Agent Newspeak

Common Concepts in Agent Groups, Symmetries and Conformity in a Simple Environment

Marco Möller

Adaptive Systems Research Group
University of Hertfordshire

27.02.2008

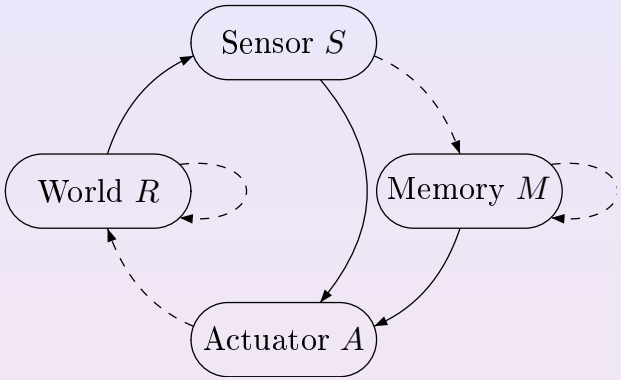
Content

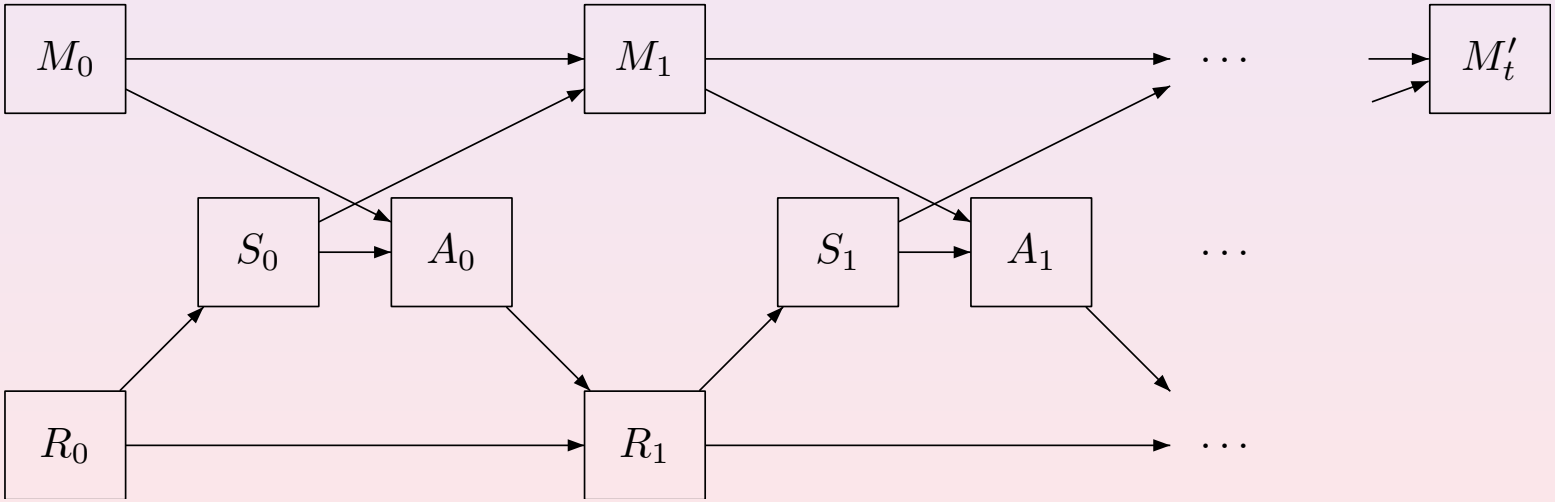
1. Scenario
   - Perception Action Loop
   - Initial Position Capturing
2. Common Concept
   - Motivation
   - Objective Common Concept
   - Subjective Common Concept
   - Results
3. Symmetry
   - Idea
   - Extrinsic Symmetry
   - Intrinsic Symmetry
   - Results

Overview

1. Scenario
   - Perception Action Loop
   - Initial Position Capturing
2. Common Concept
3. Symmetry

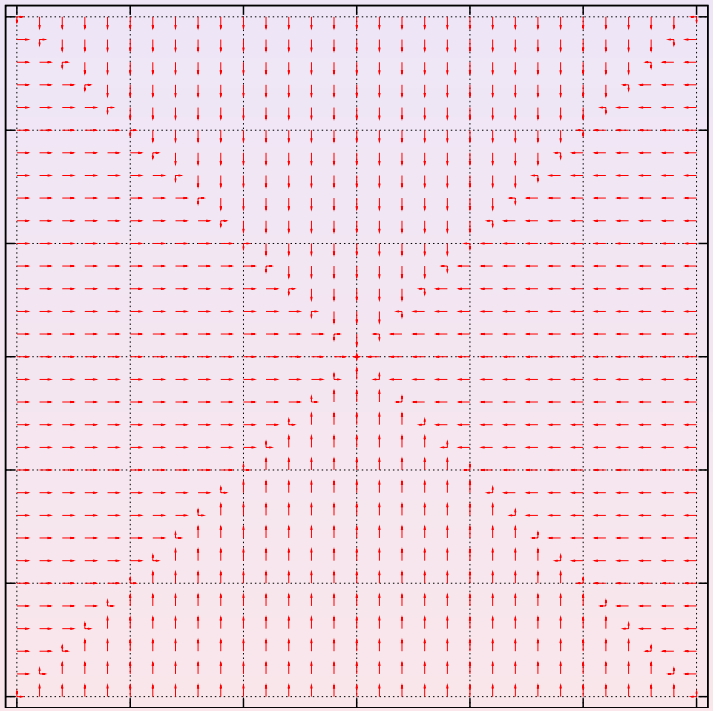
Perception-Action Loop





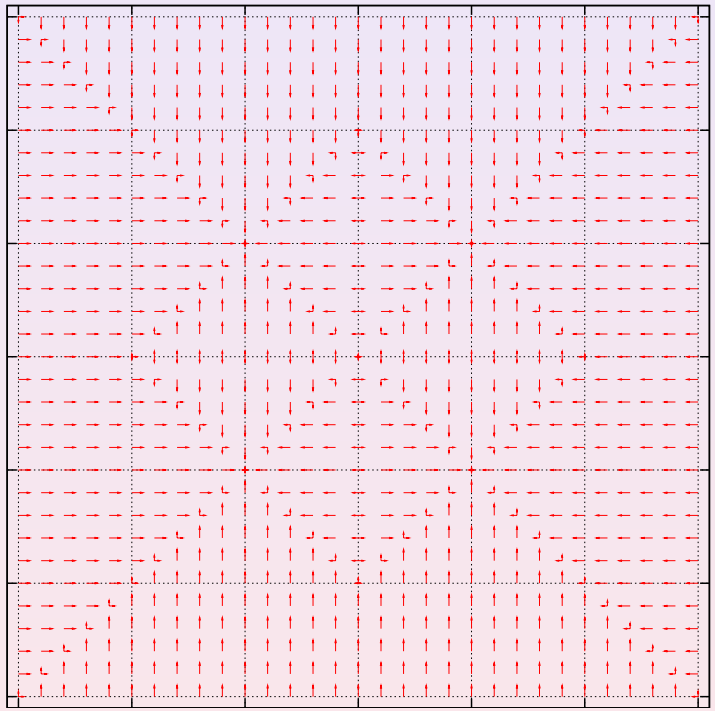
... similar to colloquium talk of Daniel Polani.

Scenarios - Sensor-Values $s \in \mathcal{S}$ for Every Place $r \in \mathcal{R}$ in the World



"s+" = single source in center

from Alexander Klyubin

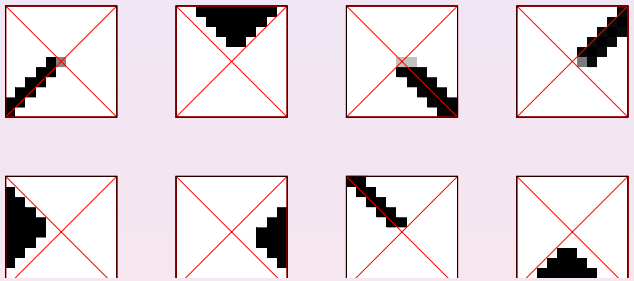


"sq" = 4 sources

new

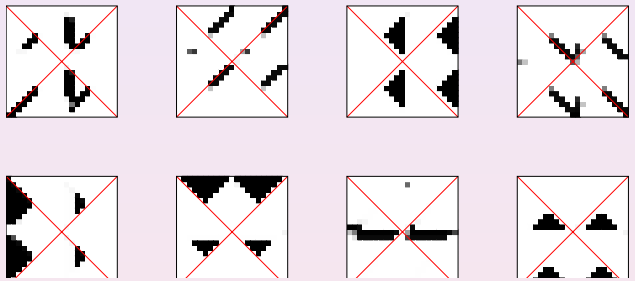
Initial Position Capturing

- Search for a controller (agent) which maximizes $I(R_0; M_{15}) \rightarrow$ Initial position capturing agent.
- Shown is the probability $p(r_0|m_{15})$ for a starting place r_0 given a memory value m_{15} (black = high probability)



single source in center - s+

$I(R_0; M_{15}) = 2.906$ bit



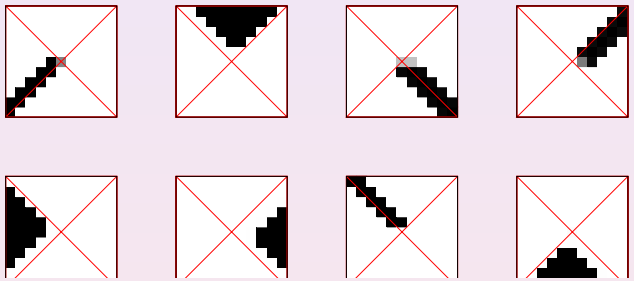
4 sources - sq

$I(R_0; M_{15}) = 2.888$ bit

- Let us call a pair (R_0, M_{15}) a *concept* iff $I(R_0; M_{15}) > 0$.
- Every memory value $m_{15} \in \mathcal{M}_{15}$ is a *symbol* what describes a region in the world R_0.
- Also our scenarios s+ / sq could be interpreted as concept: (R, S)

