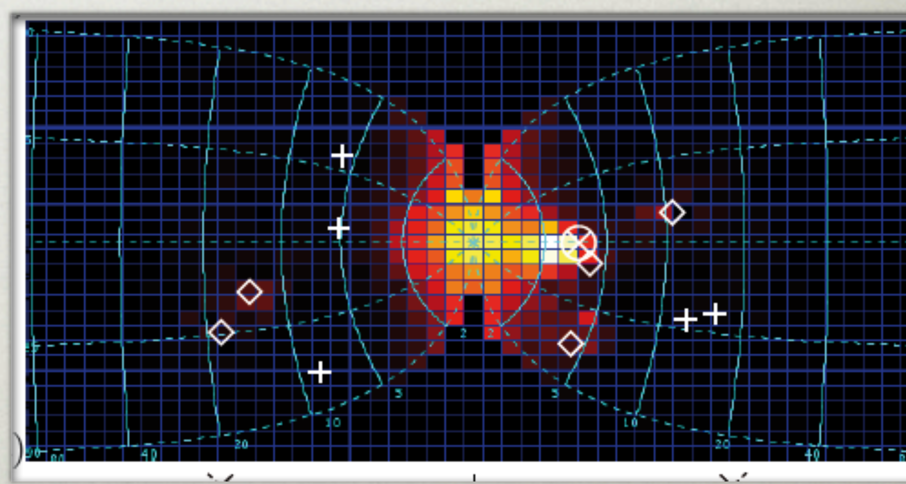


PhD Jean Laurens

PhD Francis Colas



PhD C. Moulin-Frier



PostDoc Francis Colas

Bayesian Action Perception:

Handwriting experiments

Ph.D Estelle Gilet

Gilet E, Diard J, Bessière P, 2011 Bayesian Action–Perception Computational Model: Interaction of Production and Recognition of Cursive Letters.PLoS ONE 6(6): e20387. doi:10.1371/journal.pone.0020387

Motor Equivalence?

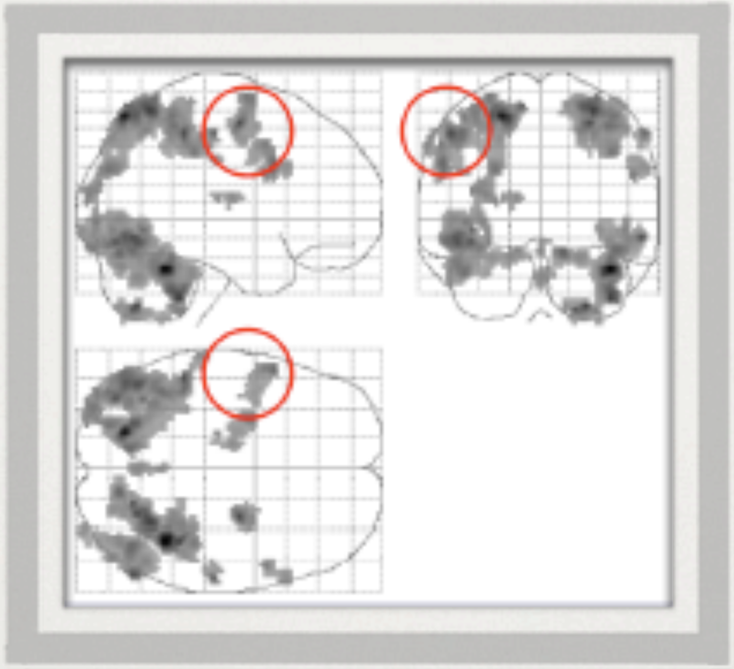
Motor Equivalence?

[Serratrice93]

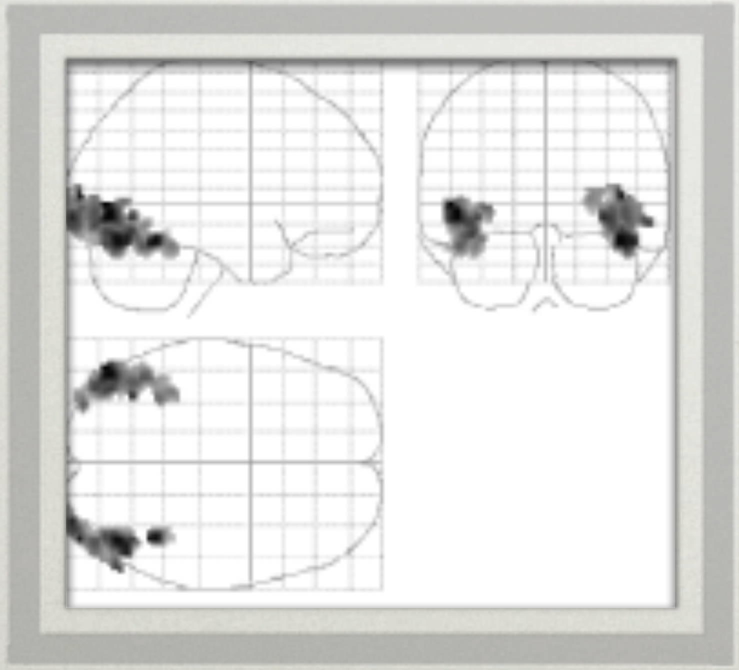
- Writer "style"
  - ➢ [Wright90]
- Common activated motor areas
  - ➢ [Wing00]

Simulation of action during perception?

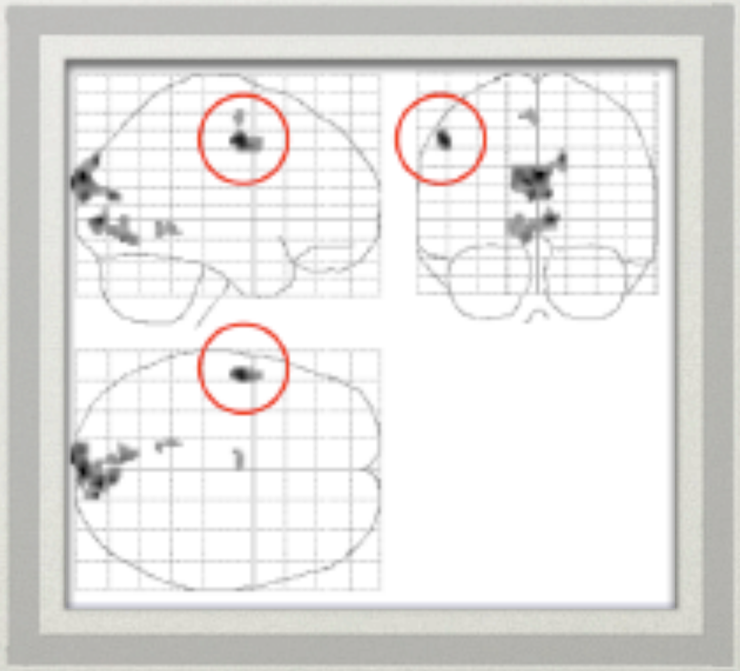
[Longcamp03]



Writing

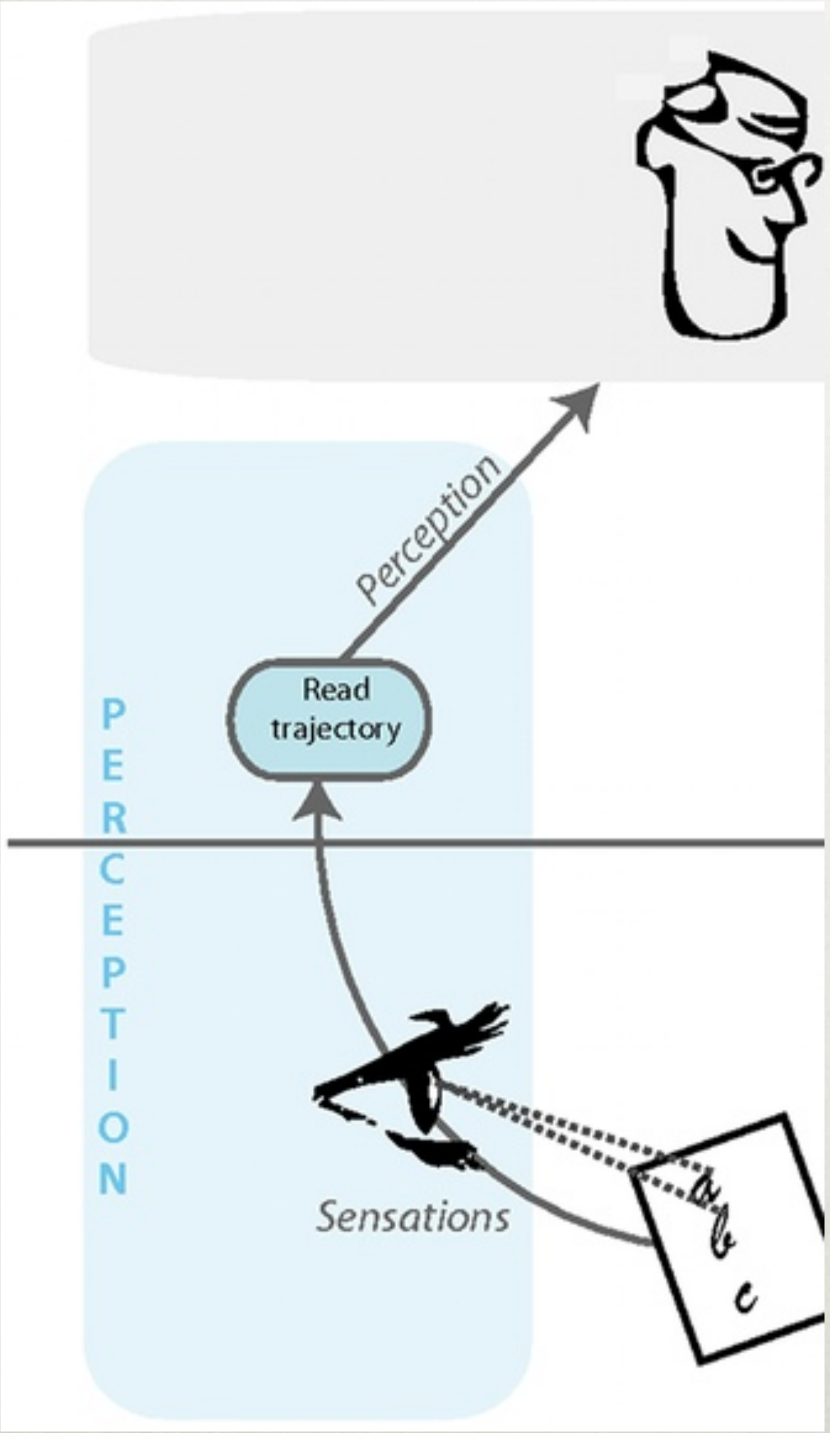


Pseudo letter reading



Letter reading

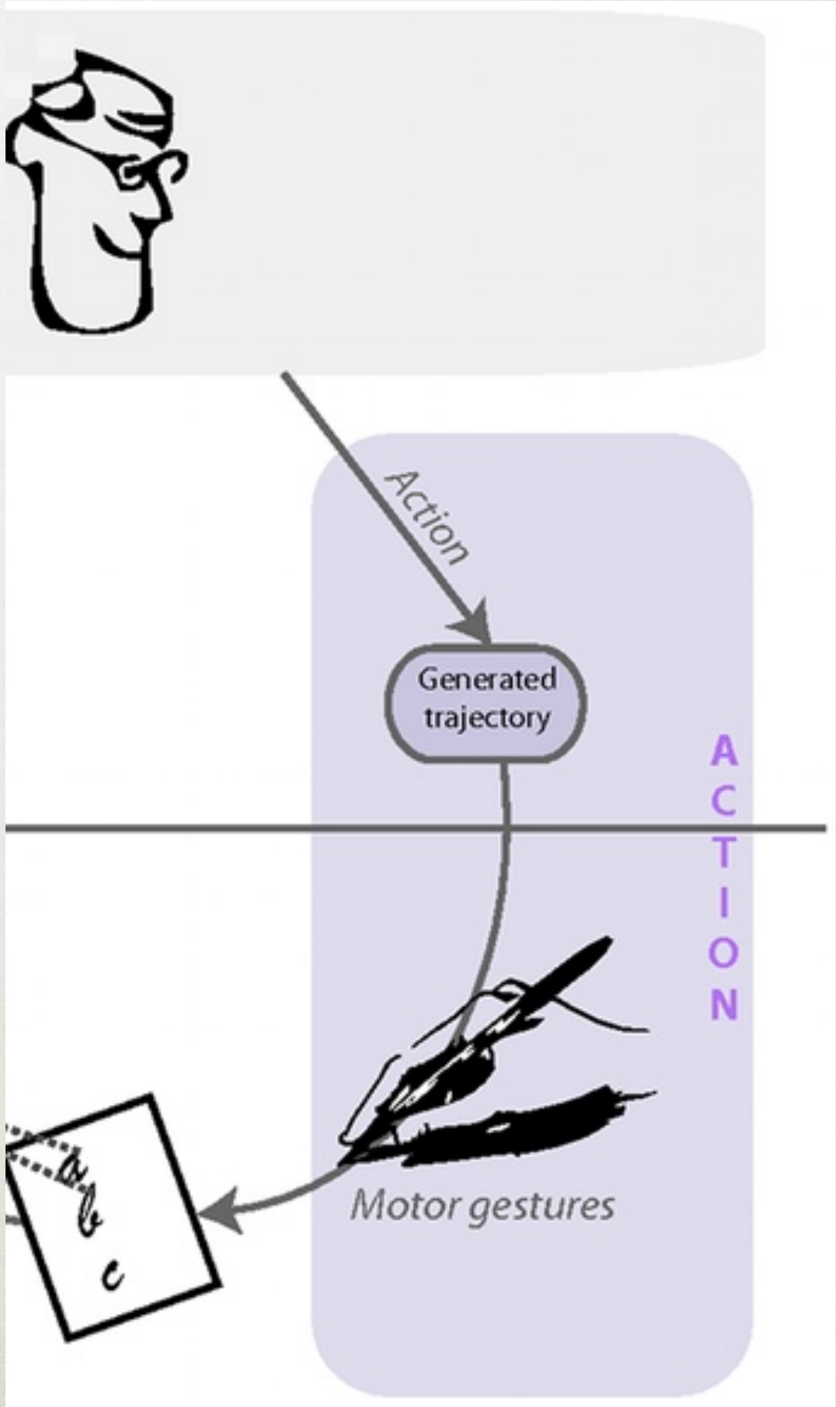
Reading



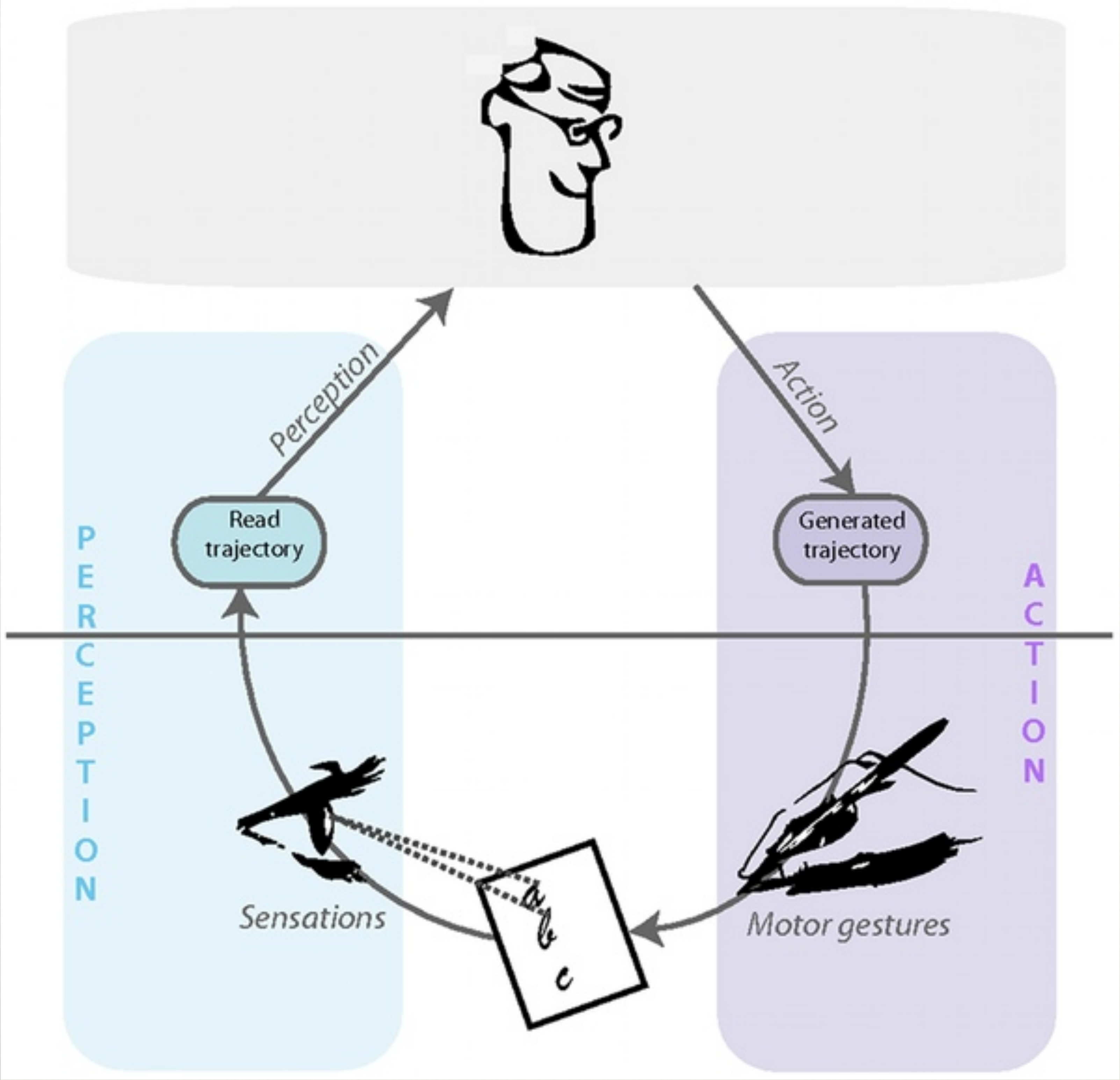
- OCR
  - [Meulenbroek96]
  - [Flash95]
- Human models
  - [Crettez98]
  - [Vuori02]
  - [Dehaene07]

Writing

- [Hinton05]
- [Meulenbroek96]
- [Flash95]



BAP model



Perception
Action
Read trajectory
Generated trajectory
PERCEPTION
ACTION
Sensations
Motor gestures
a b c

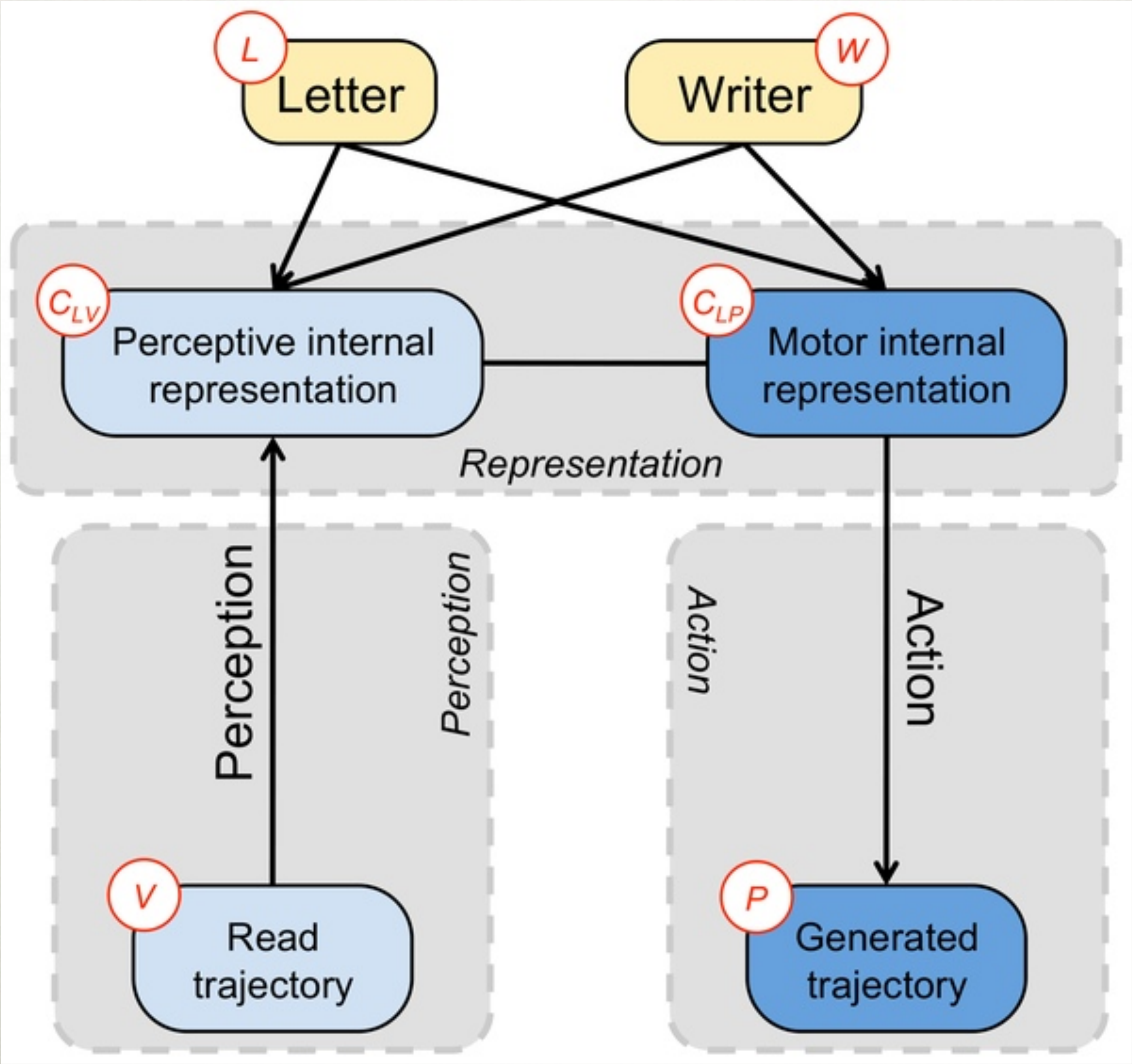
From motor perception theory to Bayesian Action Perception

