Decision-making by Probabilistic Inference: Understanding Generative Models with Hierarchical Gaussian Filter

Edoardo Mello Rella, Savina Kim, Florian Ritsert

Abstract—**Learning is our ability to update our beliefs about the world with a mixture of new and old information, basing judgements on past experiences to make future predictions and thus improve our actions. The following study takes a top-down computational approach to understanding this process consisting of generative models aimed at making inferences about the underlying mechanisms taking place in the brain. In the following, the main objective of our work is to investigate how a Hierarchical Gaussian Filter (HGF) can be used in order to better understand how types of cues-stimuli can influence output behavior. In particular we conclude that interacting with a cue-stimulus of only one type, visual-visual, can improve learning compared to a mixed cue-stimulus of audio-visual. We then explore if personal preferences can influence the outcome of an apparently unbiased experiment. And finally, in effort to set a basis for future work, the paper provides an exploration of basic ideas in constructing algorithms that can mimic the human learning and prediction process. All code can be found on https://github.com/edomel/translational_neuromodeling_project**

I. Introduction

A. What is learning?

What makes us uniquely human is our ability to adapt in a rapidly-changing environment and make decisions according to both physical and social contexts. We have trained ourselves to make decisions based upon our predictions of the environment and expected consequences about the future based on these predictions, oftentimes through trial and error. It is through these repeated trials, or in other words, our day-to-day lives, that we unconsciously construct hidden mental states in our brain. We then express these outputs physically in the form of familiar concepts such as our beliefs, emotions and new learnings.

B. Learning

Our brains are continually preparing and forecasting our interaction with environmental data and as a result, both our expected as well as unexpected interactions with the external world is then utilized for adaptation. This process of learning can be summarized simply as the process of updating our beliefs about the world with a mixture of new and old information. This allows us to base judgements on past experiences, make future predictions and thereby improve our actions [5] [8]. There are two approaches to understand learning, with either a neurobiological or computational lens, the former requiring a bottom-up approach by surveying the neuronal architecture of synapses and circuits and the later with a top-down approach with a general computational basis consisting of generative models aimed at making inferences about the underlying mechanisms. We will focus on the later.

One of the core aspects about our ability to inference is how our behavior is driven by both our prior beliefs (regardless of true or false) as well as incoming cues from the physical reality we observe. Furthermore, the sensory information we receive on which we base our decisions are typically noisy and incomplete, hence to a degree, somewhat unreliable. Therefore, to understand perception and the representations we form of the world, it is critical to take into account causal and statistical relations as a foundational basis. This is where Bayesian inference comes into play, a compelling framework for describing cognitive processes amid uncertainty.

C. Bayesian Brain Hypothesis

Recent years can be described as something of a "Bayesian revolution." The Bayesian framework has emerged to become very popular in the field of neuroscience and cognition, being utilized frequently and successfully to provide a formal description of learning, particularly with perceptual and sensorimotor learning tasks [3] [12]. This revolution has come about from a growing appreciation of Bayesian statistics and exploiting it to mathematically model problems of inference and decision-making under uncertainty. This theory blends itself to a Bayesian treatment towards incoming information, a continuous updating of conditional probabilities, also known as the "Bayesian Brain hypothesis," which aims to conceptualize brain function from initial perception to resulting behavior [6]. In the process of decision making, as more input is obtained from the world, our predictions typically become more accurate given we are becoming less uncertain. In other words, the brain is a "probabilistic prediction machine," constantly optimizing to minimize the divergence between sensory inputs and its self-generated predictions from these inputs [14]. This process can be described as "prediction error minimization" (PEM) or "predictive processing" by application of approximate Bayesian inference [9]. Utilizing this encoding strategy, the unpredictable components of a signal are fed forward to higher stages of information processing. Furthermore, this concept relates to what is known as Karl Friston's "free-energy principle" which states that PEM is an example of our brain's greater overarching demand to self-organize according to minimization of energy expenditure [6]. An optimal brain will hold and continually update a generative model of its internal and external setting with the goal of avoiding costly actions on itself through changing anatomy and on the world around it.

The Bayesian framework represents knowledge based on a distribution of differing beliefs across a range of hypotheses. This is a significant improvement from traditional models which were based simply on associativity, assuming learners were passive recipients of information [2]. Instead, current models represent how humans actively learn, adapting to the environment in effort to learn quickly and optimally, arriving at beliefs that reduce their uncertainty or make certain hypotheses more probable than others. For a better understanding, let's say the learner holds a range of hypothetical values for each descriptor. For example, as we will see later in the experiment, the association between "seeing a dog" and "seeing an owner" can fall anywhere on an infinite continuum, and the participant's knowledge is represented as a distribution of beliefs over that continuum. Even if the participant strongly believes a value of 0.7 for the associative weight between the visual stimuli, he or she also holds some belief in possible values smaller or larger. Therefore this infinite distribution of beliefs can be represented compactly with a Gaussian distribution simply with a few values, mean and variance. The greater the belief distribution, the more uncertain the participant is. Therefore while learning, the Bayesian model represents the shifting of beliefs such that whenever there is an occurrence of seeing an owner following a dog, a higher value of association becomes more believable. As a result, the distribution begins to narrow as uncertainty is lessened and the learner becomes more convinced about their association.

However, our knowledge of how these probabilistic computations are structured in the brain still remains rudimentary at best. In this paper, we aim to further our understanding of these concepts of statistical inference and learning. This paper is divided into the following sections: the first half is dedicated to providing a theoretical background and deeper dive into Bayesian Inference followed by the second half which will focus the experimental results and related discussion.

