HAL open science

Influence of the Geometry of the world model on Curiosity Based Exploration

Grégoire Sergeant-Perthuis, Nils Ruet, David Rudrauf, Dimitri Ognibene, Yvain Tisserand

To cite this version:

Grégoire Sergeant-Perthuis, Nils Ruet, David Rudrauf, Dimitri Ognibene, Yvain Tisserand. Influence of the Geometry of the world model on Curiosity Based Exploration. 2024. hal-04054736v2

HAL Id: hal-04054736

https://hal.science/hal-04054736v2

Preprint submitted on 13 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Influence of the Geometry of the world model on Curiosity Based Exploration

Grégoire Sergeant-Perthuis GREGOIRESERPER@GMAIL.COM
LCQB Sorbonne Université & OURAGAN team, Inria Paris Paris, France.

Nils Ruet
CIAMS, Université Paris-Saclay, Orsay & Université d'Orléans, Orléans, France.

David Rudrauf
CIAMS, Université Paris-Saclay, Orsay & Université d'Orléans, Orléans, France.

Dimitri Ognibene
Dipartimento di Psicologia, Università Milano-Bicocca, Milan, Italy.

Yvain Tisserand
CISA, University of Geneva, Geneva, Switzerland.

Editor: *

Abstract

In human spatial awareness, 3-D projective geometry structures information integration and action planning through perspective taking within an internal representation space. The way different perspectives are related and transform a world model defines a specific perception and imagination scheme. In mathematics, such collection of transformations corresponds to a 'group', whose 'actions' characterize the geometry of space. Imbuing world models with a group structure may capture different agents' spatial awareness and affordance schemes. We used group action as a special class of policies for perspective-dependent control. We explored how such geometric structure impacts agents' behavior, comparing how the Euclidean *versus* projective groups act on epistemic value in active inference, drive curiosity, and exploration behaviors. We formally demonstrate and simulate how the groups induce distinct behaviors in a simple search task. The projective group's nonlinear magnification of information transformed epistemic value according to the choice of frame, generating behaviors of approach toward an object of interest. The projective group structure within the agent's world model contains the Projective Consciousness Model, which is know to capture key features of consciousnes. On the other hand, the Euclidean group had no effect on epistemic value : no action was better than the initial idle state. In structuring *a priori* an agent's internal representation, we show how geometry can play a key role in information integration and action planning.

Keywords: Geometric world model; Exploration; Embodied Cognitive Science; Cognitive Modeling; Perception-Action Coupling

1 Introduction

In artificial agent learning and control, intrinsic and extrinsic rewards can be combined to optimize the balance between exploration and exploitation. Intrinsic rewards in Reinforcement Learning (RL) (Hester and Stone, 2017; Merckling et al., 2022; Oudeyer et al., 2007) or terms of epistemic value in active inference (Friston et al., 2015) have been brought forth

as mechanisms mimicking curiosity and driving exploration, e.g. by integrating prediction error or uncertainty to drive actions favoring their reduction. However, efficient exploration is a computationally hard task. Recent neural planning models have increased planning flexibility and generality (Sekar et al., 2020). Yet, it is well-known that models' structures heavily impact planning performance and tractability (Geffner and Bonet, 2013) as well as learning complexity (Goyal and Bengio, 2022). A good representation of information may improve learning and search efficiency.

These issues are particularly salient for computation-heavy, highly recursive machine learning algorithms and applications, e.g. reinforcement learning (RL) in artificial agents (Bonet and Geffner, 2019) or recursive modeling method (RMM) in multi agent systems (MAS) and partially observable stochastic games (POSG) (Geffner and Bonet, 2013). Although generic neural world models can support exploration-related processes, incorporating prior knowledge that shapes internal representations to more effectively support exploration across a broad range of environments, such as 3-D environments, may enable autonomous agents to explore more complex and realistic settings on a larger scale (Goyal and Bengio, 2022). The exploration planning problem can thus be approached by considering how the structure of representation impacts exploration behaviors. In this article, the structure of representations is encoded into the geometry of the state space of an agent; we quantify the impact of changing this geometry on the behavior of the agent.

Here, we do not consider mechanisms of representation learning, e.g. in which world dynamics and action effects need to be learned and represented, as typically done in RL. We specifically consider control and execution when object locations, world states, and maps may not be known but dynamics, rewards, and action effects already are. We focus on action selection for environment exploration and mapping.

We adopt the active inference framework, i.e. an implementation of the Bayesian Brain Hypothesis aimed at generating adaptive behaviors in agents (Friston et al., 2006), that has found applications in neuroscience (Da Costa et al., 2020) and proposed for modeling molecular machines (Timsit and Sergeant-Perthuis, 2021; Timsit et al., 2021). It relies on an internal representation of the environment that an agent is driven to explore and exploit. The agent continually updates its beliefs about plausible competing internal hypotheses on the environment state. Under common sensory limitations, active inference relates to Partially Observable Markov Decision Process (POMDP) (Kaelbling et al., 1998; Ognibene et al., 2019). The epistemic value of states is a quantity that arises in active inference (Friston et al., 2015). Its maximization drives the agent's curiosity and actions.

For exploration or search in 3-D space, it is warranted to consider how geometrical principles could be embedded into efficient control mechanisms, to regularize the internal representation of information and mediate exploration under a drive of uncertainty reduction or information maximization. Geometrical considerations have previously been integrated into a variety of optimization and machine learning approaches, such as RL, active inference, and Bayesian inference (See Related Works below), but not in the specific perspective we introduce herein.

We build upon the hypothesis that 3-D internal representations of space in agents performing active inference may correspond to specific geometries, with properties that can be exactly analysed. More specifically, we consider how different first person perspectives may relate to each other, through transformations of a world model, as a specific

perception and imagination scheme for agents. This entails considering the action of geometrical groups of transformations (in the mathematical sense of the concept in Group theory; see Section 3.1) on the spatial distribution of information experienced and encoded by agents internally. The question is whether such group action could contribute to information gain estimation and maximization, as an internal planning or perspective-dependent control mechanism. Certain geometrical groups might imply internal representations, policies, value functions and principles of action that are particularly relevant for search and exploration. More specifically, we wish to compare how different groups impact the quantification of epistemic value. We then wish to characterize how the optimization of action from those different groups may yield different exploration behaviors. We contrasted two separate toy models of an agent performing a simple search task using active inference based solely on epistemic drives. One model used Euclidean geometry and the other projective geometry for the agent's internal space. We compared the two models in terms of resulting exploration behaviors and effects on epistemic value.

We chose to compare Euclidean versus projective geometry based on previous work, leveraging psychological research on the phenomenology of spatial consciousness and its role in the control of behaviors (Rudrauf et al., 2017, 2020, 2022, 2023). This research suggested that 3-D projective geometry plays a central role in human cognition and decision-making by shaping information representation and subsequent drives. It also offers a mechanism of changes of points of view and perspectives on a world model, including for perspective taking in social cognition, which is critical for the development of strategies of action planning in humans (see Rudrauf et al. 2022, 2023). We used Euclidean geometry as a standard baseline geometry for comparison (Ognibene and Demiris, 2013). Our geometrical rationale implies a different understanding of how agents' actions in their environment (here in the behavioral sense of the term) are implemented and selected compared to usual active inference. Agents' actions, such as navigation and approach-avoidance behaviors, can naturally be seen as dual to internal changes of perspective, i.e. group actions, in their representation space. We thus used group actions as a predictive model of actual behavioral actions.

