Modelling Visual Decision Making Using a Variational Autoencoder

Tyler Malloy
Carnegie Mellon University

Modelling Visual Decision Making Using a Variational Autoencoder

Tyler Malloy (mallot@rpi.edu)
Department of Cognitive Science
110 8th Street, Troy, NY 12180

Chris R. Sims (simsc3@rpi.edu)
Department of Cognitive Science
110 8th Street, Troy, NY 12180

Abstract

Due to information processing constraints and cognitive limitations, humans necessarily form limited representations of complex visual stimuli when making utility-based decisions. However, it remains unclear what mechanisms humans use to generate representations of visual stimuli that allow them to make predictions of utility. In this paper, we develop a model that seeks to account for the formation of representations in utility-based economic decision making. This model takes the form of a β-variational autoencoder (β-VAE) trained with a novel utility-based learning objective. The proposed model forms representations of visual stimuli that can be used to make utility predictions, and are also constrained in their informational complexity. This representation modelling approach shares common features with related methods, but is unique in its connection to utility in economic decision making. We show through simulation that this approach can account for several phenomena in human economic decision making and learning tasks, including risk-averse behaviour and distortion in the calculation of expected utility.

Keywords: Cognitive Modelling, Decision Making, Information Theory

Introduction

In the context of decision making, a representation refers to the internal mental state of an agent, encompassing features from the external environment that are relevant to the decision task and the agent's objectives. The mechanisms of this representation formation must depend on the task being performed, as decision makers should seek to efficiently represent task-relevant information while ignoring or abstracting across irrelevant information. One important class of tasks which we study in this paper is that of economic decision-making based on visual stimulus.

Neuroeconomics has sought to further the understanding of the neurological underpinnings of economic decision making, though relatively little work in this area has focused on the mechanisms behind representation formation from visual stimuli. One potential mechanism for modelling this cognitive process is the variational autoencoder (VAE), a method that learns informationally limited representations of input that can be used to form lossy reconstructions (Pu et al., 2016). In this paper, we present an extension of the VAE framework that produces task-relevant representations of economic decision tasks, and predicts human-like decision making.

Traditional variational autoencoder models incorporate an information constraint based on the structure of the neural network that implements them (e.g., limiting the number of nodes in a hidden layer), but the capacity of the autoencoder is not easily controlled. β-VAEs are a variant incorporating an additional parameter that controls this information bottleneck, encouraging the model to learn more informationally compact representations. The novelty of our work lies in the application of β-VAEs onto economic decision making. Our model also differs from related methods in machine learning through the use of a novel loss function that balances stimulus reconstruction error and loss in expected utility. The model predicts at a qualitative level the decision making of an individual who has limited information processing capacity, but is otherwise rational in seeking to maximize expected utility.

The main feature of β-VAEs in limiting the amount of information used to form internal representations shares a connection to other information theoretic methods such as the information bottleneck approach and rate-distortion theory (Burgess et al., 2017). The latter has been used to conceptualize human perception as optimizing task performance subject to a constraint on channel capacity (C. R. Sims, 2016).

Rate-distortion theory has also been used to model generalization of perception as resulting from the encoding of perceptual information in a efficient way that can then be used to generalize over novel experience (C. R. Sims, 2018). This effect has been shown to produce biases in statistical and categorical learning from visual features that mimic effects present in human perception (C. J. Bates, Lerch, Sims, & Jacobs, 2019). The model described in this paper seeks to achieve the same goals as these previous methods while also producing an internal representation of a perceived stimulus and a method of explicitly estimating task-relevant utility based on these representations.

Related Methods

β-VAEs have previously been used to model the formation of task-relevant representations in visual categorization and change detection tasks (C. Bates & Jacobs, 2019). In the work of Bates et al., the model consisted of a β-VAE which formed internal representations of stimulus, and a decision module that completed the task being performed based on these representations. Results from this experimentation demonstrated a categorical bias in reconstruction depending on which task is being modelled. This supports the use of the β-VAE framework in modelling the formation of task-relevant visual stimulus representations.

The model and experimentation presented in this paper is an extension of this method onto the domain of economic decision modelling from visual stimulus, which requires adjustments in model structure and training. This is an important extension, as utility is the basis of economic decision making in cases where agents have access to utilities and probabilities required to determine optimal actions. Utility can also serve as the basis for models of human reinforcement learning used to make predictions of human decisions in learning tasks (Niv, 2009; Niv et al., 2015; Collins & Frank, 2012). The ability of extending the proposed model into the domain of reinforcement learning modelling will be further investigated in the discussion section.