From motor perception theory to Bayesian Action Perception

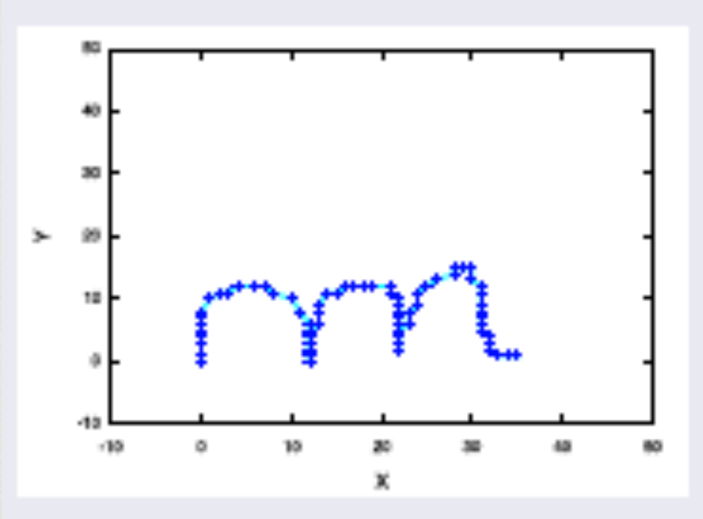
Motor perception theory
[Liberman57]

Perception for action control theory
[Schwartz01]

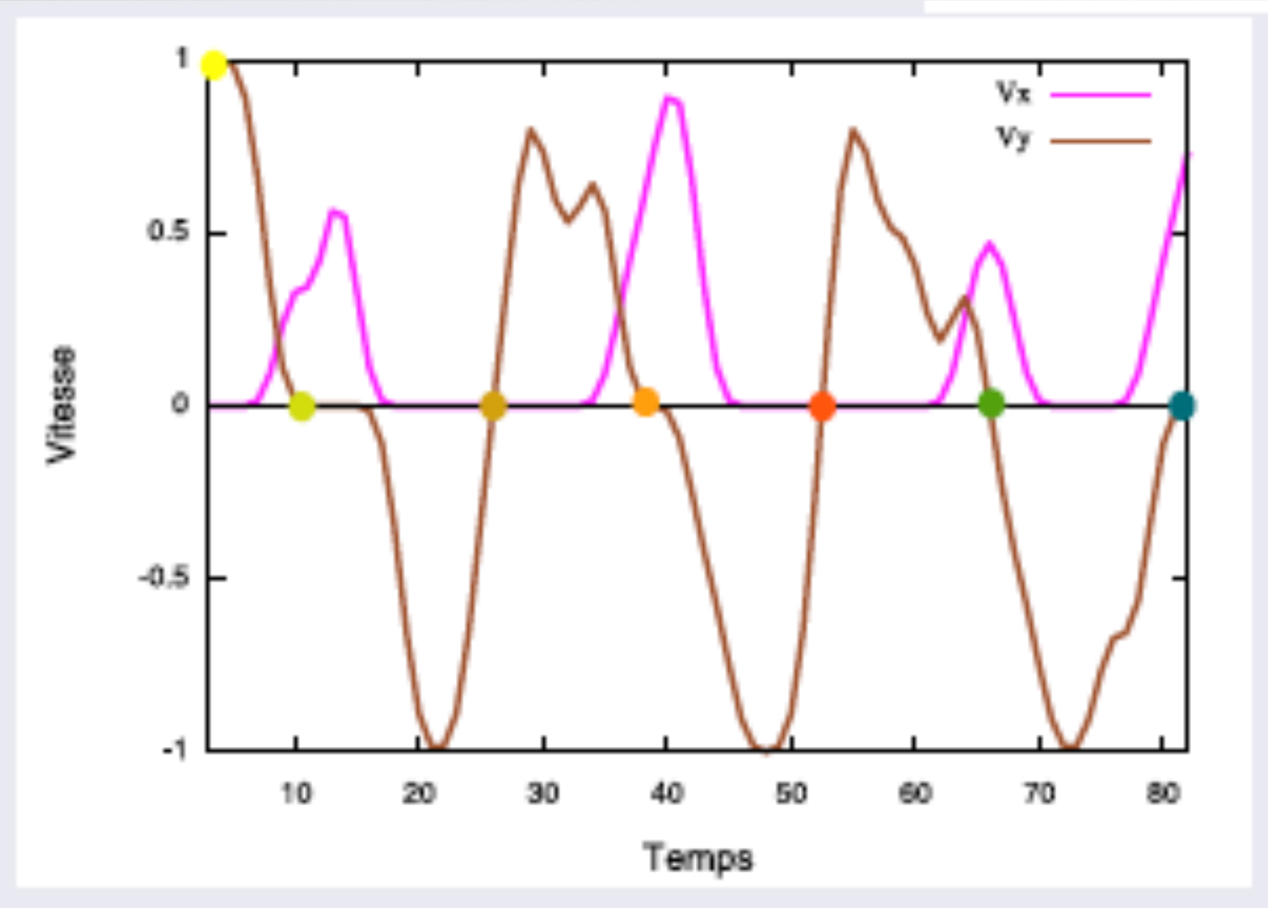
A common space for motor and perception internal representation

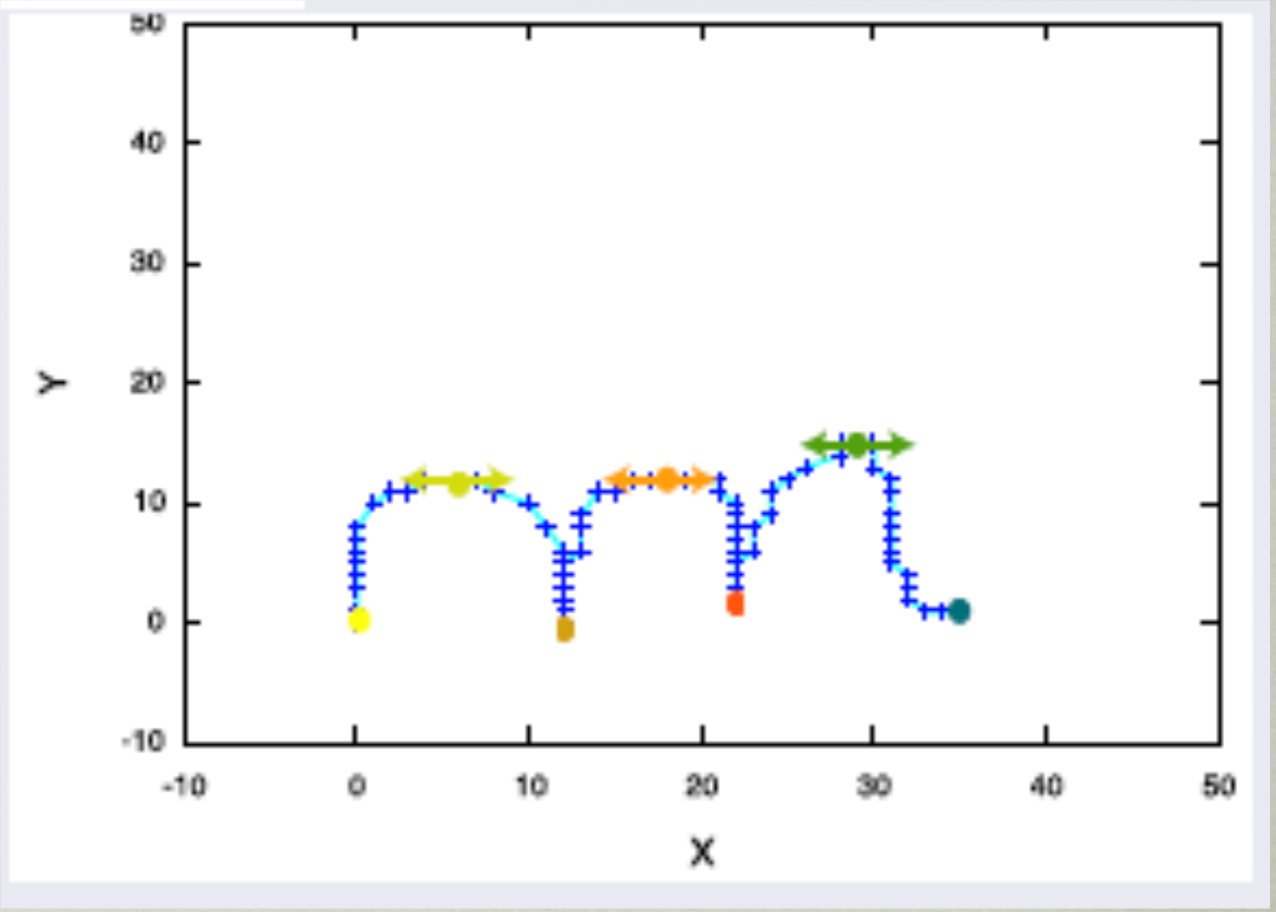


Common features for both representations

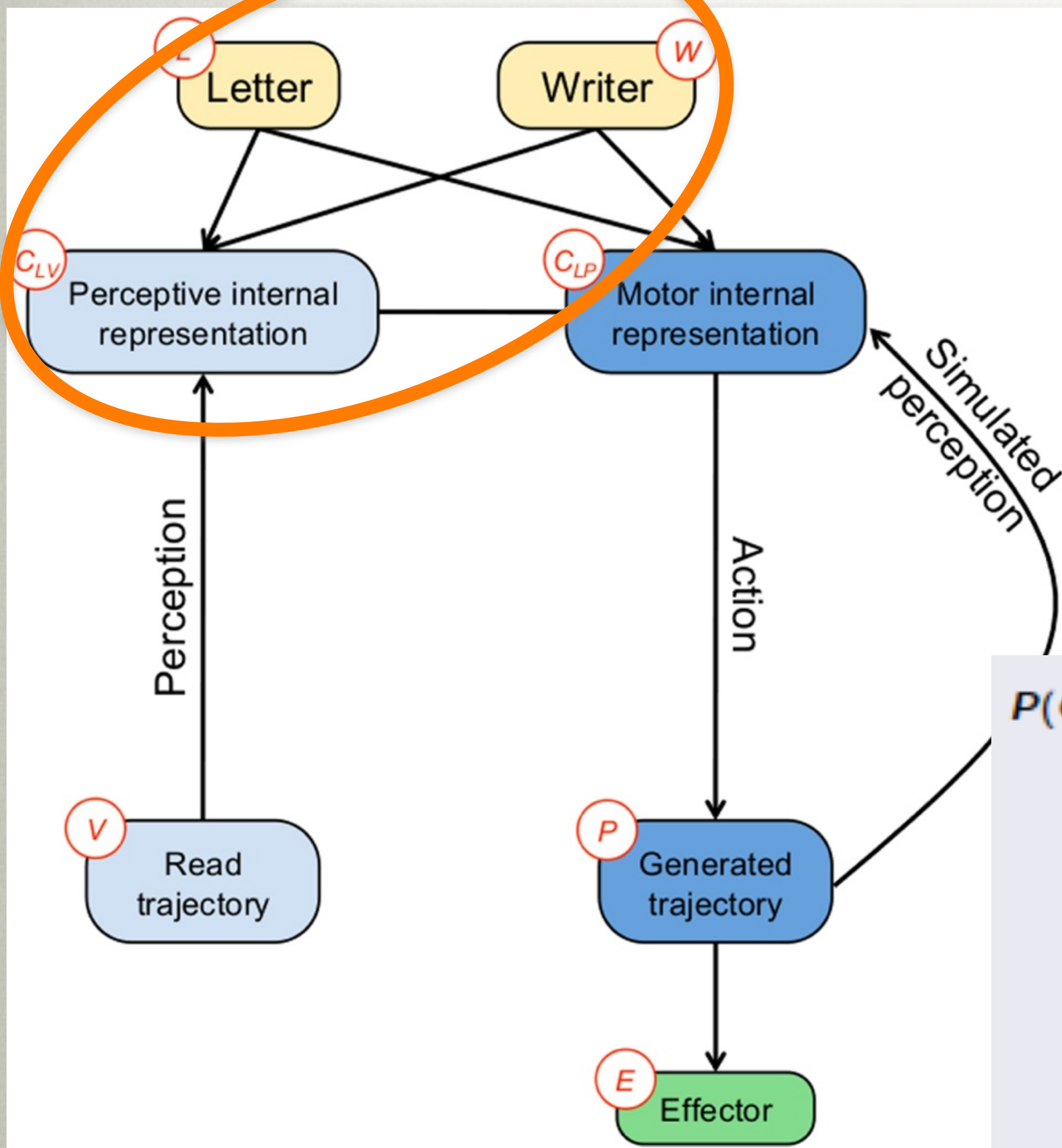


$$\frac{dx}{dt}(t) = 0 \quad \vee \quad \frac{dy}{dt}(t) = 0$$





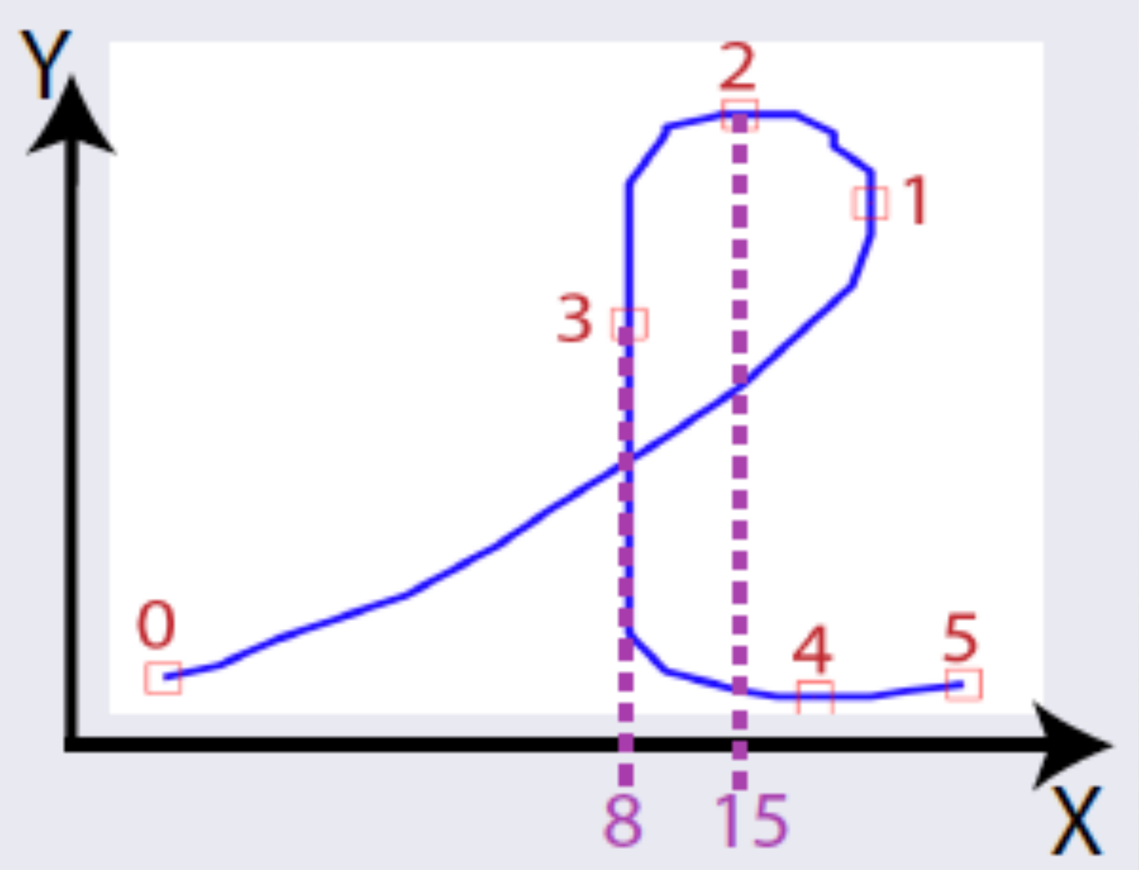
Here come the probabilities

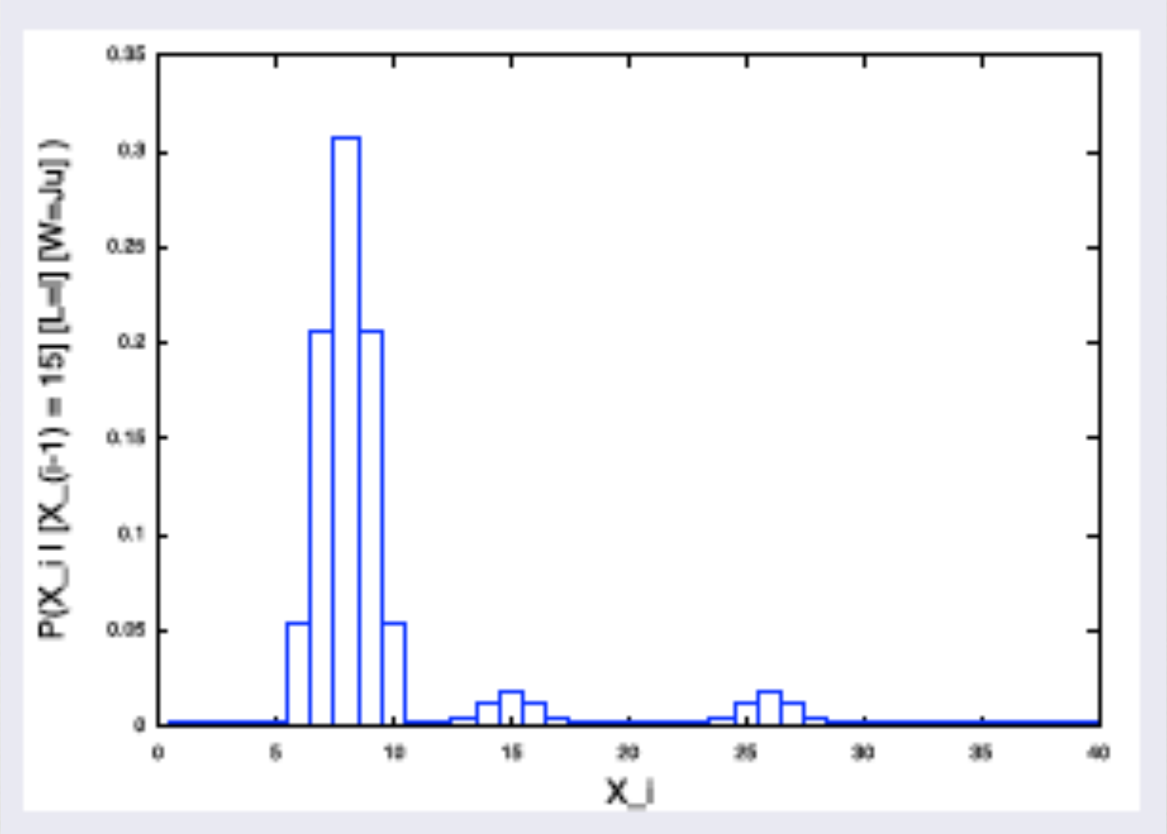


$$
\begin{aligned}
& P(C_L^{0:N}\ L\ W) \\
= \quad & P(C_L^{0:N} \mid L\ W)\ P(L)\ P(W) \\
= \quad & \left(
\begin{array}{l}
\left(
\begin{array}{l}
P(C_{Lx}^{0} \mid L\ W)\ P(C_{Ly}^{0} \mid L\ W) \\
P(C_{L\dot{x}}^{0} \mid L\ W)\ P(C_{L\dot{y}}^{0} \mid L\ W)
\end{array}
\right) \\
\prod_{n=1}^{N} \left(
\begin{array}{l}
P(C_{Lx}^{n} \mid C_{Lx}^{n-1}\ L\ W)\ P(C_{Ly}^{n} \mid C_{Ly}^{n-1}\ L\ W) \\
P(C_{L\dot{x}}^{n} \mid C_{L\dot{x}}^{n-1}\ L\ W)\ P(C_{L\dot{y}}^{n} \mid C_{L\dot{y}}^{n-1}\ L\ W)
\end{array}
\right) \\
P(L)\ P(W)
\end{array}
\right)
\end{aligned}
$$