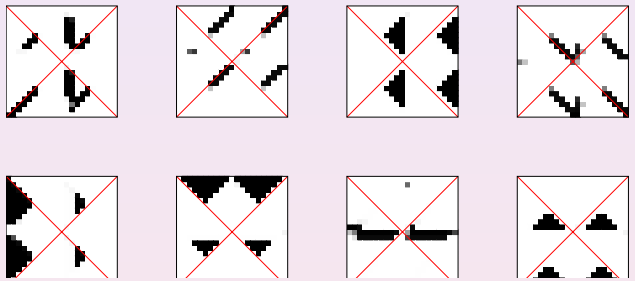
Initial Position Capturing

- Search for a controller (agent) which maximizes $I(R_0; M_{15}) \rightarrow$ Initial position capturing agent.
- Shown is the probability $p(r_0|m_{15})$ for a starting place r_0 given a memory value m_{15} (black = high probability)



single source in center - s+

$I(R_0; M_{15}) = 2.906$ bit



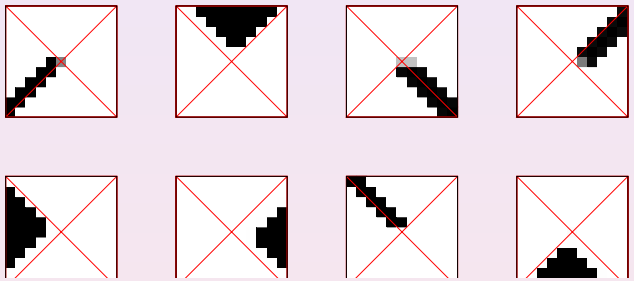
4 sources - sq

$I(R_0; M_{15}) = 2.888$ bit

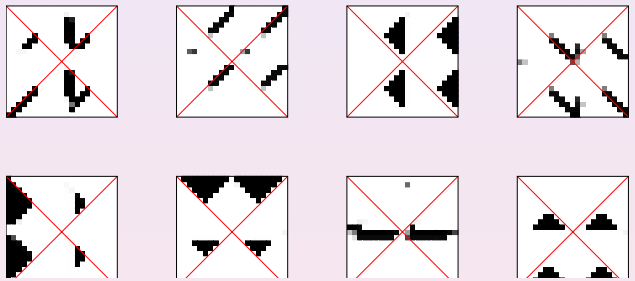
- Let us call a pair (R_0, M_{15}) a *concept* iff $I(R_0; M_{15}) > 0$.
- Every memory value $m_{15} \in \mathcal{M}_{15}$ is a *symbol* what describes a region in the world R_0.
- Also our scenarios s+ / sq could be interpreted as concept: (R, S)

Initial Position Capturing

- Search for a controller (agent) which maximizes $I(R_0; M_{15}) \rightarrow$ Initial position capturing agent.
- Shown is the probability $p(r_0|m_{15})$ for a starting place r_0 given a memory value m_{15} (black = high probability)



single source in center - s+
$I(R_0; M_{15}) = 2.906$ bit

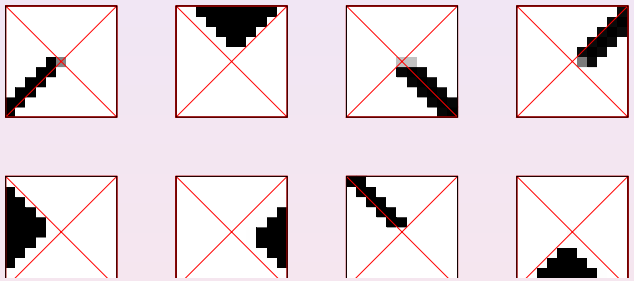


4 sources - sq
$I(R_0; M_{15}) = 2.888$ bit

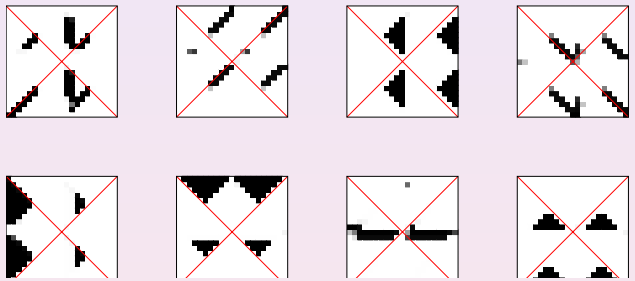
- Let us call a pair (R_0, M_{15}) a *concept* iff $I(R_0; M_{15}) > 0$.
- Every memory value $m_{15} \in \mathcal{M}_{15}$ is a *symbol* what describes a region in the world R_0.
- Also our scenarios s+ / sq could be interpreted as concept: (R, S)

Initial Position Capturing

- Search for a controller (agent) which maximizes $I(R_0; M_{15}) \rightarrow$ Initial position capturing agent.
- Shown is the probability $p(r_0|m_{15})$ for a starting place r_0 given a memory value m_{15} (black = high probability)



single source in center - s+
$I(R_0; M_{15}) = 2.906$ bit



4 sources - sq
$I(R_0; M_{15}) = 2.888$ bit

- Let us call a pair (R_0, M_{15}) a *concept* iff $I(R_0; M_{15}) > 0$.
- Every memory value $m_{15} \in \mathcal{M}_{15}$ is a *symbol* what describes a region in the world R_0.
- Also our scenarios s+ / sq could be interpreted as concept: (R, S)

Overview

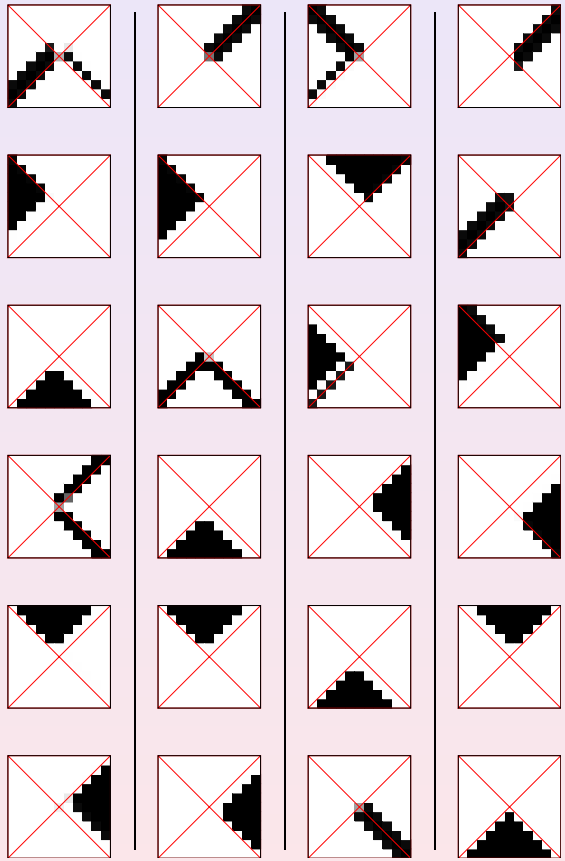
1 Scenario

2 Common Concept
- Motivation
- Objective Common Concept
- Subjective Common Concept
- Results

3 Symmetry

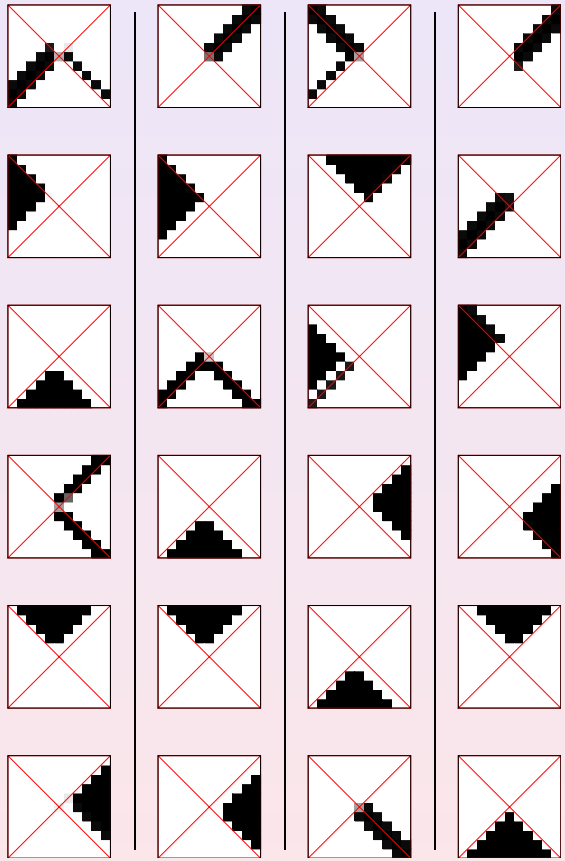
Why Common Concepts?

- 4 different good concepts for the initial position capturing agents are shown →
- Intuitively they all looking very similar
- Information theory do not consider permutations
- Difference:
  - Different diagonals where found
  - Even if is was same diagonal, they are not perfectly matching (≈ translation)
- How to combine all symbols of all concepts? ⇒ Common Concepts



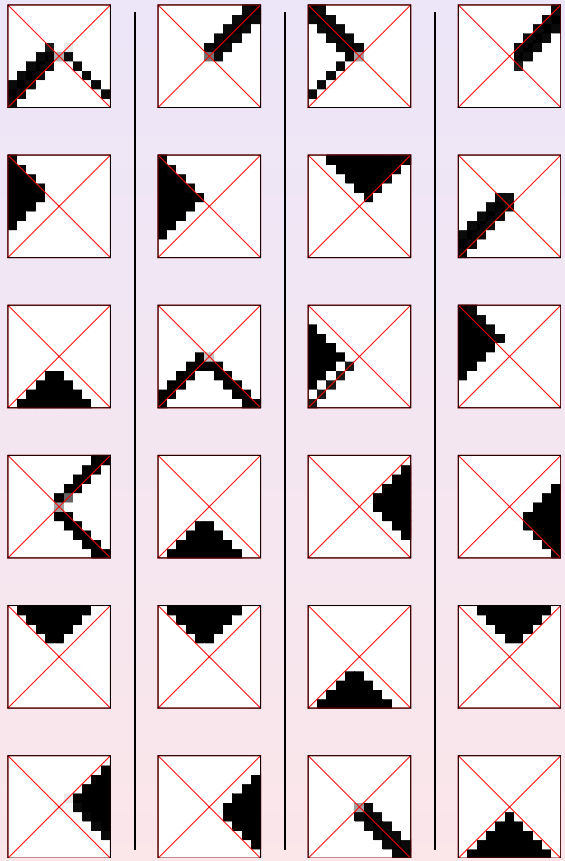
Why Common Concepts?

- 4 different good concepts for the initial position capturing agents are shown →
- Intuitively they all looking very similar
- Information theory do not consider permutations
- Difference:
  - Different diagonals where found
  - Even if is was same diagonal, they are not perfectly matching (≈ translation)
- How to combine all symbols of all concepts? ⇒ Common Concepts



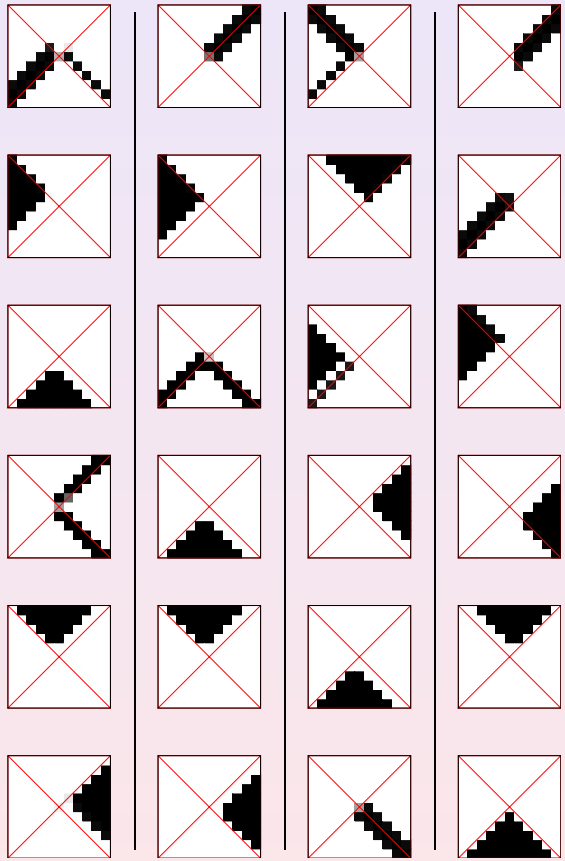
Why Common Concepts?

