Predictive Processing Facilitates Skill Acquisition by an Autonomous Agent

Bruno Lara*[1], Alejandra Ciria[2], Guido Schilacci[3], and Oswald Berthold[3]

[1] Cognitive Robotics Group, Center for Science Research, Universidad Autónoma del Estado de Morelos, Cuernavaca, México
[2] Psychology Faculty, Universidad Nacional Autónoma de México, México
[3] Adaptive Systems Group, Department of Computer Science, Humboldt-Universitaet zu Berlin, Berlin, Germany

Abstract. We examine an approach to autonomous acquisition of motor skills and behaviour using computational models. Our focus is on the role of *predictive coding* of sensorimotor data. Based on the apparent prevalence of predictive coding in animal's brains, we follow the hypothesis that predictive coding provides systematic and unique characteristics facilitating learning and exploration.

1 Introduction

Motor control for autonmous agents that need to learn some things about the relations between observable quantities.

Our work is motivated by wanting to better understand how animals such as ourselves are able to change, adjust and invent behaviour as efficiently as they do. We assume such understanding to be relevant for building a theory of robots, guidelines for the synthesis of artificial systems interacting with the real world.

Existing research has shown the problem to be quite challenging and varied and many such variations have been examined. The task is to try and systematize the phenomena by identifying functional principles that are at work over wide ranges of organisms and mechanisms (Pfeifer and Bongard, 2007).

We examine a small set of such principles in combination: Predictive coding, internal modelling, and top-down information flow. We start with the principle of Predicitive Coding (PC) which is a coding scheme encountered in the analysis of nervous systems. The principle says that of all input only that part of the input to a functional unit, implemented by a layer or a module, which has not been predicted is allowed to pass

* bruno.lara@uaem.mx

through for further processing. This quantity, usually a vector, is the prediction error (PE).

The prediction error quantifies what the system doesn't know already and is worthy of additional predictive modelling. PE combines external (input complexity conditioned on output) and internal (model failures) factors. Knowing the relation of model change to PE change allows to infer particular sensorimotor contexts as driven by external factors. Longterm PE behaviour can modulate adaptation, while shortterm PE generally drives activity.

Next is internal modelling (IM). IM is closely related to PE because any prediction in fact requires a model. Such a model needs to map the information from the past to the scope of prediction, usually one or a few timesteps (1s).

Top-down information flow then says that experience, the apex of decision making, is being driven or initiated by a prediction about the next state. Only unpredicted (unmodelled) parts of the next post-prediction measurements are being passed on.

We can take these three principles for the design of the proposed models:

– Discuss the change of terminology or dichotomies regarding:
  - predictions/motor commands,
  - forward models/inverse models/controllers/policies
  - bottom level black box: different actual forward model (FM) representations via SOMs, kNN, RLS, GP, ...
  - see discussion with antonio re: "it's just labels", underneath it is a raw network of signal flows putting observable quantities into relation to each other, from which apparent goal-driven behaviour (teleology) emerges
– Proprioceptive space: any embodied agent has to have a lowest level motor system. This is usually represented by a vector of motor angles, motor velocities, motor torques, etc but essentially it can be any given kind of controller. A controller in this sense implies a measurement and a prediction defined in the same space. This is the point at which we want to engange in interaction with the system.
– The forward model is part of a larger building block, a generative model, mapping from control (goals) to to sensory consequences (Friston, 2011).
– The main flow of information is top-down, not bottom-up, and the forward flow of sensory information is replaced by the forward flow of

prediction error (Clark, 2015). FIXME: define top, botton, forward, etc?