Learning succession of control points

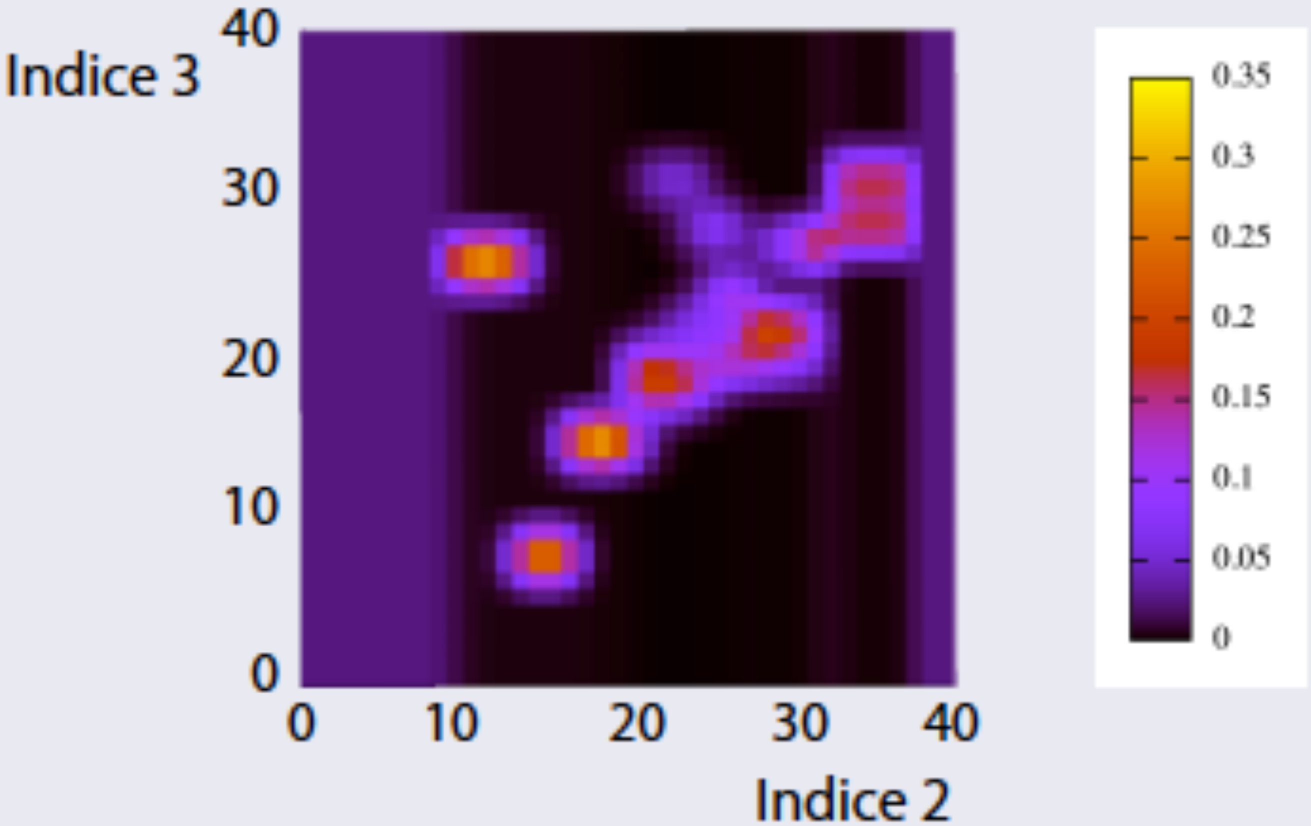
$$P(C^3_{Lx} \mid [C^2_{Lx} = 15]\ [L = l]\ [W = Julienne]) = \frac{p_i + \alpha}{N + k\alpha}.$$

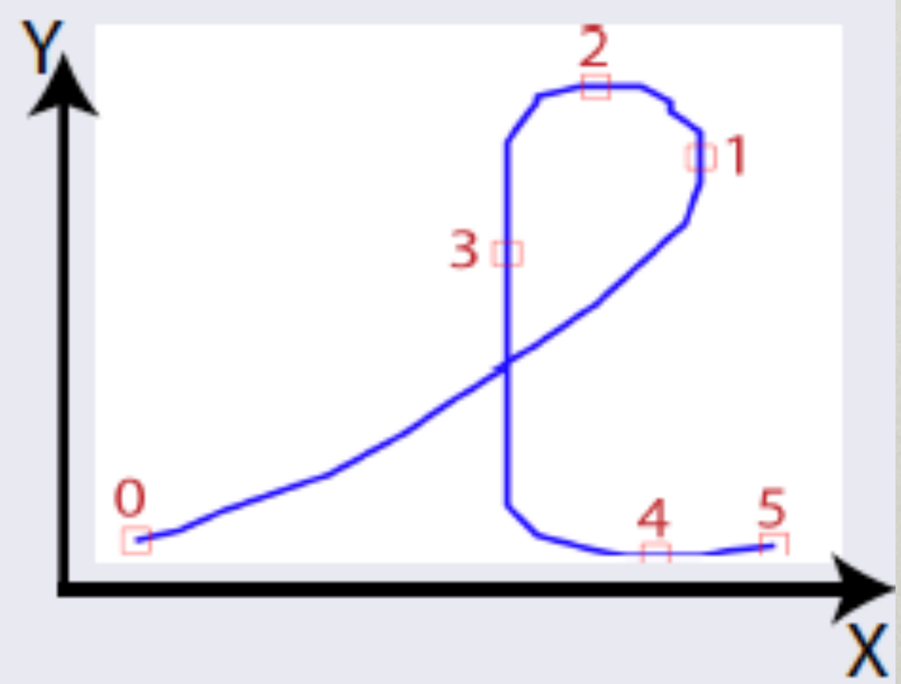




Learning succession of control points

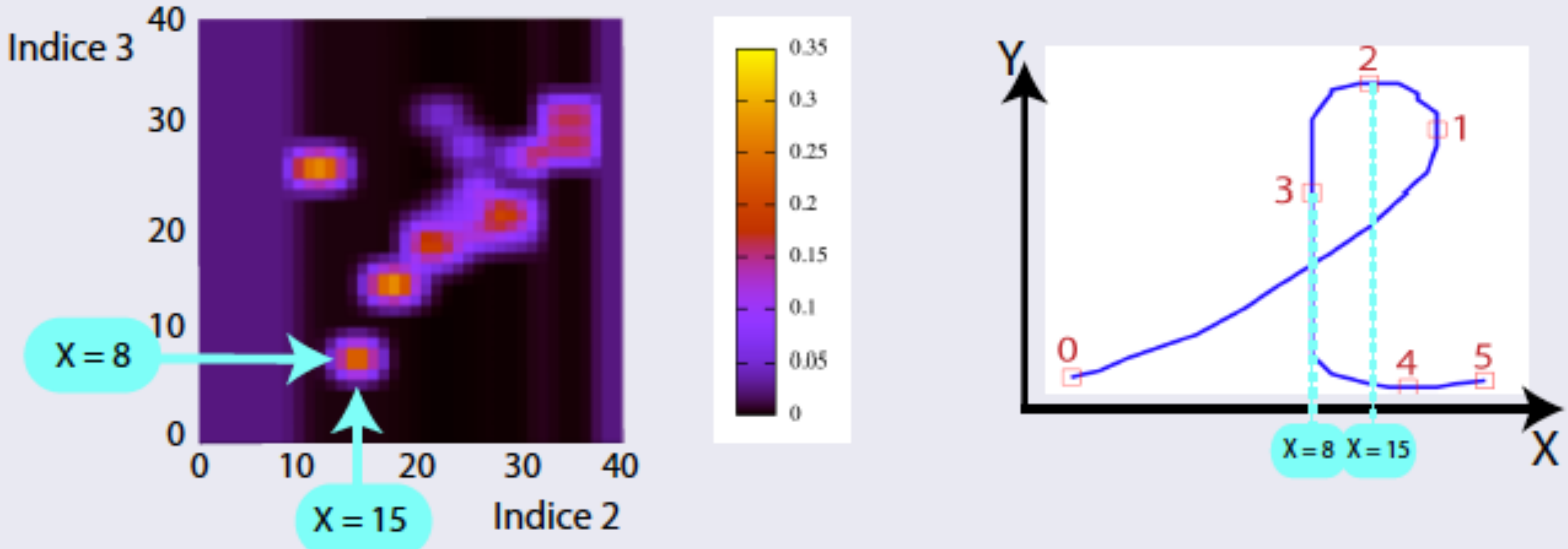
$$P(C^3_{Lx} \mid C^2_{Lx}\ [L = l]\ [W = Julienne]) \quad (2)$$





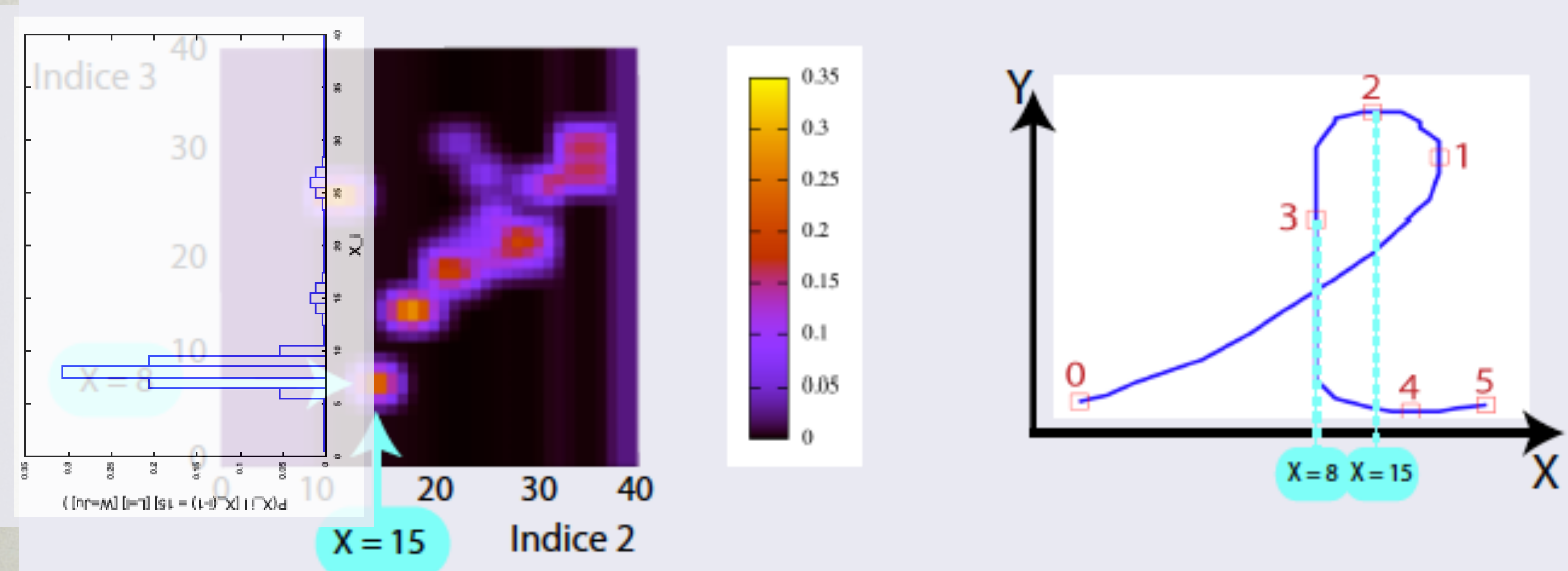
Learning succession of control points

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



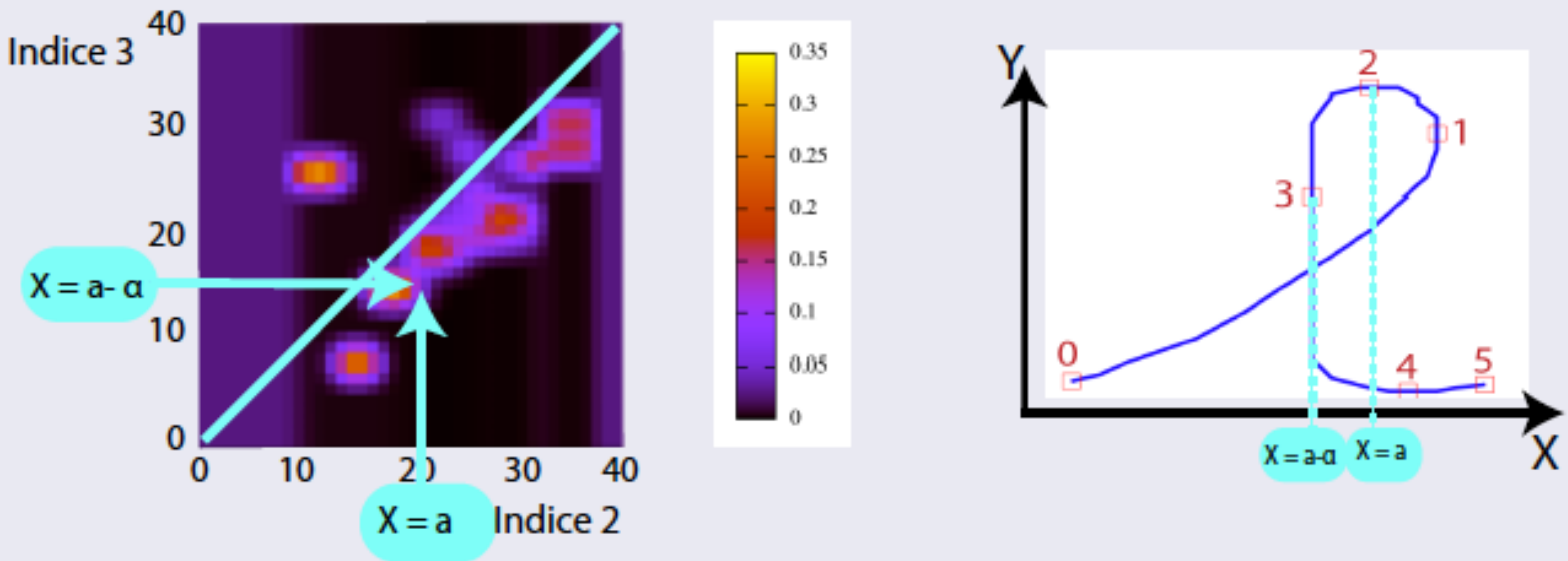
Learning succession of control points

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



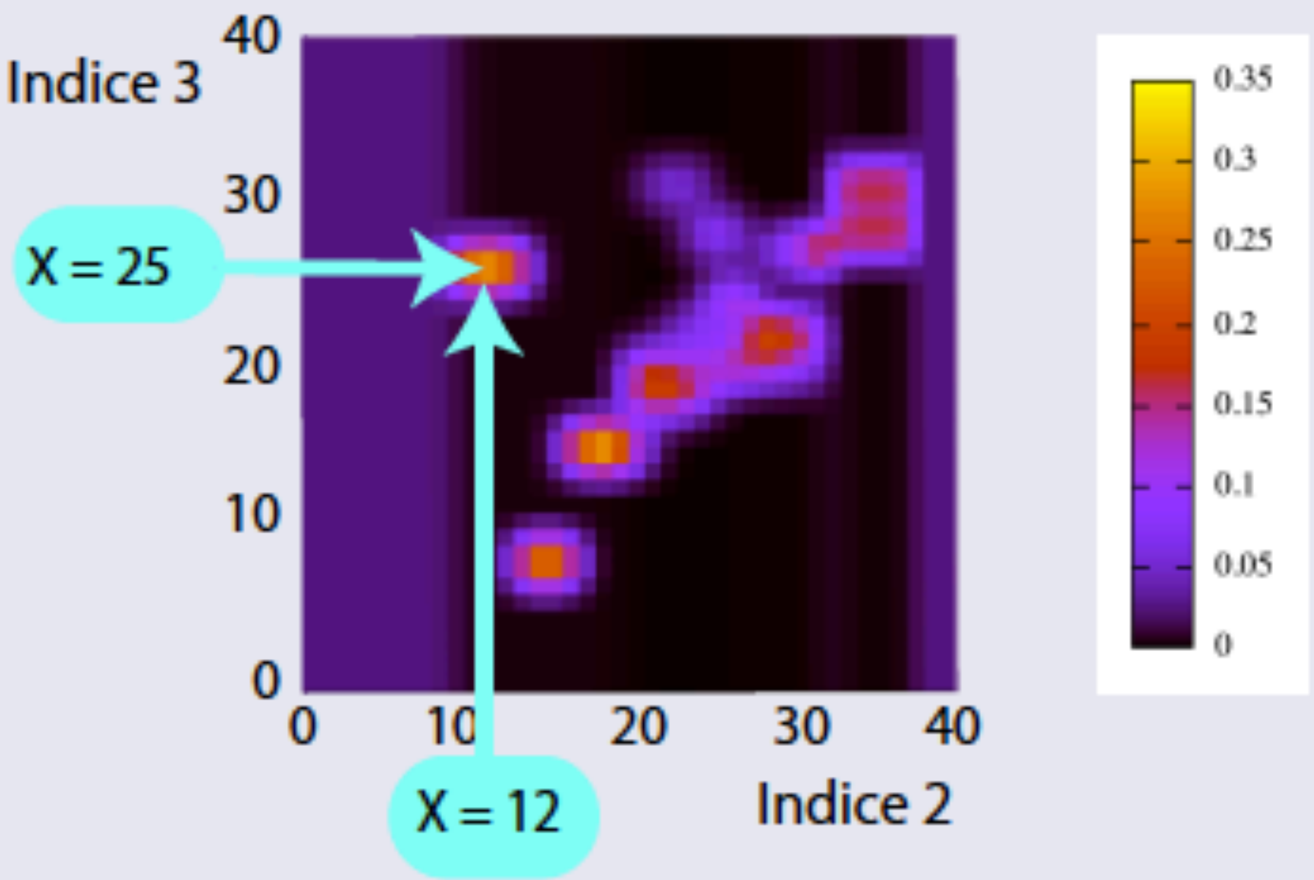
Learning succession of control points

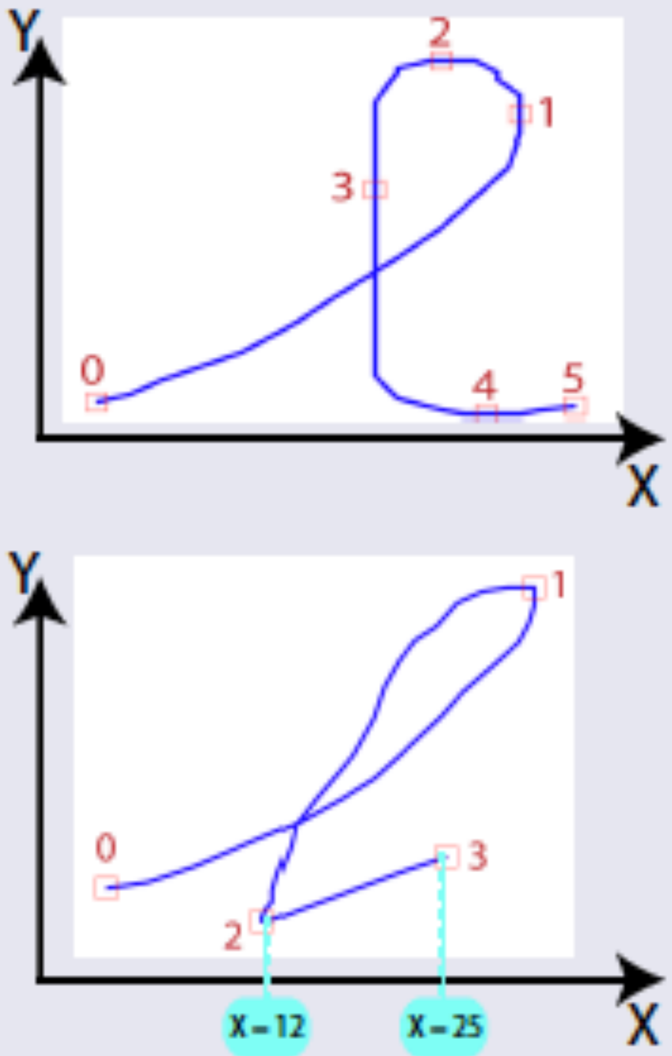
$$P(C^3_{Lx} \mid C^2_{Lx}\ [L = l]\ [W = Julienne]) \qquad (2)$$