- 4 different good concepts for the initial position capturing agents are shown →
- Intuitively they all looking very similar
- Information theory do not consider permutations
- Difference:
  - Different diagonals where found
  - Even if is was same diagonal, they are not perfectly matching (≈ translation)
- How to combine all symbols of all concepts? ⇒ Common Concepts



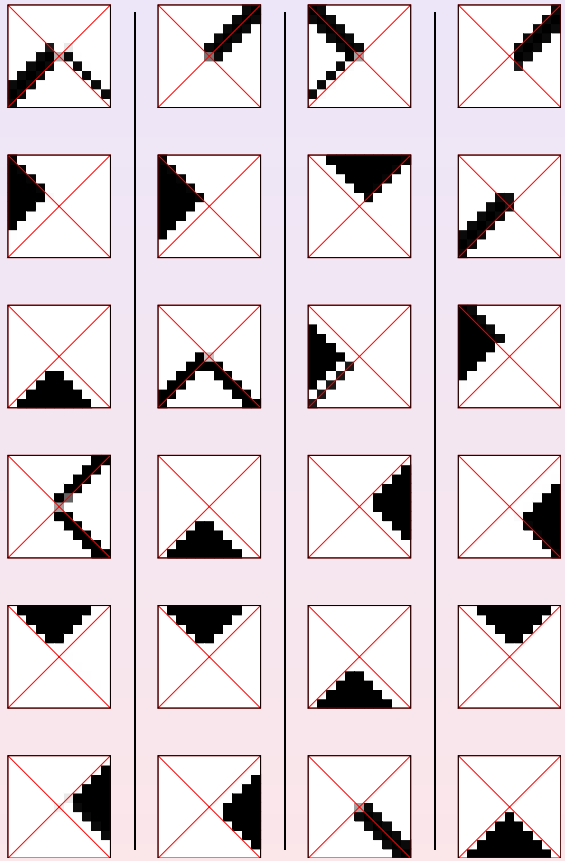
Why Common Concepts?

- 4 different good concepts for the initial position capturing agents are shown →
- Intuitively they all looking very similar
- Information theory do not consider permutations
- Difference:
  - Different diagonals where found
  - Even if is was same diagonal, they are not perfectly matching (≈ translation)
- How to combine all symbols of all concepts? ⇒ Common Concepts



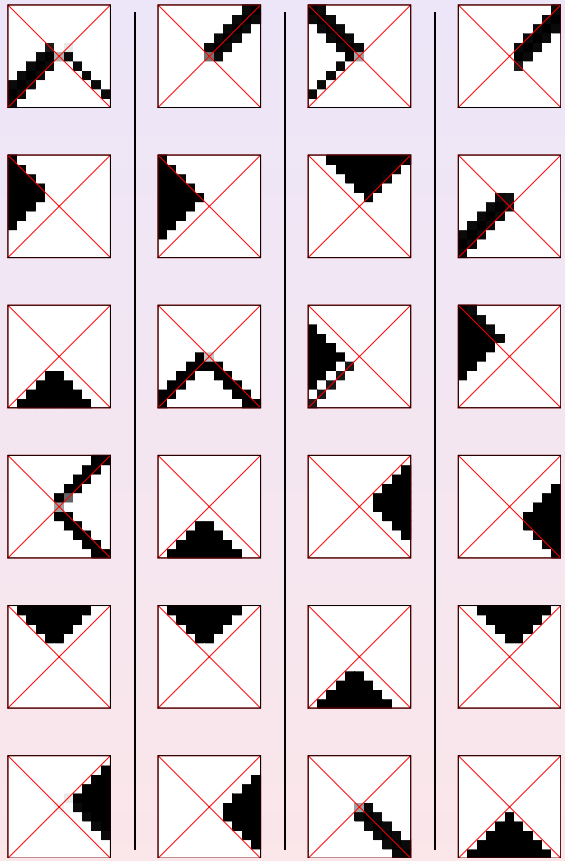
Why Common Concepts?

- 4 different good concepts for the initial position capturing agents are shown →
- Intuitively they all looking very similar
- Information theory do not consider permutations
- Difference:
  - Different diagonals where found
  - Even if is was same diagonal, they are not perfectly matching (≈ translation)
- How to combine all symbols of all concepts? ⇒ Common Concepts



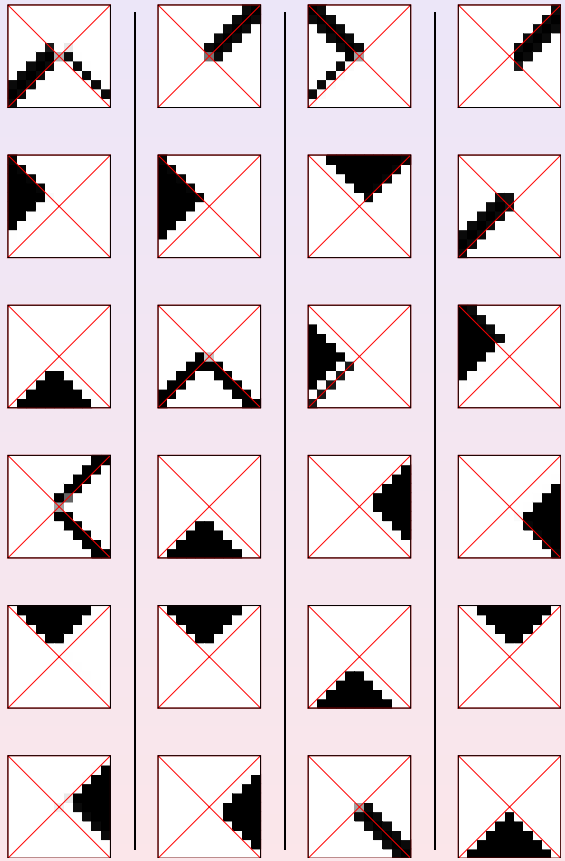
Why Common Concepts?

- 4 different good concepts for the initial position capturing agents are shown →
- Intuitively they all looking very similar
- Information theory do not consider permutations
- Difference:
  - Different diagonals where found
  - Even if is was same diagonal, they are not perfectly matching (≈ translation)
- How to combine all symbols of all concepts? ⇒ Common Concepts

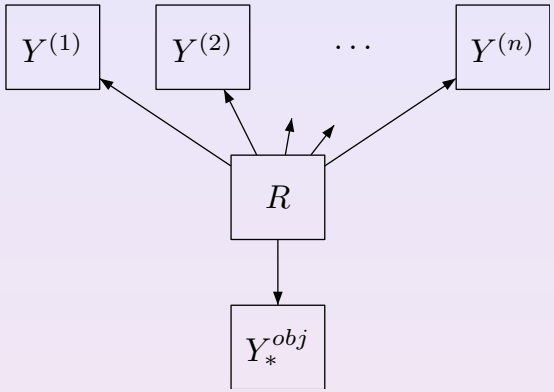


Why Common Concepts?

- 4 different good concepts for the initial position capturing agents are shown →
- Intuitively they all looking very similar
- Information theory do not consider permutations
- Difference:
  - Different diagonals where found
  - Even if is was same diagonal, they are not perfectly matching (≈ translation)
- How to combine all symbols of all concepts? ⇒ Common Concepts



Objective Common Concept



- A deterministic mapping $R \longrightarrow Y_*^{obj}$ which maximized the utility

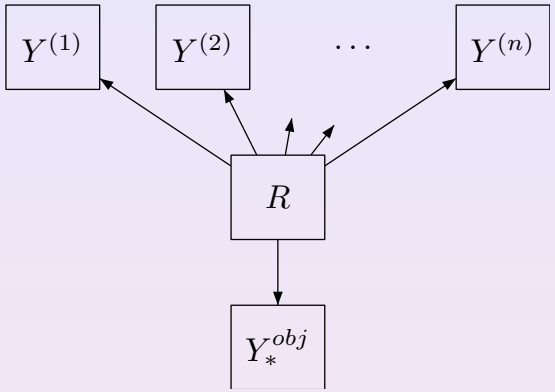
$$\max \stackrel{!}{=} \sum_i \left[I\left(Y_*^{obj}; Y^{(i)}\right) - \alpha \cdot I\left(R; Y_*^{obj}\right) \right]$$

$$\equiv \left[\sum_i I\left(Y_*^{obj}; Y^{(i)}\right) \right] - \alpha \cdot n \cdot I\left(R; Y_*^{obj}\right)$$

  defines the *objective common concept* $\left(R, Y_*^{obj}\right)$.

- Objective $\longrightarrow$ If you know R, everything about Y_*^{obj} is clear.

Objective Common Concept



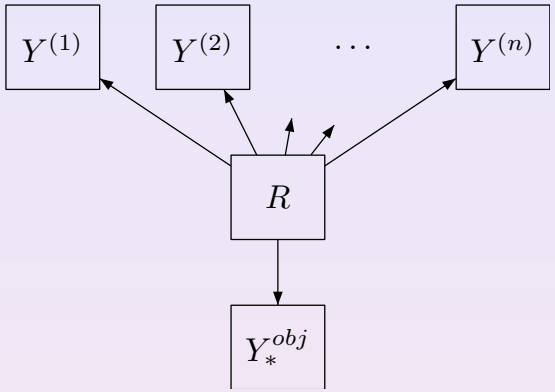
- A deterministic mapping $R \longrightarrow Y_*^{obj}$ which maximized the utility

$$\begin{aligned} \max \quad &\stackrel{!}{=} \quad \sum_i \left[I\left(Y_*^{obj}; Y^{(i)}\right) - \alpha \cdot I\left(R; Y_*^{obj}\right) \right] \\ &\equiv \quad \left[\sum_i I\left(Y_*^{obj}; Y^{(i)}\right) \right] - \alpha \cdot n \cdot I\left(R; Y_*^{obj}\right) \end{aligned}$$

  defines the *objective common concept* $\left(R, Y_*^{obj}\right)$.

- Objective $\longrightarrow$ If you know R, everything about Y_*^{obj} is clear.