II. Theoretical Background

A. Bayes' Theorem

After observing some input from the external world, the question is, how should one update his or her beliefs regarding the objects that could have produced the evidence? We are constantly faced with an inference problem where our brain must infer the cause of its sensory inputs based upon the inputs themselves [7]. The following section will provide an overview of the popular Bayesian framework. Accordingly, quantities of interest are treated as random variables and our beliefs about their value are represented as probability distributions over these variables. Bayes' Theorem states the following:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \tag{1}$$

Bayes' Theorem states that our belief in a candidate structure or hypothesis (θ is the unknown) after observations y from the environment, the "posterior" $p(\theta|y)$ is proportional to the product of two terms:

- the "prior" $p(\theta)$ which is our prior belief before having observed data
- the "likelihood" $p(y|\theta)$ which is how likely the observation would have been produced by the candidate structure θ assigning probabilities to different instances of sensory evidence given these hypotheses.

When we learn, the brain updates these priors to posteriors.

B. Generative Models with Bayesian Inference

Adopting a generative model m first requires formalizing the generation of observable data y from the environment (i.e. the sensory input stream when crossing a street from our visual, vestibular and auditory systems) as a function of parameters θ, or in other words, how the data should look like given hidden causes ("latent variables"). Latent variables cannot be observed directly but rather symbolize abstract features that allow us to perceive and structure input; for example, the speed of a bike or anticipated occurrences when crossing a street, which are compact representations of critical aspects of the environment which may be more reliable than individual local observations often burdened by noise; together this forms a stable global interpretation.

When using these models against empirical data, going from consequences to causes (versus causes to consequences when predicting observations), the aim is to estimate the unknown parameters from the observations, which can be represented as a posterior distribution over the unknowns:

$$p(\theta|y,m) = \frac{p(y|\theta,m)p(\theta|m)}{p(y|m)} \tag{2}$$

However, this leads us to questions such as accounting for where relevant priors or likelihoods come from and furthermore, how are these processes implemented in the brain's neural networks? [1] The following aims to answer this fundamental question.

C. Hierarchical Predictive Coding

A defining feature of Bayesian inference is that the brain is constantly updating its hierarchical generative model of sensory inputs to make inferences on the causal structure of the environment. This process of belief updating and estimating environmental volatility is based on hierarchically related prediction errors which are weighted by their precision [9] [5]. One way to answer this question is therefore to understand Bayesian inference in terms of precision-weighted prediction error minimization (PEM) and how it is executed in the brain. Prediction error can be defined as a mismatch between prior expectation and reality [4]. Therefore it fluctuates inversely with the hypothesis likelihood; the greater the likelihood, the smaller the prediction error generated.

The next question is then how much should the prior be updated given the prediction error generated? We look to uncertainty levels by providing weights to the reliability of the data. This can be calculated by the inverse of the variance of both density functions, in other words, the relative precisions of the Gaussian distributions [14]. The ratio then determines the learning rate where the more precise the priors are relative to incoming information, the less one learns and therefore the less the prediction error influences the posterior and vice versa. In summary, as one learns more about a topic, his or her priors will become less uncertain (more precise) and the prediction errors will be weighted less in formulating inference.

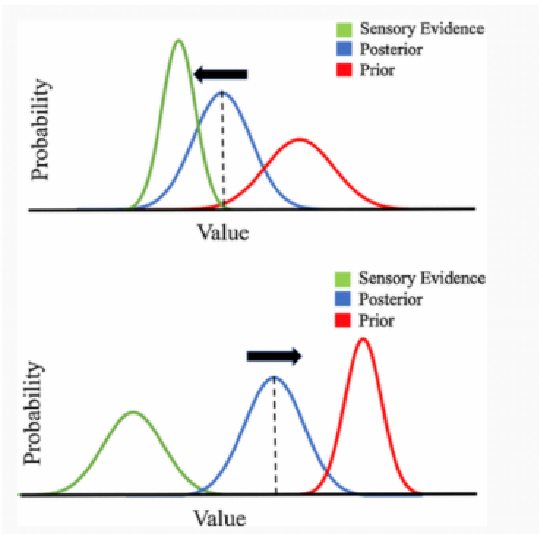

Fig. 1: Posterior beliefs are a function not just of the means but also the precisions of the sensory evidence and prior distribution. On the top graph, the comparatively greater precision of the sensory evidence has a proportionately greater influence on the posterior than the prior. On the bottom, this is reversed. *Sourced from Williams, 2018.*

It is also important to note the increased complexity necessary to apply this concept to real-world perception. For example, it is impractical to stubbornly hold onto a prior expectation about where a lecture will be held if an announcement was made about a room change. The volatility of the real world requires this system to be flexible and able to fine-tune the learning rate as conditions change over time - this brings us to hierarchical Bayesian inference.

III. Hierarchical Gaussian Filter (HGF)

A generative model is hierarchical when it contains multiple levels of latent variables and can be characterized by a hierarchy of states that evolve over time as Gaussian random walks where each step size is determined by the following higher level of hierarchy. This brings us to the flexible Bayesian learning model used in this experiment, the hierarchical Gaussian filter (HGF). The HGF is unique given its flexible learning rate while continuously adapting to environmental volatility, allowing for more meticulous representation of the learning process.
The HGF works to extract regularities across various time scales from the non-linear time series of sensory input [11][8]. This concept is based on a hierarchy of hypotheses spaces which capture regularities found in the external world at increasingly greater spatial and temporal scale. For example, the hypotheses at level two behave as the "priors" for the level below (level one) and as "sensory evidence" for the level above (level three). Going back to PEM, this method is repeated up the hierarchy and as a result, the higher level hypotheses can help guide inferences in lower levels. Furthermore, this structure allows one to make expectations for the precisions of sensory evidence across multiple levels which allows for a flexible learning rate. For example, a subject learns that the reliability of an auditory evidence is differentiable according to one's location such that hearing a car horn while crossing a street versus while looking out an office window, and thereby adjusts the precision of this evidence accordingly. This allows us to be responsive not only to the obvious environment structure but also to a context-dependent noise or uncertainty while it is relayed as incoming sensory evidence.
To analyze participant behavior, we utilized a HGF model where the weight given to prediction errors is encoded in the participant-specific learning rate, which is high in cases of abundant environmental uncertainty and vice versa. The HGF enables us to quantify different levels of perceptual uncertainty perceived by the participant. The Translational Algorithms for Psychiatry-Advancing Science (TAPAS) package with the HGF toolbox was utilized for implementation and inversion of the HGF. Further details can be found on their website (http://www.translationalneuromodeling.org/tapas/).