The approach allowed us to formally study and demonstrate how the geometry governing the internal representation space may directly impact the computation of epistemic value and ensuing exploratory behaviors. Projective geometry *versus* Euclidean geometry demonstrated remarkable properties of information integration for motion planning under epistemic drive.

2 Related Works

2.1 Representation of Space and Exploration, in the Context of Machine Learning and Control

The integration of geometrical mapping in machine learning has been proposed to reduce the high-dimensionality of input spaces and provide efficient solutions for action selection and navigation. Seminal neurally inspired models used projections on 2-D manifolds for representation learning of complex spatial information and self-motion effects (Arleo et al., 2004). The impact of changes of perspective in exploration has long been of interest (Ognibene and Demiris, 2013). Ferreira et col. (Ferreira et al., 2013) proposed an internal 3-D egocentric, spherical representation of space, to modulate information sampling and

uncertainty as a function of distance, and control a robot attention through Bayesian inference. This was a seminal example of how geometrical rationale could suggest solutions to integrate perception and action planning.

Exploration methods must often maintain high-resolution representations of space to maximize information gain following action. This may hinder exploration efficiency, in particular in large-scale environments. 3-D topological representations of ambient space have been proposed as part of an abstract planning scheme, showing promising improvements of exploration efficiency (Yang et al., 2021).

Active vision principles, combined with curiosity-based algorithms and RL, were applied to the learning of saliency maps in the context of autonomous robots' navigation (Craye et al., 2016). The approach yielded promising optimization solutions to both adaptive learning of task-independent, spatial representations, and efficient exploration policies, which could serve as prior to support long-term, task-oriented, utility-driven RL mechanisms (Craye et al., 2016) (see also Ognibene and Baldassare, 2014; Sperati and Baldassarre, 2017).

Complex control tasks with continuous state and action spaces have been solved using deep reinforcement learning (DRL) with joint learning of representations and predictions. Such approach may entail non-stationarity, risks of instability and slow convergence, in particular in control tasks with active vision. Separating representation learning and policies' computations may mitigate the issues, but may also lead to inefficient information representations. Merckling et col. (Merckling et al., 2022) have sought to build compact and meaningful representations based on task-agnostic and reward-free agent-environment interactions. They used (recursive) state representation learning (SRL) while jointly learning a state transition estimator with near-future prediction objective, to contextually remove distracting information and reduce the exploration problem complexity. Positive outcomes were maximized through inverse predictive modeling, and prediction error was used to favor actions reducing uncertainty, which improved subsequent performance in RL tasks. The authors emphasized that dealing with partial observability through memory and active vision may require new solutions to both representation learning of hidden information and exploration strategies.

Uncertainty-based methods using intrinsic reward and exploration bonuses to plan trajectories have been criticized for inducing non-stationary decaying surprise, and for being hard to structure and optimize (Guo et al., 2021). Maximum State-Visitation Entropy (MSVE) was introduced to maximize state exploration uniformity, but optimization has been often challenging for large state spaces. Guo et col. (Guo et al., 2021) have introduced Geometric Entropy Maximization, which leverages geometry-aware entropy based on Adjacency Regularization (AR) and a similarity function, in order to optimize the MSVE problem at scale.

Geometrical constraints considered across these related works were not integrated into a global model, and were somewhat *ad hoc*. They pertained to a lower level of processing than the one we are concerned with here. However, they emphasize the current needs and challenges for integrating geometry in learning, control and navigation.

Methods and algorithms combining computer vision, machine learning and optimization, e.g. for robotic planning, have integrated group theoretic concepts to obtain, for instance, invariance to rotation and translation in image processing (Lee and Moore, 2004; Qin et al., 2019; Meng et al., 2017). Likewise, the leveraging of geometrical concepts, in

Deep learning, e.g. for learning manifolds and graphs, has been growing in recent years, demonstrating very promising results for representation learning (Gerken et al., 2021; Cao et al., 2022). The approach introduces combinatorial structures to leverage prior knowledge of geometry on the data of interest, e.g. applying 'convolutional Neural Networks' to non-Euclidean space. However, the Euclidean group E_3, or more specifically SE_3 (see Lee and Moore, 2004), which includes translations and rotations, but excludes reflections, or simply SO_3, the 3-dimensional rotation group (Gerken et al., 2021), are the groups being typically considered.

Here, in addition to the Euclidean group, we also consider $PGL(\mathbb{R}^3)$, the projective general linear group in 3-D, which acts on a projective space through projective transformations. The projective group is central to computer vision, for instance to generate 2-D images from 3-D information, but is used in such context in a restricted manner. We sought approaches based on cognitive science, considering spatial cognition and its relations to action at a higher level of integration, which does not reduce to the visual modality, but instead assume the mapping of multimodal information on a supramodal internal space of representation.

2.2 Projective Consciousness Model (PCM) and Active Inference

It has been shown that geometrically constrained active inference can be used as a framework to understand and model central aspects of human spatial consciousness, through the Projective Consciousness Model (PCM) (Rudrauf et al., 2017, 2022). According to this model, consciousness accesses and represents multimodal information through a Global Workspace (Dehaene et al., 2017) within which subjective perspectives on an internal world model can be taken. The process contributes to appraise possible actions based on their expected utility and epistemic value (Rudrauf et al., 2022). In publications on PCM (Rudrauf et al., 2017; Williford et al., 2018; Rudrauf et al., 2022, 2023; Williford et al., 2022), it was hypothesized that such internal representation space is geometrically structured as a $3-D$ projective space, denoted $P_3(\mathbb{R})$. Changes of perspective then correspond to the choice of a projective transformation ψ, i.e. an action from PGL_3. A projective transformation can also be modeled as a linear isomorphism $M_\psi \in GL_4(\mathbb{R})$ up to a multiplicative constant. The model yieled an explanation for the Moon illusion (Rudrauf et al., 2020) with falsifiable predictions on how strong the effect should be depending on context; as well as for the generation of adaptive and maladaptive behaviors, consistent with developmental and clinical psychology (see Rudrauf et al., 2022). Though essential in integrative spatial cognition, notably for understanding multi-agent social interactions, perspective taking is rarely integral to existing models of consciousness or formally implemented (Koch et al., 2016; Kleiner and Tull, 2021; Mashour et al., 2020; Dehaene et al., 2017; Merker et al., 2022). The PCM assumes that projective mechanisms of perspective changes are integral to the global workspace of consciousness, both in non-social and social contexts. The advantages of mechanisms of perspective taking for cybernetics remains to be fully formulated (see Rudrauf et al., 2022).

3 Methodology

The experiment we considered is that of an agent, denoted as a, which is looking for an object O in the 'real world', the 3-D Euclidean space $E_3 := \mathbb{R}^3$. The set of moves of the agent is denoted M. The position of O is denoted $o \in E_3$. The agent 'represents' the position of the object O inside its 'internal world model'. We consider 'internal world models', spaces denoted W, that are such that there is a group acting on them; we call such spaces, group structured world models. This group accounts for the change of coordinates that each movement of the agent induces when the positions of the object are expressed in the agents reference frame. We consider two spaces in particular:

1. *Euclidean case*: W is the 3-D vector space, $W = \mathbb{R}^3$
2. *Projective case*: W is the 3-D projective space, denoted as $P_3(\mathbb{R})$

We will denote $\mathcal{B}(W)$ the Borel σ-algebra of the respective topological spaces.