The function and motivation behind β-VAE models shares a close connection with the information bottleneck approach (Burgess et al., 2017; Alemi, Fischer, Dillon, & Murphy, 2017). This method has been applied to modelling cognitive mechanisms that share similarities with economic decision making, such as predictive inference (Still, 2014) and information-constrained behaviour (Lai & Gershman, 2021; Malloy & Sims, 2020). One key feature of our proposed model is that it makes predictions on the formation of task-relevant representations under information constraints. Previous methods applying the information bottleneck approach to decision modelling have either not included representation formation, or done so in a task that did not involve utility predictions.

Within the field of economic decision modelling, sub-optimality in human decision making is understood within the frameworks of bounded rationality (Simon, 1990; Camerer, 1998) and rational inattention (C. A. Sims, 2003; Mackowiak, Matejka, Wiederholt, et al., 2020). Models developed under these frameworks can be used to predict how humans make decisions relative to their information processing limitations. As with previously discussed methods, these too do not explicitly model the formation of task-relevant representations of stimuli.

In this paper we present experimentation utilizing our proposed model resulting in similar predictions of sub-optimality in decision making as these related methods, while additionally modelling visual representation formation. As with all cognitive models based in neural networks, these representations are a metaphor for the contents of human cognition. However, through its novel structure and training method our model makes implications for how constrained representation formation can lead to sub-optimal performance.

Modeling Representations using β-VAEs

β-Variational Autoencoders

Variational autoencoders consist of a neural network which compresses an input into a lower dimensional representation that is then expanded back into a reconstruction of the input. The first half of this network structure is referred to as the encoder, while the second half is referred to as the decoder. Both portions of the network are trained simultaneously as the network takes in some input and produces an output, and through training learns to reconstruct the input as faithfully as possible.

The closeness of this reconstruction is defined by a *loss function* which determines how similar the reconstruction is to the original input. Typically in VAEs the loss function is an error between the model input and output, such as the mean-squared-error.

In β-VAEs, an additional parameter (β) is introduced to control the information capacity of the lower dimensional representation, which results in an adjustable information bottleneck (Higgins et al., 2017; Mathieu, Rainforth, Siddharth, & Teh, 2019). The loss function used to train a β-VAE is as follows:

$$\mathcal{L}(\theta,\phi;x,z,\beta) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - \beta D_{KL}\big(q_\phi(z|x)||p(z)\big) \tag{1}$$

In the above, ϕ represents the parameters of the encoder $q_\phi(z|x)$, which defines the probability distribution over latent representations z given the stimulus x. Additionally, $p_\theta(x|z)$ can be understood analogously with the decoder, as it defines the probability that a stimuli x can be produced from the latent representation z. The desired decoder is one where $p(x|z) \approx p(x|v,w)$ where v are the conditionally independent generative factors responsible for producing the stimulus, and w are the conditionally dependent factors. When $p(x|z) \approx p(x|v,w)$ the latent representation z is an adequate representation of the generative factors responsible for producing the stimuli x, as the probability of observing the original data given the latent representation is maximized. For a more complete description, see (Higgins et al., 2017).

The first term in this loss function represents the reconstruction error between the model input and output. The second term $D_{KL}\big(q_\phi(z|x)||p(z)\big)$ represents the informational complexity of the internal representations that the model generates. When $\beta = 0$ the model seeks only to minimize reconstruction error, and as β increases the amount of information used in internal representations decreases. The loss-function used in VAEs corresponds to $\beta = 1$ and as β increases a more constrained information bottleneck is applied.

The usefulness of the β-VAE method in modeling human decision making over the traditional VAE approach is in its ability to adjust the information constraint on the latent representation. For modeling human decision making, it should be possible to determine an individual participant's information processing constraint and fit the β parameter to match that. This suggests that the proposed β-VAE model might better capture the way that individuals form representations of decision making tasks given their individual capacity for storing and processing information.

