Bursting with Error:
Dealing with Unexpected Prediction Error in Babybots

Bachelor Thesis in Artificial Intelligence by
Erwin de Wolff
s4244907

Supervised by
Johan Kwisthout[1]

[1] Donders Institute for Brain, Cognition and Behaviour

Abstract

Our research builds on and adds to the prediction processing theory, specifically to model updating and model revision. We propose a hyperparameter approach to model updating at the computational level, designed to add precision to beliefs, which we tested via computer simulations. We also introduce the concept of unexpected high prediction error, which can be used as a signal to revise the generative model of agents. The latter links two of the proposed methods the brain uses to reduce future prediction error.

1   Introduction

In recent years, predictive processing has emerged as one of the main theories about the workings of the brain. The theory describes the brain as a generative model that acts as a minimisation machine. The brain makes predictions about the world, then resolves any prediction error resulting from that prediction, a difference between the prediction and the observation, such that less prediction error will be generated in the future. This theory is part of the free energy principle (Friston, 2010).

A important question is how exactly the brain resolves prediction error. A variety of possibilities have been proposed, ranging from active inference to changing the level of detail. Our work builds on the formalisation at Marr's computational level of explanation by Kwisthout, Bekkering, and van Rooij (2016). Much of the work involved in predictive processing uses Bayesian networks as their generative model. The model that we used is a causal Bayesian network, which adds the assumption that arrows in the network denote causal relations.

It is currently an open question when, as well as if, these proposed methods of handling prediction error should be used. By adding to one of these methods, *model updating* and testing its workings in one trivial and one non-trivial simulation, we aim to provide more insight in how these solutions to prediction error might be linked. In particular, a link between model updating and *model revision* is proposed, based on the occurrence of unexpected and high prediction error. By verifying whether this particular prediction error was a single occurrence or was part of a new, persistent pattern, we gain valuable information as to whether the world has altered, and therewith if model revision is a better solution than model updating.

1.1   Model updating

Model updating adjusts the beliefs about the world without changing the model or factors that play into it. Rather, conditional probabilities change as contrary evidence is observed. For example, an agent might have the belief that the change of

it being sunny at any given day is quite high. Given many rainy days in a row, we can expect this belief to have shifted to a more balanced prediction, or even biased towards downfall. In terms of our model, model updating is equivalent to updating the probabilities in the probability tables of the variables in our Bayesian network. We will use the term beliefs when talking about these probabilities.

When reasoning about beliefs in this generative model, an important conceptual distinction must be made between reducible, irreducible and unexpected uncertainty (Angela and Dayan, 2005). Reducible uncertainty is the part of our prediction that is uncertain due to lack of experience or knowledge, and is linked to the precision of a prediction. One might not know that a die tends to have a uniform distribution, for example. In this case, gaining new experiences is a valuable goal. Irreducible uncertainty originates from the world itself and cannot be overcome. It is inversely proportional to the precision of the prediction error. No matter how many times a fair die is rolled (ignoring the erosion and similar changes such repeated rolling would have on it), the probabilities of it landing on any given side do not change. In other words, the complexity of their outcome is irreducible. In this case, it would be a fruitless endeavour to acquire more knowledge, as it will not result in a better understanding of the world. Unexpected uncertainty arises when the observation is very different from a prediction with high precision. This signals an exceptional observation.

Single probabilities, as are the norm in standard Bayesian networks approaches, do not provide information on whether a belief is based on no or little experience or on a large sets of experience. Therefore, it cannot be analysed how much of the complexity of the prediction arises from reducible complexity. Without that knowledge, it is difficult to quantify how much an agent should update its beliefs should contrary (or even consistent) evidence be observed. Of course an agent can always apply the same change whatever the certainty, but this is not an efficient approach, as we will make our case now.

If beliefs are modified quite drastically each time, an agent adepts quickly to the world, but is unlikely to develop stable view of it. The irreducible complexity of the outcome of a coin toss would lead it to constantly switch its belief about the bias of the coin, whereas humans would quite quickly decide that the coin is more or less balanced, with the inherent probability of the system taken into the equation.
If on the contrary only a minor update is performed each time, beliefs are more likely to approach the actual probability in the world, but they will take much longer to learn. This again does not mirror human behaviour, where strong conclusions (false though they may be) are derived from what is often a small set of data. A perfect value between the two approaches might exists, but finding this value becomes a model-specific parameter to solve.

1.2 Model Revision

Model revision acts by changing the structure of the model to reduce prediction error. By assuming that previously irrelevant factors are now relevant, or the opposite, prediction error can be explained. Building on the previous example, one could take into account whether the sky is clouded or not to improve the predictions. This would effectively double the amount of predictions that can be made, over which model updating can be applied. In terms of our model, model revision is equivalent to adding or removing variables and causal connections.[1]

To those familiar to probabilistic networks it is apparent that by introducing a new factor upon which an existing one becomes dependent, the size of the probability table multiplies. Specifically, by making a variable dependent on an additional factor, one multiplies the number of possible input combinations that variable can have by the number of values that additional factor has. If the prediction of sunny or downfall was previously only based on a categorical variable of temperature, and now also takes into account the cloudiness, there are twice as many cases in our model (see figure 1).

Model revision is not feasibly applied whenever prediction error emerges however. Assume that each time prediction error emerges the agent expands the generative model to resolve that error. That would mean that each memory would be linked to its own specific sets of variable values. This generative model will grow to an enormous number over an average human lifespan and would not at all allow for generalisation. Secondly, systems that contain any degree of irreducible uncertainty cannot be explained away by model revision. Adding more variables does not solve the irreducible, probabilistic property of the system, as even with the exact same variable values the outcome cannot be predicted.[2]

Of course, model revision is not always needed. Instead, one can imagine that in general, a model is updated via model updating and that the structure of the model is only changed in (increasingly) rare situations of importance. This is a 'generalise first, add later' sort of approach. If (human) cognition worked in such a manner, then looking at the prediction error might provide us insights into how this system works.