- Sensory prediction errors ($Err_{e(t)}$, $Err_{p(t)}$) are required for online state estimation (inference) and for optimizing (learning) the forward model (Friston, 2011), or any kind of activity.
- Descending propioceptive predictions are equivalent to motor commands. Movement quash error at the level of spinal reflexes (Pickering and Clark, 2014). Propioceptive predictions elicit motor actions, so motor commands are replaced by those propioceptive predictions (Clark, 2015). $S^*_{p(t)} = M_{(t)}$.
- Motor commands become descending predictions or propioceptive sensations, while their exteroceptive homologous become corollary discharges (Friston, 2011).
- Motor behaviours are predictions of propioceptive patterns. Motor control is just more top-down sensory prediction (Clark, 2015).
- Action: results from our own predictions concerning the flow of sensations (Pickering and Clark, 2014).
- High-level plans and intentions (prior beliefs) associate with sensory consequences and actions flow. This is, high level states into predicted sensory patterns (Pickering and Clark, 2014).

2 Related

Long history, many models: control, development, RL, neuroscience, prediction and prediction error (residual), ...

3 Proposed Models

3.1 Basis building blocks

First two basic models and their difference in terms of learning dynamics. Do any dynamic properties support or contradict biological motor learning?

3.2 M1

The basic building block (BBB) contains an adaptive forward model (predictor) whose output is the BBB's output. We assume direct feedback from the module that the output is connected to. The BBB' inputs are an explicit goal and the prediction error with respect to the last prediction at $t-1$. These inputs taken together encode the state. Thus the forward model should in principle be able to pick up on different relations depending on the state. Graphical representation is given in Figure 1. The closed-loop dynamics of this model are described in detail in subsection 5.1.

FIXME: Fundamental limitation: can only learn monotonic relationships so far. Show to fail, indicate amendment that can solve non-monotonic cases (by observing PE over time, by using reward modulation, ...).

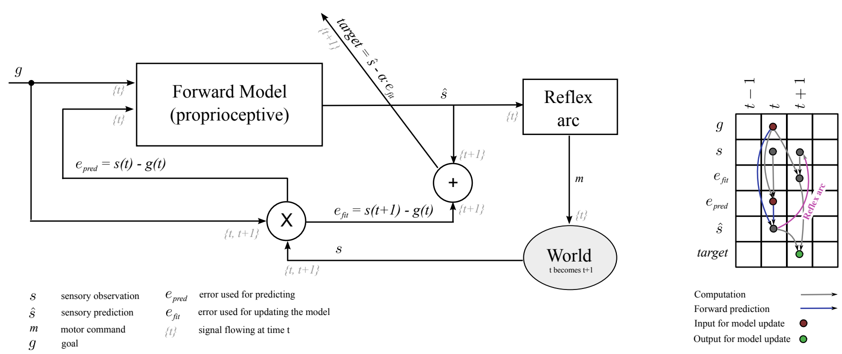


Fig. 1. Basic model M1.

3.3 M2

This modification of BBB M1 receives only the error as an input. The model's graph is shown in Figure 2. This model will not be able to distinguish between situations demanding different predictions. FIXME: It can be shown to fail when such a situation is produced by experiment design.

3.4 E2P

A third building block is the intermodal map, specifically mapping between an exteroceptive and the proprioceptive modality (thus E2P). We propose to realize this slightly differently from M1 / M2. The E2P module is bootstrapped by passively observing corresponding exteroceptive and proprioceptive data and learning those correspondences. Once bootstrapped, the E2P can be used to predict possibly ambigous proprioceptive configurations from the exeroceptive sensations they are believed to produce, meaning which have been observed in the past to do so.

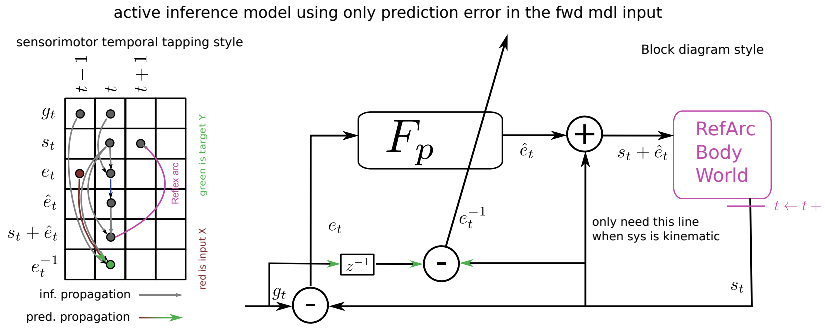


Fig. 2. Basic model M2.