In Section 3.2, we explain how the 'real world' and the 'internal world model' are related to one another in both the Euclidean and Projective case. Figure 1 illustrates the setup of the toy model and main transformations considered. The agent's internal beliefs about the position of the object are encoded by a probability measure on W that the agent updates through observations. The agent explores its environment through the computation of an epistemic value, the maximization of which captures curiosity-based exploration. In Section 3.3, we explain how epistemic value is defined for group structured internal representations. In Section 3.4 we give the details of the exploration algorithm.

3.1 Group Structured World Model

Let us first recall what a group is.

Definition 1 (Group, §2 Chapter 1 Lang, 2012)**.** A group is a set G with an operation $. : G \times G \to G$ that is associative, such that there is an element $e \in G$ for which $e.g = g$ for any $g \in G$, and any $g \in G$ has an inverse denoted g^{-1} defined as satisfying, $g.g^{-1} = g^{-1}.g = e$.

We call a group structured world model, a world model provided with a group action; we now make this statement formal.

Definition 2 (Group structured world model)**.** W is a group structured world model for the group G when there is a map $h : G \times W \to W$ denoted as $h(g, x) = g.x$ for $g \in G$ and $x \in X$, such that,

1. $(g.g_1).x = g.(g_1.x)$ for all $g, g_1 \in G$, $x \in W$
2. $e.x = x$, for all $x \in W$

In the *Euclidean case* the group structured world model, W, is the 3-D vector space $\mathbb{R}^3$; it is structured by the group of invertible matrices $GL_3(\mathbb{R})$. In the *Projective case*, the group structured world model, W, is the projective space $P_3(\mathbb{R})$; it is structured by the group of projective linear transformations $PGL(\mathbb{R}^3)$.

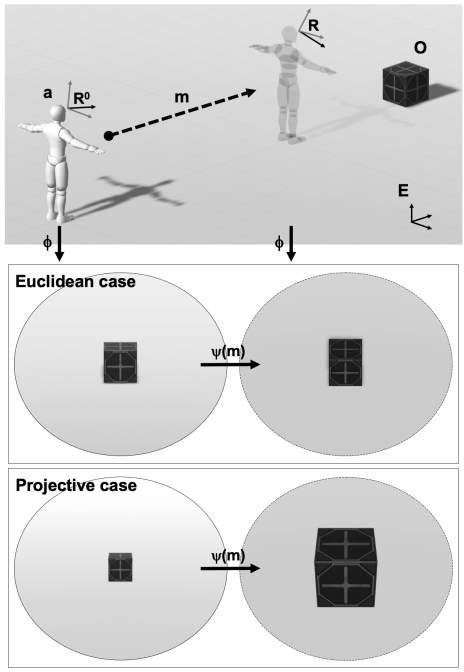

Figure 1: Toy model setup and main transformations

Upper-tier. Agent a simulates move m in Euclidean space E. $\mathcal{R}^0$ and $\mathcal{R}$ are its frames in E before and after the move, oriented toward object O. Vertical arrows indicate transformations ϕ from the external to the internal space. *Lower-tier.* Rendering of the effect of the internal group action $\psi(m)$ corresponding to move m in the Euclidean *versus* projective case. (Made with Unity).

3.2 Relating the 'Real World' to the 'Internal World Model'

We assume that the 'real world' is the 3-D Euclidean space, E_3. We assume that the 'real world' comes with with an Euclidean frame $\mathcal{R}_E$, i.e. a point $\mathcal{C}$ and three independent vectors e_0, e_1, e_2. This frame is used to set up the experiment: the configurations of the object and agent across time are encoded in this frame; it is fixed once and for all before starting the experiment. Therefore we now identify E_3 with $\mathbb{R}^3$, $\mathcal{C}$ with $(0,0,0)$ and e_0, e_1, e_2 with the respective basis vectors, $(1,0,0), (0,1,0), (0,0,1)$. The agent, denoted as a, is modeled as a solid in the 'real world'; it has its own Euclidean frame (the solid reference frame), $\mathcal{R} := (\mathcal{P}, u_0, u_1, u_2)$, with $\mathcal{P}$ the center of a and u_0, u_1, u_2 three unitary vectors that form a basis.

In the *Euclidean case*, the map that relates E_3 and its group structured world model, W, is the affine map, $\phi_{\mathcal{R}}$, that changes the coordinate in $\mathcal{R}_E$ to coordinates in $\mathcal{R}$.

In the *Projective case*, this map is a projective transformation. The choice of such a projective transformation is dictated by Proposition A.1 (Rudrauf et al., 2022). Let us now recall some facts about that transformation.

Let for any $(x, y, z) \in \mathbb{R}^3$,

$$\rho(x, y, z) = \left(\frac{x}{\gamma z + 1}, \frac{y}{\gamma z + 1}, \frac{z}{\gamma z + 1} \right) \tag{1}$$

with $\gamma \in \mathbb{R}_+$ a strictly positive parameter.

The (projective) transformation $\phi^p_{\mathcal{R}}$, from E_3 to W, which relates the 'real world' to the 'internal world model' in the projective case, is posed to be $\phi^p_{\mathcal{R}} := \rho \circ \phi_{\mathcal{R}}$.

Proposition 1. *When the agent a makes the move $m \in M$, its solid reference frame changes from $\mathcal{R}$ to $\mathcal{R}^m$. In the* Euclidean case *this move induces an invertible affine map, from the 'internal world model' to itself. In the* Projective case *it induces a projective transformation, $\psi_m \in PGL(\mathbb{R}^3)$.*

We denote 1_U or $x \to 1[x \in U]$ the characteristic function of subset U, i.e. the functions that is equal to 1 for $x \in U$ and 0 elsewhere.

Remark 1. In both cases there is a dense open subset, U, of W which is in continuous bijection with $\mathbb{R}^3$. From the Lebesgue measure dx on $\mathbb{R}^3$, we define the following measure on W, $d\lambda := 1_U \, dx$. In what follows we do not raise this technical point anymore and simply refer to $d\lambda$ as the Lebesgue measure on W.

3.3 Beliefs, Policies and Epistemic Value

3.3.1 Beliefs

The agent a keeps internal beliefs about the position of the object represented in its 'internal world model'; these beliefs are encoded by a probability measure $Q_X \in \mathbb{P}(W)$, where $\mathbb{P}(W)$ denotes the set of probability measures on W. Probability measures will be denoted with upper case letters and their densities with lower case letters. These beliefs are updated according to noisy sensory observations of the position of O. 'Markov Kernels' can be used to formalize noisy sensors. Let us recall their definition.

A 'Markov Kernel' Π from Ω_1 to Ω is a map $\Pi : \Omega \times \Omega_1 \to [0,1]$ such that for any $\omega_1 \in \Omega_1$, $\sum_{\omega \in \Omega} \Pi(\omega|\omega_1) = 1$, i.e. a map that sends any $\omega_1 \in \Omega_1$ to a probability measure $\Pi_{|\omega_1} \in \mathbb{P}(\Omega)$.