---

[1]Some would include the adding or removing of values within a variable within model revision, but we use the term 'changing level of detail' for that approach (Kwisthout, Bekkering, and van Rooij, 2016)

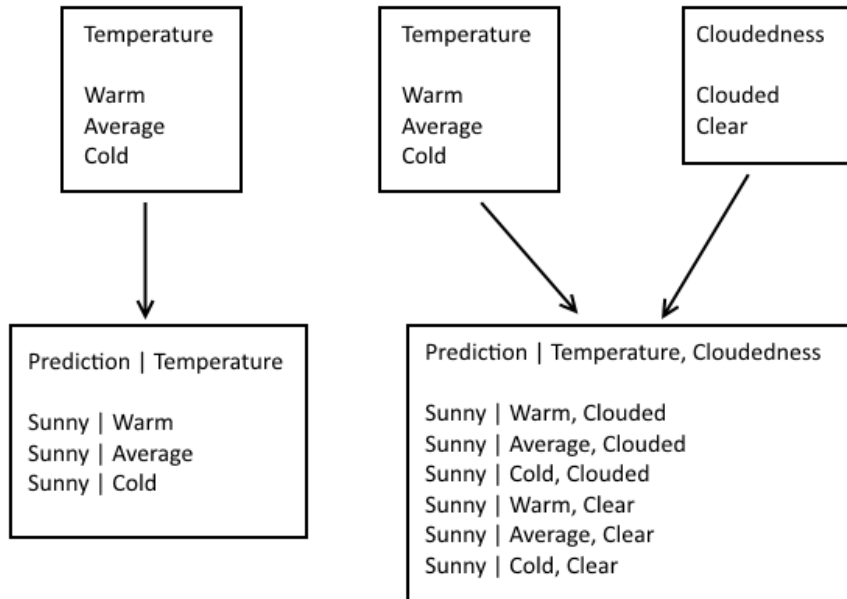[2]This is a claim at the cognitive level, not the physical.

Figure 1: *Example of increasing number of variables. Entries for Downfall have been omitted for size considerations. The left model is the agent's before including cloudiness in the evaluation of weather prediction. The right model is after model revision has occurred.*

Our work intended to find cognitively plausible answers to the issues raised above. In particular, our research tried to answers two questions:

- *How can the scale of updates of beliefs be determined in model updating?*

- *What link, if any, exists between model updating and model revision?*

The answer to the second question in particular was based on cognitive biases found in humans.

2   Hyperparameters

Hyperparameters offer a good solution to the problem of what the scale of updates of beliefs should be. By representing agent beliefs not as a singular value but rather as a function, we can describe in detail how certain that agent is about them, or in other words how much precision predictions have (Friston, 2008). We used a beta distribution in our work as the function of the hyperparameter, denoted as $\text{Beta}(\alpha, \beta)$ for any real numbers α and β.[3] This function acts as our probability

---

[3]Beta distributions work as a conjugate prior for Bernoulli or binomial distributions. For a general function for any multivariate distribution, a Dirichlet distribution is the general alternative. The workings of this function were not investigated but should be similar to the beta.
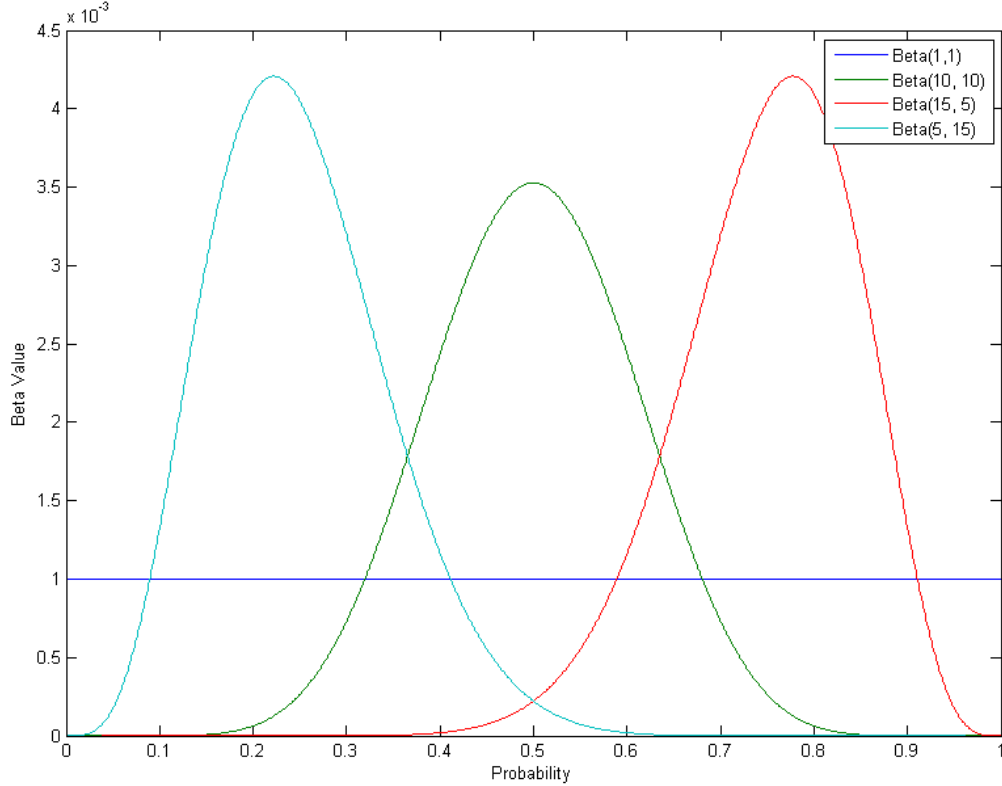
Figure 2: *Examples of beta-distributions. Beta(1,1) is the baseline in our research.*

density function for binomial belief variables (examples of which are the outcome of a coin, odd or even in a die, betting on the outcome of a drawless two player game etc.) (MacKay, 2003).