3.5 Extended models connecting modalities

Merge with alejandra's two versions (explicit update rule).

This model includes both propio and extero (Figs 3 and 4). The only difference is how the predictive error is calculated for the propioceptive information ($Err_{Pi(t)}$).

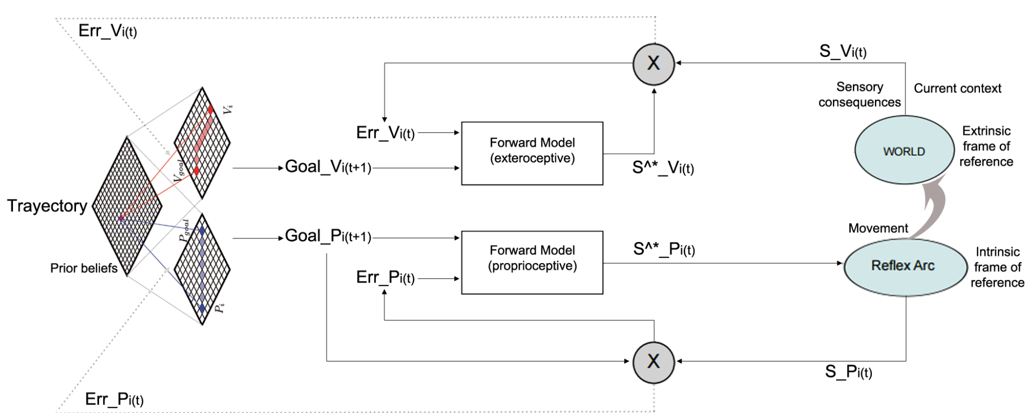


Fig. 3. *Model A*

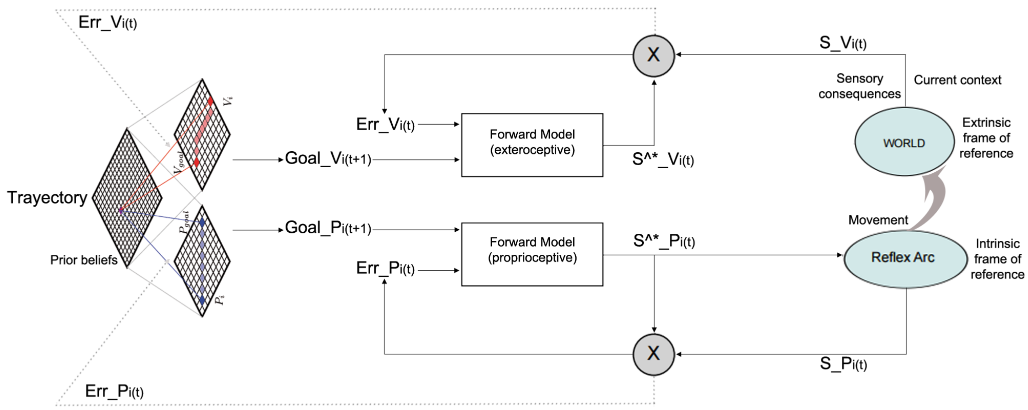


Fig. 4. *Model B*

3.6 Case Study Models

In the example to simulate, we use the visual modality as exteroceptive information.

The flow of the experiment, once trained would be as follows:

1. The visual starting position of the hand is located in space and becomes V_i
2. A visual target is located and becomes V_{goal}.
3. This two visual positions in space are mapped to the MMR map which codes for trajectories in the 2D space of the robot. For each pair of positions in visual space exists a pair of nodes in motor space. This can be seen as previous knowledge and does the required mapping between the two spaces. A version of this is already implemented. (see Section 3.7)
4. With these positions in both spaces, the flow of the experiment goes down to the forward models.
5. Movement in the intrinsic frame of reference, as already implemented in the propioceptive model, has consequences on the environment which cause changes in the extrinsic frame of reference.