Objective Common Concept



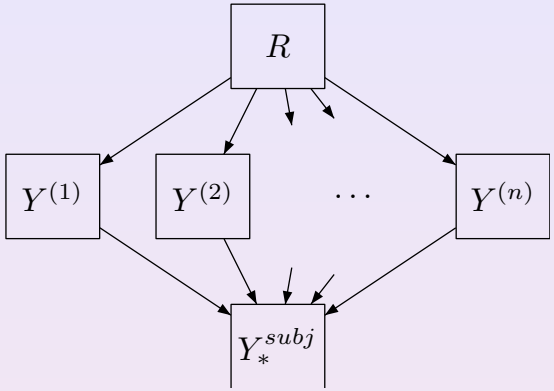
- A deterministic mapping $R \longrightarrow Y_*^{obj}$ which maximized the utility

$$\begin{aligned} \max \quad &\stackrel{!}{=} \quad \sum_i \left[I\left(Y_*^{obj}; Y^{(i)}\right) - \alpha \cdot I\left(R; Y_*^{obj}\right) \right] \\ &\equiv \quad \left[\sum_i I\left(Y_*^{obj}; Y^{(i)}\right) \right] - \alpha \cdot n \cdot I\left(R; Y_*^{obj}\right) \end{aligned}$$

  defines the *objective common concept* $\left(R, Y_*^{obj}\right)$.

- Objective $\longrightarrow$ If you know R, everything about Y_*^{obj} is clear.

Subjective Common Concept

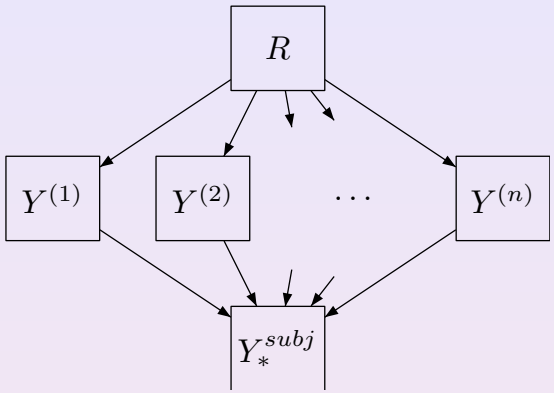


- A deterministic mapping $Y^{(1)} \times \ldots \times Y^{(n)} \longrightarrow Y_*^{subj}$ which minimized the utility

$$\min \quad \stackrel{!}{=} \quad I\left(Y^{(1)}; \ldots; Y^{(n)} | Y_*^{subj}\right) \equiv \left[\sum_{i=1}^{n} H\left(Y^{(i)} | Y_*^{subj}\right)\right] - H\left(Y^{(1)}, \ldots, Y^{(n)} | Y_*^{subj}\right)$$

  defines the *subjective common concept* $\left(R, Y_*^{subj}\right)$.
- Subjective $\longrightarrow$ the "naming" of a place in $r \in \mathcal{R}$ could be very non-deterministic and depends on the "opinion about the world" of every agent.
- *Superstition* of a concept is $H(Y|R)$ and is a measurement for this "subjectivity".

Subjective Common Concept



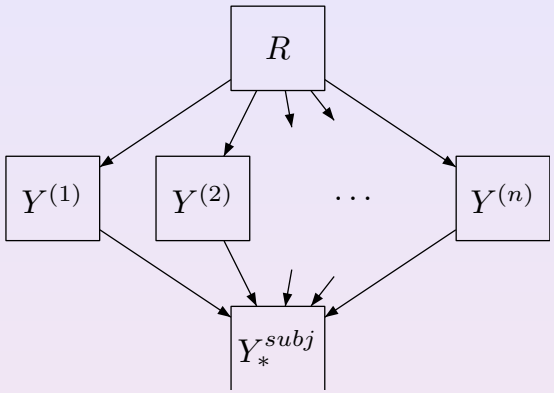
- A deterministic mapping $Y^{(1)} \times \ldots \times Y^{(n)} \longrightarrow Y_*^{subj}$ which minimized the utility

  $$\min \quad \stackrel{!}{=} \quad I\left(Y^{(1)}; \ldots; Y^{(n)} | Y_*^{subj}\right) \equiv \left[\sum_{i=1}^{n} H\left(Y^{(i)} | Y_*^{subj}\right)\right] - H\left(Y^{(1)}, \ldots, Y^{(n)} | Y_*^{subj}\right)$$

  defines the *subjective common concept* $\left(R, Y_*^{subj}\right)$.

- Subjective $\longrightarrow$ the "naming" of a place in $r \in \mathcal{R}$ could be very non-deterministic and depends on the "opinion about the world" of every agent.
- *Superstition* of a concept is $H(Y|R)$ and is a measurement for this "subjectivity".

Subjective Common Concept

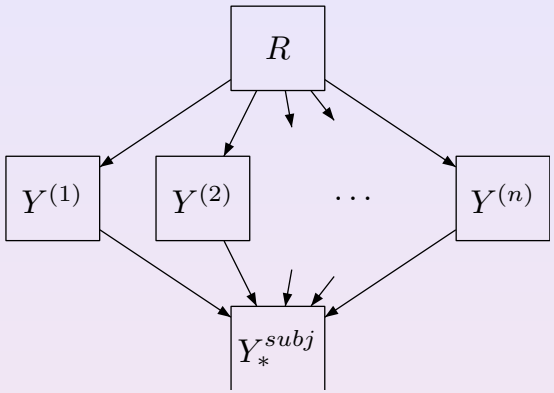


- A deterministic mapping $Y^{(1)} \times \ldots \times Y^{(n)} \rightarrow Y_*^{subj}$ which minimized the utility

  $$\min \quad \stackrel{!}{=} \quad I\left(Y^{(1)}; \ldots; Y^{(n)} | Y_*^{subj}\right) \equiv \left[\sum_{i=1}^{n} H\left(Y^{(i)} | Y_*^{subj}\right)\right] - H\left(Y^{(1)}, \ldots, Y^{(n)} | Y_*^{subj}\right)$$

  defines the *subjective common concept* $\left(R, Y_*^{subj}\right)$.
- Subjective $\rightarrow$ the "naming" of a place in $r \in \mathcal{R}$ could be very non-deterministic and depends on the "opinion about the world" of every agent.
- *Superstition* of a concept is $H(Y|R)$ and is a measurement for this "subjectivity".

Subjective Common Concept

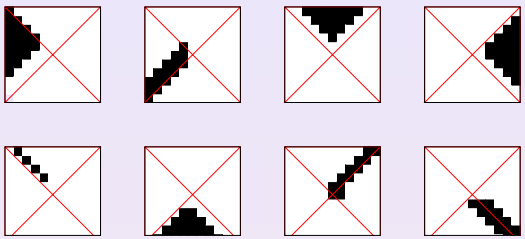


- A deterministic mapping $Y^{(1)} \times \ldots \times Y^{(n)} \rightarrow Y_*^{subj}$ which minimized the utility

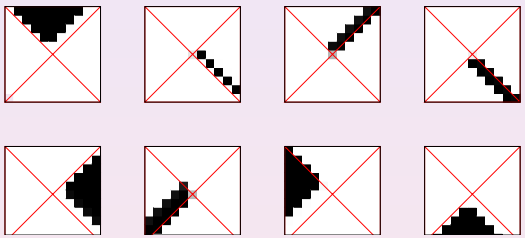
  $$\min \quad \stackrel{!}{=} \quad I\left(Y^{(1)}; \ldots; Y^{(n)} | Y_*^{subj}\right) \equiv \left[\sum_{i=1}^{n} H\left(Y^{(i)} | Y_*^{subj}\right)\right] - H\left(Y^{(1)}, \ldots, Y^{(n)} | Y_*^{subj}\right)$$

  defines the *subjective common concept* $\left(R, Y_*^{subj}\right)$.
- Subjective $\rightarrow$ the "naming" of a place in $r \in \mathcal{R}$ could be very non-deterministic and depends on the "opinion about the world" of every agent.
- *Superstition* of a concept is $H(Y|R)$ and is a measurement for this "subjectivity".

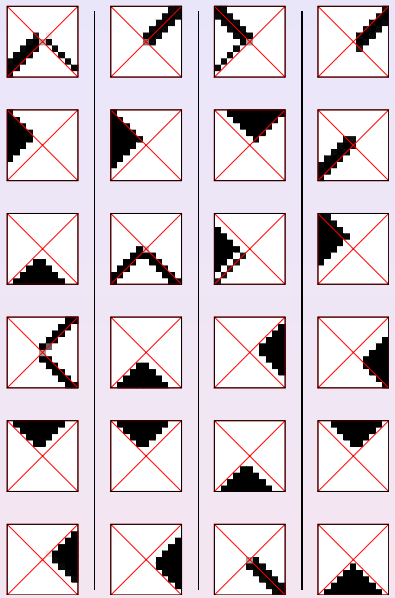
Common Concept Example



Objective common concept



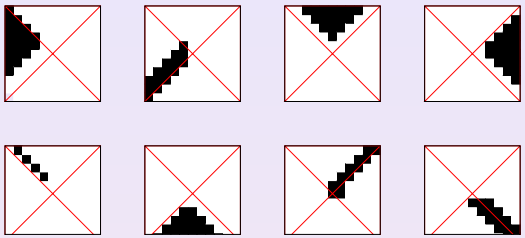
Subjective common concept



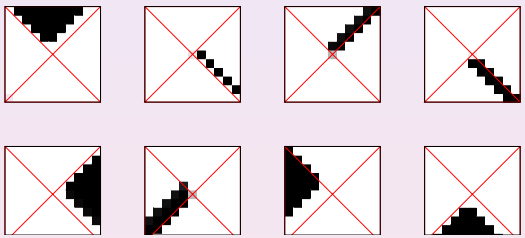
4 Input Concepts

- Almost similar: Crutchfield-distance $D(M_*^{obj}, M_*^{subj}) = H(M_*^{obj}|M_*^{subj}) + H(M_*^{subj}|M_*^{obj}) = 0.38$ bit is very small.
- Less superstition: $H\left(M_*^{subj}|R_0\right) = 0.03$ bit.
- Subjective common concept needs exponentially resources (memory, computation-time, search-space size) in number of input concepts and there size.

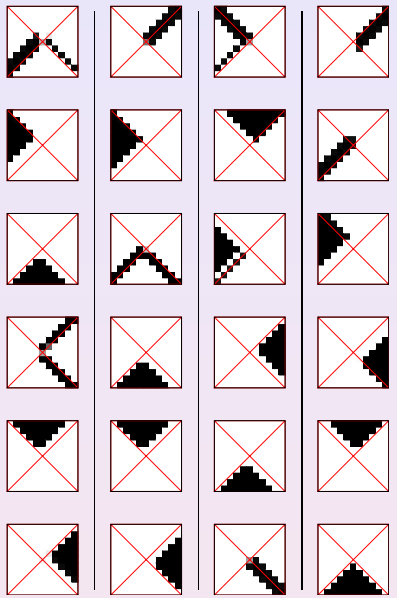
Common Concept Example



Objective common concept



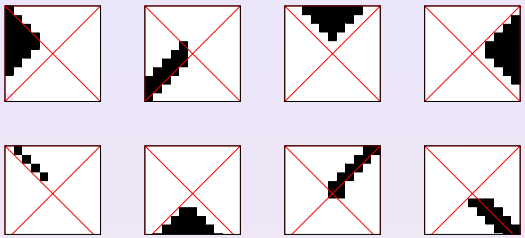
Subjective common concept



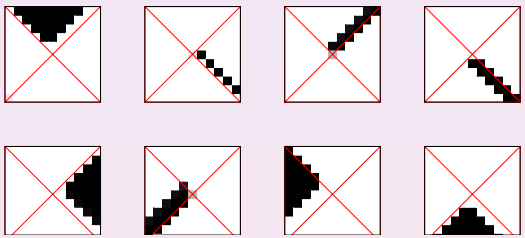
4 Input Concepts

- Almost similar: Crutchfield-distance
  $D(M_*^{obj}, M_*^{subj}) = H(M_*^{obj}|M_*^{subj}) + H(M_*^{subj}|M_*^{obj}) = 0.38$ bit is very small.
- Less superstition: $H\left(M_*^{subj}|R_0\right) = 0.03$ bit.
- Subjective common concept needs exponentially resources (memory, computation-time, search-space size) in number of input concepts and there size.