If the curve is slanted to the left or right, the belief tends towards failure or success, respectively. The singular value of the belief of the outcome, as used in Bayesian inference, is defined as the average of the beta distribution values:

$$P(x = 1) = \frac{\alpha}{\alpha + \beta} \tag{1}$$

where x denotes a Bernoulli or binomial variable and α and β are the the parameters of the beta distribution. Examples of how α and β influences the shape of this distribution are shown in figure 2.

As mentioned the beta distribution also contains information about the precision of the belief. As we have a continuous function between 0 and 1, there is almost always a non-zero value for each probability. In other words, an agent believes a spectrum of possible probabilities, with some being more likely than others. A broad, low peak indicates lack of certainty about the actual belief, or an uncertain precision. A narrow, high peak represents a more certain belief, or more certain

precision. In general, we formally define the uncertainty of the prediction as the entropy (Shannon, 2001) of the hyperparameter. Our prior probability is Beta(1,1), which corresponds to a uniform distribution over all probabilities, which can be seen as complete ignorance in which all possible probabilities are judged to be equally likely, as well as possible.[4] The entropy is maximal for this prior, thus taking to mean the most uncertain state.

2.1   Learning Hyperparameters

Since we are operating with a Causal Bayesian network as our generative model, a learning approach in line with the probabilistic foundations thereof is desirable. Specifically, agents using these hyperparameters would still like to make use of Bayes' rule to update their beliefs.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \tag{2}$$

Where A and B are any factors of the Bayesian network, in the agent's case beliefs. Bayesian updating is defined for beta distributions as well. If our dataset consists of n datapoints, which are labelled either successes ($s \leq n$) or failures ($f = n - s$), as beta distributions are binomial representations, the formula for the posterior probability for a parameter a given its prior and the evidence is:

$$Posterior(a = p|s, f) = \frac{Prior(a; \alpha_{prior}, \beta_{prior})\mathcal{L}(s, f|a)}{\int_0^1 Prior(a; \alpha_{prior}, \beta_{prior})\mathcal{L}(s, f|a)da} \tag{3}$$

Where $\mathcal{L}$ is the likelihood function given the evidence, and the integral is a normalisation step to ensure the probabilities sum to 1. This can be rewritten as:

$$Posterior(a = p|s, f) = \frac{a^{s+\alpha_{prior}-1}(1 - a)^{f+\beta_{prior}-1}}{\int_0^1 (a^{s+\alpha_{prior}-1}(1 - a)^{f+\beta_{prior}-1})dx} \tag{4}$$

In other words, one can find the posterior probabilities by adding the successes to the α parameter and the failures to the β parameter, then calculating the values and normalising to sum to 1.

The definition as described spans a single iteration. Online learning requires a repeatable approach, since agents need to function to their best understanding at every timestep as new data is observed in the world. The simplest adaptation from this one-step approach to an online approach rests on the observation that successes

---

[4]The use of Beta(1,1) is not an undisputed decision in general, but since the focus of the research was on the development of hyperparameters and entropy thereof, this prior was the most appropriate pick (as well as in line with Thomas Bayes' suggestion (Bayes, 1991)).

or ones are added to the α parameter, and the failures or zeroes to the β. Since any values for α and β result in a probability density function, it is perfectly possible to increment this step by step, as opposed to all at once. This guarantees that given the same data, both approaches end up with the same belief. Thus, the final beta distribution shape is independent of the order of the evidence. While this is apparent, it is not so easily identified what happens to the intermediary beliefs and prediction errors between beginning (Beta(1,1)) and end.

2.2 Development in Simple Experiment

To get a better insight into the development of hyperparameters and their link to prediction error a simple coin-flip simulation was ran. The agent's model is a trivially simple network consisting of one variable: coin outcome. In each timestep, the agent makes a prediction over the outcome and makes an observation. We then determine firstly the prediction error, secondly the next increment of the hyperparameter function, and thirdly the weighted prediction error, which we define as the Kullback-Leibler divergence between the prior and posterior model in a single iteration. The Kullback-Leibler divergence is defined as:

$$D_{KL}(P||Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \tag{5}$$

where P is the observed distribution and Q the predicted distribution.
These distributions for the prior and posterior model are extracted from the hyperparameter via the average.

Three datasets were used in this experiment, each consisting of 100 coinflip outcomes: 50 of these flips were tails and 50 were heads. The difference between the sets was the permutation of those outcomes. The first sets is the *homogeneous* or *paired* set, where each tails outcome is followed by a heads outcome. This is the perfect pattern of an even distribution, and should elicit the lowest overall error. The second dataset is the *sorted* set, where the first 50 outcomes are tails, and the rest is heads. The final dataset is a *random* dataset, where the outcomes are shuffled by a pseudo-random algorithm.

The result can be seen in figure 3. Here predictions were based on the outcome of a single coinflip. As predicted, the final curve shape is identical for all datasets. However, the intermediate distributions vary greatly. Being the most balanced, the homogeneous set quickly approaches the true probability and deviates very little from it. We see this reflected in the prediction error. Initially the model tends towards failure, leading to a higher prediction error when heads is observed. As the probability converges on the 50% threshold, the prediction errors become more and

more stable at the value 1. This is also reflected in the weighted prediction error, the changes between prior and posterior belief only decrease in size. The sum of the error is lowest for this dataset.