We can talk of two different versions of this, as we can use directly just the initial and goal situations in both maps or in a more precise way use intermediate points to perform micro-predictions to correct the trajectory.

The two models shown have the same difference as the previous one, this is, the use of the prediction for obtaining the predictive error.

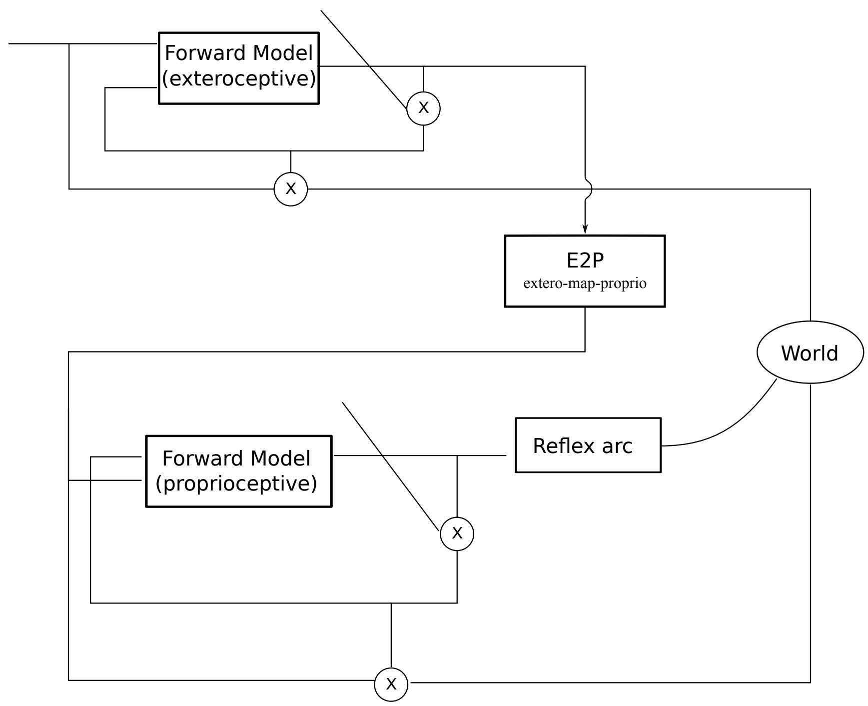


Fig. 5. Hello grpahicxs

3.7 Propio-Extero Mapping Model

There would be here, the part that Guido was implementing on the forming of trajectories.

4 Notes

4.1 control theory parallels

i don't see this as a problem at all, to the contrary, it means if you arrive at a model setup you can actually look to adaptive control and translate the solution to your formulation.

in terms of the philosophy see email from *<2016-10-11 Di>* so it's clearly about a different issue.

4.2 merge bruno's initial draft

4.3 opt's notes

- for proprio-only show 3 figures: 1) behaviour without any learning, which fails to precisely reach the goals because proprio predictions don't "match" proprio states (nonlinear deformations), 2) behaviour during learning, see how we actually move closer to the goal over a few timesteps, 3) behaviour while still learning but going to places we have already been: no error, no learning / weight changes
- additional plots: original timeseries, model sweep response plot

5 Experiments

Notes for experiments:

- deploy models to different robotic systems: Nao, SimpleArm, Pointmass, two-wheeled
- timeseries analysis
- learning transient (arbitrarily fast, see Tin + Poon)
- explicitly map out the acquired models by sweeping the inputs and record their outputs
- what happens under perturbation?
- long-term adaptation effects: are we only plastic or are we also stable? do new adaptations overwrite previous ones? does a given approximator become saturated (just think how critical weight initialization in neural networks can be) in terms of its adaption capacity?, can be enforced via decay schedules of learning rate etc