In our proposed model, the β-VAE represents a Working Memory Module (WMM) of an agent when they are making a decision based on visual stimulus. However, because stimulus representations should be domain specific due to information processing constraints, we must train this model

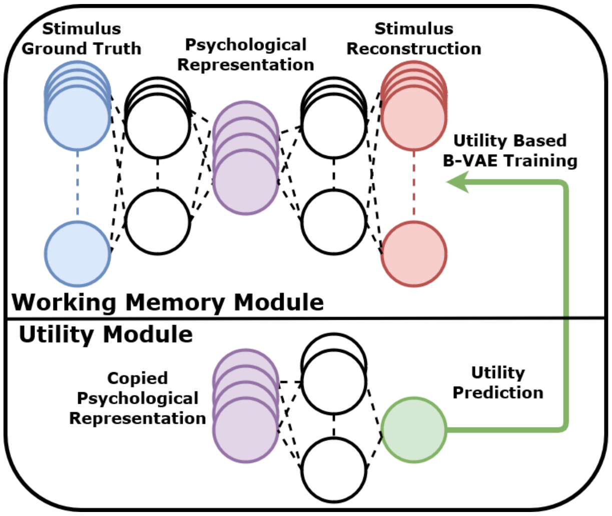

Figure 1: The Working Memory Module is a β-VAE structure that learns to reconstruct the stimulus ground truth and Utility Module prediction accuracy. Colors highlight the stimulus ground truth presented to a decision maker (blue), the internal representation that they use as working memory of the decision making problem (purple), the reconstructed stimulus (red), and the predicted utility of the stimulus (green).

to not only make accurate reconstructions, but also to allow for accurate utility predictions. This is done through the addition of a utility prediction module that allows for utility based training of the WMM.

Utility Based Training

As we are interested in modeling task-relevant representation formation and decision making, the training method of the proposed β-VAE model is adjusted to incorporate the utility of the learned representation. An additional Utility Module learns to predict the utility associated with a stimulus based on the internal representation of a stimulus that is learned by the WMM (Fig 1).

This utility module consists of a neural network that takes as input a copy of the psychological representation of a stimulus, and outputs a prediction of the utility associated with that stimulus. The network is fully connected with 2 layers of 64 units, and the output is trained based on a mean squared error loss between the prediction and the ground truth utility. This utility module is trained alongside the WMM on the same data, with the additional ground truth utilities. The utility predicted estimate is then fed into the loss function of the WMM which is trained to balance the accuracy of the stimulus reconstruction and the utility prediction as follows:

$$\mathcal{L}(r,S) = \mathcal{L}(\theta,\phi;x,z,\beta) + \upsilon\big(r(Z) - \mathbb{E}[S]\big)^2 \quad (2)$$

Where $r(Z)$ is the utility prediction output of the Utility Module, and $\mathbb{E}[S]$ is the ground truth expected utility associated with the stimulus input to the WMM. Using this altered training method, the WMM learns to reconstruct the stimulus accurately, while reducing the squared error between the utility prediction and the stimulus ground truth utility. This structure is similar to the β-VAE based method described in (C. Bates & Jacobs, 2019), though our model explicitly predicts expected utility associated with a stimulus and uses that prediction within the WMM learning objective. These alterations allow for both predictions of utility, as well as working memory representation formation to take utility into account.

Adjusting the utility-loss weight parameter υ controls the relevance of the expected utility in calculating the loss of the model's reconstruction. For example, when the υ parameter is 0, the model learns to reconstruct stimulus as faithfully as possible without accounting for the accuracy of the utility module prediction. As υ increases, the model learns to prefer representations that allow for more accurate utility predictions. Comparing the predicted decisions of the β-VAE model with human selections in economic decision making tasks can allow for a better understanding of how humans balance task-relevance and memory reconstruction accuracy to the original stimulus when forming representations.

Economic Decision Making

Maximum Expected Utility

Expected utility is defined for a decision alternative $\mathbf{x}$ based on the different outcomes that can occur as a result of selecting that alternative $[x_1, x_2, \ldots, x_i]$, the utility of those outcomes $[u(x_1), \ldots, u(x_i)]$ and the probability of those outcomes occurring $[p(x_1), \ldots, p(x_i)]$ given the option that was selected by the agent. This results in the following equation for expected utility:

$$\mathbb{E}[u(\mathbf{x})] = \sum_{i=1}^{n} p(x_i)\, u(x_i) \quad (3)$$

The proposed model takes as input a single option within in a decision problem X, corresponding to outcomes and outcome probabilities, and reconstructs that input as faithfully as possible given the information constraint. However some features of the stimulus are more relevant for maximizing utility than others, so the utility-loss method is introduced to incorporate the difference between the predicted utility and the ground truth.