The sorted set slants heavily towards failure (tails) initially, in line with the observations. Then, the curve slowly returns to the same ending value. We can see the shift in observations clearly from the spike in both normal and weighted prediction error. This set has the highest sum of errors.

The random dataset, arguably the most likely to occur in natural systems, varies more in terms of prediction error, as groups of subsequent equal observations lead to temporary low prediction error. These groups are not of sufficient impact on the belief that a spike appears in the weighted prediction error however, which develops far more similar to the homogeneous set than the sorted. The random dataset's error lays somewhere between that of sorted and homogeneous.

This experiment was ran with predictions made over a single coin outcome. There is no reason to limit a prediction in this manner, as one can easily imagine situations where more than one coin must be taken into account. Additional experiments were run where predictions were made over multiple coin flips. The result thereof are discussed later.

3   Unexpected High Prediction Error

The prediction error spike in the sorted distribution (appearing both in normal and weighted error) is easily linked to the switching from failure (tails) only to successes (heads) only. Indeed, if a series of coin flips like in the example would occur in a real life setup, it is unlikely that humans would accept this to be the result of chance, despite the series not violating any axioms of probability.[5] Rather, observers might describe alternative hypotheses: one rigged coin was swapped for another, a magnetic field influences the coin to a certain outcome (and was switched to a different mode) or they were subject to a magician's trick. These alternative hypotheses all share that additional information of the world is taken into account. To fully explain their observations, the model needed more variables than just the coin itself. This idea exists predictive processing as *model revision*.

In order to generalise this observation however, the question of when such a switch from parameter learning to model revision should take place needs answering. Our proposal is that *unexpected high prediction error* provides a theoretically plausible

---

[5]It would seem equally unlikely that the homogeneous series would be seen as a result of random processes, not design. However, since this series elicits no unexpected high prediction error, the further results and theory do not apply for it.
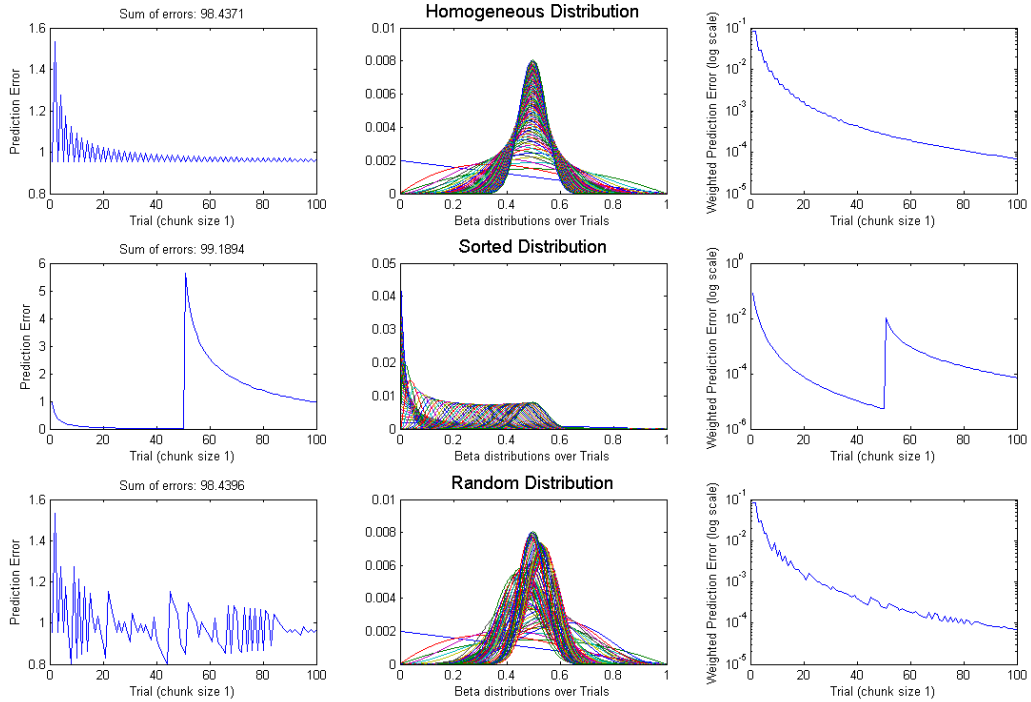
Figure 3: *Overview of the coinflip experiment. Plotted from left to right are: the prediction error per prediction, the evolution of the hyperparameter and the weighted prediction error, for each of the three datasets. Points of interest are the sum of all the errors, denoted above the left column, the high spike midway in the sorted dataset and the logarithmic scale of the weighted prediction error plots.*

answer. Unexpected high prediction error (henceforth UHPE) is defined as prediction error greater than a certain constant k. This k can be interpreted as an upper bound on expected surprise. In the current definition, it is a parameter specific to the model. If an UHPE occurs, it means that an observation was made of which the probability, given the belief, lies in the tail end of the spectrum. There are multiple ways in which this can happen. It could be that an observation was the opposite of prediction based on a very slanted belief (winning a lottery), but could also be a very unlikely series of observations, regardless of how slanted the belief is (f.e. throwing an odd number on twenty subsequent die rolls).

It should be noted that this definition is only of several, and only looks at the prediction error. There is a good argument to be made to base this k on the weighted prediction error, since this is influenced not just by the observations, but also by the precision of the belief. The jagged nature of the prediction error with the random dataset (figure 3) is not nearly as strong in the weighted prediction error, suggestion

too that the latter is perhaps a more reliable metric. Changing the approach thusly does lead to the problem that the value of k must be updated constantly as changes in the model necessarily become smaller over time, due to the growth of α and β. We chose the former approach, using standard prediction error, over the latter as it is simpler and the results are meant to be taken more as a proof-of-concept than final formalisation.