A. *Theory of hierarchical generative model*

The hierarchical generative model we introduce here deals with states and inputs that are discrete, accounting for both deterministic and probabilistic relationships between the participant's perceptual state and environmental occurrences. As shown in Mathys et al., 2011 referenced in the following, for our model the participant is only interested in a binary state of the environment (i.e. whether they will see an image of an owner or not) therefore the environmental state x_1 at time k is denoted by a $binary x_1(k) \in 0, 1$ where x_1 determines the probability of sensory input u (3).
With HGF, a binary outcome is generated by sampling using a Bernoulli distribution shown in (4) as the empirical prior density where the parameter s is a sigmoid function of the tendency x_2 (5). At time k, x_2 is normally distributed around its value at the previous time point $x_2(k-1)$, assuming that may change over time as a Gaussian random walk.

Furthermore, the dispersion of the random walk, the variance represented as $exp(kx_3 + w)$ is then determined by the state x_3 and parameters k and w which can be subject-specific. The state x_3 also determines the log-volatility of the environment. In other words, the tendency x_2 of the owner to appear performs a Grandom walk with volat Furthermore, the dispersion of the random walk, the variance represented as $exp(kx_3 + w)$ is then determined by the state x_3 and parameters k and w which can be subject-specific. The state x_3 also determines the log-volatility of the environment. In other words, the tendency x_2 of the owner to appear performs a Grandom walk with volat Furthermore, the dispersion of the random walk, the variance represented as $exp(kx_3 + w)$ is then determined by the state x_3 and parameters k and w which can be subject-specific. The state x_3 also determines the log-volatility of the environment. In other words, the tendency x_2 of the owner to appear performs a Grandom walk with volatility $exp(kx_3+w)$ (6). At the final level we set the volatility of x_3 to theta, a constant parameter which is subject-specific (7). To summarize, now with the full priors on the parameters $p(k, w, theta)$, the full generative model is represented in (8) [10].

$$p(u|x_1) = u^{x_1}(1-u)^{1-x_1} \tag{3}$$

$$p(x_1|x_2) = s(x_2)^{x_1}(1-s(x_2))^{1-x_1} = Bernoulli(x_1; s(x_2)) \tag{4}$$

$$s(x) \equiv \frac{1}{1+exp(-x)} \tag{5}$$

$$p(x_2^{(k)}|x_2^{(k-1)}, x_3^{(k)}) = \mathcal{N}(x_2^{(k)}; x_2^{(k-1)}, exp(\kappa x_3^{(k)} + \omega)) \tag{6}$$

$$p(x_3^{(k)}|x_3^{(k-1)}, \theta) = \mathcal{N}(x_3^{(k)}; x_3^{(k-1)}, \theta) \tag{7}$$

$$\begin{gathered} p(u^{(k)}|x_1^{(k)})p(x_1^{(k)}|x_2^{(k)})p((x_2^{(k)}|x_2^{(k-1)}, x_3^{(k)}, \kappa, \omega))) \\ p(x_3^{(k)}|x_3^{(k-1)}, \theta)p(x_2^{(k-1)}, x_3^{(k-1)})p(\kappa, \omega, \theta) \end{gathered} \tag{8}$$

The probability at each level of the generative model is determined by the parameters and variables at the next highest level; we assume the default of three levels. The levels are related by determining the step size of the random walk (volatility or variance) with the highest step size being constant parameter theta.

IV. Motivation

The goal of this paper is to investigate the following question:

A. *Does the nature of the cue-stimulus result in behavioral bias in a sensory-motor learning task?*

For the following experiment, we acquired behavioral data from 12 participants in a visual-visual and visual-audio learning paradigm. Each participant was presented with a binary or audio visual cue and then asked to predict whether they would see a visual stimulus or not. The complexity of the task is the fact that the probabilities that govern the cue-stimulus contingencies are adjusted over time. Further details about the experimental design are described in the subsequent section. In the following, we elaborate further on how our experiment aims to resolve this question and our experimental hypotheses.

In the first set of 115 trials, a visual cue of a dog or cat image was presented to participants who were then tasked to predict whether a subsequent visual stimulus would occur. In the second set, similarly a visual cue was presented however with a circle or square image. The third set however, the cue was auditory such that the participant was presented with either a higher or lower tone sound. For all three sets, the stimulus remained visual. Comparing HGF parameter fitting between the different sets between two different visual cues and between visual versus auditory cues aims to shed light onto how the nature of the cue potentially affects behavior. For visual-visual, one hypothesis is that more "interesting" cues such as a dog vs. square images might have a differentiating effect. For visual-auditory, given the subsequent stimulus is also visual, a visual cue could serve better for more accurate prediction abilities through the usage of complementary visual pathways in the brain.

B. *Methods*

Behavioral Task

12 subjects participated in the study. The subjects are asked to predict whether they will see an image on a subsequent display based upon the visual or auditory binary input they received on the previous screen. The experiment consisted of three sets which differed in the nature of the visual or auditory cue as well as the resulting visual stimulus. The setup is displayed in Figure 1. The visual cues consisted of two possible animal images (cat or dog) and two possible shape images (square or circle) and the auditory cues consisted of two possible tones differentiable by high and low pitch.

