Towards agents with human-like decisions under uncertainty [1]

Nuno Marques (nunocm@gmail.com)
Francisco Melo (fmelo@inesc-id.pt)
Samuel Mascarenhas (samuel.fm@gmail.com)
João Dias (joao.dias@gaips.inesc-id.pt)
Rui Prada (rui.prada@gaips.inesc-id.pt)
Ana Paiva (ana.paiva@inesc-id.pt)
INESC-ID, Instituto Superior Técnico
Av. Prof. Cavaco Silva, TagusPark
2780-990 Porto Salvo, Portugal

Abstract

Creating autonomous virtual agents capable of exhibiting human-like behaviour under uncertainty is becoming increasingly relevant, for instance in multi-agent based simulations (MABS), used to validate social theories, and also as intelligent characters in virtual training environments (VTEs). The agents in these systems should not act optimally; instead, they should display intrinsic human limitations and make judgement errors. We propose a Belief-Desire-Intention (BDI) based model which allows for the emergence of uncertainty related biases during the agent's deliberation process. To achieve it, a probability of success is calculated from the agent's beliefs and attributed to each available *intention*. These probabilities are then combined with the intention's utility using Prospect Theory, a widely validated descriptive model of human decision. We also distinguish risk from ambiguity, and allow for individual variability in attitudes towards these two types of uncertainty through the specification of indices. In a travelling scenario, we demonstrate how distinct, more realistic agent behaviours can be obtained by applying the proposed model.

Keywords: Intelligent agents; Decision making; Cognitive biases

Introduction

Uncertainty is a natural part of our world. No one can claim to know everything, no one can predict the future. We deal with uncertainty on our everyday lives and our behaviour is constantly influenced by it, even if we do not always realize it. However, in the context of virtual agents, uncertainty has usually been seen as a problem that the agent must overcome (eg. planning Peot & Smith, 1992), and thus most existing systems are aimed at achieving *optimal* agent behaviour under these conditions.

Our approach is different, in which we acknowledge the often sub-optimal, even "irrational" behaviour of humans when confronted with uncertain situations. These decision biases and judgement errors have been extensively studied and are supported by a wealth of empirical evidence (eg. Kahneman & Tversky, 1979; Camerer & Ho, 1994). We propose an agent model based on the classical Belief-Desire-Intention (BDI) paradigm, which seeks to integrate in the agent's deliberation process these deviations from rational behaviour.

Agents with the aforementioned characteristics can be specially useful for Multi-Agent Based Simulations (MABS) (Davidsson, 2001). In these systems, human behaviour is modelled at the individual (agent) level, and the resulting structure is analysed after it emerges from the agent interactions. Typically, MABS have been used to validate social theories (eg. Davidsson, 2002). The inclusion of uncertainty is of special importance in market simulations, as it strongly impacts the decisions of the agent (Arthur, 1991). From socio-cultural research, the *Uncertainty Avoidance* dimension of human cultures, identified by Hofstede (Hofstede, 2001), is another example where these agents could be used in the context of MABS. Our solution is also relevant for use in *serious games*, particularly *virtual training environments*. As these simulation often focus on social and communication aspects (eg. Johnson & Valente, 2009; Kim et al., 2009), it is increasingly important to embed the virtual characters with human-like behaviour.

This paper is organized as follows. We start by giving a possible definition of uncertainty and describing Prospect Theory, and follow with work related to ours. Then we present the model, and demonstrate it using an example scenario. Finally we discuss future improvements.

Background

In tackling the effects of *uncertainty*, one should first have an accurate definition of the term. However, this is not an easy task because different research fields or problem approaches use it with different meanings.

One important step is distinguishing *uncertainty* from the closely related concept of *risk*. In a decision context, the later refers to choices involving *known* chances (eg. a spin of a roulette wheel). However, *uncertainty* arises in a decisions involving *personal opinions* (eg. betting on what football team will win a game). Moreover, uncertainty has distinct facets (Smithson, 2008): *epistemic randomness* or *risk uncertainty* is the subjective counterpart of risk, and is usually represented by subjective probabilities; *ambiguity*, which results from overlapping beliefs (i.e, strong reasons to *believe* and *not believe*) or uncertainty about probabilities (second order uncertainty); and *vagueness*, reflected by fuzzy statements (eg. "John is tall" — what does "tall" mean?).