3.1 Dealing with UHPE

Having determined when we speak of UHPE, the question is how an agent deals with this. 'Nothing' would be a perfectly plausible answer, but does not make use of the insights of what causes this phenomenon. Rather, the agent would like information on whether this UHPE originated from an unlikely, single occurrence or if the world has changed. To this end, the agent will want to postpone updating its beliefs, repeat a similar sets of action as those that elicited the UHPE, then evaluate the outcome. If the agent would not postpone updating until after this check, this would create a different baseline for new observations to trigger an UHPE, which would not translate to a similar assessment.

If the observations from this repeated pattern are as the agent expected initially (i.e. the repeated pattern does not elicit an UHPE), the agent concludes that the UHPE was caused by an exceptional observation, and may or may not use this to observation to perform model updating on. Choosing to update based on the exception ensures that beliefs are based on all observations, while choosing not to use it for updating is a direct attempt at reducing immediate prediction error.[6]

If the repeated observations still (or again) elicit an UHPE, the agent concludes that something that is not yet detailed in its generative model has changed. The better method to deal with this prediction error is model revision, as it would take (increasingly) many trials to adjust the beliefs to the new values using model updating, and there is no guarantee that once stable, the world won't change back to the previous state. To reasonably avoid future prediction error, the agent should include the previously unused variable in its model. How the agent determines what this variable is and how it relates to other, known variables in the model is an unanswered question. It is also not clear what these new beliefs should look like: similar to the old belief or back to an uninformed state? We won't go into the answer to these questions, restricting the topic to the question of when this switch in approach takes place.

---

[6]If the agent updates towards the tail end of the spectrum from the mean, the expected prediction error in following visits might be higher than if it does not update. In that case, by ignoring current error future error is avoided. Whether these sort of considerations should be included, or are indeed true, is another question altogether.

Of course this approach is not without its assumptions. For starters, there is no reason why two consecutive UHPE's could not result from the probabilistic nature of a model. As unlikely as a single roll yahtzee is, this does not change anything about the probabilities of the next outcome. However, the likelihood of consecutive UHPE's is very small. One can define an UHPE as an observation of which the probability of occurrence is a probability function on k. To get two of these, one would need to square this probability. Since the chance of an UHPE is, by definition low, the square of this is even lower. Given $P(UHPE) = F(k) = 0.05$, the probability of two UHPE's would be 0.0025, or one in four hundred. In simple models this is well within expectation to happen during an average lifetime of a human, but in greater models this is less and less likely, since we also need to take into account the chance that all other variables have the same value.

Secondly and conversely, finding that the repeat pattern does not elicit an UHPE is not proof that the world has not changed. It could very well happen that after the world changed, an unexpected series of events occurs. In our model, the agent would then not resort to model revision. It is however, quite apparent that if the agent were to continue to operate in the same model, an new UHPE would occur rather quickly, with model revision following suit.

The rest of this research builds on the premise that UHPE's can signal model revision as explained above.

3.2   UHPE and Human Behaviour

One of the main inspirations for the UHPE and how one should deal with it came from the mobile paradigm (Rovee-Collier and Kupersmidt, 1978). In this experiment, a baby laid in a crib. As the baby kicked its limbs, the frequency thereof was recorded. After establishing the baseline a ribbon was attached to a mobile and one of the limbs of that baby. This resulted, presumably because the baby wanted the mobile to move, in an increase in kicking frequency in that limb, and only that limb. When the ribbon was cut after this increase, the kicking first spiked in what is called an extinction burst. Only then did the kicking frequency decrease.

A possible explanation for this extinction burst is that the baby is testing its supposed knowledge in rapid succession in order to falsify it. In that regard, we can imagine other situations where this kind of behaviour would emerge. Suppose that we close a door behind us while talking to somebody. If, unbeknownst to us the door locked itself shut, we can imagine that upon trying to open it again the agent would experience high surprise. This could be followed by a rapid series of jerks in an attempt to open the door. Only then will the agent conclude that something

must have changed in the world, most likely that the door must have locked itself.

Other examples include rereading a message containing unexpected news, refreshing a webpage multiple times when the internet connection is lost, trying the switch to a broken lamp multiple times, taking multiple small bites from a spoilt sandwich, sliding your foot back and forth across an unseen layer of ice and so on.

Such scenarios, in predictive processing terms, follow the structure of model updating, UHPE, repeat pattern and model revision, where the repeated behaviour is the falsifying step. When the agent modifies its model to account for a change in the world (door fell in its lock, internet is down...) the prediction error can be explained. From there on out the agent will have to learn how this new factor influences its beliefs.

4   Babybots

Having gathered an understanding of how hyperparameters evolve over time and how UHPE's can be used as an indicator for model revision, it is informative to test these findings on a non-trivial model. Our research used the babybots as used in earlier research by Otworowska, Zaadnoordijk, de Wolff, Kwisthout, and van Rooij (2016). The babybot is a simulation of the mobile paradigm discussed earlier, and simulates the movement of a mobile that is dependent on the movement of the baby's limbs. The model used by the babybot is a causal Bayesian network consisting of hypothesis, intermediate and prediction nodes (Kwisthout, Bekkering, and van Rooij, 2016). The hypothesis nodes in this network are a motor command for each limb of the baby (MS_LeftLeg, MS_RightLeg, MS_LeftArm, MS_RightArm). The intermediary nodes are the previous and current limb position for each of the limbs (Prev, Cur for all combinations). We have only a single prediction node, Mobile Movement, which is dependent on all previous and current limb positions. A simplified network with the values of each variable is shown in figure 4.

The previous limb position is the current limb position in the previous timesteps. Initially these positions are all at the lowest position. A movement can, in its lowest position, be given the motor signal 'down'. In that case the limb will stay in its position. Similarly for a high position and command 'up'. All the probabilities are known and deterministic except for the influence of the limb positions on mobile movement. These deterministic probabilities are also not subject to any change during the experiment.