The utility value of the original stimulus $\mathbb{E}[S]$ in the utility-based loss function in Eq.2 is equal to the true expected utility of the original stimulus. This differs from the expected utility of the reconstruction stimulus $r(Z)$ due to the information constraint applied to the internal representation of stimuli. The calculation of these ground-truth utilities are specific to the task being performed which will be fully detailed in the experimentation section.

Sub-optimal Decision Making

A well-studied form of sub-optimality in human economic decision making is risk-aversion, which is characterized by the undervaluing of risky prospects and overvaluing of safe prospects, relative to their true expected utility (Pratt, 1978; Holt & Laury, 2002). Traditionally, this phenomenon is accounted for by introducing an adjusted utility function that treats outcomes differently based on their value or probability (Rabin & Thaler, 2001). An example of this approach is Cumulative Prospect Theory (CPT) (Kahneman & Tversky, 1979) which has been used to model risk-aversion like effects in economic decision making (Schmidt & Zank, 2008). CPT can account for the effect of risk-aversion by weighting the utility of an outcome based on its value or probability, such as reducing the weight for outcomes that are unlikely and increasing it for likely outcomes (Schmidt & Zank, 2008).

In the following section on experimentation, we will demonstrate that our proposed model makes similar predictions of risk-averse behaviour in an economic decision making task. Importantly, the input to this model will be a visual stimulus representation of a decision making task. This makes it unique from related methods like CPT which take in as input the probabilities and outcomes associated with different options in a decision making task. Additionally, the β-VAE module within our model allows for the formation of psychological representations and stimuli reconstructions of these input that are not present in previous models of risk-aversion in economic decision making.

Experimentation

Previous methods have shown that β-VAEs can be used to produce task-relevant biases in representation formation similar to what can be expected from humans based on their behaviour (C. Bates & Jacobs, 2019). Through experimentation, we show similarly human-like behaviour when modelling visual stimulus representation formation in a utility-based economic decision making task. This is done by showing a risk-aversion effect present in the utility predictions of our model that correspond to expectations of human decision making in similar tasks. For all β-VAE models described in this paper, the model structure, hyper-parameters, and training procedure follow the original implementation described in (Higgins et al., 2017), apart from the β information constraint which is adjusted to compare different information processing constraints as described in the following sections.

Decision Making Task

We examine the behavior of our model in a "marble jar" selection task. In this task, the agent is presented with a choice between two jars of 16 marbles, where the contents of each jar are fully visible. After selecting a jar, one marble is randomly sampled from the chosen jar, and the agent receives a reward based on the color of the (randomly) selected marble.

To compare the impact of information constraints and utility-weight parameters on choice behavior, we vary these parameters and report the utility prediction and reconstruction accuracy of models at the end of training. All models are trained on 1K epochs of 1K stimuli and utility values. Marble jar stimuli are generated using a Dirichlet distribution $Dir([2,4,6,8])$ for grey, green, red and blue colored marbles, which have utilities 8, 4, 2, 1 respectively.

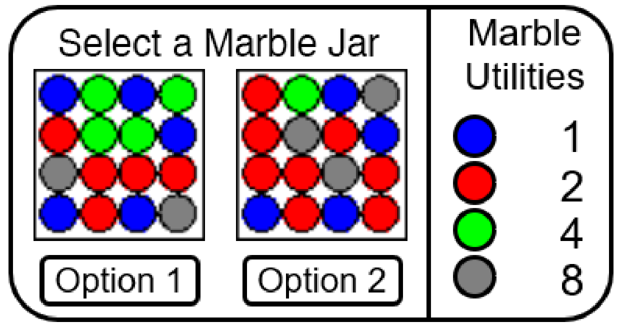

Figure 2: An example of the marble jar selection task. The decision maker chooses one of the marble jars and a single marble chosen at random will be given to the agent with the goal of maximizing their observed utility. Each color marble has a different utility and each marble jar has an expected utility and utility variance which are the mean and variance of the marble utilities. Marble ratios are defined by a Dirichlet distribution $Dir([2,4,6,8])$ for grey, green, red, blue.

The task described in Figure 2 will be used to demonstrate how our model predicts human-like decision making, specifically risk-aversion, through the use of our β-VAE model trained with the utility-based learning objective.

