

C5

Information theoretic models of curiosity in hierarchical models of the world



Fabian Sinz



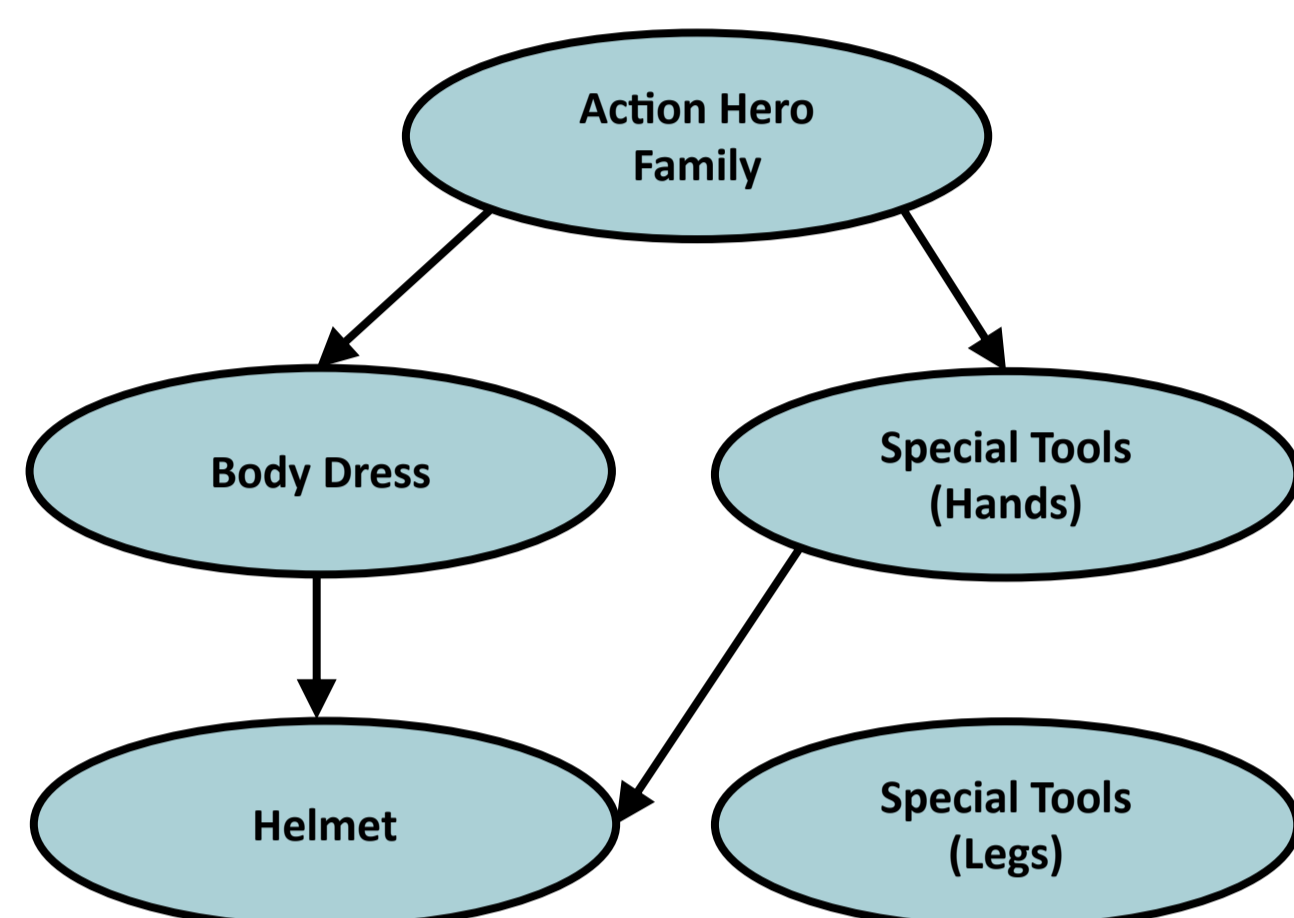
Nivedita Mani

Motivation

- Curiosity has been described in information theoretic terms which can be quantified in generative models of the world¹.
- Which information theoretic factors drive curiosity most, is still an open question.