In accordance with the set-up in the previous work, the babybot uses the concepts of exploration and exploitation. Exploration is meant to reduce uncertainty, or increase
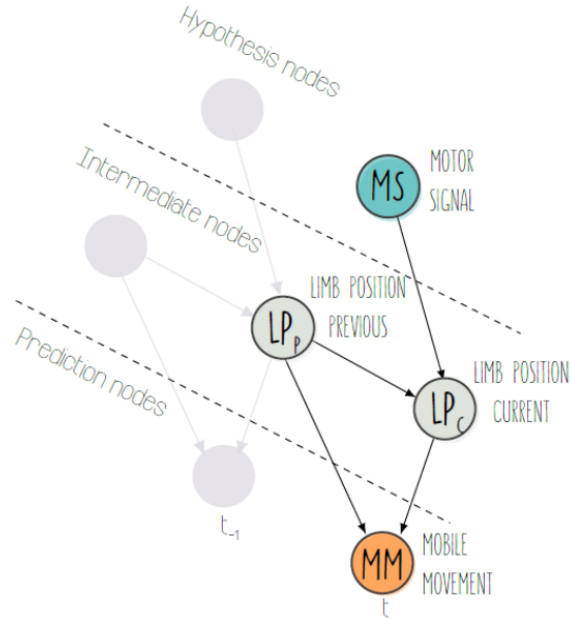
Figure 4: *The babybot model (Otworowska, Zaadnoordijk, de Wolff, Kwisthout, and van Rooij, 2016). This causal Bayesian network depicts only a single limb: the complete network had one variable for each of the four limbs, with the exception of mobile movement, which was influenced by all eight limb variables.*

precision for future predictions (Friston, Rigoli, Ognibene, Mathys, Fitzgerald, and Pezzulo, 2015). Therefore, uncertain beliefs are prioritized when exploring. As the entropy of the hyperparameter is the measure of uncertainty, exploration picks the motor commands via:

$$argmax_{\text{HypothesisNodes}}(H(P(\text{PredictionNode}|\text{HypothesisNodes}))) \qquad (6)$$

Thus exploration picks the most uncertain combination of limbs commands. One could also define a likelihood to all combinations dependent on the entropy of that state, with more uncertain states becoming more likely, but this not was not done.

Exploitation aims to make use of gathered knowledge to avoid uncertainty in the present (Friston, Rigoli, Ognibene, Mathys, Fitzgerald, and Pezzulo, 2015). This means that the most likely probability of the prediction node given the hypothesis is chosen, via:

$$argmax_{\text{HypothesisNodes}}(P(\text{PredictionNode} = true|\text{HypothesisNodes})) \qquad (7)$$

Again, one could apply a similar likelihood approach with exploitation, but this was not done.

13

There is an intuitive tradeoff between exploration and exploitation. Naturally, exploiting on the bases of no knowledge is not an efficient approach. So when there is much uncertainty (as would be case in an unfamiliar environment), exploration should be preferred over exploitation. Conversely, when there a great deal of precision, it makes little sense for the agent to subject itself to higher prediction error if it can be avoided with high precision. Thus, in familiar situations an agent should primarily exploit. Our babybot explores for a fixed number of iterations first, only to switch to exploitation indefinitely thereafter.

An important property that the prediction error should have is that it is relatively stable, the reason being that it should not elicit a UHPE when nothing has changed in the world.[7] Since k is a free parameter in our current formalisation, we can analyse for different values of k when the prediction error would lead to an UHPE. Ideally this should occur when the ribbon is cut, or shortly after. Since the how of model revision was not investigated, the simulation will only mention when k is exceeded.

4.1 Experimental Parameters

In our experiments we used 25000 iterations or cycles (meaning prediction over single outcome and matching observation) to explore, then another 25000 iterations to exploit. Afterwards, the ribbon was 'cut' and another 2000 iterations of exploitation were added. These numbers were not analytically determined, and certainly not found via comparison to human numbers, but were empirically found to have a good tradeoff between desired results and speed. All entries in the probability table of *Mobile Movement* were linked to a Beta(1,1) baseline. To find an UHPE, a value of $k = 10$ was used. This was also not found through analysis but empirically tested to work well. The repeated pattern to falsify an UHPE, as described earlier, was not implemented in the simulation.

4.2 Babybot Results

The results of the simulation are summarised in figure 5 and figure 6. The kicking frequency and prediction error were averaged in groups of 100 cycles for clarity. In actuality, both the prediction error and limb movement varied substantially more than showcased.

---

[7]This is not unlikely to happen eventually in full cognition, but should not happen in an experimental setting such as this, where factors are very still limited.

The limb that was attached to the ribbon was moved more often than the other three limbs, which remained roughly at chance level. Due to the nature of our experiment, there was no smooth increase from the explorative baseline to the full exploitation, which is inconsistent with data gathered from infants. Kicking frequency went down after the ribbon was cut, but not substantially.

The average error during exploitation was higher than during exploration. This is the inverse of what one would expect given the definition of exploitation. There did seem to be a greater deal of consistency in the prediction error than in exploration, which would agree with the nature of the two strategies.