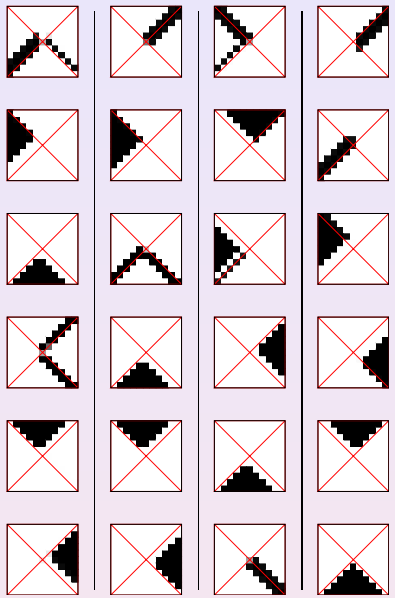
Common Concept Example



Objective common concept



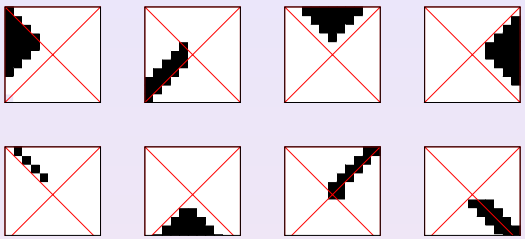
Subjective common concept



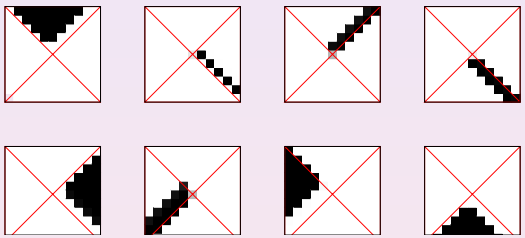
4 Input Concepts

- Almost similar: Crutchfield-distance
  $D(M_*^{obj}, M_*^{subj}) = H(M_*^{obj}|M_*^{subj}) + H(M_*^{subj}|M_*^{obj}) = 0.38$ bit is very small.
- Less superstition: $H\left(M_*^{subj}|R_0\right) = 0.03$ bit.
- Subjective common concept needs exponentially resources (memory, computation-time, search-space size) in number of input concepts and there size.

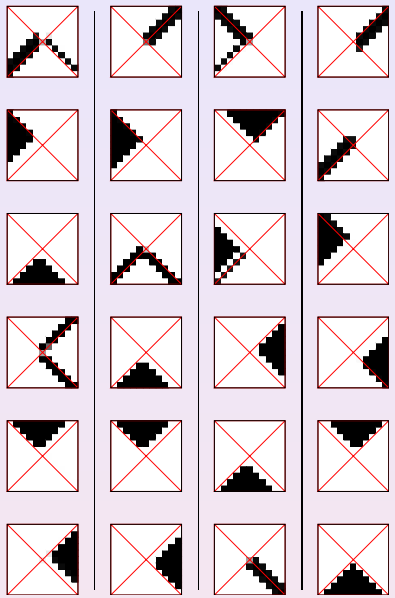
Common Concept Example



Objective common concept



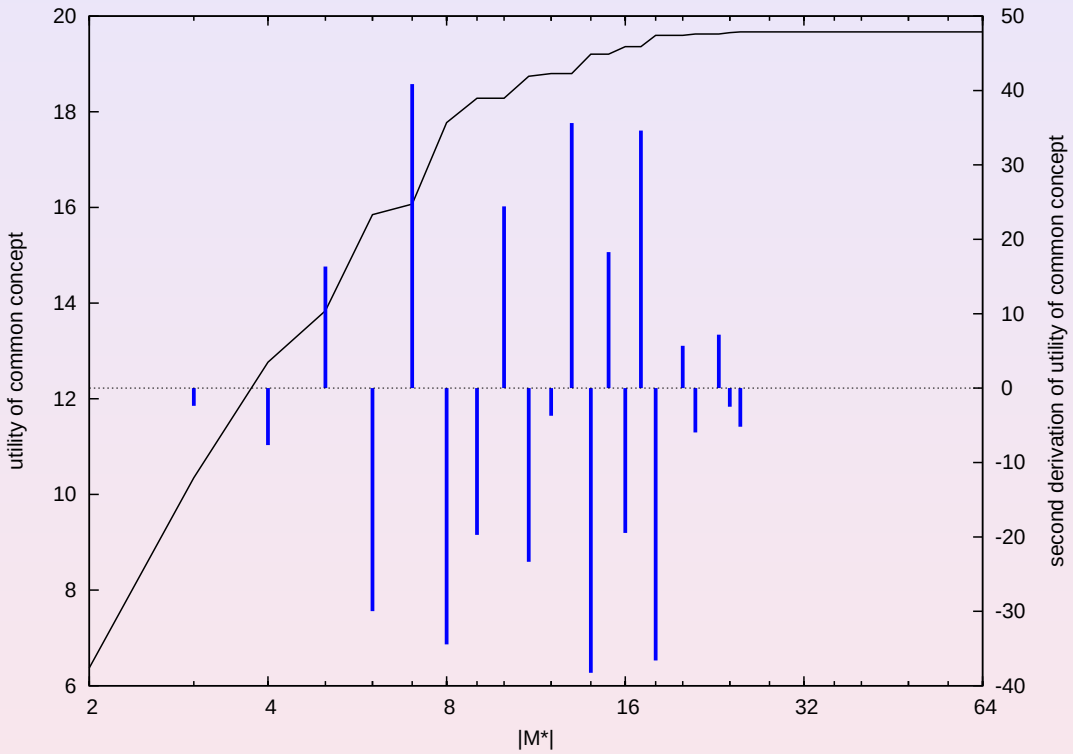
Subjective common concept



4 Input Concepts

- Almost similar: Crutchfield-distance $D\left(M_*^{obj}, M_*^{subj}\right) = H\left(M_*^{obj} | M_*^{subj}\right) + H\left(M_*^{subj} | M_*^{obj}\right) = 0.38$ bit is very small.
- Less superstition: $H\left(M_*^{subj} | R_0\right) = 0.03$ bit.
- Subjective common concept needs exponentially resources (memory, computation-time, search-space size) in number of input concepts and there size.

Optimal Concept Size



Utility of objective common concept (10 input concepts with scenario s+ and memory size $|\mathcal{M}| = 8$) depending on concept size $|\mathcal{M}_*|$ and its second derivative. Memory sizes where this derivative becomes negative are good choices.

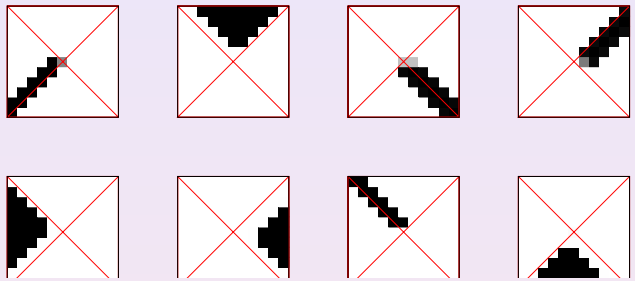
Overview

1 Scenario

2 Common Concept

3 Symmetry
- Idea
- Extrinsic Symmetry
- Intrinsic Symmetry
- Results

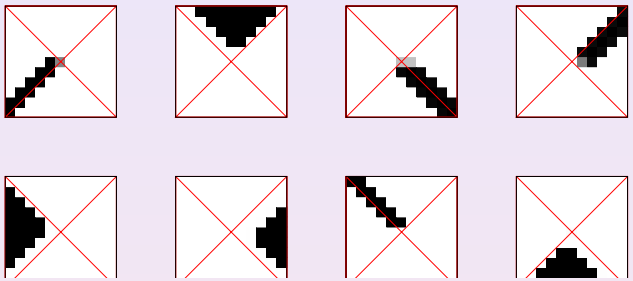
Idea for Symmetry



1. Define a set of operations on a concept → *symmetry operation*.
2. Define a utility function for every such symmetry operation. E.g.:
   1. Define a other concept where all "operated concepts" are compared to.
   2. Define how this comparison will be measured.

- We have analyzed two such symmetry methods

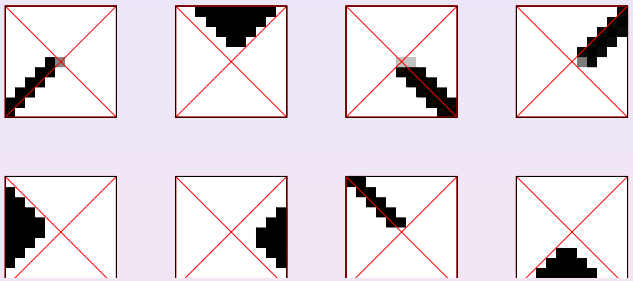
Idea for Symmetry



1. Define a set of operations on a concept $\longrightarrow$ *symmetry operation*.
2. Define a utility function for every such symmetry operation. E.g.:
   1. Define a other concept where all "operated concepts" are compared to.
   2. Define how this comparison will be measured.

- We have analyzed two such symmetry methods

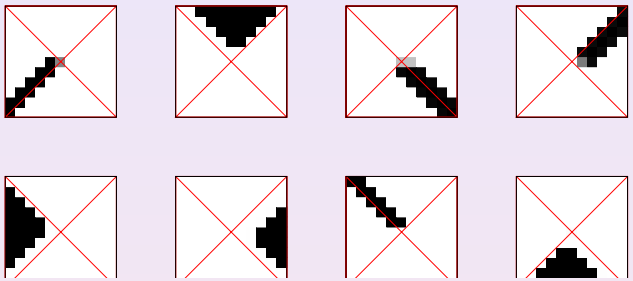
Idea for Symmetry



1. Define a set of operations on a concept $\longrightarrow$ *symmetry operation*.
2. Define a utility function for every such symmetry operation. E.g.:
   1. Define a other concept where all "operated concepts" are compared to.
   2. Define how this comparison will be measured.