Preliminary work:

- Investigations of information theoretic normative theories for vision^{2,3}
- Framework to fit normative neural sampling models to neuronal data⁴
- Children actively elicit information about specific aspects of their environment⁴



Objectives

- Develop an experimental paradigm to quantify and test different information theoretic drivers of curiosity
- Examine whether curiosity is driven by the desire to learn (improve model) or to reduce uncertainty (given the current model)
- By specifying the information-theoretic terms that underlie the drive of curiosity, the project addresses the question **How are we curious?**
- By examining factors that drive curiosity, it also addresses the question **Why we are curious?**



Can information-theoretic ideas of novelty and uncertainty in internal hierarchical generative models of the world capture how we are curious?

Methods

- Develop treasure hunt game where kids can uncover robots/action figures buried in sand. This allows us to control the information about each figure available to the players. The composition of the figures is determined by a hierarchical generative model.
- Because we have the ground truth model and know exactly what is visible to each player in each moment of the game, we can compute all the information theoretic quantities at each time point.
- Given a particular hypothesis about an information theoretic that drives curiosity, we can generate an arrangement of figures with a prediction which figure will be uncovered first.

Hypotheses:

- Curiosity is driven by the impulse to reduce uncertainty about the state of the world or the model of the world.
- Curiosity is initially driven by desire to learn (improve the model of the world) followed by desire to reduce uncertainty (about the state of the world, given the model)
- Predictable features of objects do not trigger curiosity.



Fig. 1: Impression of the planned computer game generated by DALLE-2.

Cross-project collaborations

- Key collaborations with projects examining the **factors that trigger curiosity**
- Synergies and cross-fertilization with projects that focus on drivers of curiosity in **complex environments A2, A3, B3, B4**
- Shared focus on role of curiosity in **sampling of perceptual information B1, C1, C4**

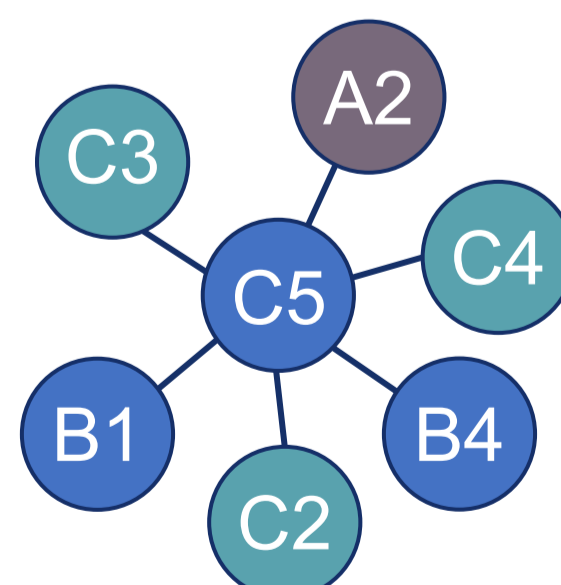


Fig. 2: Some key collaboration partners of doctoral researcher working on Project C5

Potential PhD projects

1. Characterizing curiosity in terms of information theory in hierarchical generative models of game worlds.
2. Is curiosity driven by informativeness about the state of world or about our model of the world?
3. Model driven causal interventions to direct curiosity.

References

1. Sinz, F. & Bethge, M. (2013). Temporal adaptation enhances efficient contrast gain control on natural images. *PLOS Computational Biology*, 9(1), e1002889.
2. Sinz, F. & Bethge, M. (2013). What is the limit of redundancy reduction with divisive normalization? *Neural Computation*, 25(11), 2809-2814.
3. Mani, N., Ackermann, L. (2018). Why do children learn the words they do? *Child Development Perspectives*, 12(4), 253-257.
4. Shrinivasan, S., Lurz, K.-K., Restivo, K., Denfield, G., Tolias, A. S., Walker, E. Y., Sinz F. (submitted to *NeurIPS* conference 2023). Taking the neural sampling code very seriously: A data-driven approach for assessing generative models of the visual system.