An UHPE was found at $k = 10$ at iteration 50019 and 50033, roughly around the time of the ribbon cut. This prediction error was far higher than the average at that point which was roughly 3.25. This is indicative of the volatile nature of the prediction error. Nevertheless, for this particular value of k the UHPE correctly denoted a change in the world (the ribbon being cut). This spike in prediction error was similar in appearance to that in the coinflip experiment.
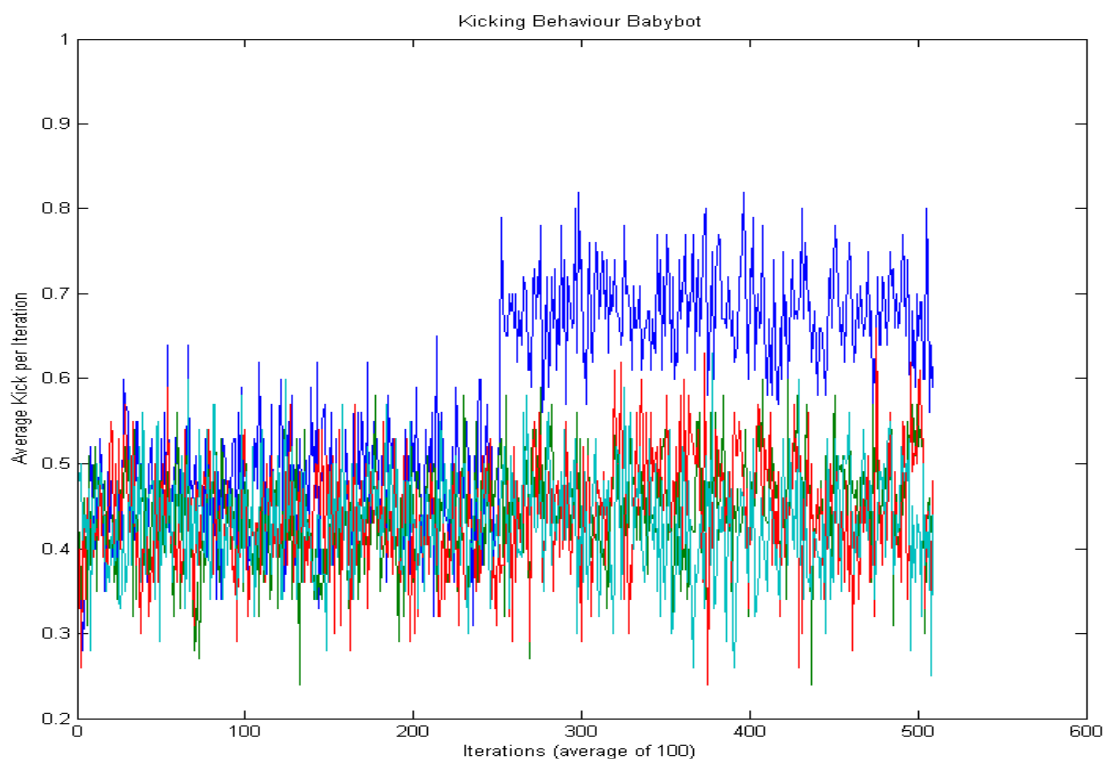

Figure 5: *Depiction of the kicking frequency of the babybot. Blue denotes the limb that is attached to the ribbon. The probability of movement is 0.5 for any limb if motor commands were chosen randomly. The values are averages of a 100 cycles.*
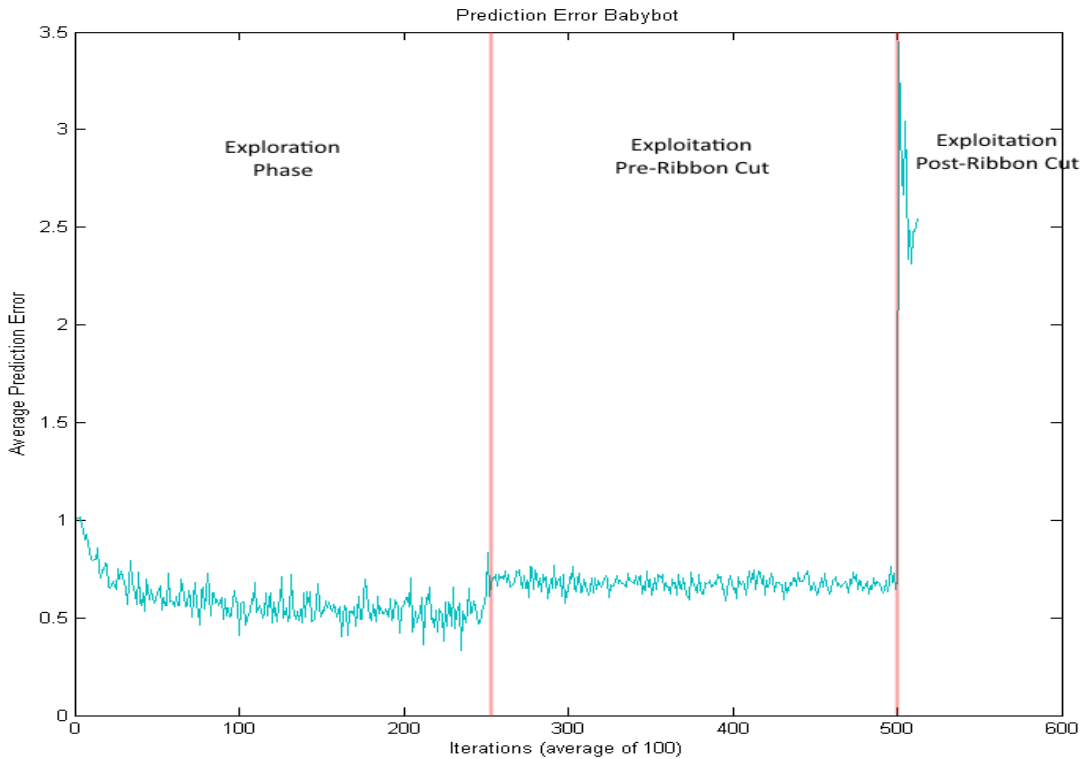
Figure 6: *Development of prediction error over time. The three distinct phases of the experiment are sectioned. Values shown are averages of a 100 cycles.*

5 Conclusion

Our research found that the attached limb increased its kicking frequency during exploitation, though average prediction error was higher than during exploration. An UHPE was found after the ribbon cut.