- We have analyzed two such symmetry methods

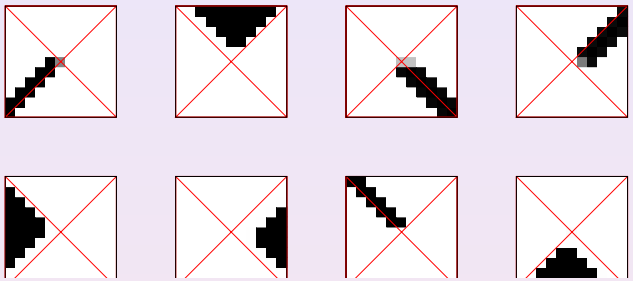
Idea for Symmetry



1. Define a set of operations on a concept $\longrightarrow$ *symmetry operation*.
2. Define a utility function for every such symmetry operation. E.g.:
   1. Define a other concept where all "operated concepts" are compared to.
   2. Define how this comparison will be measured.

- We have analyzed two such symmetry methods

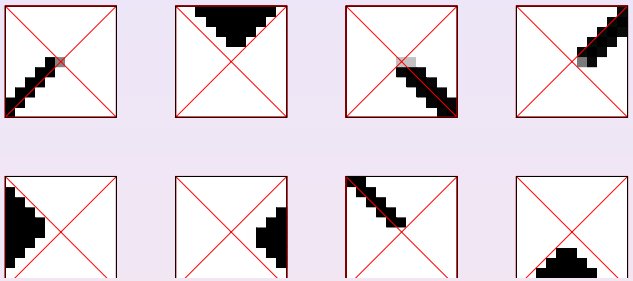
Idea for Symmetry



1. Define a set of operations on a concept $\longrightarrow$ *symmetry operation*.
2. Define a utility function for every such symmetry operation. E.g.:
   1. Define a other concept where all "operated concepts" are compared to.
   2. Define how this comparison will be measured.

- We have analyzed two such symmetry methods

Idea for Symmetry

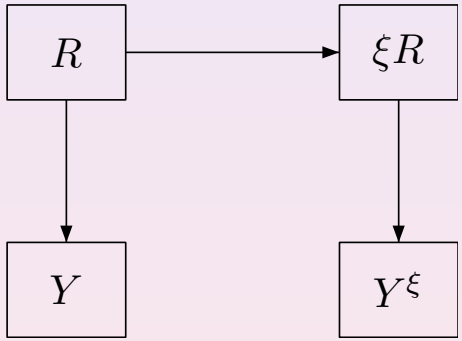


1. Define a set of operations on a concept $\rightarrow$ *symmetry operation*.
2. Define a utility function for every such symmetry operation. E.g.:
   1. Define a other concept where all "operated concepts" are compared to.
   2. Define how this comparison will be measured.

- We have analyzed two such symmetry methods

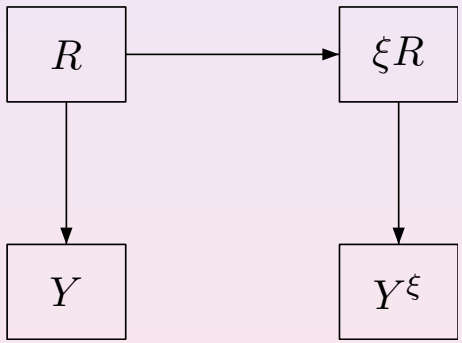
Extrinsic Symmetry

- Consider a concept $(R = \mathbb{Z}^2, Y)$ with e.g. $Y = M$ or $Y = S$.
- Perform on R all (infinite-grid) symmetry operations $\xi^{\theta,\varphi,x_0,y_0} : \mathbb{Z}^2 \to \mathbb{Z}^2$ with $\xi^{\theta,\varphi,x_0,y_0} := \xi^{x_0,y_0}_{\text{trans}} \circ \xi^{\varphi}_{\text{rot}} \circ \xi^{\theta}_{\text{mir}}$
  translation $\xi^{x_0,y_0}_{\text{trans}}$ by $(x_0, y_0) \in \mathbb{Z}^2$.
  mirroring $\xi^{\theta}_{\text{mir}}$ by $\theta \in \{+1, -1\}$
  rotation $\xi^{\varphi}_{\text{rot}}$ by $\varphi \in \{0°, 90°, 180°, 270°\}$
- The "operated world" ξR is defined by:
  $$\Pr(\xi R = r' | R = r) := \begin{cases} 1 & r' = \xi(r) \\ 0 & \text{else} \end{cases}.$$
- Map $R' = \xi R$ to Y^ξ like on R was mapped to Y:
  $\Pr(Y^\xi = y | \xi R = r) := \Pr(Y = y | R = r)$.
- We define the *utility for a extrinsic symmetry* operation $\xi^{\varphi,\theta,x_0,y_0}$ for the concept (R, Y) as $I(Y; Y^\xi)$.



Extrinsic Symmetry

- Consider a concept $(R = \mathbb{Z}^2, Y)$ with e.g. $Y = M$ or $Y = S$.
- Perform on R all (infinite-grid) symmetry operations $\xi^{\theta,\varphi,x_0,y_0} : \mathbb{Z}^2 \to \mathbb{Z}^2$ with $\xi^{\theta,\varphi,x_0,y_0} := \xi^{x_0,y_0}_{\text{trans}} \circ \xi^{\varphi}_{\text{rot}} \circ \xi^{\theta}_{\text{mir}}$
  translation $\xi^{x_0,y_0}_{\text{trans}}$ by $(x_0, y_0) \in \mathbb{Z}^2$.
  mirroring $\xi^{\theta}_{\text{mir}}$ by $\theta \in \{+1, -1\}$
  rotation $\xi^{\varphi}_{\text{rot}}$ by $\varphi \in \{0°, 90°, 180°, 270°\}$
- The "operated world" ξR is defined by:
  $$\Pr(\xi R = r' | R = r) := \begin{cases} 1 & r' = \xi(r) \\ 0 & \text{else} \end{cases}.$$
- Map $R' = \xi R$ to Y^ξ like on R was mapped to Y:
  $\Pr(Y^\xi = y | \xi R = r) := \Pr(Y = y | R = r)$.
- We define the *utility for a extrinsic symmetry* operation $\xi^{\varphi,\theta,x_0,y_0}$ for the concept (R, Y) as $I(Y; Y^\xi)$.



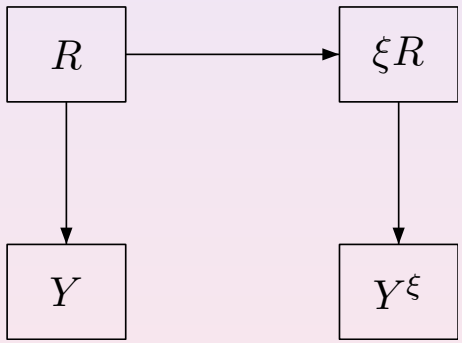
Extrinsic Symmetry

- Consider a concept $(R = \mathbb{Z}^2, Y)$ with e.g. $Y = M$ or $Y = S$.
- Perform on R all (infinite-grid) symmetry operations $\xi^{\theta,\varphi,x_0,y_0} : \mathbb{Z}^2 \to \mathbb{Z}^2$ with $\xi^{\theta,\varphi,x_0,y_0} := \xi^{x_0,y_0}_{\text{trans}} \circ \xi^{\varphi}_{\text{rot}} \circ \xi^{\theta}_{\text{mir}}$

  translation $\xi^{x_0,y_0}_{\text{trans}}$ by $(x_0, y_0) \in \mathbb{Z}^2$.

  mirroring $\xi^{\theta}_{\text{mir}}$ by $\theta \in \{+1, -1\}$

  rotation $\xi^{\varphi}_{\text{rot}}$ by $\varphi \in \{0°, 90°, 180°, 270°\}$



- The "operated world" ξR is defined by:
$$\Pr(\xi R = r' | R = r) := \begin{cases} 1 & r' = \xi(r) \\ 0 & \text{else} \end{cases}.$$
- Map $R' = \xi R$ to Y^{ξ} like on R was mapped to Y:
$\Pr(Y^{\xi} = y | \xi R = r) := \Pr(Y = y | R = r)$.
- We define the *utility for a extrinsic symmetry* operation $\xi^{\varphi,\theta,x_0,y_0}$ for the concept (R, Y) as $I(Y; Y^{\xi})$.

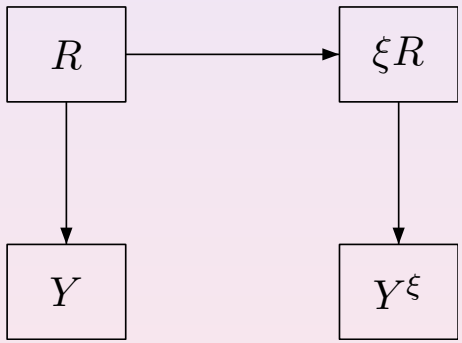
Extrinsic Symmetry

- Consider a concept $(R = \mathbb{Z}^2, Y)$ with e.g. $Y = M$ or $Y = S$.
- Perform on R all (infinite-grid) symmetry operations $\xi^{\theta,\varphi,x_0,y_0} : \mathbb{Z}^2 \to \mathbb{Z}^2$ with
  $\xi^{\theta,\varphi,x_0,y_0} := \xi^{x_0,y_0}_{\text{trans}} \circ \xi^{\varphi}_{\text{rot}} \circ \xi^{\theta}_{\text{mir}}$

  translation $\xi^{x_0,y_0}_{\text{trans}}$ by $(x_0, y_0) \in \mathbb{Z}^2$.

  mirroring $\xi^{\theta}_{\text{mir}}$ by $\theta \in \{+1, -1\}$

  rotation $\xi^{\varphi}_{\text{rot}}$ by $\varphi \in \{0°, 90°, 180°, 270°\}$
- The "operated world" ξR is defined by:
  $$\Pr(\xi R = r' | R = r) := \begin{cases} 1 & r' = \xi(r) \\ 0 & \text{else} \end{cases}.$$
- Map $R' = \xi R$ to Y^ξ like on R was mapped to Y:
  $\Pr(Y^\xi = y | \xi R = r) := \Pr(Y = y | R = r)$.
- We define the *utility for a extrinsic symmetry* operation $\xi^{\varphi,\theta,x_0,y_0}$ for the concept (R, Y) as $I(Y; Y^\xi)$.