The topic of how humans choose (or should choose) under

[1]This work was partially supported by the Portuguese Fundação para a Ciência e Tecnologia under project PEst-OE/EEI/LA0021/2011. (INESC-ID multiannual funding) through the PIDDAC Program funds. Additionally, it was funded by the National Project SEMIRA (ERA-Compl/0002/2009), and European Project eCute (ICT-5-4.2 257666) projects.

uncertainty has been extensively studied over the last centuries. Decision making theories which seek to predict the optimal choice, such as the classical Expected Utility theory (EU), are called *normative*. However, people do not generally obey the axioms of normative theories (some examples of violations are described in the following section). Given our goal of achieving human-like behaviour, we focus on theories seeking to describe how humans *actually* act. Within these, decision behaviour has been observed to differ when the subject is offered a description of available choices (*decisions from description* paradigm), versus when he can learn by direct experimentation (*decisions from experience* paradigm), as shown by Hertwig, Barron, Weber, & Erev, 2004. As we will see, the solution proposed in this paper assumes that the agent learns by asking and not by experimentation, and therefore we restrict ourselves to the former category.

Prospect theory

The most validated *descriptive* theory of human decision is called Prospect Theory (PT) (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). Some of the decision biases it addresses are:

Framing Effects: there is evidence that the framing of options (in terms of gains or losses) significantly impacts the choices people make (Tversky & Kahneman, 1986);

Nonlinear preferences: the idea that a risky prospect is linear in outcome probabilities has been proved false, most prominently by the Allais paradox (Allais, 1979);

Source dependence: as demonstrated by the Ellsberg paradox, people's decisions depend not only on the degree of uncertainty but also on its source; this phenomena has been explained both from an *ambiguity aversion* (people dislike ambiguity, Ellsberg, 1961) and from a *competence hypothesis* perspective (people prefer a bet on their area of competence when compared to equivalent bet based on objective probabilities, Heath & Tversky, 1991);

Fourfold pattern of risk: empirical studies indicate that people are generally risk averse for high probability gains and low probability losses, and risk seeking for low probability gains and high probability losses (Tversky & Kahneman, 1992).

These biases are accounted by PT by assuming a *framing phase* prior to the actual evaluation; a *value function* (v) which distort utilities; and a *weighting function* (π) which distorts probabilities. Our focus is on the biases created by these functions, and how to integrate them in the BDI model, as the modelling of framing effects has already been explored in a context similar to ours (Ito & Marsella, 2011). In PT, the valuation attributed to a prospect (i.e, a gamble) f, which has n possible outcomes $x_i, i = 1...n$, each with utility U_i and probability p_i, is given by:

$$V(f) = \sum_i v(U_i)\pi(p_i)$$

The choice for the specific value and weighting functions is arbitrary, as long as they obey certain properties. The value function should reflect the effects of *diminishing sensitivity* (variations in utility are less perceived the further they are from the reference point), and thus be concave for gains and convex for losses (S-shape). Furthermore, it should be steeper for losses than for gains, reflecting the phenomena of *loss aversion*.

The weighting function transforms a probability, and should also reflect the effects of diminishing sensitivity. However, in this case there are two boundaries ($p = 0$ and $p = 1$), and thus the resulting function is inverse S-shaped. The curvature of these functions reflect the individual propensity to decision biases, which is usually accounted for by assuming parametrized functional forms.

Our integration of PT in the BDI model, as shown later, is restricted to choices involving prospects with at most two outcomes. Therefore, both Prospect Theory and its more recent development, Cumulative Prospect Theory (CPT) (Tversky & Kahneman, 1992), coincide. We are presenting the original formulation of the theory. It is also important to note that, although PT is originally based on studies where probabilities were objectively stated, and thus related to decisions under *risk*, its fundamental properties were also verified in decisions under *risk uncertainty* (Tversky & Fox, 1995).

Related Work

In Pezzulo's et. al. proposal, measures of *ignorance* (what the agent does not know), *contradiction* and *uncertainty* (difference between opposing beliefs) are computed by the agent, and used in the decision process using custom rules (Pezzulo, Lorini, & Calvi, 2004).