Modelling Results

Using the example decision task shown in Figure 2 we can see how risk-aversion could be demonstrated by a decision maker. The first and second options have total utilities of 46 and 48 and variances of 5.26 and 6.66 respectively. While the second option has a slightly higher expected utility, it also has a higher variance which may impact the choice of the decision maker. In this example, a bias in choosing option 1 over option 2 would reflect an instance of risk-aversion, as the decision maker is preferring certainty in outcome over a increase in expected utility.

risk-aversion: We can use this decision making task to investigate how the utility weight parameter υ and information constraint parameter β impact the decisions of our model. We additionally include for comparison MEU and CPT calculations for these probabilities. Because of the flexibility of CPT, a wide range of possible values for predicted utility are possible, and these values are selected to reflect the risk-averse effect observed in human decision making.

Model	Utility Predictions	Recon. Error
MEU (normative)	(**46**, **48**)	N/A
β-VAE + Utility	(48.2, 42.1)	1611.4736
β-VAE	(45.6, 47.8)	1089.0422
VAE + Utility	(46.6, 47.8)	2810.2334
VAE	(41.3, 44.8)	1803.3821
CPT	(47, 45)	N/A

The results shown in the table above indicate that the β-

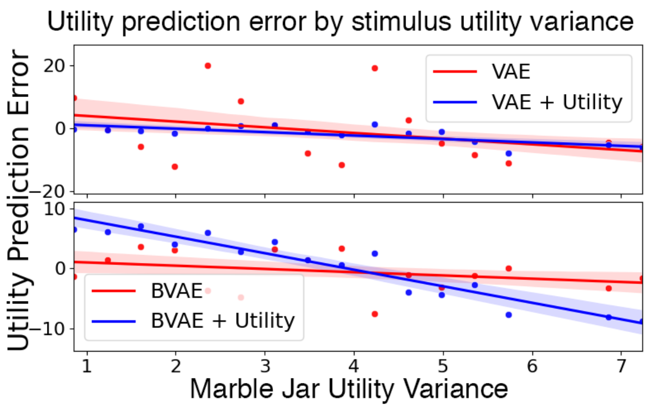

Figure 3: Utility prediction error based on marble jar utility variance of an ablation of β-VAE and VAE models with and without utility based training. Points represent mean utility predictions of all marble jars with the same utility variance. Lines represent a linear regression of all predicted utilities, calculated with the Seaborn Python library (Waskom, 2021). VAE models have $\beta = 1$, β-VAE models have $\beta = 100$, utility models have $\upsilon = 1000$ and non-utility models have $\upsilon = 0$.

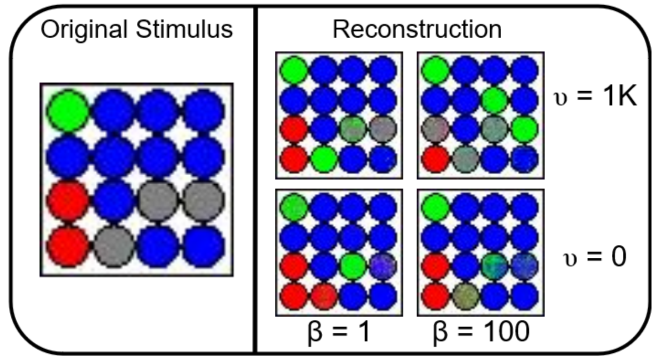

Figure 4: Comparisons of original stimulus and reconstruction for an example marble jar with the same ablation of different model types as described in Figure 3. Marble jar original stimulus is from the constructed set of marble jars with utility 42 that were not part of the original training data set.

VAE + Utility model demonstrates a risk-aversion effect for the example stimulus in Figure 2. Although the second jar has higher expected utility, the model values the first jar more highly. This effect can also be observed using a cumulative prospect theory model, as it has decision weight parameters that can be adjusted to produce a similar effect. However, it is important to note that the CPT model would need to fit an individual parameter for each possible outcome, whereas our proposed model is parameterized only with the utility-weight and information constraint. Additionally, CPT functions by altering the utility maximization method, whereas our approach assumes decision makers maximize utility, but doing so with limited information processing ability.

Comparing the reconstruction errors for each model demonstrates the improved generalizability to out of training stimuli afforded by the β-VAE models, which is one main justification for their use (Burgess et al., 2017). Additionally, each model trained with the utility prediction method has a lowered reconstruction accuracy. This corresponds with the expectation of a model with limited information processing, as information used to represent utility can lower the amount of information available to accurately reconstruct the original stimulus. An interesting result is that the loss-aversion like effect is observed as a result of the differences between differently colored marble proportions and utilities, as opposed to an imposed preference as is the case with the CPT model.