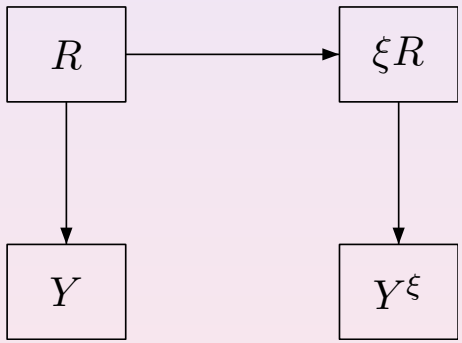
Extrinsic Symmetry

- Consider a concept $\left(R = \mathbb{Z}^2, Y\right)$ with e.g. $Y = M$ or $Y = S$.
- Perform on R all (infinite-grid) symmetry operations $\xi^{\theta,\varphi,x_0,y_0} : \mathbb{Z}^2 \to \mathbb{Z}^2$ with $\xi^{\theta,\varphi,x_0,y_0} := \xi^{x_0,y_0}_{\text{trans}} \circ \xi^{\varphi}_{\text{rot}} \circ \xi^{\theta}_{\text{mir}}$

  translation $\xi^{x_0,y_0}_{\text{trans}}$ by $(x_0, y_0) \in \mathbb{Z}^2$.

  mirroring $\xi^{\theta}_{\text{mir}}$ by $\theta \in \{+1, -1\}$

  rotation $\xi^{\varphi}_{\text{rot}}$ by $\varphi \in \{0°, 90°, 180°, 270°\}$



- The "operated world" ξR is defined by:

$$\Pr\left(\xi R = r' | R = r\right) := \begin{cases} 1 & r' = \xi\left(r\right) \\ 0 & \text{else} \end{cases}.$$

- Map $R' = \xi R$ to Y^{ξ} like on R was mapped to Y: $\Pr\left(Y^{\xi} = y | \xi R = r\right) := \Pr\left(Y = y | R = r\right)$.
- We define the *utility for a extrinsic symmetry* operation $\xi^{\varphi,\theta,x_0,y_0}$ for the concept (R, Y) as $I\left(Y; Y^{\xi}\right)$.

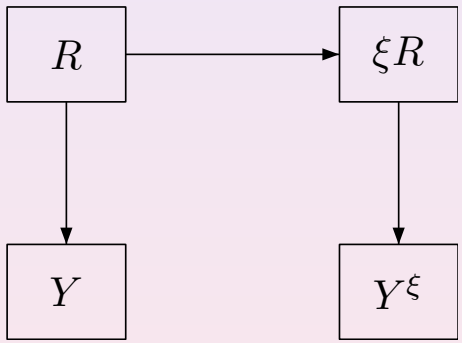
Extrinsic Symmetry

- Consider a concept $\left(R = \mathbb{Z}^2, Y\right)$ with e.g. $Y = M$ or $Y = S$.
- Perform on R all (infinite-grid) symmetry operations $\xi^{\theta,\varphi,x_0,y_0} : \mathbb{Z}^2 \to \mathbb{Z}^2$ with $\xi^{\theta,\varphi,x_0,y_0} := \xi^{x_0,y_0}_{\text{trans}} \circ \xi^{\varphi}_{\text{rot}} \circ \xi^{\theta}_{\text{mir}}$

  translation $\xi^{x_0,y_0}_{\text{trans}}$ by $(x_0, y_0) \in \mathbb{Z}^2$.

  mirroring $\xi^{\theta}_{\text{mir}}$ by $\theta \in \{+1, -1\}$

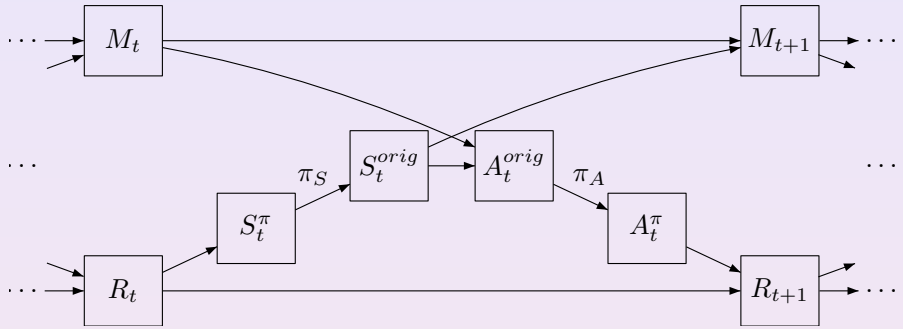
  rotation $\xi^{\varphi}_{\text{rot}}$ by $\varphi \in \{0°, 90°, 180°, 270°\}$



- The "operated world" ξR is defined by:
$$\Pr\left(\xi R = r' | R = r\right) := \begin{cases} 1 & r' = \xi\left(r\right) \\ 0 & \text{else} \end{cases}.$$
- Map $R' = \xi R$ to Y^{ξ} like on R was mapped to Y: $\Pr\left(Y^{\xi} = y | \xi R = r\right) := \Pr\left(Y = y | R = r\right)$.
- We define the *utility for a extrinsic symmetry* operation $\xi^{\varphi,\theta,x_0,y_0}$ for the concept (R, Y) as $I\left(Y; Y^{\xi}\right)$.

Intrinsic Symmetry

- Consider a "permutation of the embodiment" of the agent:

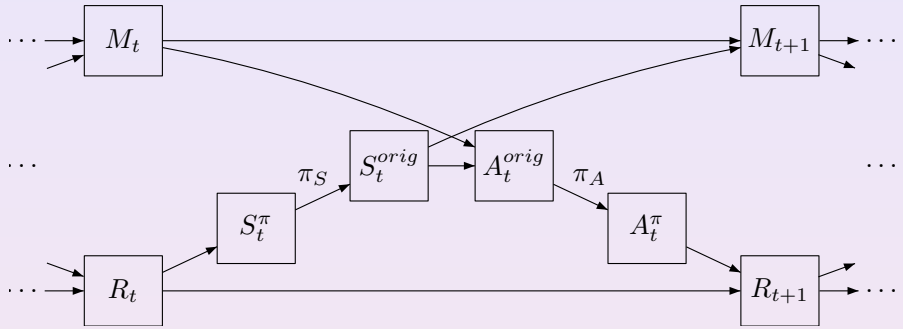


- With each permutation pair (π_S, π_A) the agent generates a different concept (R_0, M_{15}^{π}).
- Lets consider a group of similar (but not permuted) agents and there concepts $\left\{\left(R_0, M_{15}^{(1)}\right), \ldots, \left(R_0, M_{15}^{(n)}\right)\right\}$.
- Calculate based on this concepts a (objective) common concept (R_0, M_*).
- We define the *utility of a symmetry operation* π as the comparison of the according concept (R_0, M_{15}^{π}) with a (objective) common concept (R_0, M_*) generated by a whole group of similar "unchanged" agents

$$I\left(M_*; M_{15}^{\pi}\right).$$

Intrinsic Symmetry

- Consider a "permutation of the embodiment" of the agent:

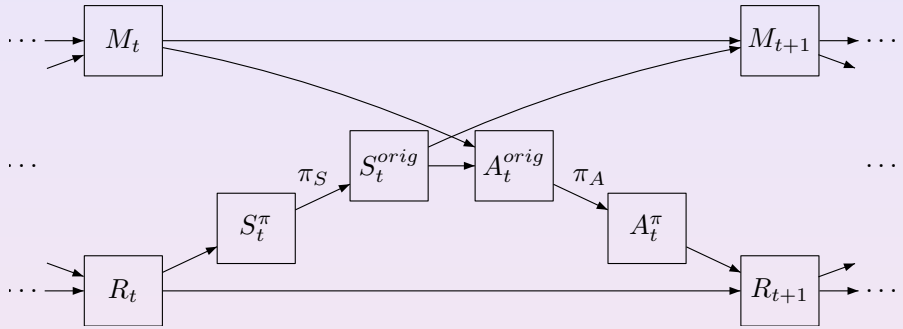


- With each permutation pair (π_S, π_A) the agent generates a different concept (R_0, M_{15}^{π}).
- Lets consider a group of similar (but not permuted) agents and there concepts $\left\{\left(R_0, M_{15}^{(1)}\right), \ldots, \left(R_0, M_{15}^{(n)}\right)\right\}$.
- Calculate based on this concepts a (objective) common concept (R_0, M_*).
- We define the *utility of a symmetry operation* π as the comparison of the according concept (R_0, M_{15}^{π}) with a (objective) common concept (R_0, M_*) generated by a whole group of similar "unchanged" agents

$$I\left(M_*; M_{15}^{\pi}\right).$$

Intrinsic Symmetry

- Consider a "permutation of the embodiment" of the agent:

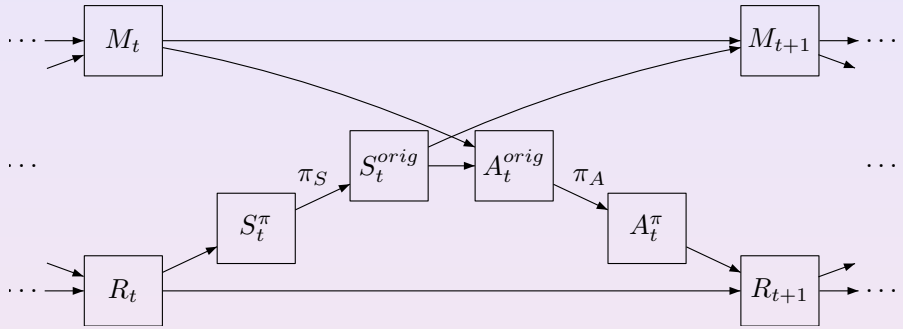


- With each permutation pair (π_S, π_A) the agent generates a different concept (R_0, M_{15}^{π}).
- Lets consider a group of similar (but not permuted) agents and there concepts $\left\{\left(R_0, M_{15}^{(1)}\right), \ldots, \left(R_0, M_{15}^{(n)}\right)\right\}$.
- Calculate based on this concepts a (objective) common concept (R_0, M_*).
- We define the *utility of a symmetry operation* π as the comparison of the according concept (R_0, M_{15}^{π}) with a (objective) common concept (R_0, M_*) generated by a whole group of similar "unchanged" agents

$$I\left(M_*; M_{15}^{\pi}\right).$$

Intrinsic Symmetry

- Consider a "permutation of the embodiment" of the agent:

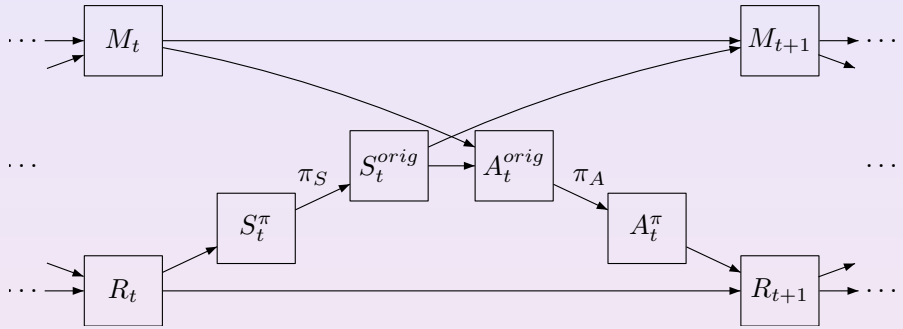


- With each permutation pair (π_S, π_A) the agent generates a different concept (R_0, M_{15}^{π}).
- Lets consider a group of similar (but not permuted) agents and there concepts $\left\{\left(R_0, M_{15}^{(1)}\right), \ldots, \left(R_0, M_{15}^{(n)}\right)\right\}$.
- Calculate based on this concepts a (objective) common concept (R_0, M_*).
- We define the *utility of a symmetry operation* π as the comparison of the according concept (R_0, M_{15}^{π}) with a (objective) common concept (R_0, M_*) generated by a whole group of similar "unchanged" agents

$$I\left(M_*; M_{15}^{\pi}\right).$$

Intrinsic Symmetry

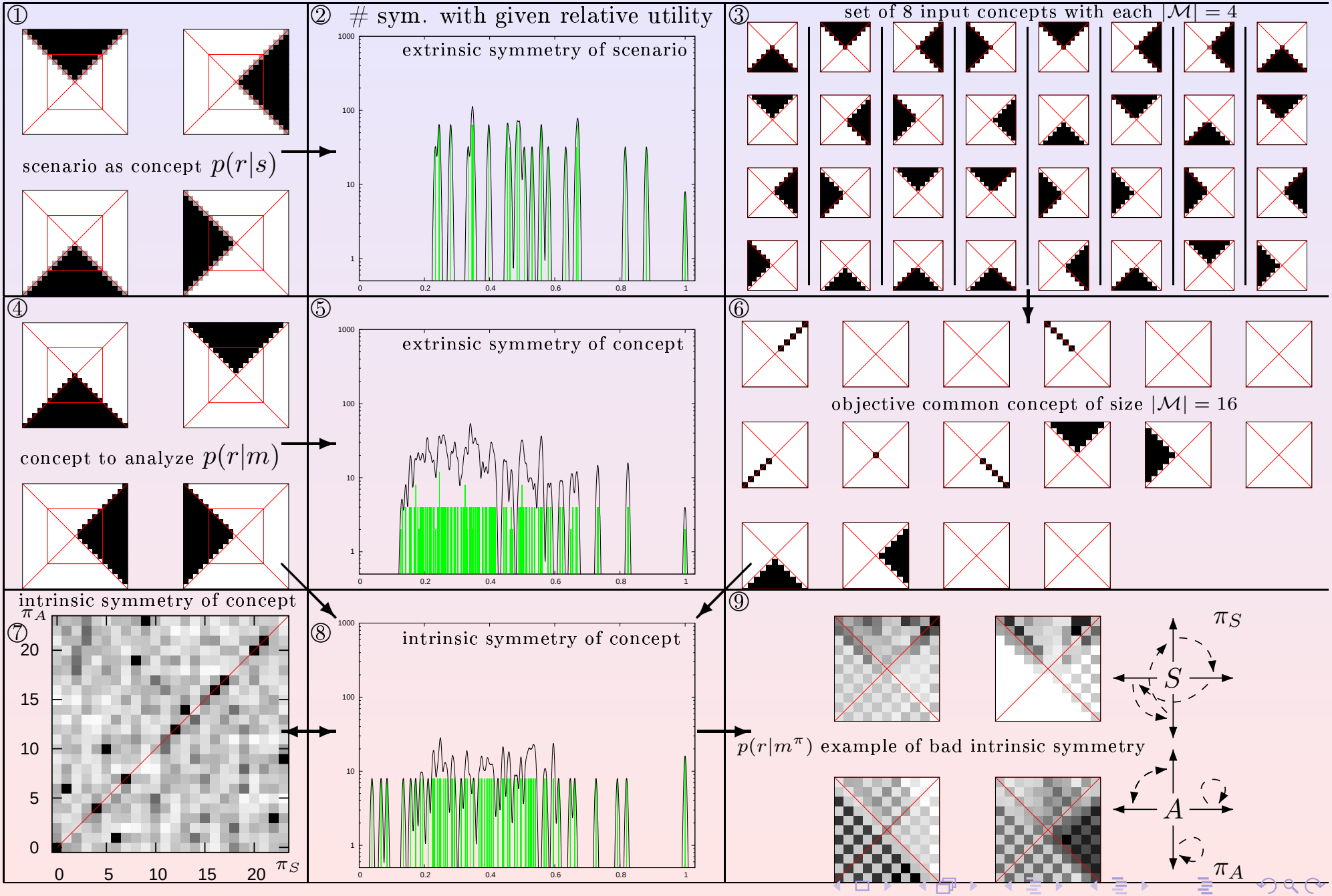
- Consider a "permutation of the embodiment" of the agent:



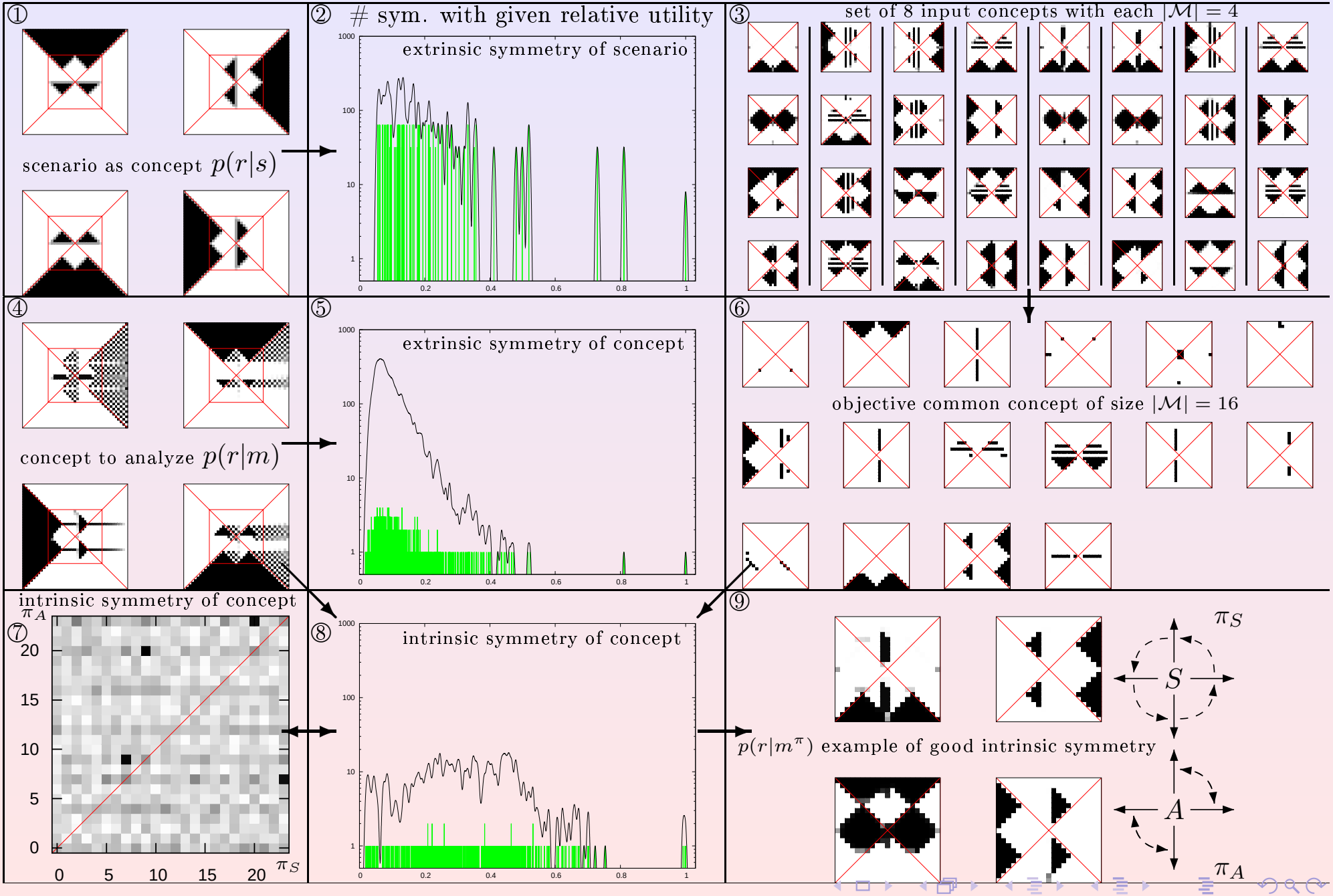
- With each permutation pair (π_S, π_A) the agent generates a different concept (R_0, M_{15}^{π}).
- Lets consider a group of similar (but not permuted) agents and there concepts $\left\{\left(R_0, M_{15}^{(1)}\right), \ldots, \left(R_0, M_{15}^{(n)}\right)\right\}$.
- Calculate based on this concepts a (objective) common concept (R_0, M_*).
- We define the *utility of a symmetry operation* π as the comparison of the according concept (R_0, M_{15}^{π}) with a (objective) common concept (R_0, M_*) generated by a whole group of similar "unchanged" agents

$$I\left(M_*; M_{15}^{\pi}\right).$$

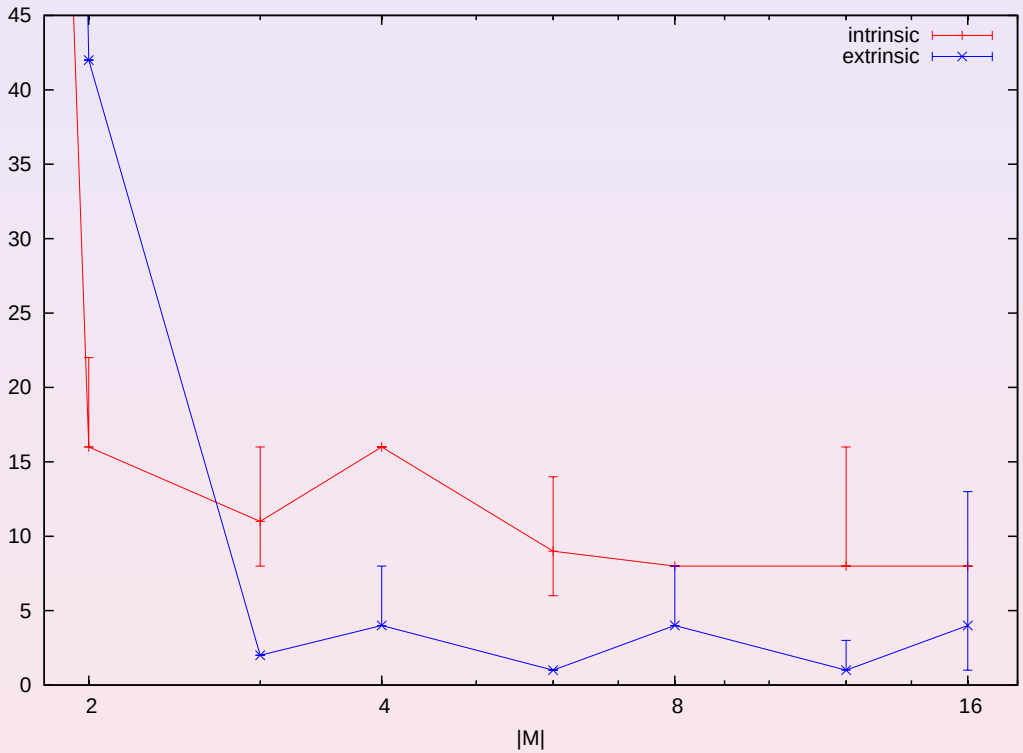
Results Scenario s+



Results Scenario sq



Number of Symmetries vs. Agent Memory Size



Amount of good (82,5% | 85% | 87.5% of max utility) intrinsic and extrinsic symmetries in memory size $|\mathcal{M}|$ dependence.

Questions?

http://www.xkcd.org/166