FAtiMA-PSI (Dias & Paiva, 2005; Mascarenhas, Dias, Prada, & Paiva, 2010) is an architecture geared towards the creation of believable virtual characters. It has a strong focus on emotional aspects and human motivations. It already represents some forms of uncertainty, as stochastic action outcomes and estimations of goal success based on past observations. This architecture also provides several parameters which allow an author to define agents with different personalities. However, it does not model unreliable perceptions - the environment is considered completely observable, and the decision process is based on EU theory.

The graded BDI model and abstract architecture (g-BDI, Casali, Godo, & Sierra, 2009) extends the classical BDI model by allowing uncertainty to be represented in the agent's mental attitudes. g-BDI permits not only uncertain (graded) beliefs, but also desires and intentions. Graded desires correspond to degrees of preference (or rejection) over states of the world, and graded intentions represent the preference over specific ways (plans) to achieve desires. The formalization of the g-BDI allow different *contexts* to operate each in its own logic. Thus, the belief context (BC), for example, can use probability measures to represent uncertainty.

The Contextually-Based Utility (CBU) model (Ito &

Marsella, 2011) combines principles of cognitive appraisal theories with decision theoretic notions, with the main purpose of capturing framing effects with greater accuracy. For each possible goal outcome, a contextual utility value is calculated using two salient features: *pleasantness*, the outcome's intrinsic attractiveness or unattractiveness; and *congruence*, how much achieving the goal contributes to the agent's expectations. A decision weight is also computed using the outcomes' probabilities. These three measures are transformed by an S-shaped function which models the effects of diminishing sensitivity in relation to a variable reference point, and are then linearly combined to obtain a goal's final valuation.

The work presented above share with ours the purpose of achieving human-like agent behaviours in decisions under uncertainty. However, almost none of them apply a validated descriptive decision theory. The exception is CBU, which is not an agent model in itself. Thus, our approach differs in what we consider fundamental requisites of our solution: 1) capturing widely validated findings on human decision behaviour; and 2) being a generally applicable agent model.

Model

We chose the BDI model because its folk psychology roots are consistent with our goal of modelling human-like behaviours, and also due to its flexibility and wide application. An overview of the proposed model is shown in figure 1. In the present section, each component is explained in detail.

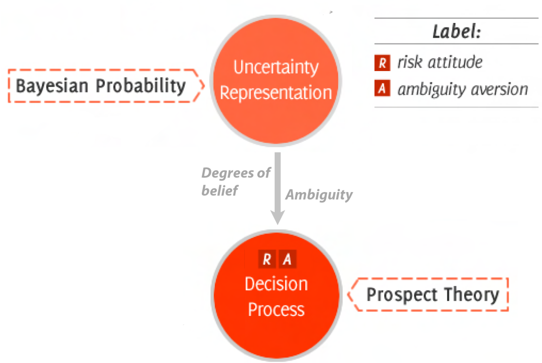

Figure 1: Model overview

Uncertainty representation

This component deals on how to represent uncertainty in the agent beliefs. We used bayesian probability, as it is the most developed model of uncertainty representation. We assume a world defined by a set of crisp (true/false) *propositions*, represented by upper-case letters, for example O="Hotel is open". Exhaustive and mutually exclusive subsets of propositions are called *variables*. A probability distribution over the values of each variable forms the agent's belief state, i.e, his opinions on the actual value of propositions. Thus, if $\Theta = \{A_1, \dots, A_n\}$ is a variable, then $\sum_i P(A_i) = 1$, where $P(A_i)$ denotes the subjective probability of A_i being *true*.

In order to focus on the behavioural consequences of uncertainty, we made two simplifying assumptions: 1) all variables are conditionally independent; and 2) the agent can only be uncertain about *static* propositions (propositions whose value never change, because no *actions* exist with such effects). We expect to address these limitations in future work.