We conclude that hyperparameters offer a consistent and informative alternative to single value beliefs. The added notion of precision allows for a natural diminishing effect of new observations on the beliefs, where earlier data causes more change between prior and posterior beliefs than later data. This property leads to eventual stability without sacrificing quick approximation of an agent's beliefs. Our research has also shown that learning of both trivial and non-trivial models is possible using model updating with hyperparameters.

The prediction error being higher on average during exploitation than exploration is inconsistent with the theory. It is unclear whether this was the result of code error, model specific properties or chance. However, the prediction error can still be used to find highly improbable observations, which could point to a changed

16

world. These UHPE's can be used to signal model revision, pointing at a possible link between model updating and the former.

6 Discussion & Future Work

The babybot took a lot of iterations to learn the mobile paradigm. This is inconsistent with observations in humans, where similar patterns emerge in a much smaller timespan. A possible suggestion to align this better is to make the model updating step more volatile. This can be done by changing how α and β are updated. Alternatively, one could also spread the activation by not choosing values for all hypothesis nodes, but only those of interest. If those hypothesis nodes that are not manipulated are interpreted to have their values at equal probabilities, more than just one of the table entries of a prediction node is updated. This could drastically increase the rate of learning in simple cases. It is not clear however whether this approach would have an adverse effect on more complicated situations (f.e. both arms must go up, both legs down for the babybot).

Currently, our babybot uses a very implausible method to decide whether to explore or exploit. A more natural approach would have a very low, initial exploitation probability, that smoothly increases as the precision of the predictions goes up (i.e. entropy lowers). Preferably such an approach would make use of precision.

Related to the previous point is to introduce the notion that actions cost energy. Introducing a bias towards conserving energy might produce more humanlike behaviour. If at any time multiple hypotheses arise that have the same probability of causing a certain (wanted) outcome of a prediction variable, that hypothesis which moves fewer limbs is preferred.

Having found a good indicator of world change, a natural next step is to investigate how model revision finds alternative explanations and determines the causal links with regard to the existing model. In addition, although we proposed a mechanism that signals that a variable should be added to the generative model, there is no reverse step. How and when the removal of a redundant factor takes place is the second half of the question, that we did not explore. Perhaps this is not necessary, as model updating might remove the effect this redundant variable has in the model. Still, the existence of that causal relation is open for interpretation: does the agent still believe there to be a causal effect, or is it no longer seen as a variable that should be taken into account?

Finally, the falsification step following an UHPE needs formal defining. Ideally, this formalisation should lead to a rapid pattern of actions that use the current model to check whether the world operates as thought. It is possible that extra parameters need to be added to allow this check.

7 Alternatives Approaches

At various stages during research alternative approaches were considered. The first of these was a different updating step in the hyperparameter. In the current set-up, precision of belief is always increased: no evidence will cause uncertainty to come back. This is valid when considering a lot of data all at once, but perhaps less so if this evidence is processed in small batches, after which new beliefs arise.

We briefly looked into ways of implementing this uncertainty factor into our model. The most simple was a simple discount factor γ for which $0 \leq \gamma \leq 1$ that the prior α and β parameters are multiplied by before each iteration. This guarantees that after there is a point where more evidence does no longer increase precision. Thus, an agent will only have a certain belief if it recently saw a series of observations that conform to it. Additionally, after each update the older evidence reduces in importance, which leads to a recency factor. We chose to use the statistically accepted definition of Bayesian updating in our simulations instead of this γ approach, but adoption of this idea may lead to more flexible beliefs that filter out old, redundant observations.

We also looked at an alternative strategy of picking between exploration and exploitation rather than the threshold strategy described. Specifically, a softmax approach was used.

$$P(\text{Exploit}) = \frac{argmax\left(e^{P(\text{PredictionNode})/T}\right)}{\sum_i e^{P_i(\text{PredictionNode})/T}} \tag{8}$$

where T is the sum of all standard deviations of the hyperparameters. The idea of this approach was that as the more evidence was gathered, the more peaked the hyperparameter distribution would become, leading to a smaller standard deviation. Therefore, the temperate would drop, causing the chance of exploiting to increase naturally over time. However, test in the babybot simulation showed that the temperature did not diminish enough to reach this effect.

References

Angela and Dayan (2005), 'Uncertainty, neuromodulation, and attention', *Neuron* **46**(4), 681–692.

Bayes (1991), 'An essay towards solving a problem in the doctrine of chances. 1763.', *MD computing: computers in medical practice* **8**(3), 157.

Friston (2008), 'Hierarchical models in the brain', *PLoS Comput Biol* **4**(11), e1000211.

Friston (2010), 'The free-energy principle: a unified brain theory?', *Nature Reviews Neuroscience* **11**(2), 127–138.

Friston, Rigoli, Ognibene, Mathys, Fitzgerald, and Pezzulo (2015), 'Active inference and epistemic value', *Cognitive neuroscience* **6**(4), 187–214.

Kwisthout, Bekkering, and van Rooij (2016), 'To be precise, the details don't matter: On predictive processing, precision, and level of detail of predictions', *In press* .

MacKay (2003), *Information theory, inference and learning algorithms*, Cambridge university press.

Otworowska, Zaadnoordijk, de Wolff, Kwisthout, and van Rooij (2016), *Causal learning in the crib: A predictive processing formalisation and babybot simulation.*

Rovee-Collier, Morrongiello, A. and Kupersmidt (1978), 'Topographical response differentiation and reversal in 3-month-old infants', *Infant Behavior and Development* **1**, 323–333.

Shannon (2001), 'A mathematical theory of communication', *ACM SIGMOBILE Mobile Computing and Communications Review* **5**(1), 3–55.