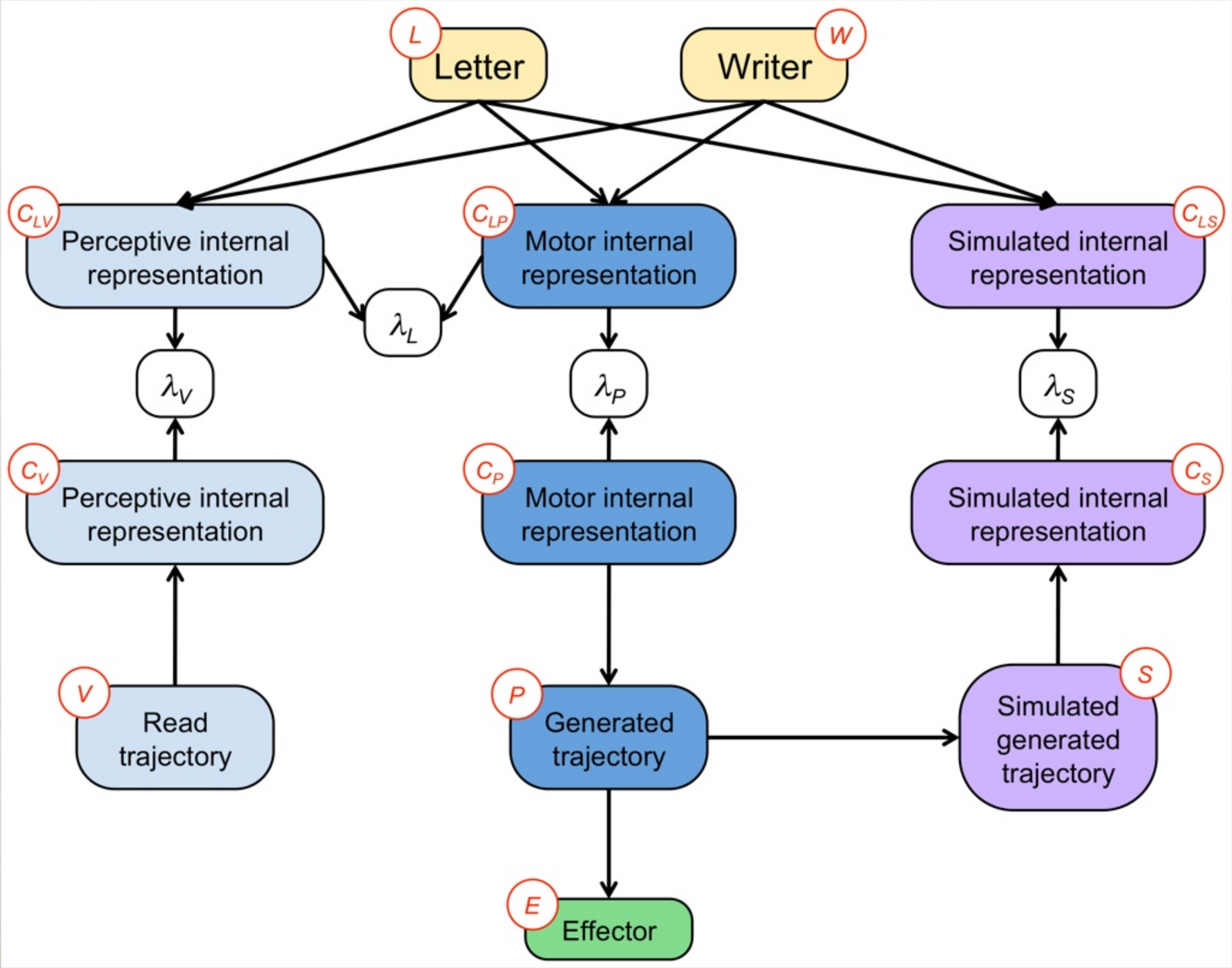
Learning succession of control points

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



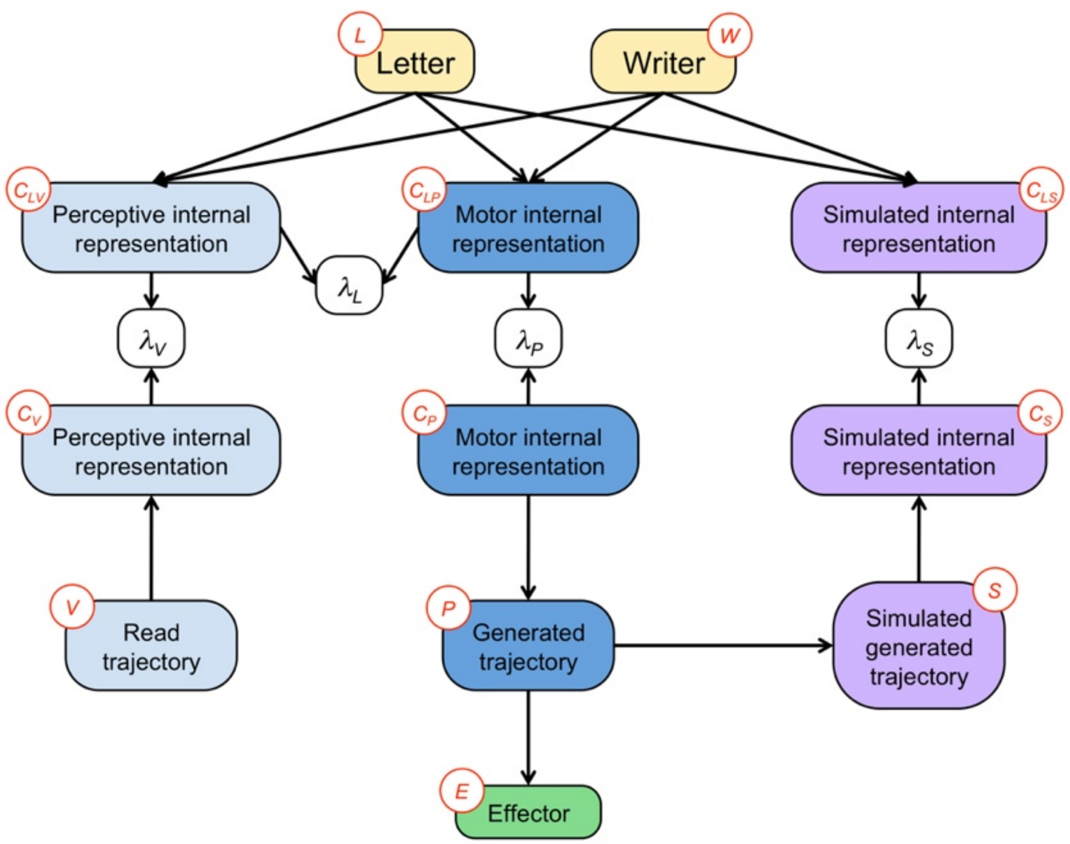


BAP MODEL



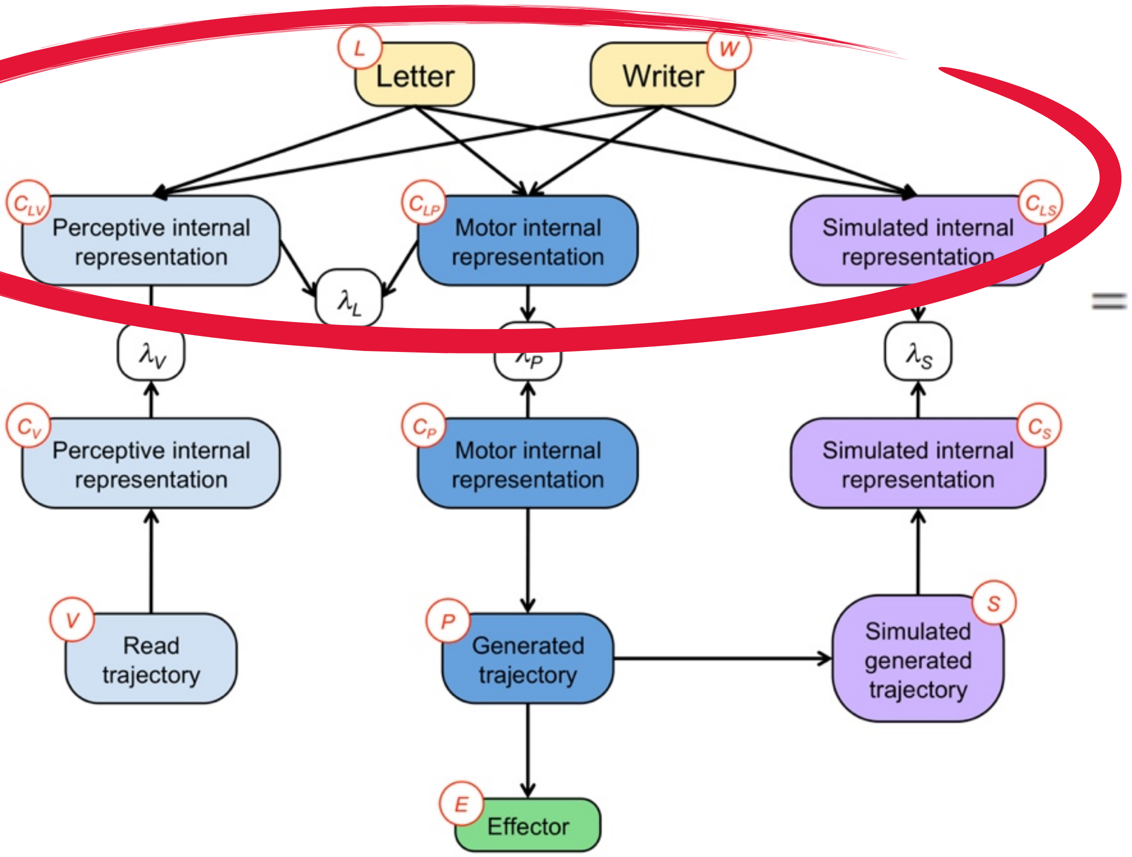
L
Letter
W
Writer
C_{LV}
Perceptive internal representation
C_{LP}
Motor internal representation
C_{LS}
Simulated internal representation
λ_L
λ_V
λ_P
λ_S
C_V
Perceptive internal representation
C_P
Motor internal representation
C_S
Simulated internal representation
V
Read trajectory
P
Generated trajectory
S
Simulated generated trajectory
E
Effector

BAP MODEL



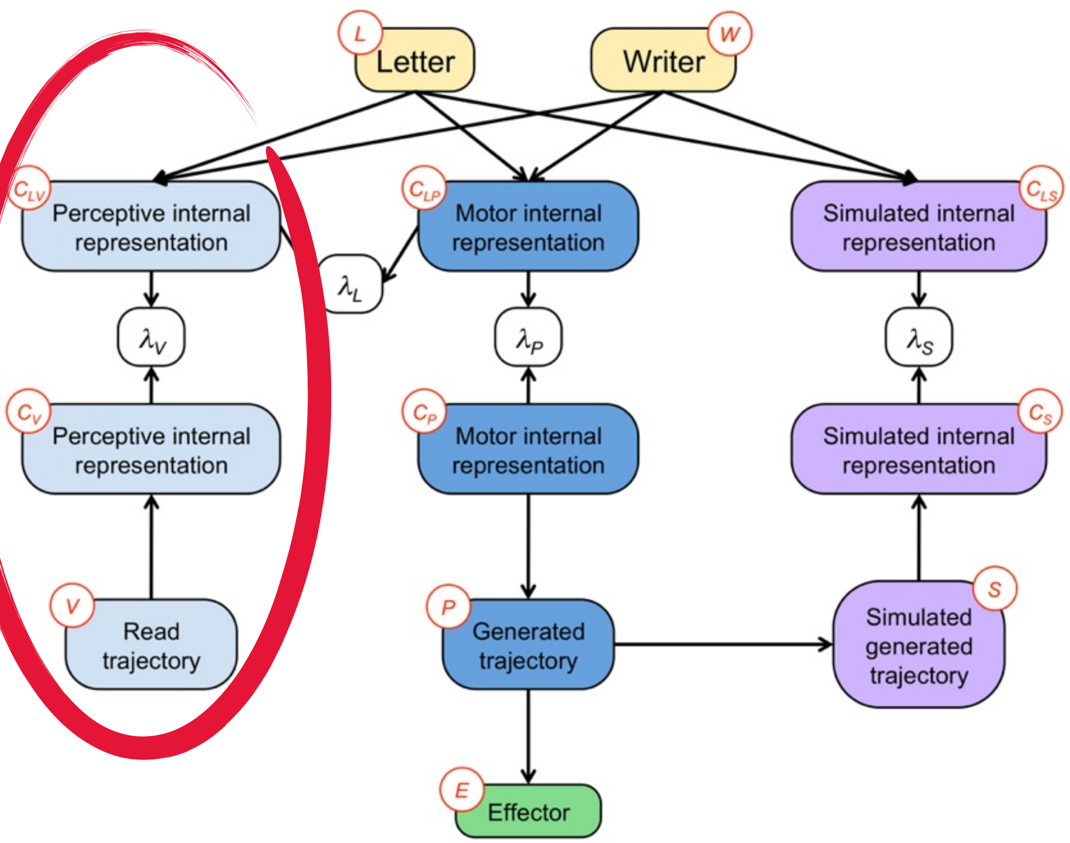
$$
\begin{aligned}
&P\left(\begin{matrix} C_{LV}\ C_{LP}\ C_{LS}\ C_V\ \lambda_P\ \lambda_V\ \lambda_S\ \lambda_L\ E \\ C_P\ C_S\ V\ P\ L\ W\ S \end{matrix}\right) \\
&= \left(\begin{matrix}
P(C_{LV} \mid L\ W)\ P(C_{LP} \mid L\ W)\ P(C_{LS} \mid L\ W) \\
P(L)\ P(W)\ P(\lambda_L \mid C_{LV}\ C_{LP}) \\
P(\lambda_V \mid C_{LV}\ C_V)\ P(C_V \mid V)\ P(V) \\
P(\lambda_P \mid C_{LP}\ C_P)\ P(P \mid C_P)\ P(E \mid P) \\
P(\lambda_S \mid C_{LS}\ C_S)\ P(C_S \mid S)\ P(S \mid P)
\end{matrix}\right)
\end{aligned}
$$

BAP MODEL



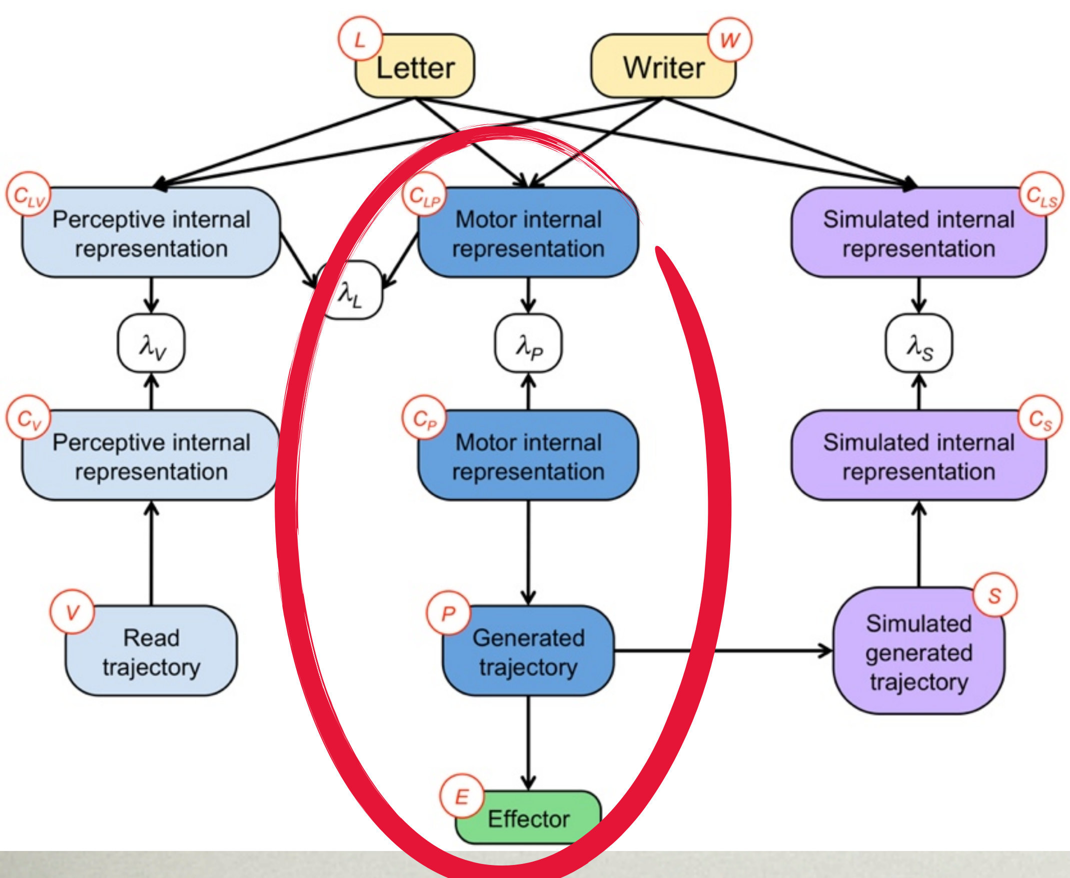
$$
\begin{aligned}
P\left(\begin{matrix} C_{LV}\ C_{LP}\ C_{LS}\ C_V\ \lambda_P\ \lambda_V\ \lambda_S\ \lambda_l\ E \\ C_P\ C_S\ V\ P\ L\ W\ S \end{matrix}\right)
= \left(\begin{matrix}
\boxed{P(C_{LV} \mid L\ W)\ P(C_{LP} \mid L\ W)\ P(C_{LS} \mid L\ W) \\ P(L)\ P(W)\ P(\lambda_l \mid C_{LV}\ C_{LP})} \\
P(\lambda_V \mid C_{LV}\ C_V)\ P(C_V \mid V)\ P(V) \\
P(\lambda_P \mid C_{LP}\ C_P)\ P(P \mid C_P)\ P(E \mid P) \\
P(\lambda_S \mid C_{LS}\ C_S)\ P(C_S \mid S)\ P(S \mid P)
\end{matrix}\right)
\end{aligned}
$$

BAP model



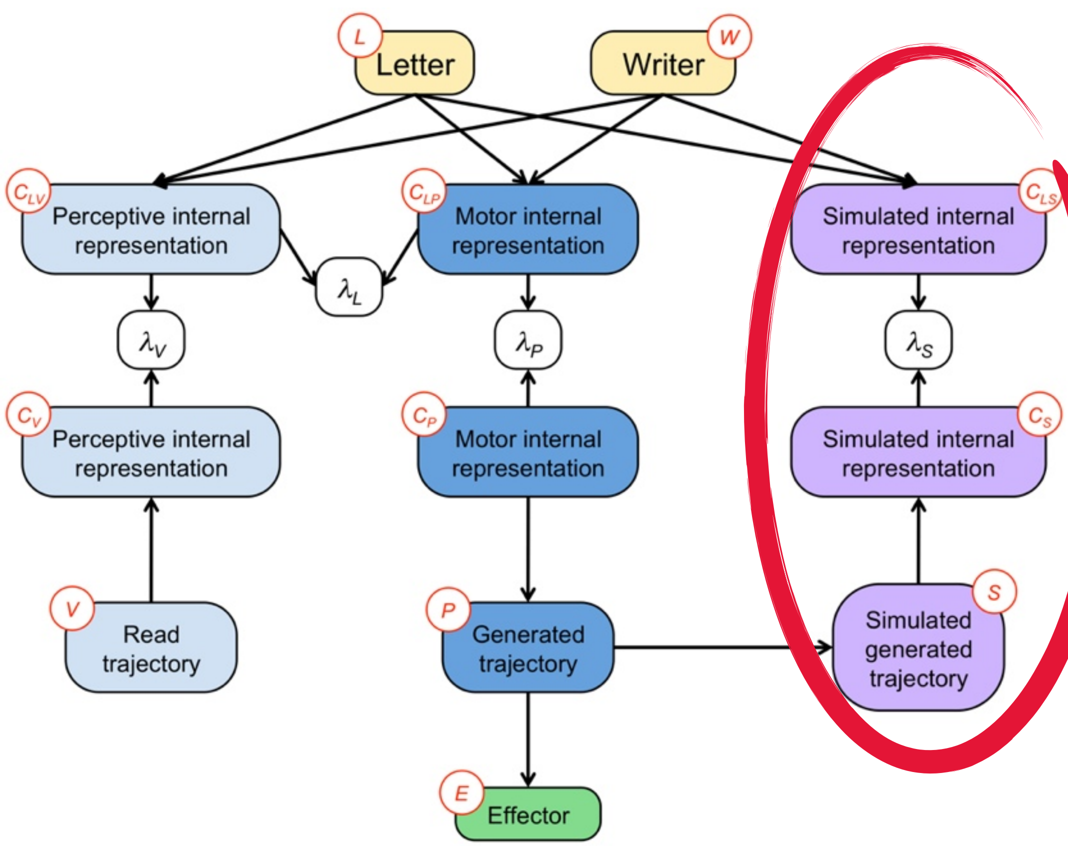
$$P\left(\begin{array}{c} C_{LV}\ C_{LP}\ C_{LS}\ C_V\ \lambda_P\ \lambda_V\ \lambda_S\ \lambda_L\ E \\ C_P\ C_S\ V\ P\ L\ W\ S \end{array}\right) = \left(\begin{array}{l} P(C_{LV} \mid L\ W)\ P(C_{LP} \mid L\ W)\ P(C_{LS} \mid L\ W) \\ P(L)\ P(W)\ P(\lambda_L \mid C_{LV}\ C_{LP}) \\ \boxed{P(\lambda_V \mid C_{LV}\ C_V)\ P(C_V \mid V)\ P(V)} \\ P(\lambda_P \mid C_{LP}\ C_P)\ P(P \mid C_P)\ P(E \mid P) \\ P(\lambda_S \mid C_{LS}\ C_S)\ P(C_S \mid S)\ P(S \mid P) \end{array}\right)$$

BAP MODEL



$$
\begin{aligned}
&P\left(\begin{matrix} C_{LV}\ C_{LP}\ C_{LS}\ C_V\ \lambda_P\ \lambda_V\ \lambda_S\ \lambda_L\ E \\ C_P\ C_S\ V\ P\ L\ W\ S \end{matrix}\right) \\
&= \left(\begin{matrix}
P(C_{LV} \mid L\ W)\, P(C_{LP} \mid L\ W)\, P(C_{LS} \mid L\ W) \\
P(L)\, P(W)\, P(\lambda_L \mid C_{LV}\ C_{LP}) \\
P(\lambda_V \mid C_{LV}\ C_V)\, P(C_V \mid V)\, P(V) \\
\boxed{P(\lambda_P \mid C_{LP}\ C_P)\, P(P \mid C_P)\, P(E \mid P)} \\
P(\lambda_S \mid C_{LS}\ C_S)\, P(C_S \mid S)\, P(S \mid P)
\end{matrix}\right)
\end{aligned}
$$