Belief revision The agent can change its opinions by making either a *direct observation* or by asking questions to other agents. In the latter case, the degree at which the agent believes what he is told depends on the evidence's *credibility*. We represent the credibility of an evidence provided by source i, asserting a proposition A, as $cr(\varepsilon_i^A) \in [0,1[$. The value cr is calculated based on the history of previously received answers from the same source. We follow the method proposed in (Pearl, 1988), which allows Bayes Rule to be applied to uncertain evidence. Assuming independent sources, the agent's beliefs are updated using the formula below:

$$P(H|\varepsilon_i^A) = \begin{cases} (cr(\varepsilon_i^A) + \frac{1-cr(\varepsilon_i^A)}{n}) \cdot \frac{P(H)}{P(\varepsilon_i^A)} & if H = A \\ \frac{1-cr(\varepsilon_i^A)}{n} \cdot \frac{P(H)}{P(\varepsilon_i^A)} & otherwise \end{cases}$$

where $P(H)$ is the prior probability of value H in a variable. When new evidence comes which asserts A, the above formula increases the belief in A while decreasing the belief in the other values of the same variable, such that they still sum to one. Note that a direct observation corresponds to $cr(\varepsilon_i^A) = 1$, and thus always leads to absolute certainty on a variable's value.

Solutions and doubts

This section serves as a bridge between the preceding (uncertainty representation) and following (decision process) components, by demonstrating how the agent's beliefs are integrated in the decision process. In the BDI model, the general behaviour of an agent is guided by the active *intention* - a commitment to achieve a goal and a plan to do it. We call each plan generated by the planning process a *solution*. We assume that each solution only contains indispensable actions (as usual in classical planning), and therefore if a single action fails, the solution also fails.

The execution of an action, in turn, is dependent on its Action Pre Conditions (APC) validity. APCs are world propositions or their negation. Thus, success in achieving a goal is ultimately dependent on the validity of APCs and, if among the APCs some correspond to uncertain propositions, they possibly invalidate the entire solution and make the goal impossible to attain. Uncertain APCs are what we call *doubts*, and they are the basic components from which uncertainty related biases will arise (see Figure 2). To distance ourselves from the problem of planning in uncertain environments, during planning doubts are considered valid and thus ignored.

Decision process

Within the deliberative process of the agent, our model deals with the evaluation of competing *solutions*, which essentially

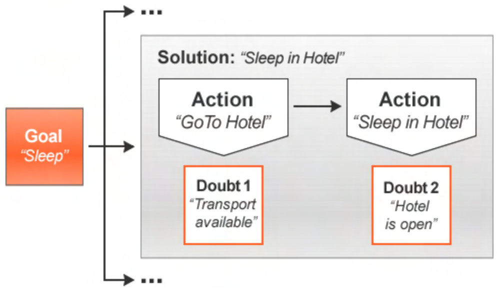

Figure 2: An example of a solution containing two doubts.

correspond to the *intentions* of the agent — possible paths to commit to and pursue. Prospect Theory (PT), the psychology based model described before, is used during this process. This theory has been shown to accurately predict the choice behaviour of humans in many real life situations, and as such it nicely fits our goal — achieving actual, and not optimal behaviour. The prospects (i.e, the solutions) only have two possible outcomes: *success* and *failure*.

Value function We assume a utility of 0 for failing, and a utility of success given by the sum of utility of each individual action within in the solution. Using the *status quo* as the reference point ($v(0) = 0$), the above definition implies that a solution is always a positive prospect. We applied Tversky and Kahneman two branch power function, but because we only have positive outcomes we can restrict to the positive part (with $\alpha = 0.88$ as estimated in Tversky & Kahneman, 1992):

$$v = U^{\alpha}, U \geq 0$$

Weighting function The weighting function $\pi(p)$ transforms the beliefs the agent holds about states of the world into the decision weights he actually utilises when making a decision. The probability assigned to the successful outcome corresponds to no invalid doubts at all (as a single doubt failure invalidates the whole solution). Thus, variables being independent, p is obtained by multiplying the probabilities of all solution doubts being valid. The probability p is then transformed by the weighting function. We use the function proposed by Wu & Gonzalez, 1999:

$$\pi = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}}$$

In the above function, the two parameters reflect two distinct cognitive biases: γ (estimated as $\gamma = 0.44$) represents the impact of *diminishing sensitivity* to probability variations through the function curvature; δ (estimated as $\delta = 0.77$) represents the *attractiveness* to gambles through the function elevation.

Individual variability towards risk uncertainty is allowed via the indirect specification of the δ parameter. Staying within the parameter values estimated by Wu & Gonzalez, 1999, we propose to substitute δ with δ^R, such that:

$$\delta^R = \delta + (0.5 - R)$$

where $R \in [0, 1]$ is a *risk aversion* index available to the scenario author. A value of R close to 1 results in a risk seeking attitude, as the agent will generally overweight probability values; on the other hand, a value close to 0 corresponds to a risk averse attitude, as the probabilities are underweighted.