For each experiment, the participant was asked to take part in three sets of trials with 115 trials per set (345 trials per experiment). In the first set, the participant was shown either an image of a dog or cat and then asked to predict whether they would see an image of an owner or not in the subsequent display. After making a decision of yes or no, the participant was shown an image of either an owner or no owner thereby confirming whether they had chosen correctly or incorrectly. In the second set, the participant was shown either an image of a square or circle and asked to predict whether they would see an image of a line or no line in the subsequent display. Similarly, the third set consisted of the participant being presented with either of two different auditory tones and asked to predict whether they would see a line or no line in the subsequent display.

Each of the three sets of 115 trials are then divided into four discrete probability blocks. The cue-stimulus likelihoods were governed by the following probabilities:

$$\begin{aligned} &P(output|input_1) = P(nooutput|input_2) = \\ &= 1 - P(nooutput|input_1) = 1 - P(output|input_2) \end{aligned} \tag{9}$$

$$p(input_1) = p(input_2)$$

The number of trials per block were varied such that the first block consisted of 30 trials, 25 trials in the second block, 32 trials in the third block followed by a final fourth block of 28 trials. The appearance of the output is related to the input that you observe with a probabilistic model. Changes in the p(output|input1) occurred after each block and the probability was modified in a discrete fashion. For set 1 (dog/cat), the probabilities changed in the following order: 0.7, 0.5, 0.2, 0.9; for set 2 (square/circle), the probabilities changed in the following order: 0.2, 0.7,0.5,0.9; and for set 3 (high tone/low tone), the probabilities changed in the following order: 0.9, 0.5, 0.2, 0.7. The order in which the different sets were presented was altered across the participants, therefore for a subset of the

12 participants, the auditory-visual cue-stimulus was presented first, followed by the square/circle visual-visual cue-stimulus and then the dog/cat visual-visual cue-stimulus in effort to remove order bias.
Subjects were instructed that the appearance of an output is related to the input initially observed with a probabilistic model that would change over the course of the task. The aim of the participants were to minimize their prediction errors, on average, across all the blocks of the task.

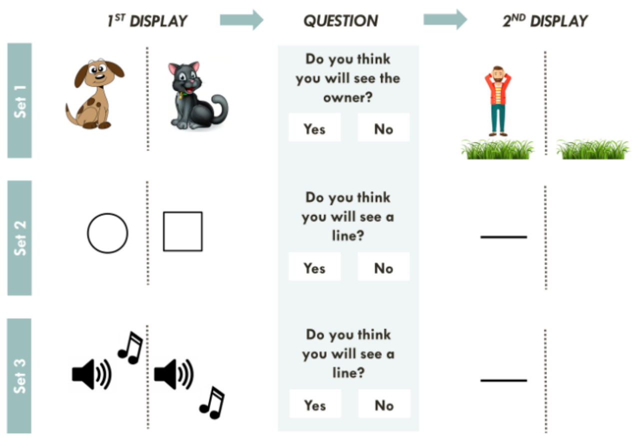

Fig. 2: Each trial consisted of a binary visual or auditory cue, followed by a response period in which the participant was asked to enter yes or no, a visual stimulus presentation and then a variable inter-trial-interval

C. Questionnaire

Prior to beginning the experiment, participants were asked a short questionnaire consisting of three of the following questions:

- How do you feel today? (5 being excellent, 1 being poor)
- How satisfied with life do you feel in general? (5 being excellent, 1 being poor)
- Do you prefer dogs or cats? (0 for dog, 1 for cat)

The aim of the questionnaire was to first, to potentially eliminate significantly under-performing participants with increasing confidence and second, to correlate the responses with experimental results while also investigating potential differences in model parameters depending on preferences.

V. Results

A. Fitting the HGF

For each participant we individually fit a 3-layer HGF model for each of the three sets separately and acquired subject-specific parameter estimations for tonic log-volatility ω_2 and ω_3, which represents the constant component of log-volatility and thereby has a modulating effect on the participant's learning rate. Figure 3 shows the relative changes for trial each set, averaged over all participants. Instead of assessing the absolute values of ω_2 and ω_3, we analyzed the variance of each from the optimal ω.

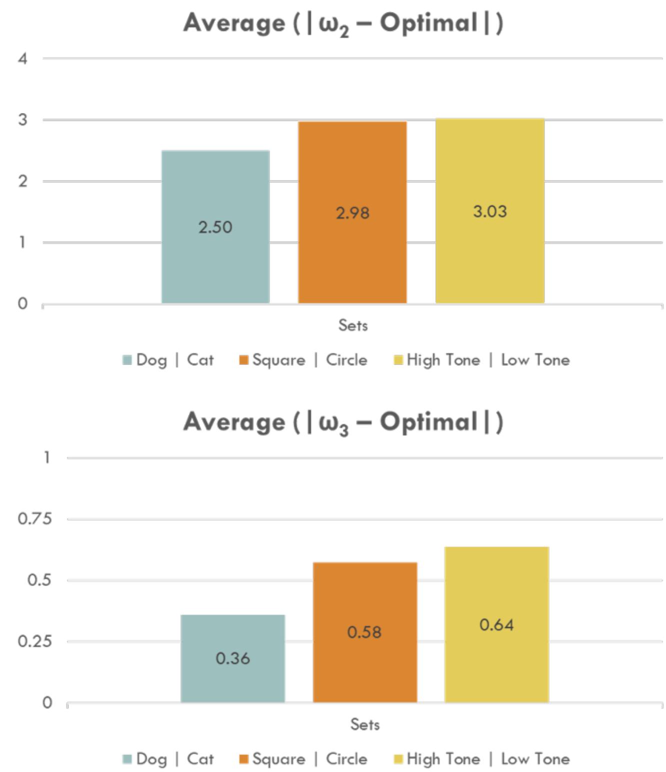

Fig. 3: Relative changes of ω_2 and ω_3 to the optimal w-value for each trial set, averaged over all participants