The uncertainty on the sensors of a is captured by a Markov kernel $P_{Y|X}$ from W to W. It is a parameter of the experiment: it is fixed before the agent starts looking for O. The couple $(P_{Y|X}, Q_X)$ defines the following probability density, $P_{X,Y} \in \mathbb{P}(W \times W)$: for any $x, y \in W$,

$$P_{X,Y}(dx, dy) := p_{Y|X}(y|x) q_X(x) dx dy \tag{2}$$

where dx is the Lebesgue measure on W. An observation of the position of the object $y^o \in W$ triggers an update of the belief Q_X to the belief with density

$$Q_{X|y^o} = \frac{p_{X,Y}(x|y^o) q_X(x) dx}{\int_{x \in W} p_{X,Y}(x|y^o) q_X(x) dx} \tag{3}$$

3.3.2 Policies

Recall that the agent has a set of moves it can make M; moves $m \in M$ are associated with the group action $\psi_m : W \to W$ (Proposition 1). The agent plans the consequence of its moves on its internal world model one step ahead: each change of frame induces the following Markov Kernel, for any $m \in M$, $A \in \mathcal{B}(W)$, and $x_0 \in W$,

$$p_{X_1|X_0,m}(A|x_0, m) = 1[\psi_m(x_0) \in A] \tag{4}$$

Each move m spreads a prior Q_X on X_0 into the following prior on X_1: $\forall A \in \mathcal{B}(W)$,

$$\psi_{m,*} Q_X(A) := \int 1[\psi_m(x_0) \in A] q_X(x_0) dx \tag{5}$$

$$= Q_X(\psi_m^{-1} A) \tag{6}$$

We chose to denote this probability measure as $\psi_{m,*} Q_X$, because it is the standard mathematical notation for the 'pushforward measure' by ψ_m. The generative model the agent uses to plan its future actions is summarized in Figure 2.

3.3.3 Epistemic Value

The following definition is a restatement of the *epistemic value* introduced in (Friston et al., 2015) in the case of the kernel $P_{Y|X} : W \to W$.

Definition 3 (Epistemic Value)**.** For any probability measure $Q_X \in \mathbb{P}(W)$, the epistemic value of this measure is:

$$C(Q_X) := \mathbb{E}_{P_Y}\left[H(P_{X|Y}|Q_X)\right] \tag{7}$$

$$= \int p_Y(y) dy \int p_{X|Y}(x|y) \ln \frac{p_{X|Y}(x|y)}{q_X(x)} dx \tag{8}$$

H is the relative entropy, also called Kullback-Leibler divergence.

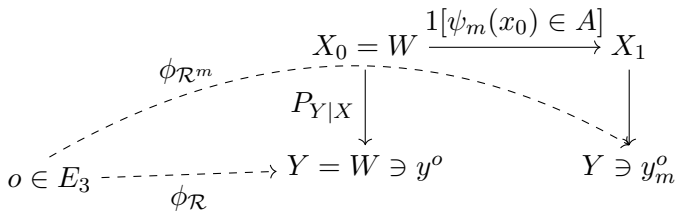

Figure 2: $m \in M$ is a move of the agent a, $1[\psi_m(x_0) \in A]$ defines the kernel induced by move m, $P_{Y|X}$ is the noisy sensor. The diagram constituted of solid arrows defines the generative model the agent uses to plan its actions. o is the position of the object in the 'real world', $y^o \in W$ is the representation of o in the 'internal world model' of a with respect to the solid reference frame $\mathcal{R}$, y^o_m is the same for the reference frame $\mathcal{R}^m$ after move m.

Reexpressing Equation 7, it becomes apparent that epistemic value is simply a mutual information:

$$C(Q_X) = \int p_{X,Y} \ln \frac{p_{X,Y}(x,y)}{p_Y(y) q_X(x)} dxdy \tag{9}$$

We propose to define the epistemic value of move m as the epistemic value of the induced prior on X_1,

$$C(m) := C(\psi_{m,*} Q_X) \tag{10}$$

3.4 Exploration Algorithm

Let us now put the previous elements together to describe the exploration behavior programmed in our agent. The agent a is initialized in a configuration of the 'real world', with solid reference frame $\mathcal{R}^0$; the object O is positioned at $o \in E_3$. a starts with an initial belief $Q^0_X \in \mathbb{P}(W)$ on the position of O. It plans one step ahead the consequence of move m; move m induces a group action $\psi_m : W \to W$ that pushes forward the belief Q^0_X to $\psi_{m,*} Q^0_X$. The agent then evaluates the epistemic value of $(P_{Y|X}, \psi_{m,*} Q^0_X)$ for each move m and chooses the move that maximizes this value, m^*. a executes the move m^* which transforms its solid reference frame $\mathcal{R}_0$ to $\mathcal{R}$. It can then observe (with its 'noisy sensors') the position of O which is $y^o := \phi_{\mathcal{R}}(o)$ in its internal world model, which triggers the update of prior $\psi_{m,*} Q^0_X$ to the distribution conditioned on the observation: $\left(\psi_{m,*} Q^0_X\right)_{|y^o_m}$. The process is iterated with this new prior. The exploration algorithm is summarized in Algorithm 1.

4 Theoretical Predictions

We wish to understand how the group by which the internal world model is structured influences the exploration behavior of the agent. The *Euclidean case* serves as the reference model; in this case the world model shares the same structure as the real world: it is the 'classical' way of modeling this exploration problem. The *Projective case* corresponds to the

Algorithm 1: Curiosity based Exploration for agent a

Data: Initialization: Q_X^0 initial belief, $\mathcal{R}^0$ initial solid reference frame of a

$Q_X \leftarrow Q_X^0$;
while *True* **do**
  $m^* \leftarrow \text{argmax}_{m \in M} C(\psi_{m,*} Q_X)$;
  $\mathcal{R} \leftarrow$ solid reference frame of a after move m^*;
  $Q_X \leftarrow \psi_{m,*} Q_X$;
  $y^o \leftarrow \phi_{\mathcal{R}}(o)$;
  $Q_X \leftarrow Q_{X|y^o}$;
end

hypothesis underlying the PCM. The following Theorem states that this experiment allows us to discriminate when the behavior of the agent is dictated by 'objective' perspectives (Euclidean change of frame) *versus* by 'subjective' perspectives (projective change of frame) on its environment.

We consider the following noisy sensor, for any $x, y \in \mathbb{R}^3$,

$$P_{Y|X}(y|x) = \frac{3}{4\pi\epsilon^3} 1[\|x - y\| \leq \epsilon] \tag{11}$$

where $\|.\|$ designates the Euclidean norm on $\mathbb{R}^3$, i.e. $\|x\|^2 = x_0^2 + x_1^2 + x_2^2$; $\epsilon > 0$ is a strictly positive real number.

Theorem (Discrimination of behavior with respect to internal representations)**.** *Let us assume that staying still is always a possible move for the agent.*

Euclidean case: *when the agent has an objective representation of its environment, given by an affine map, the agent stays still.*

Projective case: *Assume now that the set of moves M is finite; assume furthermore that after any possible move, the agent faces O, in other words, we assume that the agent knows in which direction to look in order to find the object but is still uncertain on* where *the object is exactly. If it has a 'subjective' perspectives, i.e. its representation is given through a projective transformation, it will choose the moves that allows it to approach O (for any ϵ small enough).*

Proof. The details of the proof are given in Appendix A.2. Let us here sketch the proof. The agent circumscribes a region of space in which it believes it is likely to find the object. This region corresponds to the error the agent tolerates on the measurement it makes of the position of O; we can also see it as the precision up to which the agent measures the position of O. In the *Euclidean case*, the region in which the agent circumscribes the object appears to always be of the same size, irrespective of the agent's configuration with respect to the object. Therefore not moving ends up being an optimal option and the agent will not approach the object without additional extrinsic reward. In the *Projective case*, the agent can 'zoom' on this region in order to gain more precision in measuring o; the configurations of the agent in which this region is magnified are more informative regarding the position of O and therefore preferred by the agent. The only way for the agent to actually zoom

onto this area is to approach the location it believes O is likely to be, therefore the agent will end up approaching O.