The resulting functions are shown in figure 3.

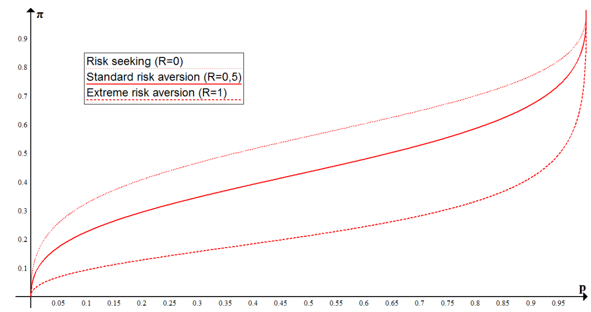

Figure 3: Weighting function at different levels of risk aversion (R).

Ambiguity

Although PT captures risk biases, ambiguity related biases are not considered. Before discussing them, we present how is ambiguity quantified in the proposed model.

Research in psychology and neuroscience proposes *information entropy* of possible meanings one attributes to a situation (or the world) as a mathematical quantification of ambiguity (Takahashi, Oono, Radford, & Others, 2007; Hirsh, Mar, & Peterson, 2012). Information entropy, also known as Shannon Entropy, measures the amount of "disorder", or uncertainty, associated with a random variable. If $\theta = \{A_1, ..., A_n\}$ is a variable, it's entropy is given by (we use e as the base of the logarithm):

$$H(\theta) = -\sum_{i}^{n} p(A_i) \log p(A_i)$$

In our model, the variables correspond the agent's belief state. The ambiguity being experienced by the agent, however, depends only on the goal under pursuit, i.e, the solution under execution. Thus, it arises from solution doubt variables:

$$\Delta(S) = \sum_i H(\theta_i), \theta_i \in \{doubts(S)\}$$

Where $\Delta(S)$ represents the ambiguity of solution S. High levels of ambiguity are caused by ignorance about the actual value of doubts — when the probability distributions are "flat", and/or there are many doubts. Lower levels of ambiguity, on the other hand, are characterized by strong beliefs (probabilities close to either 1 or 0) and/or a fewer number of doubts. Note that the amount of ambiguity is independent on what variable values are preferred, and thus distinct from risk.

Existing research seem to indicate that people are averse to this type of ambiguity, i.e, they tend to choose lower entropy over higher entropy options (Takahashi et al., 2007). In order to represent this effect, we first reduce it to the $[0,1[$ interval:

$$\Delta^* = 1 - e^{-k\Delta}$$

Where $k = 0.2$ is a normalizing factor. Then, we introduce ambiguity in the δ^R parameter defined before, such that:

$$\delta^\Delta = \delta^R(1 - A\Delta^*)$$

Where $A \in [0,1]$ is an *ambiguity aversion index*. Therefore, values of A close to 1 reinforce the aversion to solutions with high ambiguity.

Example scenario

In this section we demonstrate how the proposed model gives origin to different behaviours, by focusing on preference reversals caused by distinct attitudes towards uncertainty. The scenario being presented is deliberately simple, in order to facilitate the presentation.

Consider an author who wishes to create an uncertain travel scenario. In addition to the agents introduced further below, it contains two other entities: the city of Agra (location of the TajMahal) and the region of Ladakh (home to several beautiful Buddhist monasteries), which will be the available travel destinations. When an agent learns about them (eg. at initialization time), the associated properties are stored in the agent's memory as variables. By default, properties are certain, but uncertainty about specific entities or specific properties may be specified through the agent's configuration file.

In the presented test case, all agents will share the exact same knowledge: they are completely ignorant about finding accommodation in Ladakh, but are almost certain to find it in Agra. This is represented by two different probability distributions over the variables originated by the property *accAvailable* of Agra and Ladakh, as shown in table 1. The

Table 1: **Agent's uncertain knowledge**

Agra(accAvailable)		
Values	**Belief**	**Entropy**
True	0.9	0.3251
False	0.1	

Ladakh(accAvailable)		
Values	**Belief**	**Entropy**
True	0.5	0.6931
False	0.5	

atomic actions available to the agents in order to achieve their *Travel()* goal are specified in table 2, which in this case differ only in their utility, as the doubts are equivalent (having a place to sleep). The created solutions are shown in table 3.