B. Performance Analysis

To determine participant's performance profiles, we analyzed their responses in each of the four trial blocks that are conducted at various cue-stimulus occurrence probability distributions. This analysis was carried out for each of the three sets separately. A participant's performance score is calculated as his or her proportion of correct responses. The response's correctness is based on the underlying probability distribution; for example, in a block where seeing an image of an owner given an image of a dog initially occurs with probability greater than 0.5, the correct response is "yes." Furthermore, responses from blocks where the conditional probability was at chance level (0.5) were omitted in the performance analysis. In figure 4 the average number mistakes in the 115 trials for each set are shown.

It is possible to observe from figure 3 and 4 that, as expected, participants performed better when the cue was visual. This hypothesis seems to be confirmed as, both in terms of the total number of mistakes and by how far the parameters are from the ideal. This decrease in performance with an audio cue indicates a confirmation that it is potentially easier to learn from visual input. The other fact that could explain part of the difference is that, as reported from some participants, it is more difficult relating audio to visual cue-stimulus compared to complementary visual to visual.
Trying to compare, instead, the performances when having an "interesting" input compared to an "impersonal" one, the result is not clear. In fact the parameters are more similar to the optimal ones with cats and dogs as inputs while the number of

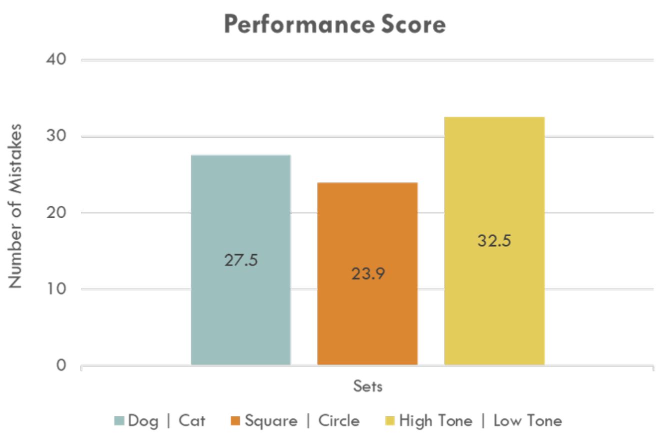

Fig. 4: Average number of mistakes made by participants in each of the three sets

mistakes is lower for figures as inputs. This therefore doesn't allow us to infer any difference in how these influence learning. Therefore we can say that, even in the case there could be a difference in performance, other aspects influenced the results greater than the difference in the input. In fact, even when the experiment was conducted by changing the order of the tasks, the number of samples was not big enough to cover the differences caused by this external factor (learning or "getting bored" from one set to the next).

VI. Data Clustering

A. Objectives

This section of our work had three main objectives:

1) **Prove the validity of the HGF model**
In order to further prove the consistency and the validity of this model we wanted to check if the classes of different parameters corresponded to different performances in terms of average number of mistakes. In particular we wanted to check if the parameters produced by the fitting could be divided into different classes and if the classes could infer any meaningful prediction on the responses given by the participants in the experiment.

2) **Check relations between the responses and the personal states of the participants**
Assuming the possibility of a correlation between the mood and more generally the personal state of the participants and their performances we wanted to see if these could be inferred from the experiment conducted. In particular, assuming that the previous point (objective 1) could be proved successfully, we tried to show some correlation using the classes found clustering the parameters computed.
Even if this is an interesting argument to validate we were not expecting to see strong evidence from the results given the participants did not show any particularly strong belief about their personal state.

3) **Demonstrate an influence of input cue in relation to personal preference on performance**
Having presented cats and dogs as visual cues in one experimental set and asking for personal preferences between the two animals, we wanted to investigate whether preference of one input could influence performance. We decided to use cats and dogs as animals that can be characterized by a strong personal preference and that could therefore accentuate the effect we were trying to see.

B. Method

For this section we exploited the unsupervised clustering method called Deterministic Annealing Clustering as a method that generally shows good results in clustering vectors of arbitrary dimensions. In particular, in order to exploit this method, the results were saved as vectors and each vector was assigned to a participant. To clarify with an example, for the first task each participant corresponded to a two dimensional vector that contained respectively the average of the ω_2 throughout the three sets of the experiment and of the ω_3.
Deterministic Annealing Clustering, explained in details in [13], is a method inspired by physics that combines random exploration and moves in the direction of a minimum using a temperature parameter. As in the solidification of materials, as the temperature decrease the moves tend to be more and more directed towards the nearest minimum making the random effect less and less significant. This, decreasing the temperature at an appropriate rate, usually achieves good results in avoiding local and sub-optimal minima in favour of better results. In particular, this method increases the classes in which the data is divided as they reach a critical temperature. This allows to avoid creating unnecessary classes automatically.
This method was applied to the three tasks in the following way:

1) Two dimensional arrays of the average ω_2 and ω_3 are the vectors to be clustered. Then from the classes found the average number of mistakes of the elements in each class are computed checking if the results are related.

2) Two approaches are utilized: first the vectors to be clustered are enriched with the information from the first two questions from the questionnaire making them four-dimensional arrays, then the previous clustering (from section 1) is kept and the score from the questionnaire are averaged in each class and compared.

3) In this last section only the parameters referring to the cat-and-dog set are used. With these ω_2 and ω_3 the same approaches used in the previous section (section 2) are executed.

C. Results

Before analyzing the results it is useful to note that the total number of classes was set to two by default as the amount of data was not high enough to allow us to obtain significant results when trying to identify more classes.