□

Remark 2. This particular choice of Markov kernel (Equation 11) allows for an explicit expression of epistemic value which simplifies the proof of the result; however we expect the result to hold for a larger class of kernels.

In the next section we present an implementation of this experiment and simulation results.

5 Simulations

5.1 Methods

Algorithm 1 is implemented in the following manner (source code available at `https://github.com/NilsRuet/effect-of-geometry-on-exploration`). Beliefs and the Markov kernel corresponding to sensors were considered to be multivariate normal distributions, that is $P_{Y|X} \sim \mathcal{N}(\mu_{Y|X}, \Sigma_{Y|X})$ and $Q_X \sim \mathcal{N}(\mu_X, \Sigma_X)$. Belief update through the action of a group was approximated using a Gaussian distribution; a projective transformation changes a Gaussian distribution into a non Gaussian one which is difficult to describe. Therefore we replace this non-Gaussian distribution with a Gaussian distribution with same mean and variance.

We assumed $\mu_{Y|X} = x$ (which implies $\mu_y = \mu_x$) and $\Sigma_{Y|X} = \epsilon^2\mathbb{I}$ where $\mathbb{I}$ is the identity matrix and $\epsilon > 0$ a positive real number. As a result, for a given observation y^o, $Q_{X|y^o}$ and $C(\psi_{m,*}Q_X)$ can be computed efficiently. The joint distribution P on X, Y is a Gaussian distribution:

$$P(x, y) = P_{Y|X}(y|x)Q_X(x)$$
$$P(x, y) \sim \mathcal{N}(\mu_{X,Y}, \Sigma_{X,Y})$$

with $\mu_{X,Y} = (\mu_X, \mu_X)$ and

$$\Sigma_{X,Y} = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XX} \\ \Sigma_{XX} & \epsilon^2\mathbb{I} + \Sigma_{XX} \end{pmatrix} \tag{12}$$

The variance of Y is $\Sigma_{YY} = \epsilon^2\mathbb{I} + \Sigma_{XX}$.

The joint distribution being Gaussian entails that the distribution of X conditioned on $y = y^o$ is also Gaussian, thus $Q_{X|y^o} \sim \mathcal{N}(\mu_{X|y^o}, \Sigma_{X|y^o})$. Applying Proposition 3.13.(Eaton, 2007) to our setting, the mean and covariance of the conditioned distribution are given by:

$$\mu_{X|y^o} = \mu_X + \Sigma_{XX}\Sigma_{YY}^{-1}(y^o - \mu_X) \tag{13}$$

$$\Sigma_{X|y^o} = \Sigma_{XX} - \Sigma_{XX}\Sigma_{YY}^{-1}\Sigma_{XX} \tag{14}$$

Epistemic value is computed using the Kullback-Leibler divergence. With full knowledge of the joint distribution, in the Gaussian case, following the expression of entropy for gaussian vectors (Chapter 12 Equation (12.39) Cover and Thomas, 2006) it is computed as:

$$C(Q_X) = I(X; Y) = \frac{1}{2} \ln \frac{(\det \Sigma_{XX})(\det \Sigma_{YY})}{\det \Sigma_{X,Y}} \tag{15}$$

The set of moves that can be selected by the agent is restricted to translations as the agent must always face the object. This constraint, as well as the choice of a simple model of noisy sensor (with homogeneous precision and resolution), was motivated by the aim of making formal demonstrations of theorems tractable, and implementations tightly related to the theoretical predictions. This is a departure from how spatial sampling typically operates in perception, e.g. decreasing resolution of the visual field with eccentricity (see (Rudrauf et al., 2022, 2023) for more realistic but less formally analyzable applications of a broader version of the PCM). However, the specific aim herein of demonstrating fundamental properties of the action of different geometrical groups on epistemic value and ensuing behaviors of approach motivated such restrictions.

The set of possible translations is composed of eight translations with the same norm, with evenly distributed angles (one of them being oriented toward the object irrespective of the position of the agent), and also contained an idle state, i.e. no translation. Here the angles correspond to the angles of the translation and not a rotation angle of the solid frame of the agent as the agent must always face the object.

We approximated the belief after the action m of a given group using a Gaussian distribution, $\psi_{m,*}Q_X \sim \mathcal{N}(\mu_m, \Sigma_m)$. The mean and covariance matrix are approximated using numerical integration:

$$\mu_m = \int xp(\psi_m^{-1}(x))\frac{1}{|\det J_{\psi_m}(\psi_m^{-1}(x))|}dx \tag{16}$$

$$\Sigma_m = \int (x-\mu_m)(x-\mu_m)^T p(\psi_m^{-1}(x))\frac{1}{|\det J_{\psi_m}(\psi_m^{-1}(x))|}dx \tag{17}$$

We ran two sets of simulations. In the first one (Figure 3 left tier), the agent started from an initial position with the object always located at a fixed position, and the algorithm was applied for 20 iterations, for both the Euclidean and Projective internal spaces. The agent started at $(0, 0)$ and the object was located at $(0, 2)$ in the world frame E_2 spanning the agent's displacement floor. If all translations were associated with epistemic values that only varied within a small range ($\pm 1e-4$) as compared to the epistemic value of the idle state, reflecting numerical imprecision, the idle state was selected (the agent did not move). The aim of this set of simulations was to compare trajectories of agents displacements through time across the two geometries. In the second set of simulations (Figure 3 right tier), the agent started at the center of a 20×20 grid of possible positions of the object.

The positions that are too close to the agent are excluded so that a sufficient effect can be observed. For each object position, we considered only the first set of translations that the agent could envision from its initial position. We computed epistemic value across the set of possible translations of the agent for each object position and for both the Euclidean and projective internal spaces. The aim of this set of simulations was to be able to systematically compare epistemic values across the two geometries for all possible positions of objects.

5.2 Results

Figure 3 left tier shows a representative example of trajectories obtained in the Euclidean *versus* projective cases. In the projective case, the agent always approached the object. In the Euclidean case, the agent always stayed idle. Figure 3 right tier shows epistemic value as

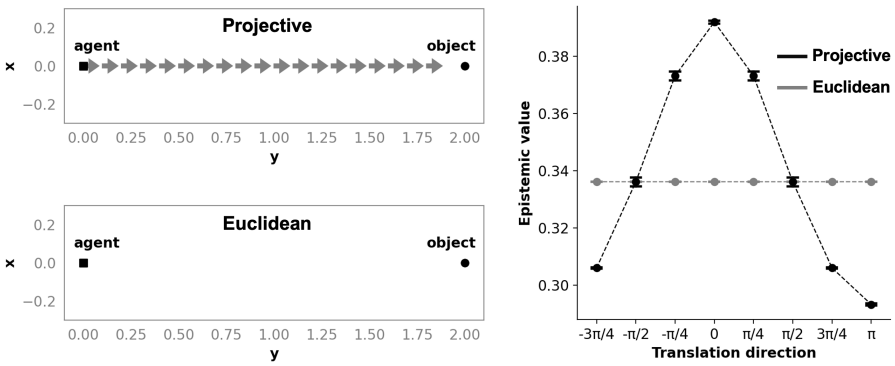

Figure 3: Simulation results

Left-tier. Trajectory of the agent for the projective *versus* Euclidean internal spaces. *Right-tier.* Epistemic value as a function of directions of translation with respect to the object direction, for the projective *versus* Euclidean internal spaces. Points are average values across comparable directions, and error-bars are standard errors.

a function of translation direction expressed in radians for both the projective and Euclidean cases. The direction of 0 radian corresponds to the object direction. We averaged epistemic values across object positions within comparable directions. In the Projective case, average epistemic value was maximal for the direction of the object, and decreased with directions farther away from it. In the Euclidean case, average epistemic value was identical across directions. Maximum average epistemic value was much higher for the projective case than for the Euclidean case.