BAP MODEL

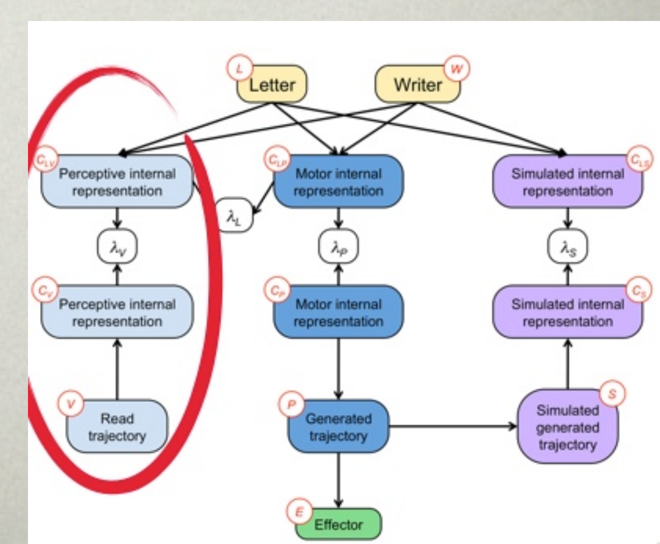


$$
\begin{aligned}
&P\left(\begin{array}{l} C_{LV}\ C_{LP}\ C_{LS}\ C_V\ \lambda_P\ \lambda_V\ \lambda_S\ \lambda_L\ E \\ C_P\ C_S\ V\ P\ L\ W\ S \end{array}\right) \\
&= \left(\begin{array}{l}
P(C_{LV} \mid L\ W)\ P(C_{LP} \mid L\ W)\ P(C_{LS} \mid L\ W) \\
P(L)\ P(W)\ P(\lambda_L \mid C_{LV}\ C_{LP}) \\
P(\lambda_V \mid C_{LV}\ C_V)\ P(C_V \mid V)\ P(V) \\
P(\lambda_P \mid C_{LP}\ C_P)\ P(P \mid C_P)\ P(E \mid P) \\
\boxed{P(\lambda_S \mid C_{LS}\ C_S)\ P(C_S \mid S)\ P(S \mid P)}
\end{array}\right)
\end{aligned}
$$

Letter recognition

knowing the scripter

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1])$$



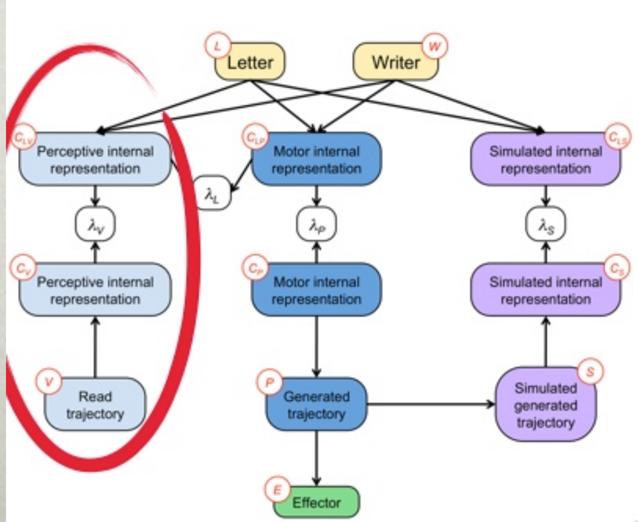
Letter recognition

knowing the scripter

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1])$$

$$\propto \begin{pmatrix} P([C_{LVx}^0 = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LVy}^0 = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LV\dot{x}}^0 = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LV\dot{y}}^0 = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \end{pmatrix}$$

$$\prod_{n=1}^{N} \begin{pmatrix} P([C_{LVx}^n = f(v_x^{0:M}, v_y^{0:M})] \mid [C_{LVx}^{n-1} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \\ P([C_{LVy}^n = f(v_x^{0:M}, v_y^{0:M})] \mid [C_{LVy}^{n-1} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \\ P([C_{LV\dot{x}}^n = f(v_x^{0:M}, v_y^{0:M})] \mid [C_{LV\dot{x}}^{n-1} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \\ P([C_{LV\dot{y}}^n = f(v_x^{0:M}, v_y^{0:M})] \mid [C_{LV\dot{y}}^{n-1} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \end{pmatrix}$$

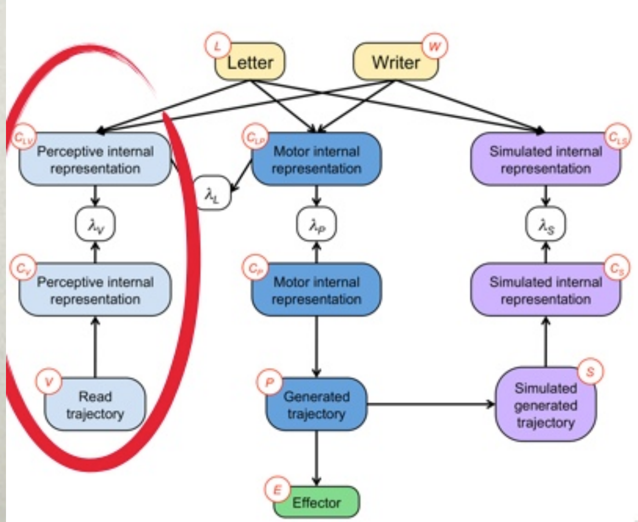


Letter recognition

knowing the scripter

	a	b	c	d	e	f	g	h	k	l	m	n	o	p	q	r	s	u	v	w	y	z	%
a	0.95	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.05
b	0	0.72	0	0	0.05	0	0	0.12	0.03	0.05	0	0	0	0	0	0	0	0	0	0	0	0	0.03
c	0	0	0.92	0	0	0	0	0	0	0	0	0	0.05	0	0	0	0	0	0	0	0	0	0.03
d	0.03	0	0	0.94	0	0	0	0	0	0	0	0	0	0	0.03	0	0	0	0	0	0	0	0
e	0	0	0	0	0.87	0	0	0	0	0	0	0	0	0	0	0.10	0	0	0	0	0	0	0.03
f	0	0	0	0	0	0.97	0	0.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
g	0	0	0	0	0	0	0.90	0	0	0	0	0	0	0	0.10	0	0	0	0	0	0	0	0
h	0	0.03	0	0	0	0.03	0	0.91	0	0.03	0	0	0	0	0	0	0	0	0	0	0	0	0
k	0	0	0	0	0	0	0	0	0.97	0.03	0	0	0	0	0	0	0	0	0	0	0	0	0
l	0	0.10	0	0	0	0	0	0.08	0	0.82	0	0	0	0	0	0	0	0	0	0	0	0	0
m	0	0	0	0	0	0	0	0	0	0	0.97	0	0	0	0	0	0	0	0	0.03	0	0	0
n	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
o	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
p	0	0	0	0	0	0	0	0	0	0	0	0	0	0.94	0	0	0.03	0	0	0	0	0	0.03
q	0	0	0	0	0	0	0.15	0	0	0	0	0	0	0	0.85	0	0	0	0	0	0	0	0
r	0	0	0	0	0	0	0	0	0	0	0	0.03	0	0.03	0	0.86	0	0.05	0.03	0	0	0	0
s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
u	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
v	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
z	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.95	0.05

93,36%

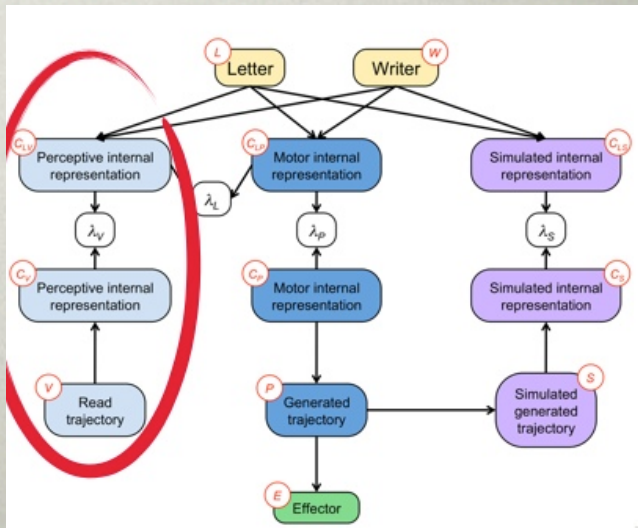


Scripter recognition
knowing the letter

$$P(W \mid [V_X^{0:M} = v_x^{0:M}]\,[V_Y^{0:M} = v_y^{0:M}]\,[L = l]\,[\lambda_V = 1])$$

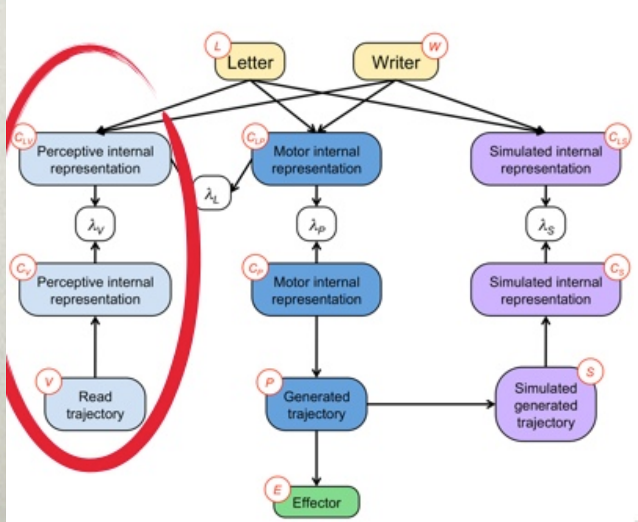
	Estelle	Julienne	Jean-Louis	Christophe
Estelle	0.76	0.03	0.07	0.14
Julienne	0.02	0.80	0.07	0.11
Jean-Louis	0	0	1	0
Christophe	0.10	0.14	0.13	0.62

79,5%



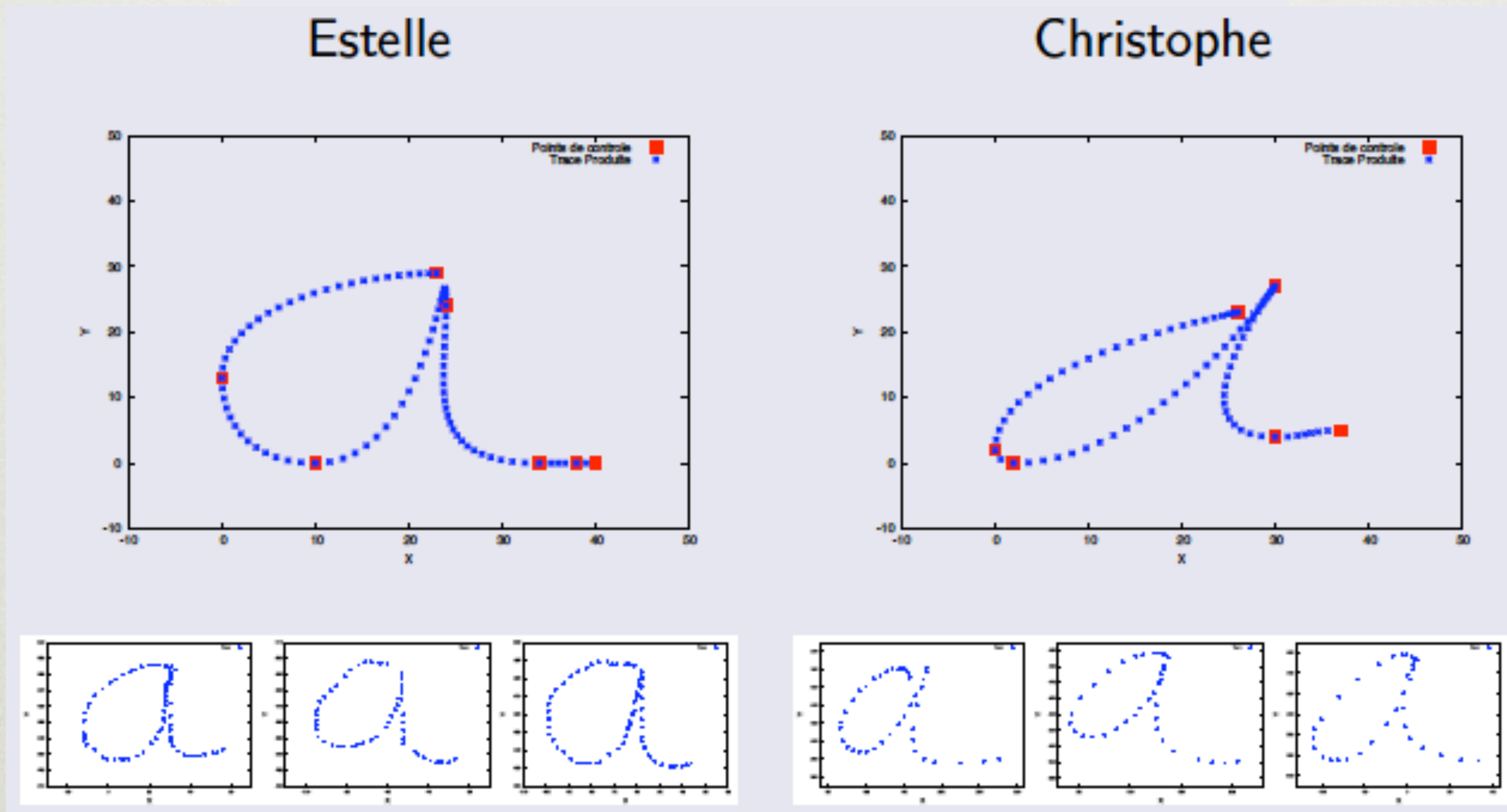
Motor control

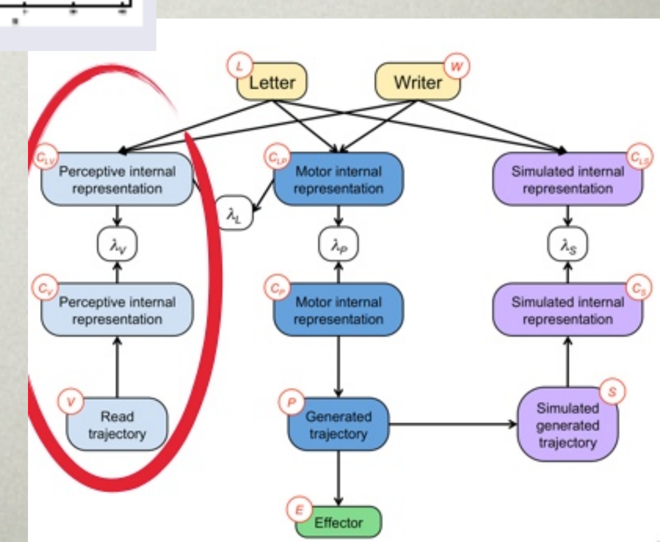
$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1])$$



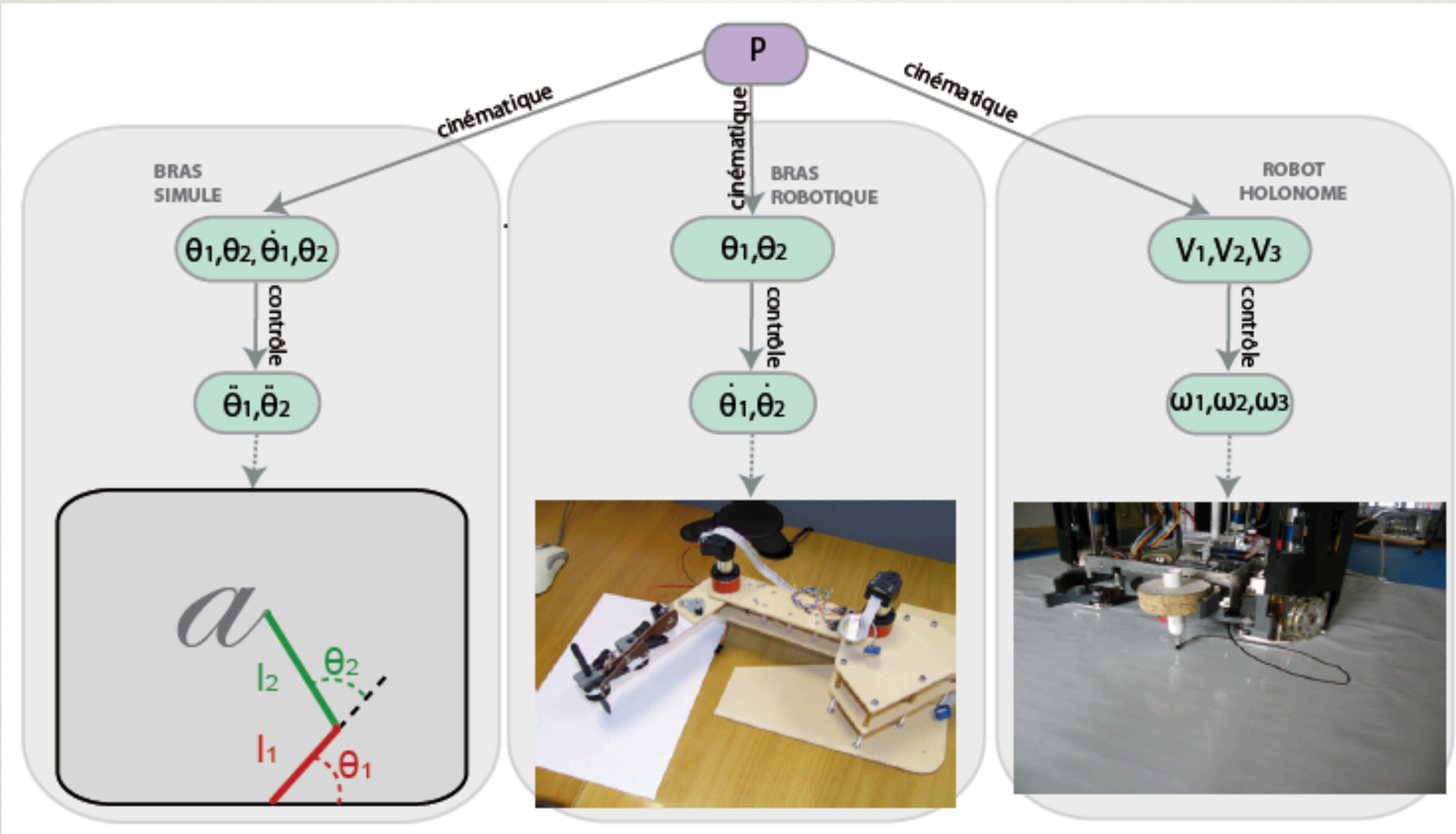
Motor control
Inter scripter variability

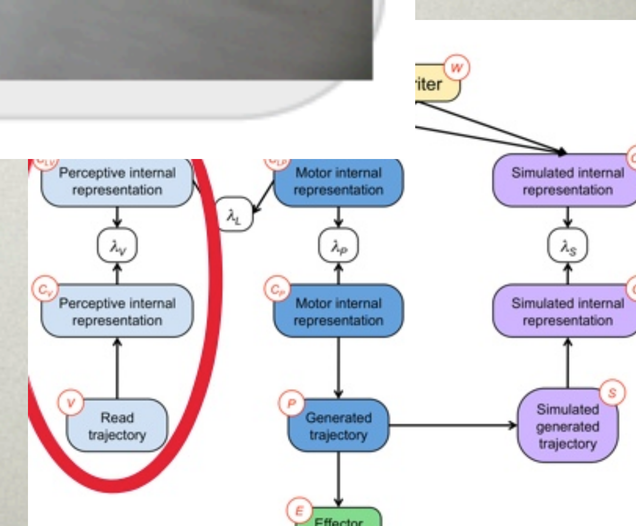
$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1])$$