Note that the difference in predicted utility is exaggerated from what a human would likely determine for this task, as the information constraint ($\beta = 100$) and utility weight ($\upsilon = 1000$) are larger than values that would better reflect human behaviour. In practice, it is possible to fit these parameters based on observations from individuals, as has been done in related methods (Malloy & Sims, 2020; Niv et al., 2015; Collins & Frank, 2012). This could result in utility predictions matching the behaviour of individual participants, though this is outside of the scope of the present work.

Utility Estimate Bias: In order to understand the basis for the risk-aversion effect, we examined how the utility training method and information constraint impact utility predictions as the risk associated with a stimulus changes. Additionally, we sought to examine the generalizability of our utility prediction module to stimuli that have not been seen previously. To allow for this, we constructed a new data set consisting of every possible marble jar with the same total utility (42) but with different variances. The following figure compares the bias in predicted utility for models with and without an information constraint and utility-based training.

These results show that the β-VAE model alone demonstrates a risk-aversion like effect, shown by a positive utility error for low variance marble jars, and a negative utility error for higher variance jars. This corresponds with our understanding of the risk-aversion effect which increases the distortion of preference as outcomes become more or less probable.

Generalization: These utility prediction error results additionally demonstrate the high generalizability of utility prediction in models trained using the utility-based learning objective (blue). Models trained without utility included in their learning objective (red) have a wider range of prediction errors above and below the regression trend. This reflects the similarly high generalizability of human economic decision making in these types of utility-based tasks.

The final comparison of different model types is in the reconstruction accuracy of a new stimulus not used on training. Figure 4 shows a stimulus not used during training. The right hand side compares stimulus reconstructions of the same 4 model ablation described previously, with and without an information constraint and utility-based training.

These reconstruction examples demonstrate the impact that information constraints and utility-based training have on generalized stimulus reconstruction. Models trained with the utility-based learning objective are better able to reconstruct the higher utility grey marbles. Interestingly, these grey marbles are not necessarily in the same location, as this is not relevant to the expected utility of a marble jar. The model thus demonstrates that its latent representations have acquired a useful invariance (the exact position of a marble in the jar). This corresponds to intuition from human perceptual memory in this type of task, as the location of marbles is irrelevant to predicting utility.

Discussion

Risk-Averse Representations

Results from utility prediction errors of our proposed model demonstrated a risk-aversion like effect. These results suggest that one aspect of risk-aversion is the formation of informationally compressed representations of visual stimulus in economic decision making tasks. However, alone these results do not fully explain the source of risk-averse representation formation. This can be better understood by considering the relative abundance of marbles and their utilities.

Marble piles with more grey marbles have higher variance, and one possible interpretation of underestimating these utilities would be a decision maker determining the utility of a stimuli with grey marbles by counting only the utility of those grey marbles and nothing else, leading to an under estimate. However, this is one possible explanation and additional comparisons would need to be made to more fully understand the precise ways in which these utility estimates are risk-averse. Generally the risk-averse behaviour should be understood as resulting from different probabilities and utilities of marbles, and how the information constraints and utility-weights impact representations of these stimuli.

An important implication of this explanation is that it would be unlikely to observe the same risk-averse behaviour in an alternate version of the marble task that is very uniform in marble probability or utility. While this is a slight weakness to our proposed model, these types of stimuli would also likely result in only a slight risk-averse behaviour in humans, since marble piles would on average have utilities much closer together. Additionally, the proposed model seeks to account for one source of risk-aversion, though others are likely to exist. It is possible that the type of risk-averse behaviour that results from forming informationally-compressed and utility-based representations only occurs when there is a considerable difference between stimuli utility.

Human Representation Formation

The modeling experiments presented in this paper sought to examine the properties of stimulus representations learned when facing constraints on the ability to encode and represent task features. To do this, we represented visual decision making in a similar manner as previous approaches, with a β variational autoencoder trained to learn internal representations, and a separate module trained to perform a task based on those representations. The novelty of our method is in its learning to explicitly predict utility based on task-relevant visual representations in an economic decision task.

Our results demonstrate that when agents face constraints on the ability to encode information veridically, systematic distortions are introduced in the representation of the probability and utility of decision alternatives. In particular, stimuli with a higher utility variance have a lower predicted utility, with the opposite being true for stimuli with a low utility variance. This corresponds to observations of human behaviour in economic decision making tasks. Importantly, our model makes similar predictions as existing methods while taking the input to be the visual task stimuli, and producing a psychological representation that can be used to reconstruct the original stimulus and make utility predictions.