6 Discussion

In this article, we introduced a generative model for environment exploration based on a first-person perspective in which actions are encoded as changes in perspective. The families of geometry for the world model encode possible 'kinds' of perspective-taking on the environment and structures the representation of sensory evidence within the world model of the agent. In other words, each family corresponds to a specific perception and imagination scheme for the agent. We encoded two such families, namely the Euclidean *versus* projective group as acting on the internal world model of the agent, i.e., within the geometric properties of this internal world model. We showed that different geometries induce different behaviors, focusing on the two cases: when the internal world model of the agent followed Euclidean geometry *versus* projective geometry. This result contributes to understanding how integrative geometrical processing and principles can play a central role in cybernetics. In our approach, the geometry of the world model links perceptual and imaginary representations with actions and behaviors.

Although beyond the scope of this article, we are interested in generalizing the approach to compare the results obtained with those that could be obtained with other models in which geometry plays a central integrative role, e.g. as in (Ferreira et al., 2013) who used an internal 3-D egocentric, spherical representation of space. Such an approach would need

to be expressed in terms of group action to make it formally comparable to our approach. Likewise, it would be interesting to compare other groups, and other more realistic models of sensors, and what types of behaviors they would induce. This would be useful to refine predictions and envision experimental designs for empirical validation in humans. We would also like to use more sophisticated settings (see for instance Rudrauf et al., 2022, 2023), even though it may be incompatible with the derivation of analytical solutions, but could lead to richer simulations and induction of behaviors.

We also wish to further examine how the geometry of a latent space intertwines with information processing. One motivation is theoretical, as we would like to assess how geometry changes learning behavior (Goyal and Bengio, 2022). In this contribution, we have discarded representation learning *per se*, as it was beyond the scope of this study focusing on planning. In future work, we intend to use deep learning to learn group structured representations. It is important to note that such approach differs from geometric deep learning (Bronstein et al., 2021; Sergeant-Perthuis et al., 2022) as we do not seek to learn equivariant representations: a group structure will only be considered for the internal world model but none will be presupposed on the observation side. Likewise, we are interested in examining how geometry may play a role in overt and covert attention.

Our contribution can simplify the design of novel agent architectures where exploratory and sensory choices of actions naturally emerge as a consequence of the internal representation and the reflection of the perceptual mechanism of the agent's embodiment. The study of world models with projective geometries was motivated by ongoing work in computational psychology aimed at reproducing features of consciousness. Projective geometry induces effects of magnification and focalization on information that appear immediately relevant for spatial attention, and more generally for contextual salience. Another motivation for this research is more practical, as we would like to use such principles to design virtual and robotic artificial agents mimicking human cognition and behaviors following (Rudrauf et al., 2022, 2023).

References

A. Arleo, F. Smeraldi, and W. Gerstner. Cognitive navigation based on nonuniform gabor space sampling, unsupervised growing networks, and reinforcement learning. *IEEE Transactions on Neural Networks*, 15(3):639–652, 2004. doi: 10.1109/TNN.2004.826221.

Blai Bonet and Hector Geffner. Learning first-order symbolic representations for planning from the structure of the state space. *arXiv preprint arXiv:1909.05546*, 2019.

Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges, 2021. URL `https://arxiv.org/abs/2104.13478`.

Wenming Cao, Canta Zheng, Zhiyue Yan, and Weixin Xie. Geometric deep learning: progress, applications and challenges. *Science China Information Sciences*, 65(2):126101, 2022.

Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, USA, 2006. ISBN 0471241954.

Céline Craye, David Filliat, and Jean-François Goudou. Rl-iac: An exploration policy for online saliency learning on an autonomous mobile robot. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4877–4884, 2016. doi: 10.1109/IROS.2016.7759716.

Lancelot Da Costa, Thomas Parr, Noor Sajid, Sebastijan Veselic, Victorita Neacsu, and Karl Friston. Active inference on discrete state-spaces: A synthesis. *Journal of Mathematical Psychology*, 99:102447, 2020. ISSN 0022-2496. doi: https://doi.org/10.1016/j.jmp.2020.102447. URL `https://www.sciencedirect.com/science/article/pii/S0022249620300857`.

Stanislas Dehaene, Hakwan Lau, and Sid Kouider. What is consciousness, and could machines have it? *Science*, 358(6362):486–492, 2017.

Morris L. Eaton. Multivariate statistics: A vector space approach. *Lecture Notes-Monograph Series*, 53:i–512, 2007. ISSN 07492170. URL `http://www.jstor.org/stable/20461449`.

João Ferreira, Jorge Lobo, Pierre Bessiere, Miguel Castelo-Branco, and Jorge Dias. A bayesian framework for active artificial perception. *IEEE transactions on cybernetics*, 43 (2):699–711, 2013.

Karl Friston, James Kilner, and Lee Harrison. A free energy principle for the brain. *Journal of Physiology-Paris*, 100(1):70–87, 2006. ISSN 0928-4257. doi: https://doi.org/10.1016/j.jphysparis.2006.10.001. URL `https://www.sciencedirect.com/science/article/pii/S092842570600060X`. Theoretical and Computational Neuroscience: Understanding Brain Functions.

Karl Friston, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo. Active inference and epistemic value. *Cognitive Neuroscience*, 6 (4):187–214, 2015. ISSN 17588936. doi: 10.1080/17588928.2015.1020053.

Hector Geffner and Blai Bonet. A concise introduction to models and methods for automated planning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 8 (1):1–141, 2013.

Jan E Gerken, Jimmy Aronsson, Oscar Carlsson, Hampus Linander, Fredrik Ohlsson, Christoffer Petersson, and Daniel Persson. Geometric deep learning and equivariant neural networks. *arXiv preprint arXiv:2105.13926*, 2021.

Anirudh Goyal and Yoshua Bengio. Inductive biases for deep learning of higher-level cognition. *Proceedings of the Royal Society A*, 478(2266):20210068, 2022.

Zhaohan Daniel Guo, Mohammad Gheshlaghi Azar, Alaa Saade, Shantanu Thakoor, Bilal Piot, Bernardo Avila Pires, Michal Valko, Thomas Mesnard, Tor Lattimore, and Rémi Munos. Geometric entropic exploration. *arXiv preprint arXiv:2101.02055*, 2021.

Todd Hester and Peter Stone. Intrinsically motivated model learning for developing curious robots. *Artificial Intelligence*, 247:170–186, 2017.

Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134, 1998.

Johannes Kleiner and Sean Tull. The mathematical structure of integrated information theory. *Frontiers in Applied Mathematics and Statistics*, 6, 2021. ISSN 2297-4687. doi: 10.3389/fams.2020.602973. URL `https://www.frontiersin.org/article/10.3389/fams.2020.602973`.

Christof Koch, Marcello Massimini, Melanie Boly, and Giulio Tononi. Neural correlates of consciousness: progress and problems. *Nature Reviews Neuroscience*, 2016.