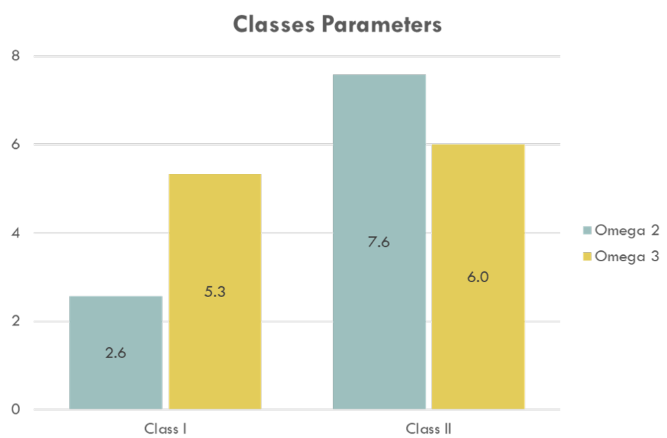

Fig. 5: Centroid values for ω_1 and ω_3 for the two different classes

In this section the answers to the three different problems posed are analyzed separately:

1) **Prove the validity of the HGF model**
As expected the results supported the hypothesis we wanted to confirm. In fact, even if there was no significant difference between the two classes in terms of ω_3 (specifically the results were -5.340 and -6.004 for the two classes with the difference caused mainly by one outlier), the difference in ω_2 was consistent. In particular, for class I $\omega_2 = -2.577$ while for class II $\omega_2 = -7.584$ with a standard deviation of the two classes respectively $\sigma = 0.869$ and $\sigma = 0.961$ as seen in figure 5. To prove the validity of the model was the average number of mistakes that for the first class was 21.43 while for the second was 37.13 (shown in figure 6) with a distance in the averages 3.3 times larger than the highest of the standard deviations and with no intersection in the interval of values of mistakes between the two classes.
The number of samples is not big enough to be considered a valid proof but the difference in the classes and the correspondence between the parameters and the mistakes made are significantly large to be a clear sign of the likely correctness of our assumption.

2) **Check relations between the responses and the personal states of the participants**
As expected, it was not possible to see any meaningful correlations between the parameters and the responses regarding the personal state. In fact the two classes had responses regarding their personal state that differed of less than 5% on average. Furthermore the two methods to identify the two classes didn't show any difference as the class members remained the same as the ones in the previous point.

Fig. 6: Average number of mistakes of the elements in each class

3) **Demonstrate an influence of input cue in relation to personal preference on performance**
Unfortunately, we were not able to reach a conclusion for this objective given that out of the twelve participants in the experiment, only one had a preference for cats over dogs. Therefore the information was not significant enough to obtain any kind of meaningful differentiating result.

D. *Future Improvements*

Overall the first improvement that could help in any section would be having more data to fit and cluster. Excluding this, the validation of the HGF seems already significant enough. Considering the other two sections, certainly having more data would be necessary to obtain any result. In particular, considering the correlation between performances and personal state, it would be interesting to conduct the experiment with participants reporting signs of some psychological disturb.
Assuming an appropriate amount of data, it would potentially be feasible to reach an explanation for the classes found and use the knowledge in order to create experiments able to observe inner-class differences. Focusing instead on the third task, if our point were to be demonstrated, the objective could be trying to quantify how much the result is biased in order to be able to exclude this term in future experiments.

VII. Method Exploration

A. *Objectives*

The main purpose of this last section is trying to evaluate different ideas of algorithms that could act on the same experiment given to the participants. These responses were compared to the human ones trying to find the similarities in terms of number of mistakes, parameters of the HGF that could have crated the same response and, possibly, the more general behaviour and trend of the responses. This intent is the basis for further research in alternative methods to model the process of responding to tasks similar to the current. In particular, it is theoretically possible that some models are more effective in describing certain agents while for others a different model could be more appropriate.

B. Methods

As previously mentioned, three methods were used to simulate the experiment and will be briefly described in the following. All three were thought thinking of how a human agent could think rationally when answering the inference question. Then a quick and approximate optimization of the parameters was run in order to get the best performances out of the methods. This is not necessarily indicative of the best choice of parameters (to simulate a human agent) but seems optimally appropriate in evaluating the methods.

1) **Non-Bayesian probability update**
The first method implemented uses a variable to store the probability of a correspondence between on input and an output (this is implemented taking into account that $p(output|input1) = p(no\ output|input2)$. Based on this variable that, as a probability, will be $p \in [0, 1]$, the prediction of an output is chosen. Then the actual response to each input is used to update the probability. This is done by summing the previously computed probability multiplied by a parameter α and the actual response (represented as a binary value) multiplied by $1 - \alpha$. To initialize the experiment the agent is considered unbiased and the probability is set to 0.5 making the first prediction random. For the simulation, the parameter was set to $\alpha = 0.9$ as it seemed to give the best results on average. This method learns in the same way, no matter if the probability is increasing or decreasing. This is something that was assumed for simplicity but also tested and checked utilizing the question about cat-dog preference.

2) **Low-pass filter inspired probability computation**
The second method is based on the same basic idea that the response can be chosen based on a variable that stores the probability for choosing one response. Also the initialization was the same as the variable and was set to 0.5 to start without any sort of bias. What changed in this method was the way of computing the probability update. In fact, as with a low-pass filter, the probability was computed using an average over the past 7 observations. As the experiment is based on binary variables and the probability is clearly in the same range, a simple average was enough. The decisions of keeping 7 variables was made as it proved to be the best choice providing both smoothness and a quick learning curve.
To add further smoothing, the an α parameter is used in the same way as the previous method. In this case, instead, it was set to 0.5 since the learning needs to be faster as averaging is already done. This addition was useful in particular when the probability is slightly above or below 0.5 as this allows avoiding unnecessary oscillations and does not significantly change the learning time when a strong probability appeared.