Table 2: Actions

Actions	**Utility**	**Doubts**
TravelTo(Agra)	-2	
Visit(Agra)	8	Agra(accAvailable)=True
TravelTo(Ladakh)	-3	
Visit(Ladakh)	13	Ladakh(accAvailable)=True

In the first case, we show the impact of distinct attitudes towards risk. The solutions are evaluated by Peter, an extremely

Table 3: Solutions

Solutions	**Util**	**Prob**	**Amb**
$\boldsymbol{S}_1$ **:** TravelTo(Agra) $\rightarrow$ Visit(Agra)	6	0.9	0.3251
$\boldsymbol{S}_2$ **:** TravelTo(Ladakh) $\rightarrow$ Visit(Ladakh)	10	0.5	0.6931

risk averse and ambiguity tolerant agent ($R = 1$; $A = 0$), and Sara, who is risk seeking and also ambiguity tolerant ($R = 0$; $A = 0$). We start with Peter's evaluation:

$$V(S_1)=v(6)*\pi(0.9)=6^{0.88}*\left(\frac{1.27*0.9^{0.44}}{1.27*0.9^{0.44}+(1-0.9)^{0.44}}\right)=\mathbf{2.01}$$

$$V(S_2)=v(10)*\pi(0.5)=10^{0.88}*\left(\frac{1.27*0.5^{0.44}}{1.27*0.5^{0.44}+(1-0.5)^{0.44}}\right)=1.61$$

And Sara's evaluation is shown below:

$$V(S_1)=v(6)*\pi(0.9)=6^{0.88}*\left(\frac{0.27*0.9^{0.44}}{0.27*0.9^{0.44}+(1-0.9)^{0.44}}\right)=3.72$$

$$V(S_2)=v(10)*\pi(0.5)=10^{0.88}*\left(\frac{0.27*0.9^{0.44}}{0.27*0.5^{0.44}+(1-0.5)^{0.44}}\right)=\mathbf{4.24}$$

As we can see, Peter avoids the higher utility solution due to its greater risk. On the other hand, Sara is not concerned, and visits Ladakh.

In the second test case, we demonstrate the effects of ambiguity. We use two other agents, both average risk averse. However, John is ambiguity tolerant ($R = 0.5; A = 0$) while Laura is ambiguity averse ($R = 0.5; A = 1$). John evaluates the solutions similarly as shown before, resulting in the valuations: $V(S_1) = 3.24$ and $V(S_2) = \mathbf{3.30}$, and therefore goes to Ladakh. In Laura's decision, the parameter δ^A includes the effects of ambiguity as shown below:

$$\Delta_1^*=1-e^{-0.2*0.3251}=0.0630$$

$$\Delta_2^*=1-e^{-0.2*0.6931}=0.1294$$

$$\delta_1^\Delta=0.77*(1-1.0*0.0630)=0.7215$$

$$\delta_2^\Delta=0.77*(1-1.0*0.1294)=0.6703$$

$$V(S_1)=v(6)*\pi(0.9)=6^{0.88}*\left(\frac{0.7215*0.9^{0.44}}{0.7215*0.9^{0.44}+(1-0.9)^{0.44}}\right)=\mathbf{3.17}$$

$$V(S_2)=v(10)*\pi(0.5)=10^{0.88}*\left(\frac{0.6703*0.5^{0.44}}{0.6703*0.5^{0.44}+(1-0.5)^{0.44}}\right)=3.04$$

For Laura, the ambiguity associated with travelling to Ladakh is too great to overcome the potential reward, and she prefers to go to Agra. As it can be seen, the effects of ambiguity are independent of risk, given that the agent's only differ in their ambiguity attitude. It should be noted that, although we fixed the uncertainty levels and varied the agent parametrizations, an equivalent preference reversal can be achieved by varying the ambiguity levels while maintaining risk and the indices.