Serge Lang. *Algebra*, volume 211. Springer Science & Business Media, 2012.

Pei Yean Lee and John B Moore. Geometric optimization for 3d pose estimation of quadratic surfaces. In *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, volume 1, pages 131–135. IEEE, 2004.

George A. Mashour, Pieter Roelfsema, Jean-Pierre Changeux, and Stanislas Dehaene. Conscious processing and the global neuronal workspace hypothesis. *Neuron*, 105(5):776–798, 2020. ISSN 0896-6273. doi: https://doi.org/10.1016/j.neuron.2020.01.026. URL `https://www.sciencedirect.com/science/article/pii/S0896627320300520`.

Zehui Meng, Hao Sun, Hailong Qin, Ziyue Chen, Cihang Zhou, and Marcelo H Ang. Intelligent robotic system for autonomous exploration and active slam in unknown environments. In *2017 IEEE/SICE International Symposium on System Integration (SII)*, pages 651–656. IEEE, 2017.

Astrid Merckling, Nicolas Perrin-Gilbert, Alex Coninx, and Stéphane Doncieux. Exploratory state representation learning. *Frontiers in Robotics and AI*, 9, 2022.

Bjorn Merker, Kenneth Williford, and David Rudrauf. The integrated information theory of consciousness: a case of mistaken identity. *Behavioral and Brain Sciences*, 45, 2022.

Dimitri Ognibene and Gianluca Baldassare. Ecological active vision: four bioinspired principles to integrate bottom–up and adaptive top–down attention tested with a simple camera-arm robot. *IEEE transactions on autonomous mental development*, 7(1):3–25, 2014.

Dimitri Ognibene and Yiannis Demiris. Towards active event recognition. In *Twenty-Third International Joint Conference on Artificial Intelligence*, pages 2495–2501, 2013.

Dimitri Ognibene, Lorenzo Mirante, and Letizia Marchegiani. Proactive intention recognition for joint human-robot search and rescue missions through monte-carlo planning in pomdp environments. In Miguel A. Salichs, Shuzhi Sam Ge, Emilia Ivanova Barakova, John-John Cabibihan, Alan R. Wagner, Álvaro Castro-González, and Hongsheng He, editors, *Social Robotics*, pages 332–343, Cham, 2019. Springer International Publishing. ISBN 978-3-030-35888-4.

Pierre-Yves Oudeyer, Frdric Kaplan, and Verena V Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE transactions on evolutionary computation*, 11(2):265–286, 2007.

Hailong Qin, Zehui Meng, Wei Meng, Xudong Chen, Hao Sun, Feng Lin, and Marcelo H Ang. Autonomous exploration and mapping system using heterogeneous uavs and ugvs in gps-denied environments. *IEEE Transactions on Vehicular Technology*, 68(2):1339–1350, 2019.

D. Rudrauf, G. Sergeant-Perthuis, O. Belli, Y. Tisserand, and G. Di Marzo Serugendo. Modeling the subjective perspective of consciousness and its role in the control of behaviours. *Journal of Theoretical Biology*, 534:110957, 2022. ISSN 0022-5193. doi: https://doi.org/10.1016/j.jtbi.2021.110957. URL `https://www.sciencedirect.com/science/article/pii/S0022519321003763`.

D Rudrauf, G Sergeant-Perhtuis, Y Tisserand, T Monnor, V De Gevigney, and O Belli. Combining the Projective Consciousness Model and Virtual Humans for immersive psychological research: a proof-of-concept simulating a ToM assessment. *ACM Transactions on Interactive Intelligent Systems*, 2023.

David Rudrauf, Daniel Bennequin, Isabela Granic, Gregory Landini, Karl Friston, and Kenneth Williford. A mathematical model of embodied consciousness. *Journal of theoretical biology*, 428:106–131, 2017.

David Rudrauf, Daniel Bennequin, and Kenneth Williford. The moon illusion explained by the projective consciousness model. *Journal of Theoretical Biology*, 507:110455, 2020.

Ramanan Sekar, Oleh Rybkin, Kostas Daniilidis, Pieter Abbeel, Danijar Hafner, and Deepak Pathak. Planning to explore via self-supervised world models. In *International Conference on Machine Learning*, pages 8583–8592. PMLR, 2020.

Grégoire Sergeant-Perthuis, Jakob Maier, Joan Bruna, and Edouard Oyallon. On non-linear operators for geometric deep learning. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 10984–10995. Curran Associates, Inc., 2022. URL `https://proceedings.neurips.cc/paper_files/paper/2022/file/474815daf1d4096ff78b7e4fdd2086a5-Paper-Conference.pdf`.

Valerio Sperati and Gianluca Baldassarre. Bio-inspired model learning visual goals and attention skills through contingencies and intrinsic motivations. *IEEE Transactions on Cognitive and Developmental Systems*, 10(2):326–344, 2017.

Youri Timsit and Grégoire Sergeant-Perthuis. Towards the idea of molecular brains. *International Journal of Molecular Sciences*, 22(21), 2021. ISSN 1422-0067. doi: 10.3390/ijms222111868. URL `https://www.mdpi.com/1422-0067/22/21/11868`.

Youri Timsit, Grégoire Sergeant-Perthuis, and Daniel Bennequin. Evolution of ribosomal protein network architectures. *Scientific Reports*, 11(1):625, January 2021.

ISSN 2045-2322. doi: 10.1038/s41598-020-80194-4. URL https://doi.org/10.1038/s41598-020-80194-4.

Kenneth Williford, Daniel Bennequin, Karl Friston, and David Rudrauf. The projective consciousness model and phenomenal selfhood. *Frontiers in Psychology*, 9:2571, 2018.

Kenneth Williford, Daniel Bennequin, and David Rudrauf. Pre-reflective self-consciousness & projective geometry. *Review of Philosophy and Psychology*, 13(2):365–396, 2022.

Fan Yang, Dung-Han Lee, John Keller, and Sebastian Scherer. Graph-based topological exploration planning in large-scale 3d environments. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 12730–12736, 2021. doi: 10.1109/ICRA48506.2021.9561830.

Appendix A. Proof of Proposition and Theorem

A.1 Proof of Proposition 1

Any 3-D affine transformation is encoded by a matrix $M = (m_{i,j}; i, j = 1..3)$ and a vector $(m_{j,4}; j = 1..3)$; let $(m^{\mathcal{R}}_{i,j}; i, j = 1..3)$ be the matrix associated to $\phi_{\mathcal{R}}$ and $(m^{\mathcal{R}}_{4,j}; j = 1..3)$ its vector.

Projective case: $\phi^p_{\mathcal{R}} = \rho \circ \phi_{\mathcal{R}}$ is the projective map with expression in homogeneous coordinates given by the matrix,

$$\begin{pmatrix} m_{1,1} & m_{1,2} & m_{1,3} & m_{1,4} \\ m_{2,1} & m_{2,2} & m_{2,3} & m_{2,4} \\ m_{3,1} & m_{3,2} & m_{3,3} & m_{3,4} \\ 0 & 0 & \gamma & 1 \end{pmatrix}$$

By construction, the transition map in the projective case, ψ^p_m, is $\phi^p_{\mathcal{R}^m} \circ {\phi^p_{\mathcal{R}}}^{-1}$; it is the composition of two projective transformations, therefore it is a projective transformation.