3) **Reward-based probability update**
The last method tried was more inspired by reinforcement learning, in other words, trying to see how a human agent could think in solving the problem if it was moved by positive and negative rewards. In order to implement this a deterministic approach was chosen. Differently from the previous cases, the response was based on two variables whose relative values determine the response to be chosen. The two variables were alternatively updated depending on the input that should have appeared. The single variable then was strongly affected by a prediction error (whose effect would even keep giving its effect for the following updates) while a correct prediction would enforce the correct belief. In order not to increase the variable values too heavily, the smallest one was always set to zero and the maximum difference between the two variables was set such that three wrong predictions could be enough to change the belief.
Design choices could be made both in the decision on how to set the maximum possible difference and on how much effect the wrong and right predictions should have. In this case, the variables were set to obtain good results even if, due to the higher complexity of the method, a good exploration of the different parameter combination could not be done.

Given these three methods, the responses were simulated giving as input to the algorithms the sequences seen by all the participants in the experiment. This was done in order to better compare the single simulated responses to the one produced conducting the experiment.

C. Results and Performance Evaluation

The responses of these simulations were evaluated first in the number of mistakes and in the parameters of the HGF that could have generated the same response. The average results from this first evaluation can be seen in figure 7 and 8 where we can see respectively the number of mistakes done on average by the algorithm and the average parameters with that the HGF could have generated the same output. It is possible to see that the performances improve as the method gets to be more complex. In particular, comparing these results to the classes found from the actual data collected in the experiment, it is possible to see that the performances of the first method are comparable to the second class while the third method has performances more similar to the first class. The fact that in this case the number of mistakes is still lower compared to the ones committed by a human performer can be explained by the fact that this method is completely deterministic.

After comparing these parameters, we tried to find some quantitative score to evaluate how similar the responses are compared to the human ones. Unfortunately, we could not identify a score which would appropriately show the similarity. We therefore show here both the response given by one of the participants to an input (figure 9) and the one given by the third method that was shown to be the most promising one to the same input (figure 10).

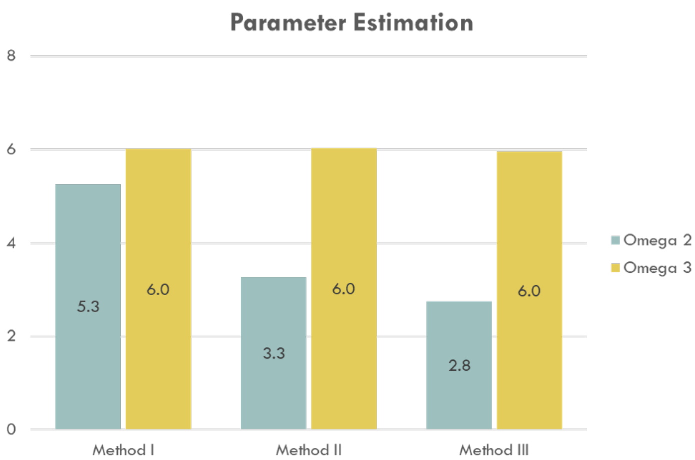

Fig. 7: HGF parameters that could have generated the same response as the methods to simulate the experiment

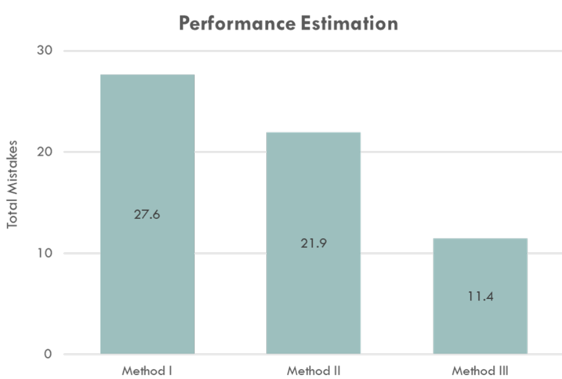

Fig. 8: Average numbers of mistakes made by each of the methods on all the trials

It can be clearly seen that the similarity is extremely high and we can therefore infer that the method used can be a good starting point for mimicking human behaviour. The different number of mistakes could be a result of the method being deterministic instead of probabilistic. Some approaches to overcome this problem are further discussed in the following section.

D. Future work

The current methods have been developed specifically to solve this problem. Future development could be based on the investigation of the following two directions:

1) It would be necessary, in order to make these methods more general, to extend these ideas so that they could be applied to more general situations. This should be feasible as the ideas behind the methods are basic and make no assumption on the experiment at hand. In fact, if we accept that in decision-making there exist a probabilistic component, then the idea of a low-pass filter can easily be applied adding some kind of memory to store probabilities if the experiment varies in the types of input. At the same time, also an algorithm that reacts to negative rewards and goes in the direction of positive ones can be applied in general contexts. Also in this case adding some memory can help making the method more general.
Considering specifically this third algorithm, a significant weakness of it is that at the moment it is a deterministic method while we have assumed that in decisions there is a probabilistic component. This problem could be overcome by inserting a probability in the decisions based on the deterministically computed parameters that determine the likelihood of each outcome. This new method would decrease the performances of the simulated agent but it would make them more realistic. Another possibility is instead computing in a probabilistic manner the parameters and from these acting in a greedy manner. And finally these methods could be combined making the algorithm probabilistic at each stage. These developments should then be tested before taking the best option.