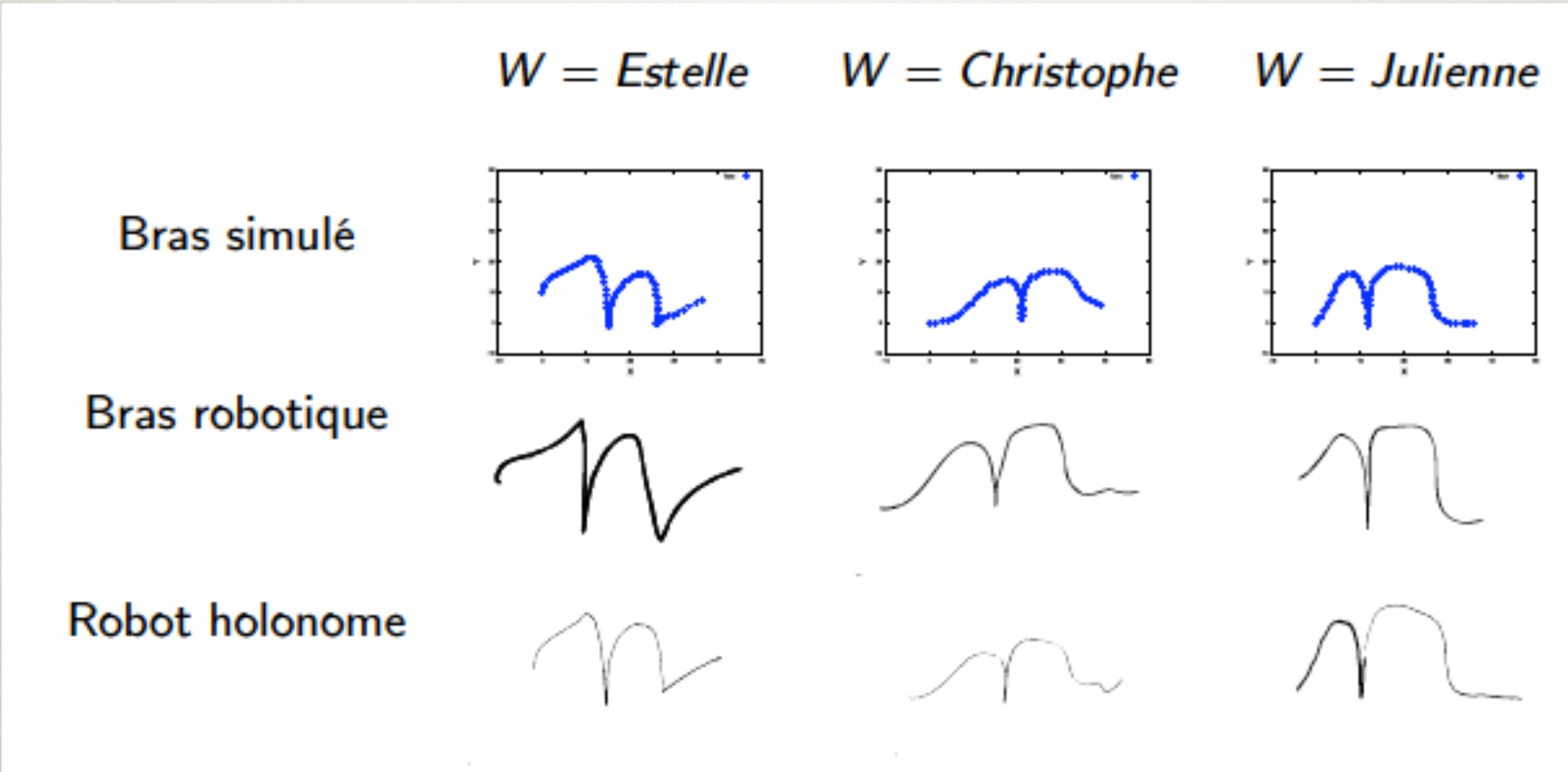


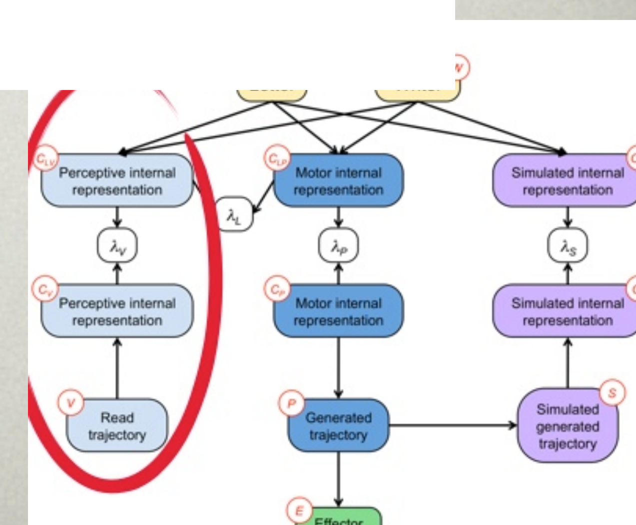
Motor equivalence





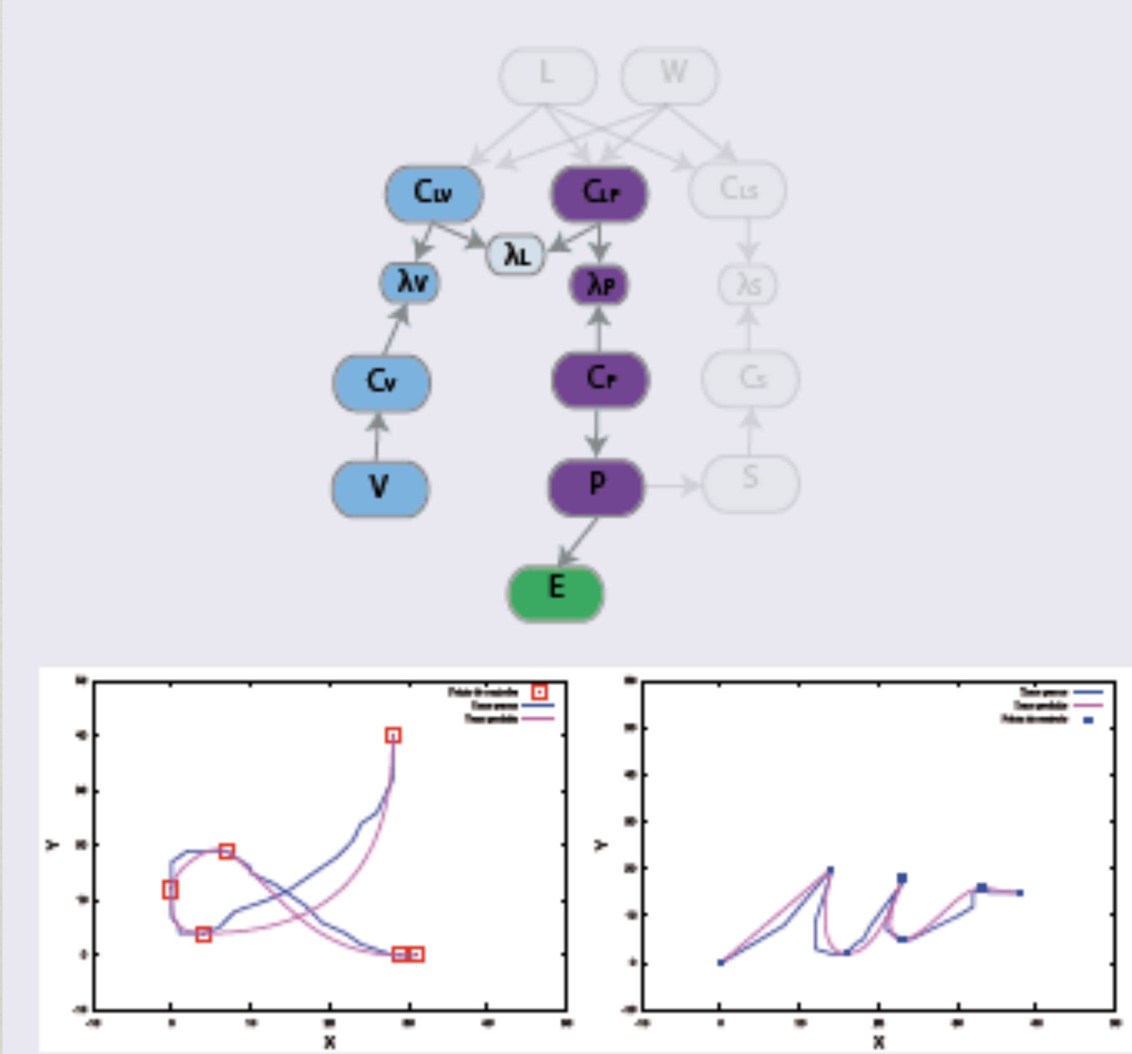
Motor equivalence



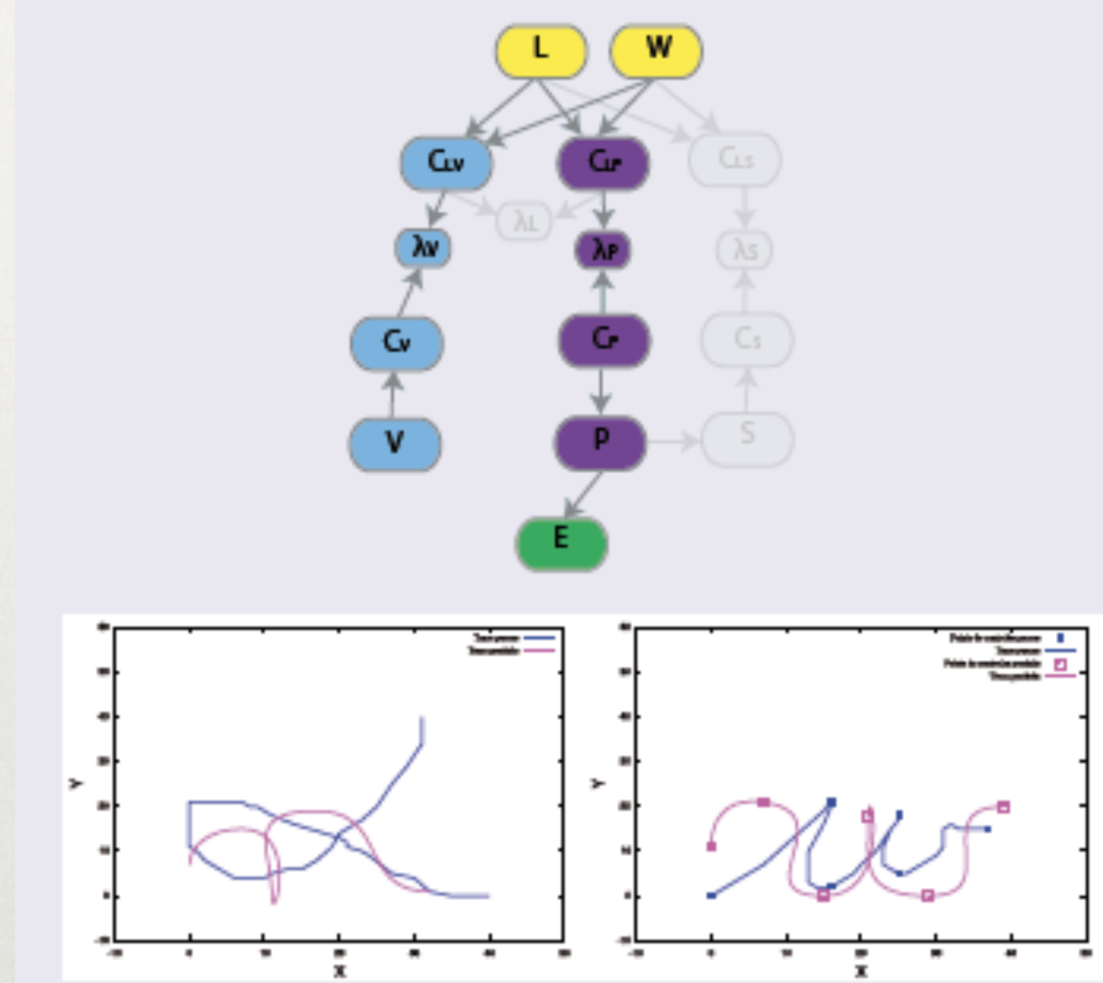


Copy

$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid V_x^{0:M}\ V_y^{0:M}\ [\lambda_V = 1]\ [\lambda_l = 1]\ [\lambda_P = 1])$

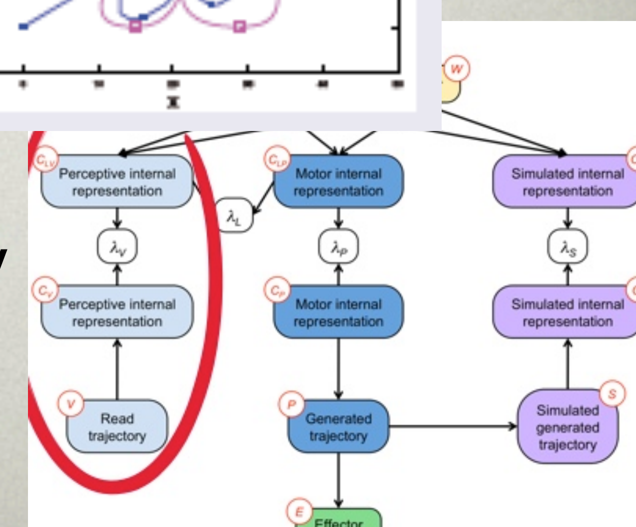


$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid V_x^{0:M}\ V_y^{0:M}\ [\lambda_V = 1]\ [\lambda_P = 1]\ [W = w])$



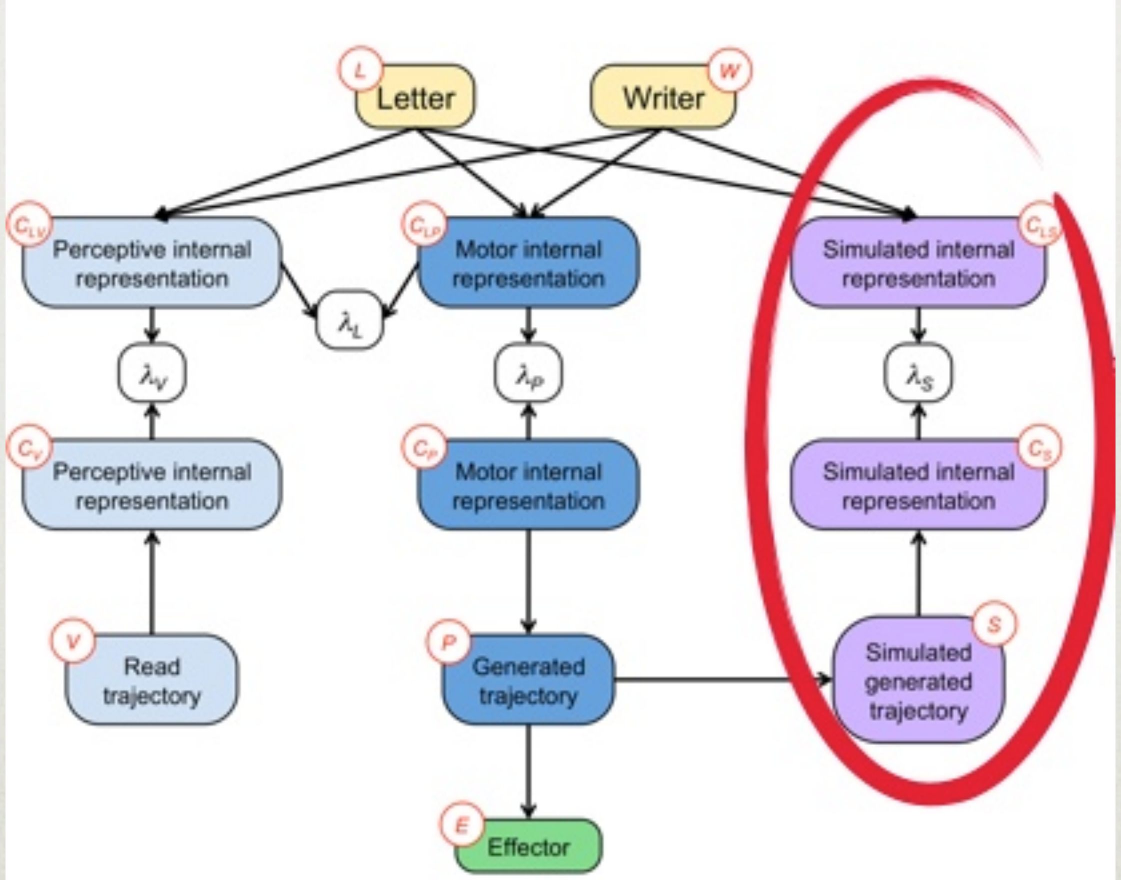
Trace copy

Letter copy

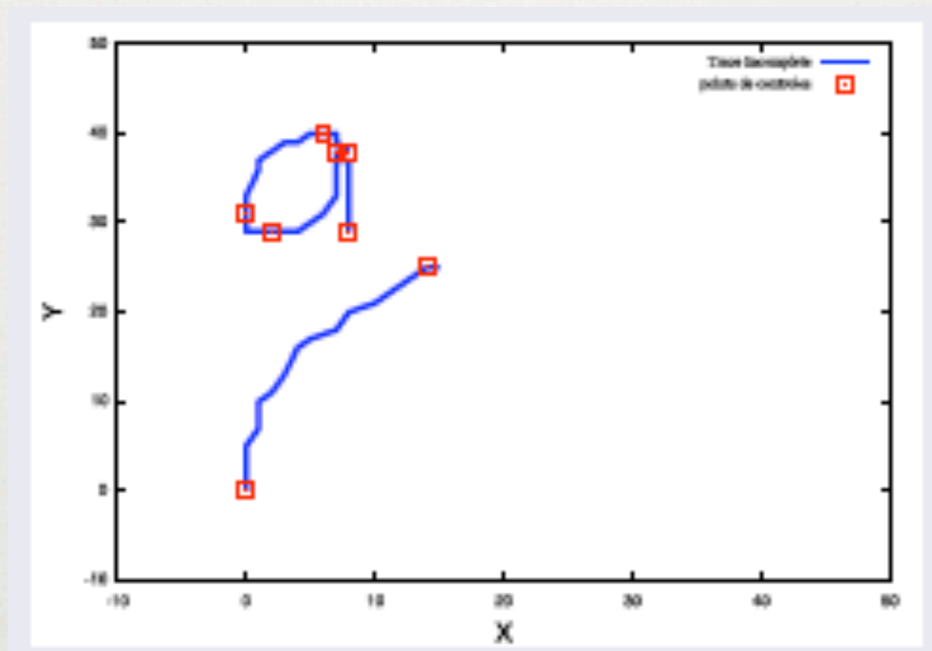


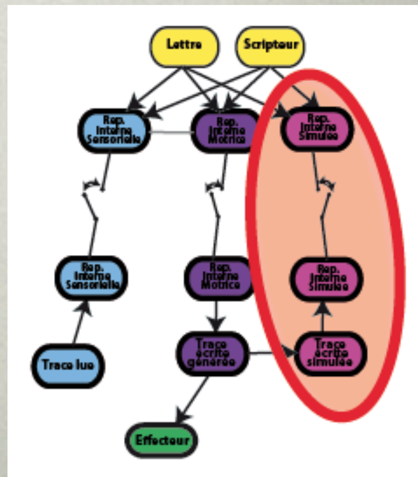
Letter recognition with motor simulation

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\,[V_Y^{0:M} = v_y^{0:M}]\,[W = w]\,[\lambda_V = 1]\,[\lambda_L = 1]\,[\lambda_P = 1]\,[\lambda_S = 1])$$

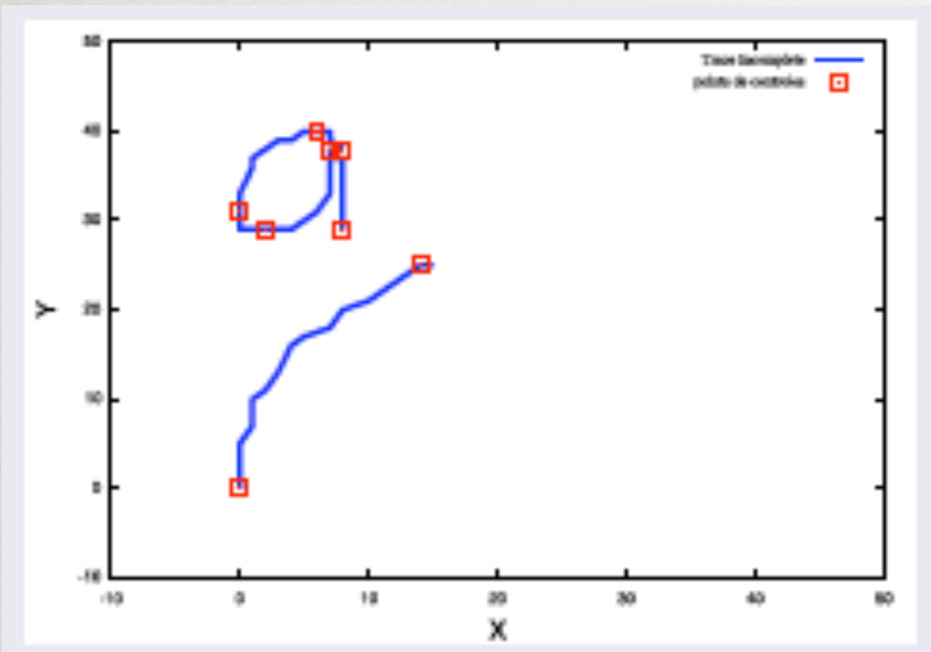


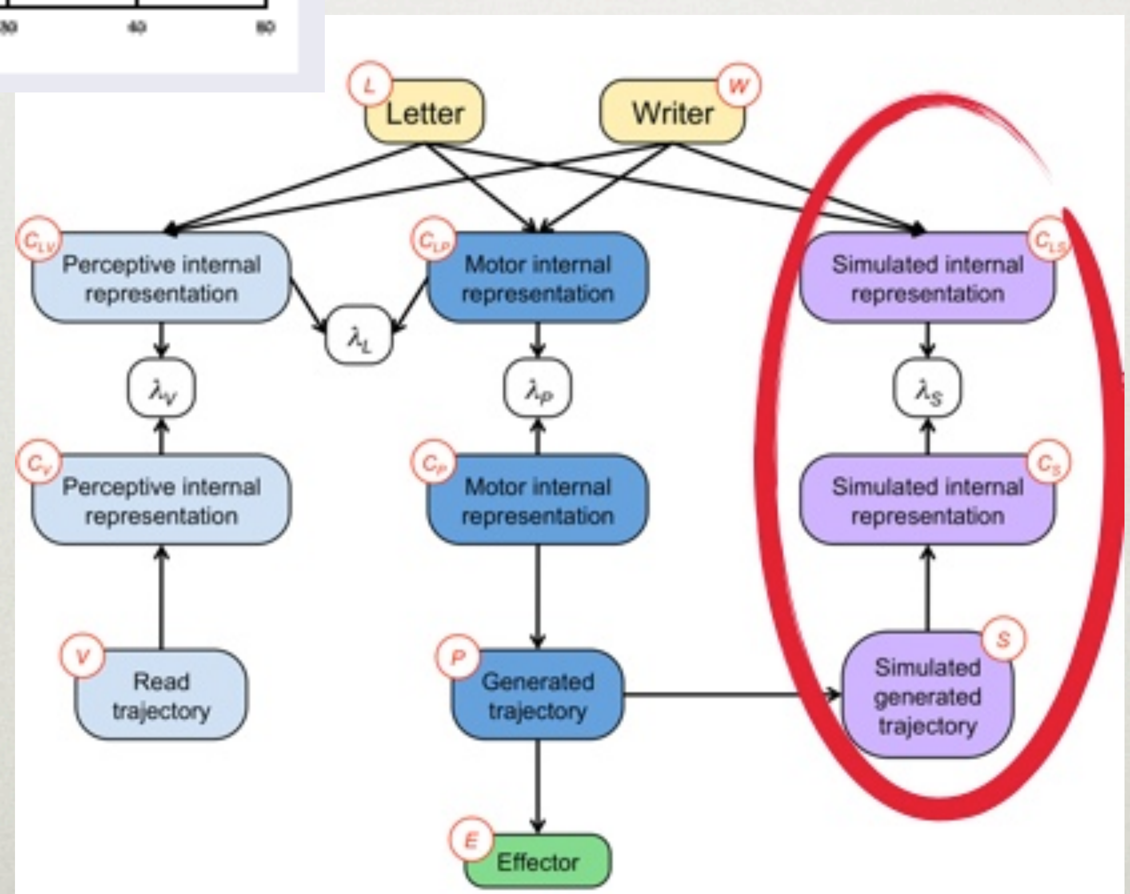
Letter recognition with motor simulation