A.2 Proof of Theorem 1

We will denote B^ϵ_y the Euclidean ball of radius 1 around $y \in \mathbb{R}^3$, i.e. $B^\epsilon_y = \{x \in \mathbb{R}^3 | \quad \|x - y\| \leq \epsilon\}$.

Lemma 1. *For any $Q \in \mathbb{P}(W)$, both in Euclidean and Projective cases, for any affine map or projective transformation $\psi : W \to W$,*

$$C(\psi_* Q) = -\int dy Q(\psi^{-1}(B^\epsilon_y)) \ln Q(\psi^{-1}(B^\epsilon_y)) \tag{18}$$

Proof.

$$C(\psi_* Q) = \frac{3}{4\pi\epsilon^3} \times \int \psi_* Q(dx_1) \int dy \, 1[x_1 \in B^\epsilon_y] \ln \frac{1[x_1 \in B^\epsilon_y]}{\int \psi_* Q(dx_1) \, 1[x_1 \in B^\epsilon_y]}$$

$$= -\frac{3}{4\pi\epsilon^3}\int dy \ln Q(\psi^{-1}(B_y^\epsilon)\int \psi_* Q(dx_1)\, 1[x_1 \in B_y^\epsilon]$$
$$= -\frac{3}{4\pi\epsilon^3}\int dy Q(\psi^{-1}(B_y^\epsilon)) \ln Q(\psi^{-1}(B_y^\epsilon)) \tag{19}$$

■

□

Proof of Theorem:

<u>Euclidean case:</u> for any set of moves M, and for any $m \in M$, ψ_m is a rotation; therefore for any $y \in W$, $\psi_m^{-1}(B_y^\epsilon) = B^\epsilon_{\psi_m^{-1}(y)}$. Then, for any prior $Q \in \mathbb{P}(W)$,

$$C(\psi_{*,m}Q) = -\frac{3}{4\pi\epsilon^3}\int dy Q(B^\epsilon_{\psi_m^{-1}(y)}) \ln Q(B^\epsilon_{\psi_m^{-1}(y)})$$
$$= -\frac{3}{4\pi\epsilon^3}\int dy Q(B_y^\epsilon) \ln Q(B_y^\epsilon) \tag{20}$$

In this case, the epistemic value is independent from the change of Euclidean frame, and not moving is a perfectly valid choice for the agent to maximize it, at each time step of the exploration algorithm (Algorithm 1).

Remark 3. The fact that staying still is a valid strategy arises as the agent assumes (or believes) that it has access to the whole configuration space of O. If it knew it had limited access to it, through for example limited sight, we expect the agent would look around until the object O would be in sight, and then stop moving.

<u>Projective case:</u>

Consider two projective transformations $\psi, \psi_1 : W \to W$, if for any $y \in W$,

$$\psi^{-1}(B_y^\epsilon) \subseteq \psi_1^{-1}(B_y^\epsilon) \tag{21}$$

then,

$$-Q(\psi^{-1}(B_y^\epsilon)) \ln Q(\psi^{-1}(B_y^\epsilon)) \tag{22}$$
$$\geq -Q(\psi_1^{-1}(B_y^\epsilon)) \ln Q(\psi_1^{-1}(B_y^\epsilon)) \tag{23}$$

This suggests that the moves that maximize epistemic value are those where ψ_m^{-1} shrinks the zone around $y^o = \rho(\phi_{\mathcal{R}}(o))$, which is the representation of O in the internal world of the agent. In particular, it means magnifying the zone around $\rho(\phi_{\mathcal{R}^m}(o))$ in the agent's new frame, $\mathcal{R}^m$, after move m. The only way to do so is to select moves that bring the agent closer to O. Let us denote $y_m^o := \rho(\phi_{\mathcal{R}^m}(o))$. Let us now make the previous argument more formal. We assume that the set of moves M is finite. Let $Q_0 = q_0 d\lambda$ be any initial prior on $W = P_3(\mathbb{R})$, at stating time $t = 0$. After one step, move m_1 is chosen and the agent updates its prior as,

$$q_1(x) \cong 1[x \in B^{\epsilon}_{y^o_{m_1}}]q_0(x) \tag{24}$$

where $\cong$ means proportional to. The prior we now consider is Q_1 denoted simply as Q. One shows that there is $\alpha > 0$, such that for all $m \in M$, and $\epsilon > 0$ small enough,

$$C(\psi_{*,m}Q) = -\frac{3}{4\pi\epsilon^3} \times \int dy\, 1[y \in B^{\alpha\epsilon}_{y^o_m}] Q(\psi_m^{-1}(B^{\epsilon}_y)) \ln Q(\psi_m^{-1}(B^{\epsilon}_y)) \tag{25}$$

$$\tag{26}$$

Let $\approx$ stand for 'approximately equal to' (equal at first order in expansion in powers of ϵ). Then from the previous statement the summand can be approximated by its value in y^o_m:

$$C(\psi_{*,m}Q) \approx -\alpha^3 Q(\psi_m^{-1}(B^{\epsilon}_{y^o_m})) \ln Q(\psi_m^{-1}(B^{\epsilon}_{y^o_m})) \tag{27}$$

Furthermore, $Q(\psi_m^{-1}(B^{\epsilon}_{y^o_m})) \approx \frac{4\pi\epsilon^3}{3} \frac{q_1(y^o)}{|\det \nabla\psi_m|(y^o)}$, where $|\det \nabla\psi_m|(y^o)$ is the absolute value of the Jacobian determinant of ψ_m at y^o. The epistemic value is maximized when $|\det \nabla\psi_m|(y^o)$ is maximized. By definition, $\psi_m = \rho \circ \phi_{\mathcal{R}^m} \circ \phi_{\mathcal{R}}^{-1} \circ \rho^{-1}$, therefore, by the chain rule of differentiation

$$\begin{aligned} &|\det \nabla\psi_m|(y^o) \\ &= |\det \nabla\rho|(\phi_{\mathcal{R}^m}(o)).|\det \phi_{\mathcal{R}^m}|(o).|\det \nabla[\phi_{\mathcal{R}}^{-1} \circ \rho^{-1}]|(y^o) \end{aligned} \tag{28}$$

Let us make explicit each terms in the previous equation. $\phi_{\mathcal{R}^m}$ is a rigid movement therefore, $|\det \phi_{\mathcal{R}^m}|(o) = 1$. $|\det \nabla[\phi_{\mathcal{R}}^{-1} \circ \rho^{-1}]|(y^o)$ does not depend on m so we can label it as a constant C. $\phi_{\mathcal{R}^m}(o)$ is the coordinate of O in the Euclidean frame $\mathcal{R}^m$; let us denote (x^m, y^m, z^m) these coordinates, i.e. $(x^m, y^m, z^m) := \phi_{\mathcal{R}^m}(o)$. Then,

$$|\det \nabla\rho|(x^m, y^m, z^m) = \frac{1}{(\gamma z^m + 1)^4} \tag{29}$$

Therefore, $|\det \nabla\psi_m|(y^o) = C\frac{1}{(\gamma z^m+1)^4}$.

As we assumed that for any move $m \in M$, the object O is always in front of the agent, then $z^m \geq 0$; in this case, z^m is also the distance of the agent to the object. Epistemic value is maximized when z^m is minimized and therefore the agent selects moves that reduce its distance to the object. Denote one of such move m^*; the argument then loops back with the new reference frame $\mathcal{R}^{m^*}$ and updated belief $q \leftarrow \psi_{m,*}q_{|y^o_m}$. ■