5.1 Basic behavioural example M1

- Basic operation: Figure 6, Figure 7, Figure 8, Figure 9 show variations on the basic operation theme. We are evaluating a k-NN and a Gaussian Process based forward model each over 200 and 1000 timesteps interaction on a one-dimensional identity system. The forward model starts off in a zero-state, the first prediction being zero. Any resulting prediction error drives an adaptation on the forward model output towards reduced error for a given goal, error input pair. We sweep this input space once at numsteps/2 and once at the end of the experiment which demonstrates the change in the overall model response.
- Compare identity system with small aberration system: now we can do the same for a 1D system while introducing aberrations in the system's response function.
- Failure under non-monotonicity: if the aberrations become systemically non-monotonic the learning process fails (FIXME: at least I predict that ;)
- How PE is low when revisiting previously observed configurations: longterm PE plot of 10000 timesteps and averaged PE in addition to instantaneous PE

5.2 Basic behavioural example M2

5.3 Basic behavioural example E2P

6 Discussion

Discuss, compare with existing approaches like std fwd/inv model pairs or reinforcement learning

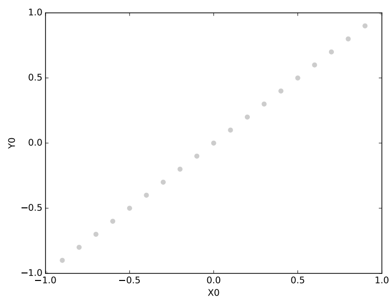
7 Summary

References

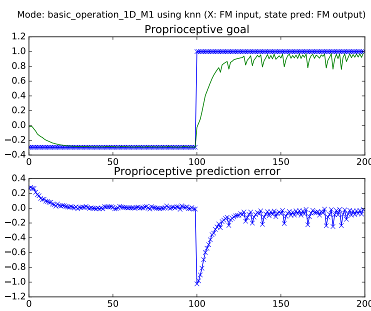
Clark, Andy (2015). “Embodied prediction”. In: *Open MIND*. Open MIND. Frankfurt am Main: MIND Group.

Friston, Karl (2011). “What is optimal about motor control?” In: *Neuron* 72.3, pp. 488–498.

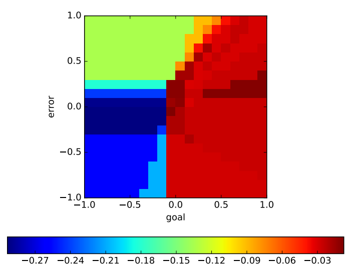
Pfeifer, Rolf and Josh Bongard (2007). *How the Body Shapes the Way We Think*. MIT Press.



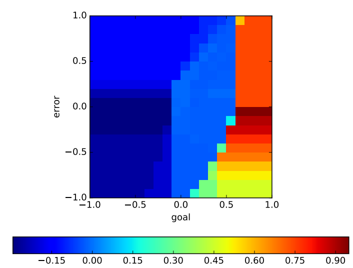
(a) System sweep response



(b) Goal, prediction, error

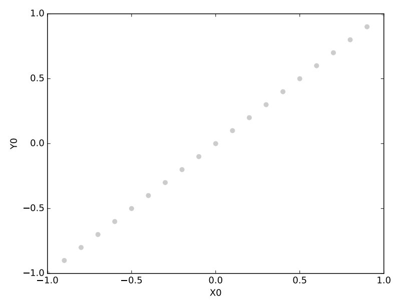


(c) Model sweep t = 100

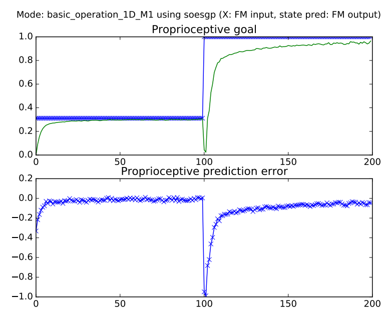


(d) Model sweep t = 200

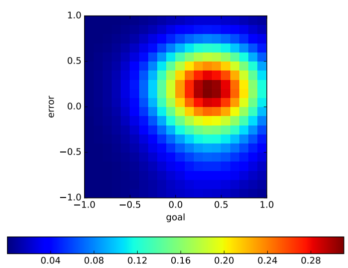
Fig. 6. Basic operation 200 timesteps knn.