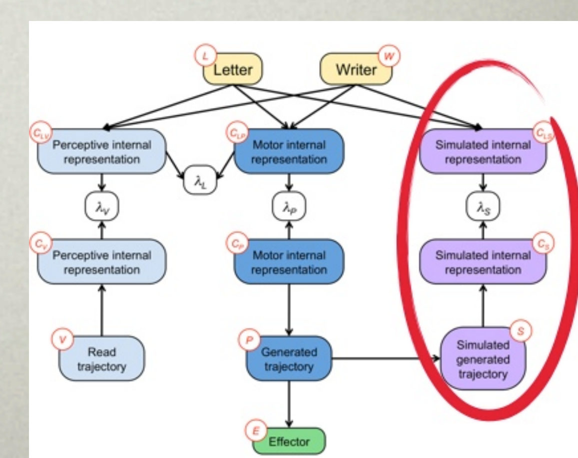




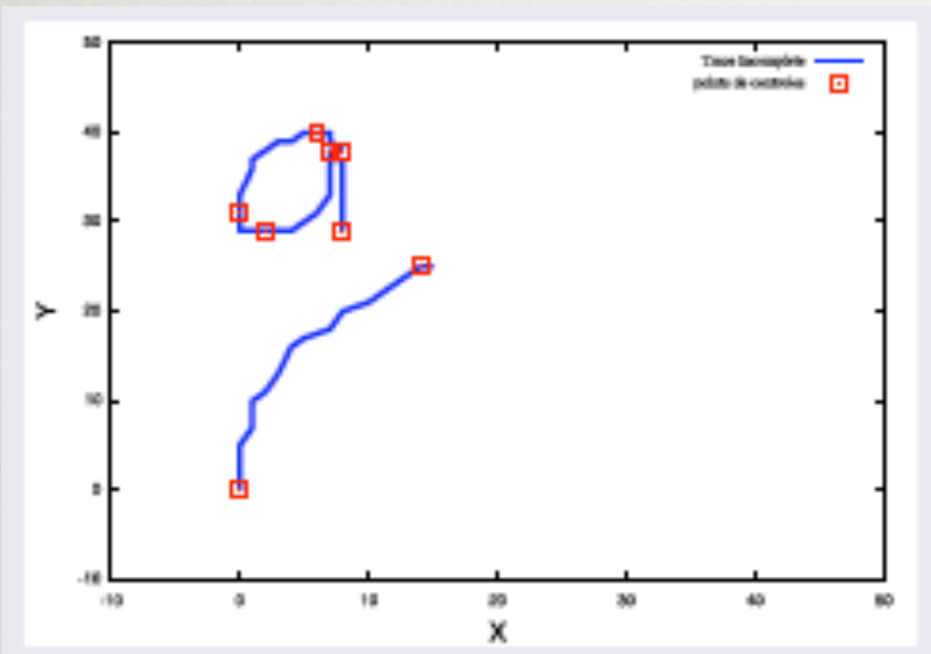
Letter recognition with motor simulation

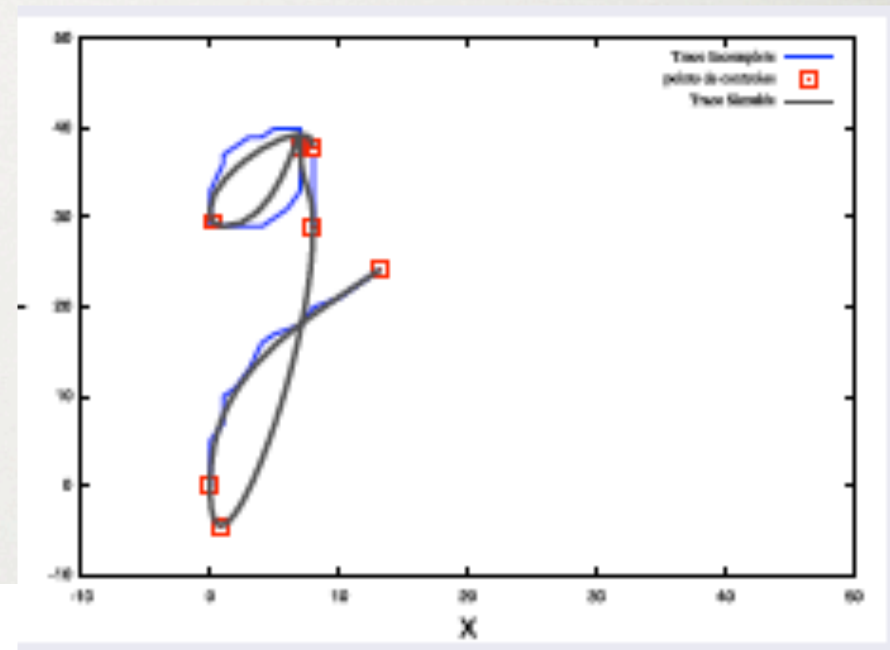


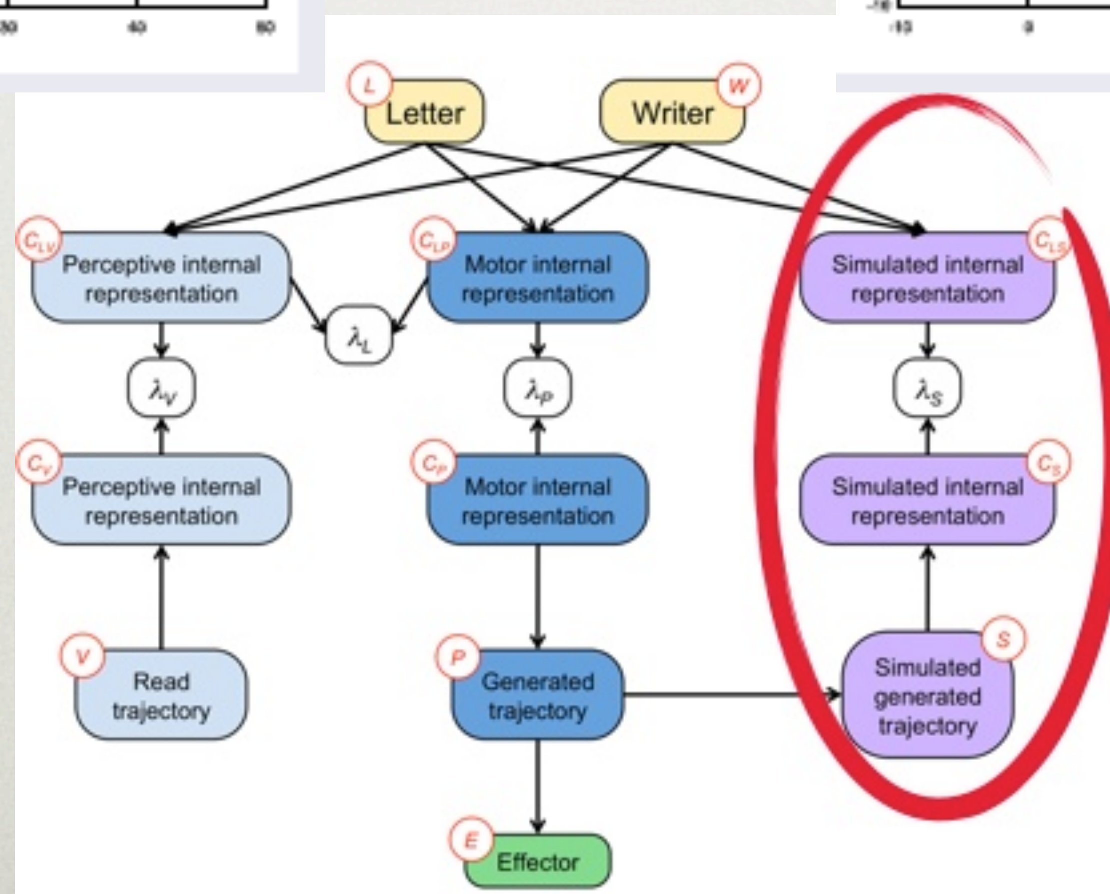


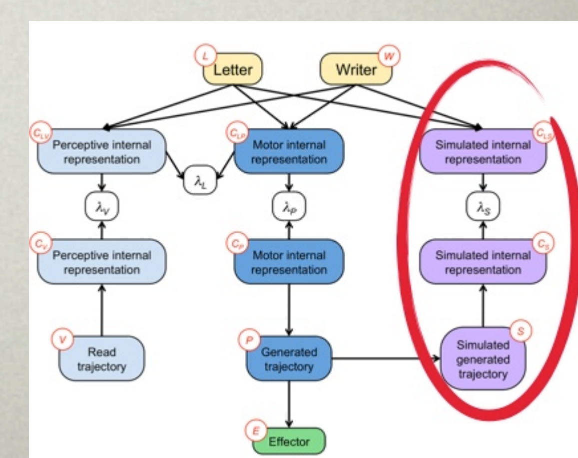


Letter recognition with motor simulation

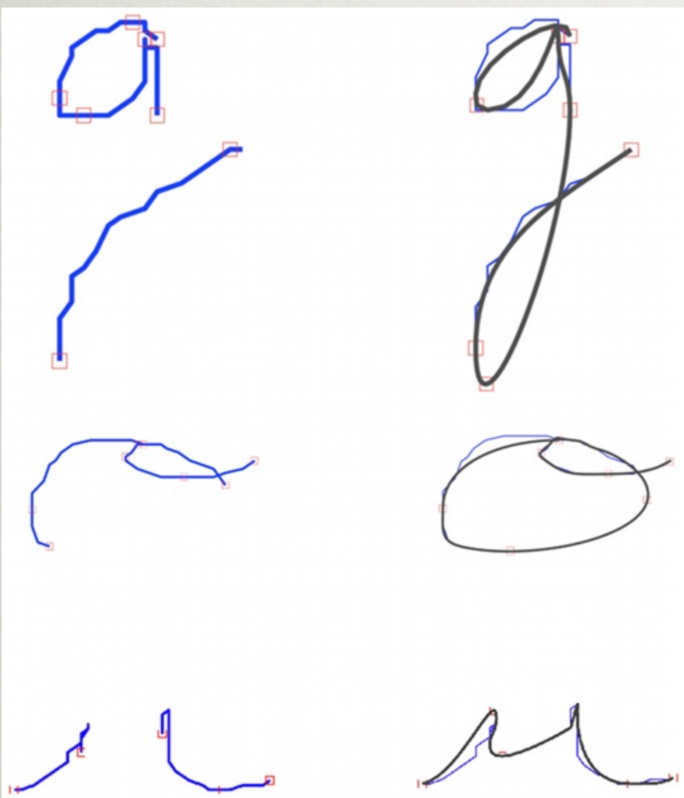






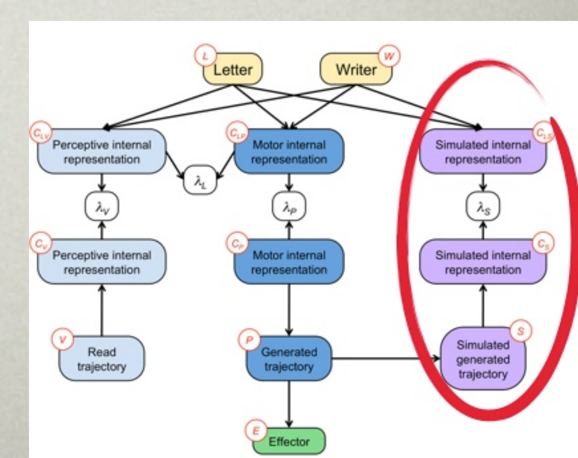


Results



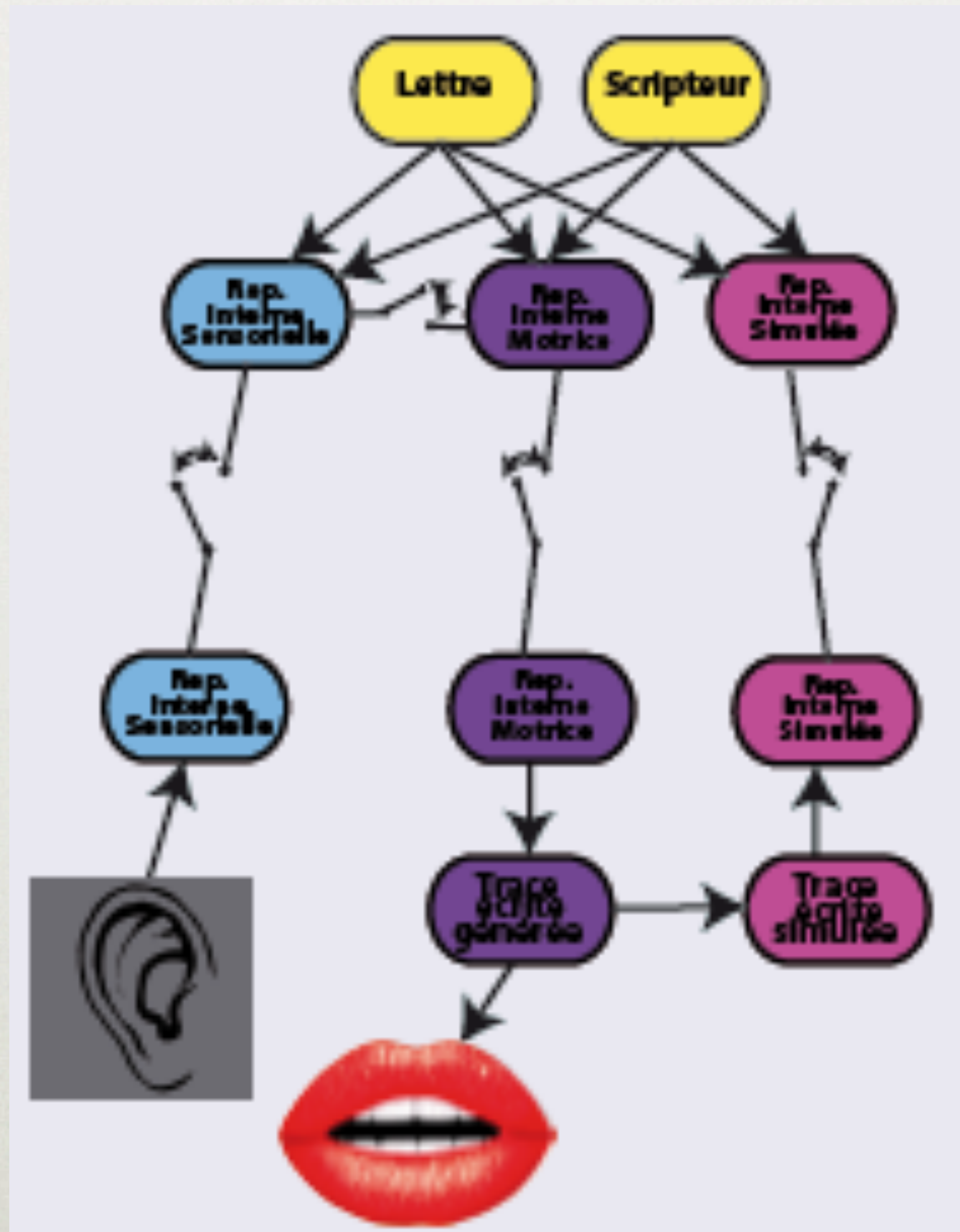
	f	g	h	k	l	m	n	o	p	q	r
With motor simulation	0	0	0	0	0	0	0	0	0	1	0
Without motor simulation	0	0.9	0	0	0	0	0	0	0	0.1	0

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)

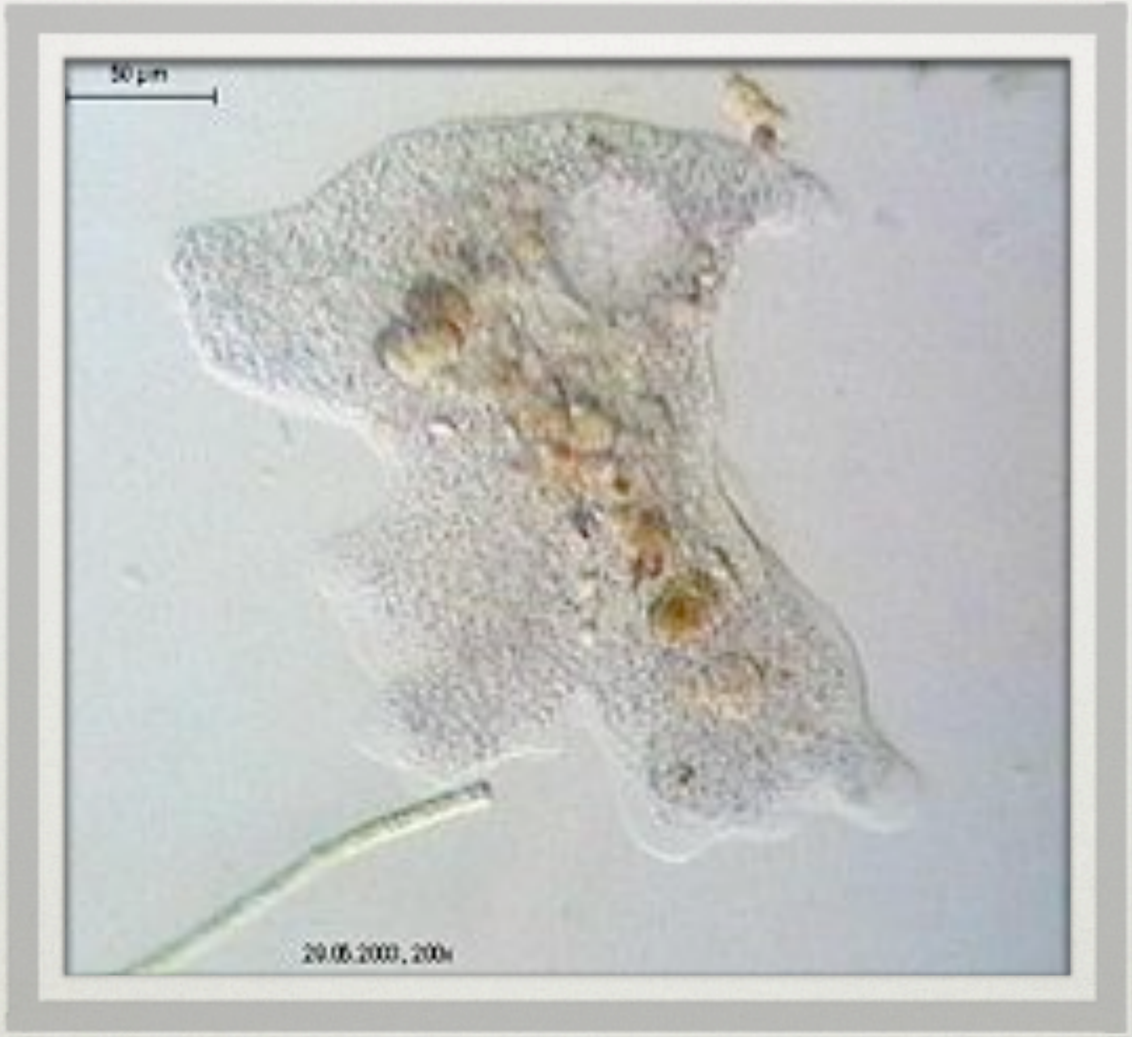


Overview

- 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?