Conclusion

In this paper, we proposed an agent model which combines the flexibility and generality of the BDI paradigm with a widely validated descriptive decision model, in order to capture decisions biases in the agent's deliberative process. The two test cases presented in the previous section showed how different parametrizations in agent's attitudes towards uncertainty have distinct behavioural consequences. A rational agent using EU, on the other hand, would always select the same solution. Furthermore, by considering risk and ambiguity as separate constructs, we expect to better capture human decisions under uncertainty.

Although the modelled risk biases are well grounded on empirical evidence, the effects of ambiguity require further validation. In particular, the magnitude at which entropy causes an aversion effect in a decision context should be further studied, so that a more accurate expression can be used in the calculation of δ^A. We also expect to address the present limitations by possibly integrating a Bayesian Belief Network (Pearl, 1988), allowing for conditional dependences between variables, and using a planner capable of dealing with partially observable domains (eg. Peot & Smith, 1992), in order to drop the *static* propositions assumption. Finally, we consider accounting for framing effects as an important step in achieving more realistic decision behaviours, and the ideas behind the CBU model (Ito & Marsella, 2011) are certainty worth exploring.

References

Allais, M. (1979). *The so-called allais paradox and rational decisions under uncertainty*. Springer.

Arthur, W. B. (1991). Designing economic agents that act like human agents: A behavioral approach to bounded rationality. *The American Economic Review*, *81*(2), 353–359.

Camerer, C. F., & Ho, T. H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of risk and uncertainty*, *8*(2), 167–196.

Casali, A., Godo, L., & Sierra, C. (2009). g-bdi: A graded intensional agent model for practical reasoning. *Modeling Decisions for Artificial Intelligence*, 5–20.

Davidsson, P. (2001). Multi agent based simulation: beyond social simulation. *Multi-Agent-Based Simulation*, 141–155.

Davidsson, P. (2002). Agent based social simulation: A computer science view. *Journal of artificial societies and social simulation*, *5*(1).

Dias, J., & Paiva, A. (2005). Feeling and reasoning: A computational model for emotional characters. *Progress in Artificial Intelligence*, 127–140.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *The Quarterly Journal of Economics*, 643–669.

Heath, C., & Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of risk and uncertainty*, *28*(1), 5–28.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*(8), 534–539.

Hirsh, J. B., Mar, R. a., & Peterson, J. B. (2012, April). Psychological entropy: A framework for understanding uncertainty-related anxiety. *Psychological review*, *119*(2), 304–20.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. Sage Publications, Inc.

Ito, J., & Marsella, S. (2011). Contextually-based utility: An appraisal-based approach at modeling framing and decisions. In *Proceedings of the twenty-fifth {AAAI} conference on artificial intelligence, august 7-11, 2011, san francisco, ca, united states* (Vol. 2, pp. 60–65).

Johnson, W. L., & Valente, A. (2009). Tactical Language and Culture Training Systems: using AI to teach foreign languages and cultures. *AI Magazine*, *30*(2), 72.

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

Kim, J., Hill, R., Jr, Durlach, P., Lane, H., Forbell, E., Core, M., ... Hart, J. (2009). BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education*, *19*(3), 289–308.

Mascarenhas, S., Dias, J., Prada, R., & Paiva, A. (2010). A dimensional model for cultural behaviour in virtual agents. *Applied Artificial Intelligence*, *24*(6), 552–574.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann.

Peot, M., & Smith, D. (1992). Conditional nonlinear planning. *... conference on Artificial intelligence planning ...*, 189–197.

Pezzulo, G., Lorini, E., & Calvi, G. (2004). How do I Know how much I don't Know? A cognitive approach about Uncertainty and Ignorance. *Proceedings of COGSCI ...*.

Smithson, M. (2008). Psychology ambivalent view of uncertainty. *Uncertainty and risk: Multidisciplinary perspectives*, 205–218.

Takahashi, T., Oono, H., Radford, M. H. B., & Others. (2007, January). Comparison of probabilistic choice models in humans. *Behavioral and Brain Functions*, *3*(1), 20.

Tversky, A., & Fox, C. R. (1995). Weighing risk and uncertainty. *Psychological review*, *102*(2), 269.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of business*, 251–278.

Tversky, A., & Kahneman, D. (1992, October). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, *5*(4), 297–323.

Wu, G., & Gonzalez, R. (1999). Nonlinear decision weights in choice under uncertainty. *Management Science*, 74–85.