Modelling Human Learning

As mentioned previously, the inclusion of utility predictions and a utility-based training method within our proposed model can allow for the modelling of reinforcement learning in humans. This can be done by adjusting the training method of the utility prediction module. In our experiments, decision makers were assumed to have knowledge of outcome probabilities and utilities. In the learning setting, these values would not be known and instead learned by making decisions and observing outcomes. Thus our utility prediction model would make a prediction and observe an outcome after the decision has been made, and update their prediction based on this observed outcome. This can be done using the standard temporal difference equation used in reinforcement learning, which is motivated by human biological processes implicit in learning (Niv, 2009; Niv et al., 2015).

References

Alemi, A. A., Fischer, I., Dillon, J. V., & Murphy, K. (2017). Deep variational information bottleneck. *International Conference on Learning Representations*.

Bates, C., & Jacobs, R. (2019). Efficient data compression leads to categorical bias in perception and perceptual memory. In *Cogsci* (pp. 1369–1375).

Bates, C. J., Lerch, R. A., Sims, C. R., & Jacobs, R. A. (2019). Adaptive allocation of human visual working memory capacity during statistical and categorical learning. *Journal of vision*, *19*(2), 11–11.

Burgess, C. P., Higgins, I., Pal, A., Matthey, L., Watters, N., Desjardins, G., & Lerchner, A. (2017). Understanding disentangling in β-vae. *NIPS Workshop on Learning Disentangled Representations*.

Camerer, C. (1998). Bounded rationality in individual decision making. *Experimental economics*, *1*(2), 163–183.

Collins, A. G., & Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? a behavioral, computational, and neurogenetic analysis. *European Journal of Neuroscience*, *35*(7), 1024–1035.

Higgins, I., Matthey, L., Pal, A., Burgess, C., Glorot, X., Botvinick, M., … Lerchner, A. (2017). beta-vae: Learning basic visual concepts with a constrained variational framework. In *International conference on learning representations.*

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, *92*(5), 1644–1655.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Lai, L., & Gershman, S. J. (2021). Policy compression: An information bottleneck in action selection. Academic Press. doi: https://doi.org/10.1016/bs.plm.2021.02.004

Mackowiak, B., Matejka, F., Wiederholt, M., et al. (2020). Rational inattention: A review. *CEPR Discussion Papers*(15408).

Malloy, T., & Sims, C. R. (2020). Modelling human information processing limitations in learning tasks with reinforcement learning. *Proceedings of the international conference on cognitive modelling*, *18*, 159–160.

Mathieu, E., Rainforth, T., Siddharth, N., & Teh, Y. W. (2019). Disentangling disentanglement in variational autoencoders. In *International conference on machine learning* (pp. 4402–4412).

Niv, Y. (2009). Reinforcement learning in the brain. *Journal of Mathematical Psychology*, *53*(3), 139–154.

Niv, Y., Daniel, R., Geana, A., Gershman, S. J., Leong, Y. C., Radulescu, A., & Wilson, R. C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. *Journal of Neuroscience*, *35*(21), 8145–8157.

Pratt, J. W. (1978). Risk aversion in the small and in the large. In *Uncertainty in economics* (pp. 59–79). Elsevier.

Pu, Y., Gan, Z., Henao, R., Yuan, X., Li, C., Stevens, A., & Carin, L. (2016). Variational autoencoder for deep learning of images, labels and captions. In *Advances in neural information processing systems.*

Rabin, M., & Thaler, R. H. (2001). Anomalies: risk aversion. *Journal of Economic perspectives*, *15*(1), 219–232.

Schmidt, U., & Zank, H. (2008). Risk aversion in cumulative prospect theory. *Management Science*, *54*(1), 208–216.

Simon, H. A. (1990). Bounded rationality. In *Utility and probability* (pp. 15–18). Springer.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, *50*(3), 665–690.

Sims, C. R. (2016). Rate–distortion theory and human perception. *Cognition*, *152*, 181–198.

Sims, C. R. (2018). Efficient coding explains the universal law of generalization in human perception. *Science*, *360*(6389), 652–656.

Still, S. (2014). Information bottleneck approach to predictive inference. *Entropy*, *16*(2), 968–989.

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021.