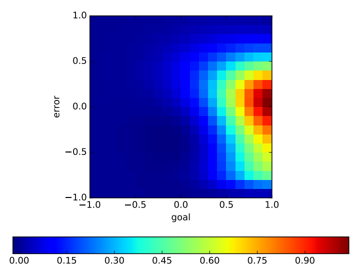
(a) System sweep response



(b) Goal, prediction, error

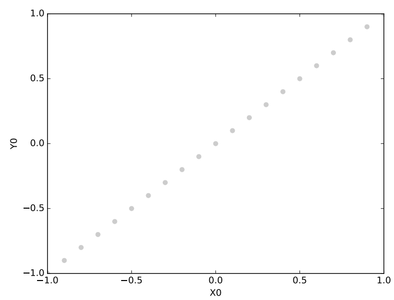


(c) Model sweep t = 100

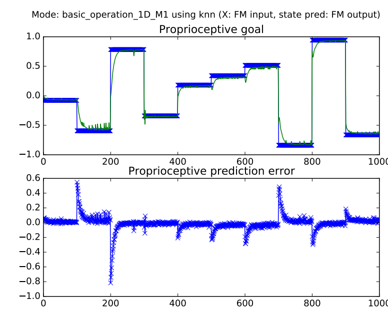


(d) Model sweep t = 200

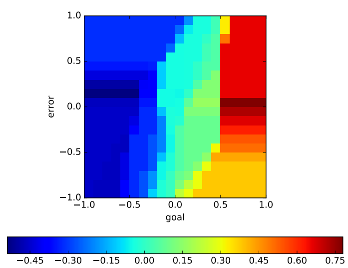
Fig. 7. Basic operation 200 timesteps GP.



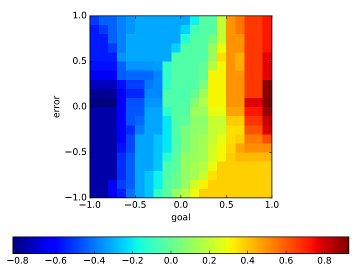
(a) System sweep response



(b) Goal, prediction, error

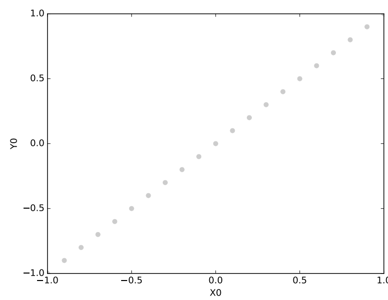


(c) Model sweep t = 500

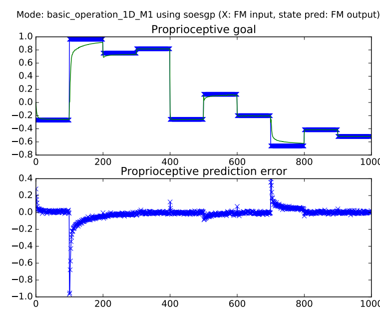


(d) Model sweep t = 1000

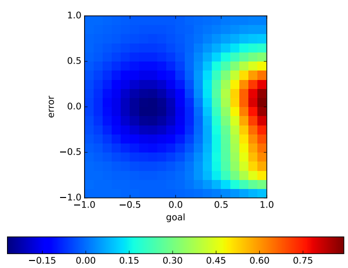
Fig. 8. Basic operation 1000 timesteps knn.



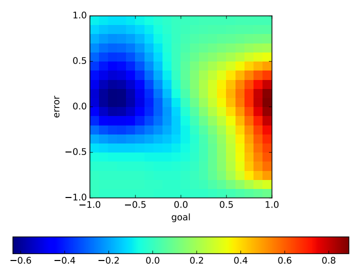
(a) System sweep response



(b) Goal, prediction, error



(c) Model sweep t = 500



(d) Model sweep t = 1000

Fig. 9. Basic operation 1000 timesteps GP.

Pickering, Martin J and Andy Clark (2014). “Getting ahead: forward models and their place in cognitive architecture”. In: *Trends in cognitive sciences* 18.9, pp. 451–456.