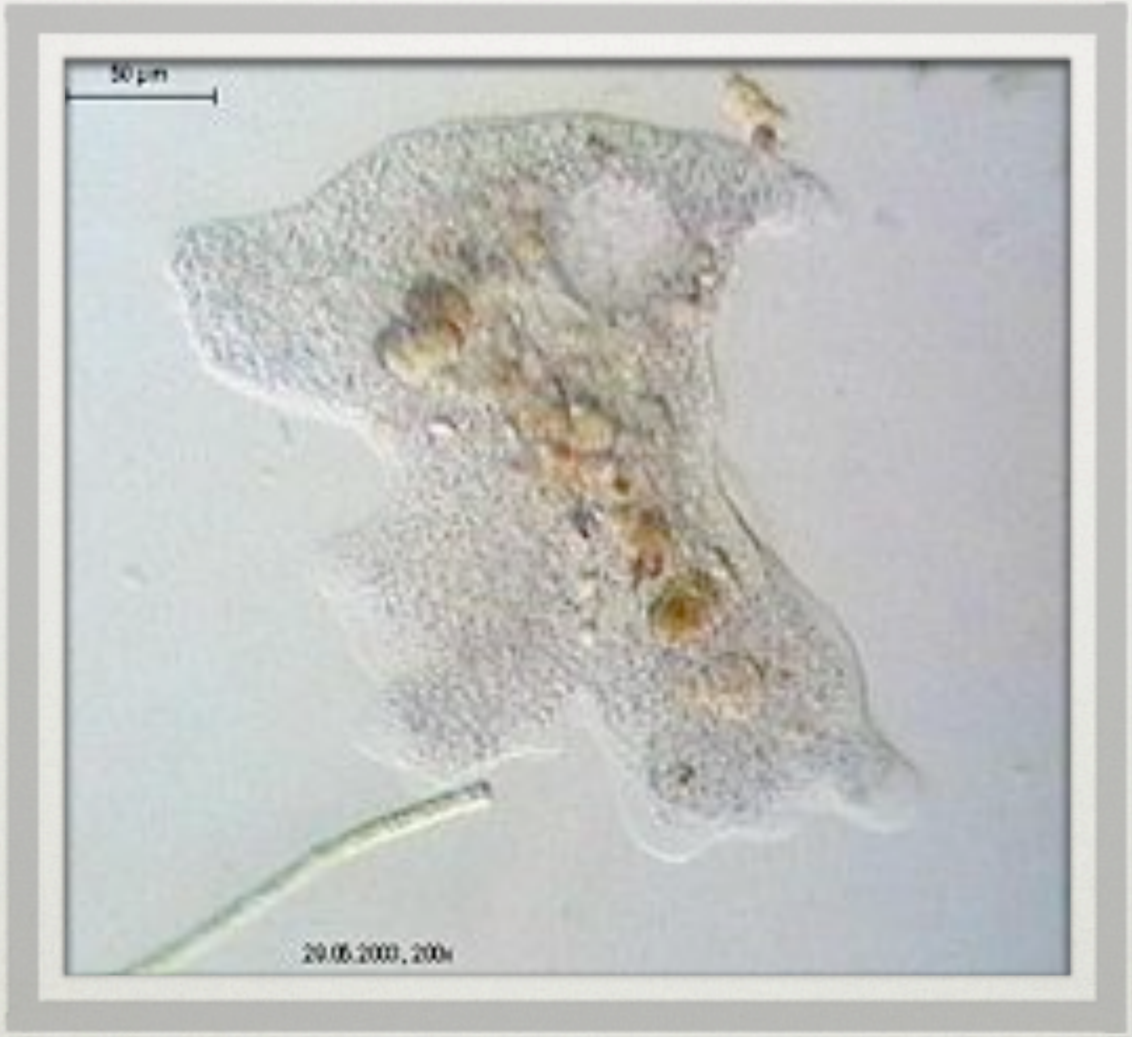
Amoeba

How is it performing probabilistic inference?



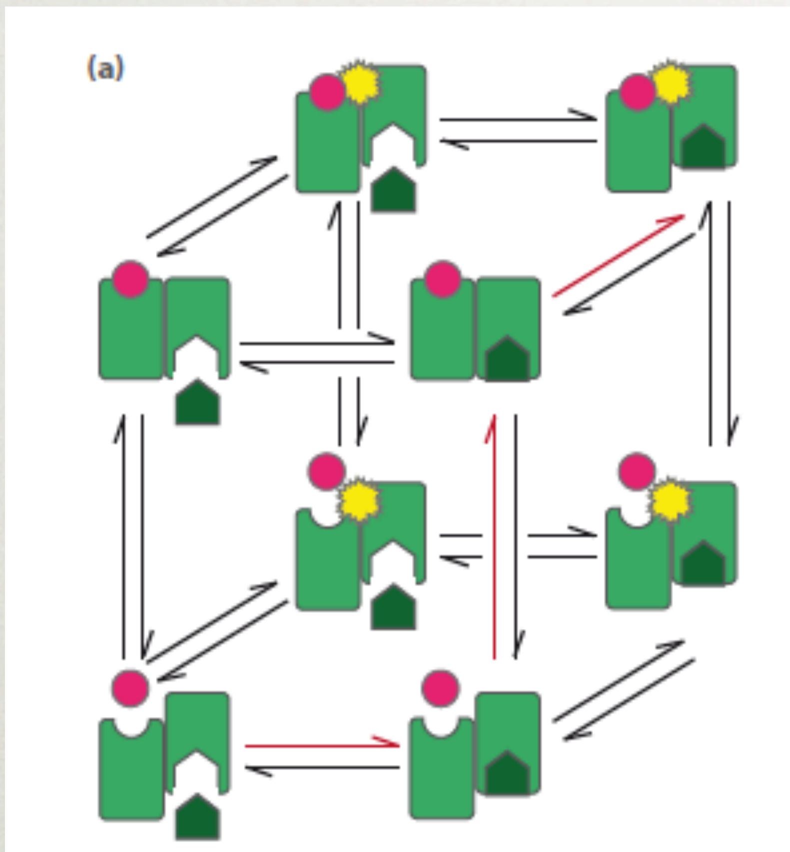
Amoeba

How is it performing probabilistic inference?



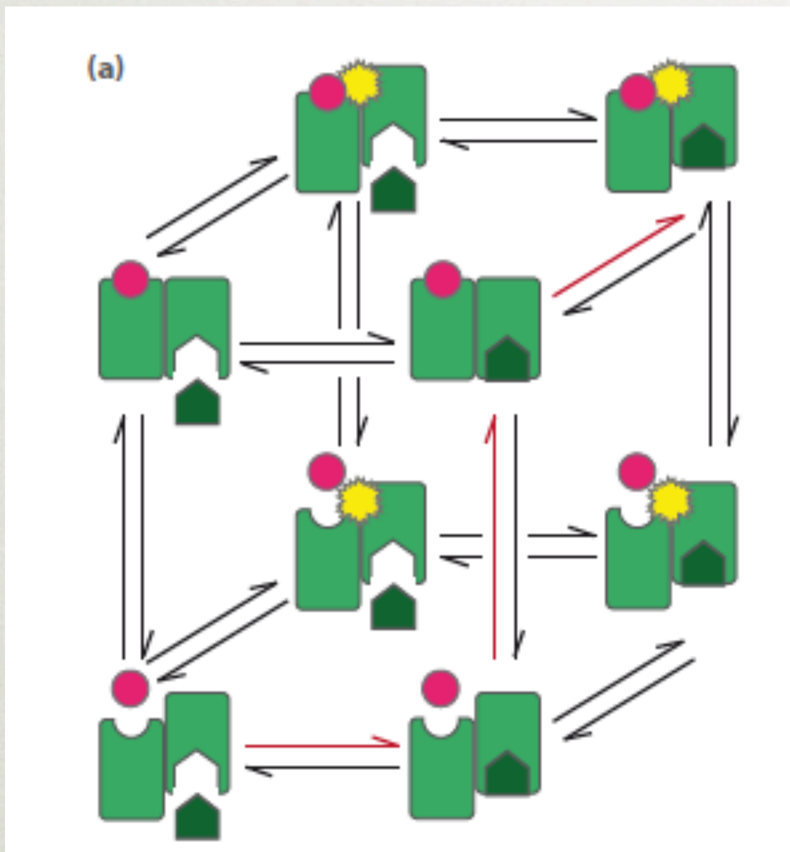
Cell signaling

8 allosteric states
2 messengers



$S_{000} + M_1 \rightarrow S_{100}, p_{000 \rightarrow 100}$	$S_{100} \rightarrow S_{000} + M_1, p_{100 \rightarrow 000}$
$S_{000} + M_2 \rightarrow S_{010}, p_{000 \rightarrow 010}$	$S_{010} \rightarrow S_{000} + M_2, p_{010 \rightarrow 000}$
$S_{010} + M_1 \rightarrow S_{110}, p_{010 \rightarrow 110}$	$S_{110} \rightarrow S_{010} + M_1, p_{110 \rightarrow 010}$
$S_{100} + M_2 \rightarrow S_{110}, p_{100 \rightarrow 110}$	$S_{110} \rightarrow S_{100} + M_2, p_{110 \rightarrow 100}$
$S_{001} + M_1 \rightarrow S_{101}, p_{001 \rightarrow 101}$	$S_{101} \rightarrow S_{001} + M_1, p_{101 \rightarrow 001}$
$S_{001} + M_2 \rightarrow S_{011}, p_{001 \rightarrow 011}$	$S_{011} \rightarrow S_{001} + M_2, p_{011 \rightarrow 001}$
$S_{011} + M_1 \rightarrow S_{111}, p_{011 \rightarrow 111}$	$S_{111} \rightarrow S_{011} + M_1, p_{111 \rightarrow 011}$
$S_{101} + M_2 \rightarrow S_{111}, p_{101 \rightarrow 111}$	$S_{111} \rightarrow S_{101} + M_2, p_{111 \rightarrow 101}$
$S_{000} \rightarrow S_{001}, p_{000 \rightarrow 001}$	$S_{001} \rightarrow S_{000}, p_{001 \rightarrow 000}$
$S_{010} \rightarrow S_{011}, p_{010 \rightarrow 011}$	$S_{011} \rightarrow S_{010}, p_{011 \rightarrow 010}$
$S_{100} \rightarrow S_{101}, p_{100 \rightarrow 101}$	$S_{101} \rightarrow S_{100}, p_{101 \rightarrow 100}$
$S_{110} \rightarrow S_{111}, p_{100 \rightarrow 111}$	$S_{111} \rightarrow S_{110}, p_{111 \rightarrow 110}$

8 allosteric states
2 messengers



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

$$P([\Omega_1=0]\wedge[\Omega_2=1]\wedge[\Omega_3=0]) = \frac{k_{000\to 010}\times m_2}{D}$$

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

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

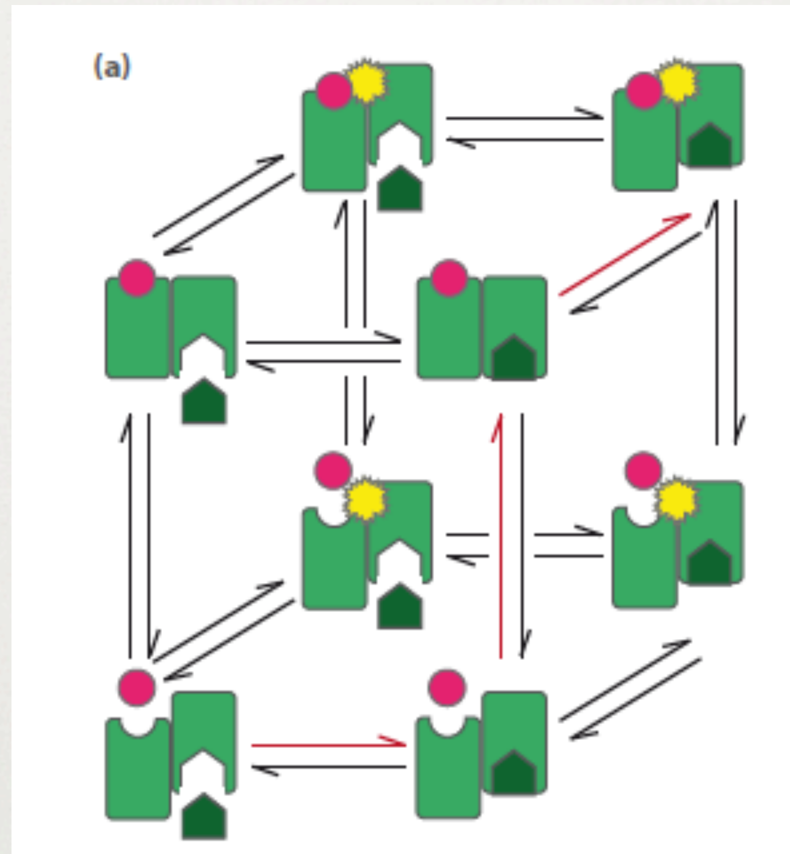
$$P([\Omega_1=0]\wedge[\Omega_2=0]\wedge[\Omega_3=1]) = \frac{k_{000\to 001}}{D}$$

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

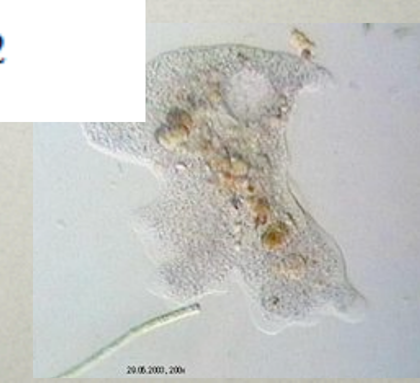
$$P([\Omega_1=1]\wedge[\Omega_2=0]\wedge[\Omega_3=1]) = \frac{k_{000\to 100}\times k_{100\to 101}\times m_1}{D}$$

$$P([\Omega_1=1]\wedge[\Omega_2=1]\wedge[\Omega_3=1]) = \frac{k_{000\to 010}\times k_{010\to 110}\times k_{110\to 111}\times m_1\times m_2}{D}$$

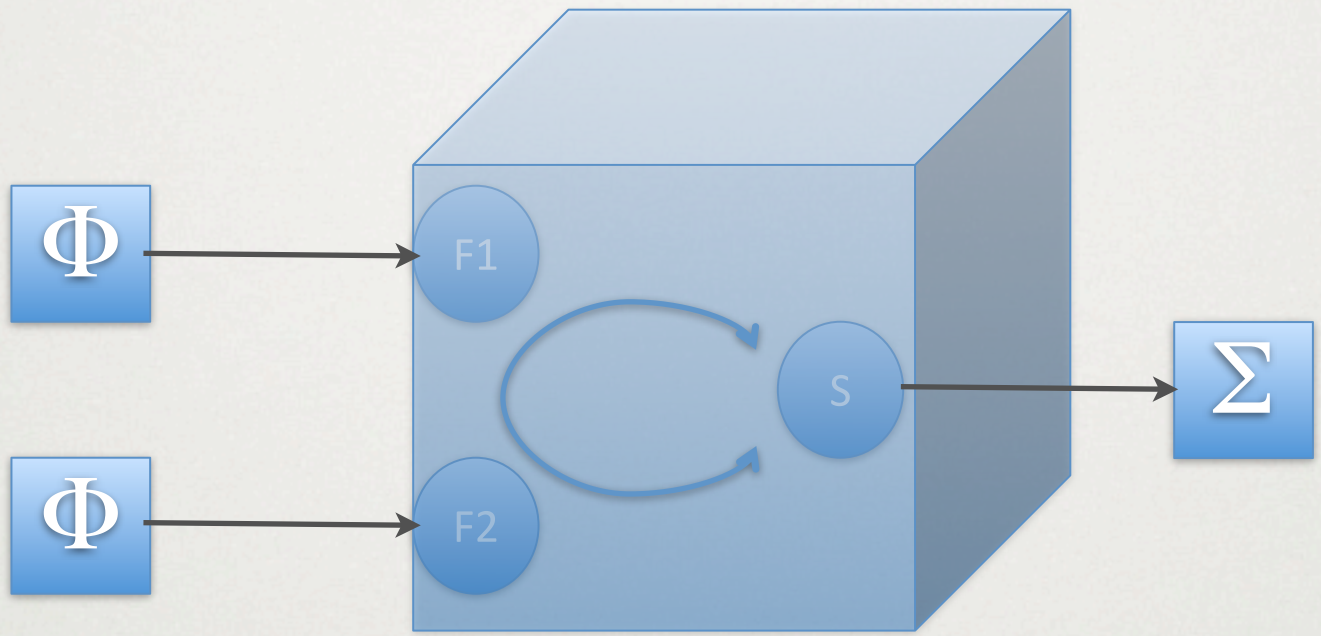
8 allosteric states
2 messengers



$$O([\Omega_3]) = \frac{\left(\begin{array}{c} k_{000\rightarrow 001} + k_{000\rightarrow 100} \times k_{100\rightarrow 101} \times m_1 \\ + k_{000\rightarrow 010} \times k_{010\rightarrow 011} \times m_2 + k_{000\rightarrow 010} \times k_{010\rightarrow 110} \times k_{110\rightarrow 111} \times m_1 \times m_2 \end{array}\right)}{1 + k_{000\rightarrow 100} \times m_1 + k_{000\rightarrow 010} \times m_2 + k_{000\rightarrow 010} \times k_{010\rightarrow 110} \times m_1 \times m_2}$$



Bayesian gate



$$\Sigma = \frac{P\left(\left[S=1\right] \mid \phi 1 \phi 2 \left[\lambda 1 = 1\right]\left[\lambda 2 = 1\right]\right)}{P\left(\left[S=0\right] \mid \phi 1 \phi 2 \left[\lambda 1 = 1\right]\left[\lambda 2 = 1\right]\right)}$$

$$= \frac{P\left(\left[f1=0\right]\left[f2=0\right]\left[S=1\right]\right) + P(011)\phi 2 + P(101)\phi 1 + P(111)\phi 1 \phi 2}{P(000) + P(010)\phi 2 + P(100)\phi 1 + P(110)\phi 1 \phi 2}$$

Bayesian gate

$$O([\Omega_3]) = \frac{\begin{pmatrix} k_{000\to001} + k_{000\to100} \times k_{100\to101} \times m_1 \\ + k_{000\to010} \times k_{010\to011} \times m_2 + k_{000\to010} \times k_{010\to110} \times k_{110\to111} \times m_1 \times m_2 \end{pmatrix}}{1 + k_{000\to100} \times m_1 + k_{000\to010} \times m_2 + k_{000\to010} \times k_{010\to110} \times m_1 \times m_2}$$

$$\Sigma = \frac{P([S=1] \mid \phi1\phi2[\lambda1=1][\lambda2=1])}{P([S=0] \mid \phi1\phi2[\lambda1=1][\lambda2=1])}$$

$$= \frac{P([f1=0][f2=0][S=1]) + P(011)\phi2 + P(101)\phi1 + P(111)\phi1\phi2}{P(000) + P(010)\phi2 + P(100)\phi1 + P(110)\phi1\phi2}$$

Bayesian biochemistry: Basic ideas

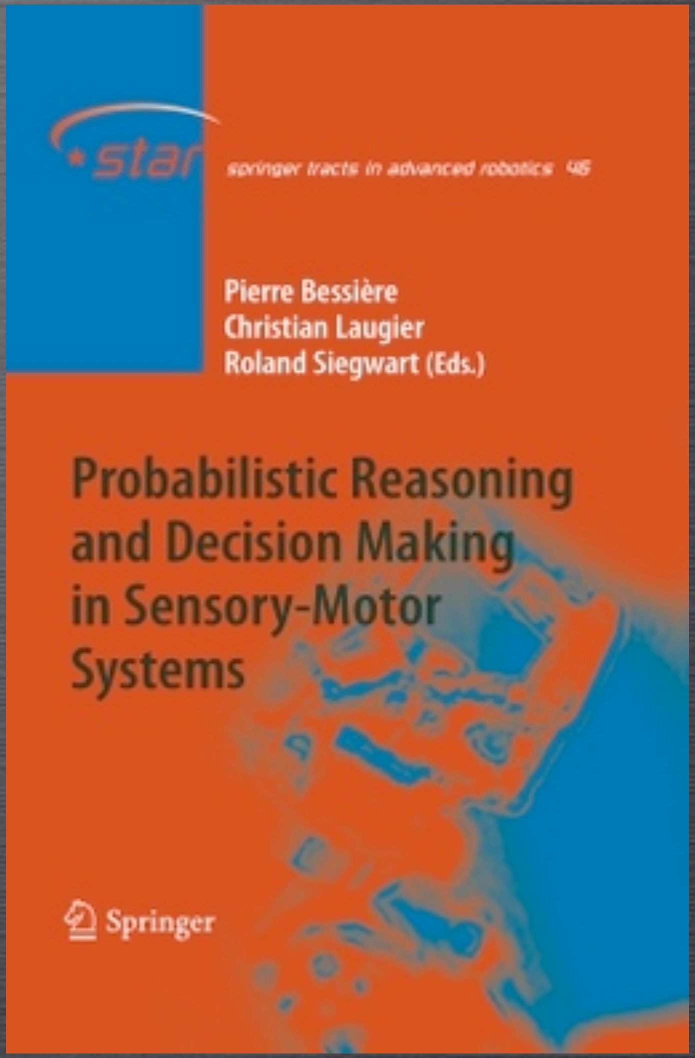
- Bayesian values -> Concentration of messengers, membrane potential & spike frequency
- Bayesian gates -> equilibrium between allosteric macromolecules & messengers
- Bayesian inference -> signal transduction.
- The interplay between local biochemical mechanisms and distant electrical propagation in 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