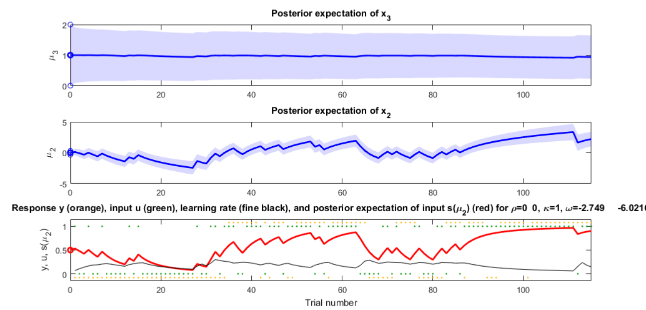

Fig. 9: Response of one of the participant obtaining in class I (with a low number of mistakes).

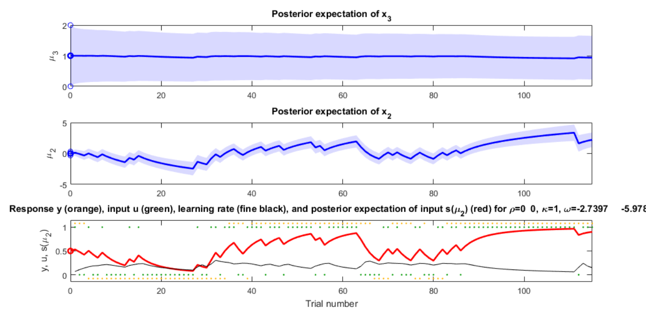

Fig. 10: Simulated response using the reward-based method applied to the same input as the agent in figure 9

2) The other area in which these methods should be developed is in the inversion of their application. In fact, at the moment they produce responses from the input-output combination they are presented. It still has to be developed rather the part that takes the responses from a real agent and the input-output combinations and thereby tries to predict the parameters that might have most likely produced the observed

output. As with the HGF framework, it could be possible for each method to predict all the parameters or on the other hand fix some of them and predict the others.

The other improvement that can be done and that is strongly correlated with previous sections of this report is allowing these methods to have a bias to represent the preference of learning one output compared to another. This still is only based on our hypothesis given the actual effect of a preference could not be significantly demonstrated with the experiment.

VIII. Conclusion

The objectives of the experiment have partially been reached. In fact the HGF validity was further proved and doing so it was also possible to observe classes of people with different parameters. With these classes it was possible to show correlation between the stimulus type and the performances in learning. Even if it was not possible to observe differences caused by different preferences, the ease of clustering this data is a promising sign for how research in solving this problem could be done in the future.

There is also potential in finding promising results by exploring new methods of simulating the experiment considered which could even give responses extremely similar to the human ones. To conclude, we can be satisfied with the results obtained given our demonstration in the validity of select hypotheses and thereby allowing us to discover and provide recommendations of ideas for future development and research.

IX. Author Contributions

Excluding the *Method Exploration* section that was mostly developed by Edoardo Mello Rella, all the other parts of the project were the results of a common work. In fact all the decisions and part of the implementation was done during the frequent meetings.

Trying to identify the single parts where each member contributed most, we can say that Savina Kim was more involved in the deepening the theoretical background of the HGF and in designing the experiment. Florian Ritsert instead, developed a GUI to make the fitting of the HGF parameters easier for the user, and contributed more in fitting the parameters and identifying the differences in performance when changing the stimulus nature. Edoardo Mello Rella coded the clustering algorithm and set the method to apply it and worked more in setting the objectives of the clustering, identifying the different classes from the fitted parameters and interpreting their relations. To conclude, all the authors contributed equally in finding participants to the experiment in order to collect the data.

References

[1] Nick Chater et al. “Bayesian models of cognition”. In: *Wiley Interdisciplinary Reviews: Cognitive Science* 1.6 (2010), pp. 811–823.

[2] J. Daunizeau et al. “Observing the observer (I): meta-Bayesian models of learning and decision-making”. In: *Plos One* 5.12 (2010), e15554.

[3] Nathaniel D. Daw and Kenji Doya. “The computational neurobiology of learning and reward”. English (US). In: *Current Opinion in Neurobiology* 16.2 (Apr. 2006), pp. 199–204. ISSN: 0959-4388.

[4] Hanneke Den Ouden, Peter Kok, and Floris de Lange. “How Prediction Errors Shape Perception, Attention, and Motivation”. In: *Frontiers in psychology* 3 (Dec. 2012), p. 548.

[5] Timothy E J Behrens et al. “Learning the value of information in an uncertain world. Nature Neuroscience, 10(9), 1214-1221”. In: *Nature neuroscience* 10 (Oct. 2007), pp. 1214–21.

[6] Karl J. Friston. “The history of the future of the Bayesian brain”. In: *NeuroImage*. 2012.

[7] H v. Helmholtz. “Handbuch der physiologischen”. In: (1867).

[8] Jakob Hohwy. “Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization”. In: *Consciousness and Cognition* 47 (Sept. 2016).

[9] Sandra Iglesias et al. “Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning”. In: *Neuron* 80 (Oct. 2013), pp. 519–30.

[10] Christoph Mathys et al. “A Bayesian foundation for individual learning under uncertainty”. In: *Frontiers in Human Neuroscience* 5 (2011), p. 39.

[11] Christoph Mathys et al. “Uncertainty in perception and the Hierarchical Gaussian Filter”. In: *Frontiers in Human Neuroscience* 8 (2014), p. 825.

[12] John O’Doherty, Alan Hampton, and Hackjin Kim. “Model-Based fMRI and Its Application to Reward Learning and Decision Making”. In: *Annals of the New York Academy of Sciences* 1104 (June 2007), pp. 35–53.

[13] Kenneth Rose. “Deterministic Annealing for Clustering, Compression, Classification, Regression, and Related Optimization Problems”. In: *Proceedings of the IEEE*. 1998, pp. 2210–2239.

[14] Daniel Williams. “Predictive coding and thought”. In: (Mar. 2018).