Modeling Human Decision Making using Extended Behavior Networks

Klaus Dorer

Hochschule Offenburg, Elektrotechnik-Informationstechnik, Germany,
klaus.dorer@fh-offenburg.de

Abstract. In their famous work on prospect theory Kahneman and Tversky have presented a couple of examples where human decision making deviates from rational decision making as defined by decision theory. This paper describes the use of extended behavior networks to model human decision making in the sense of prospect theory. We show that the experimental findings of non-rational decision making described by Kahneman and Tversky can be reproduced using a slight variation of extended behavior networks.

1 Introduction

Looking forward to the goal of RoboCup to win against the human world champion team of soccer in 2050 one could state the question whether the decision making of the robots should be human like or rational with respect to decision theory. No matter what the answer to this question is, there should be no doubt that the robots should be able to model their opponents to understand and predict their decision making. Since the opponent team will be humans we therefore have to be able to model human decision making.

In their famous work on prospect theory Kahneman and Tversky [1979] have shown that human decision making does violate the tenets of decision theory. In a series of experiments they have shown a couple of deviations to the predictions of decision theory among those are that humans overestimate low probabilities and underestimate high probabilities and that subjective utility can differ from objective utility. Daniel Kahneman was awarded the Nobel prize in economic sciences in response to this work.

Behavior Networks [Maes, 1989] were introduced as a means to combine reactive and deliberative decision making using a mechanism of activation spreading to determine the best behavior. With Extended Behavior Networks (EBNs) [Dorer, 1999a] the mechanism of activation spreading was changed so that activation is a measure of expected utility of a behavior. In this paper we show that the mechanism of activation spreading in EBNs only needs slight modifications to reproduce human decision making reported by prospect theory.

Section 2 presents a number of experiments described by Kahneman and Tversky for their work on prospect theory. Section 3 introduces EBNs and the mechanism of activation spreading. Section 4 describes how EBNs can be used to model human decision making and reports on experimental results achieved applying EBNs to the same experiments conducted by Kahneman and Tversky with humans. Section 5 concludes the paper indicating possible future work.

2 Prospect Theory

Decision theory is based on the principle of maximum expected utility. If an agent is faced with a decision between actions (prospects) of uncertain outcome it should choose the prospect that has the highest expected utility. More formal: the expected utility of a prospect $P(u_1, p_1; \ldots u_n, p_n)$ is calculated as

$$eu_P = \sum_{i=1\ldots n} p_i \times u_i \tag{1}$$

where p_i is the probability of outcome u_i and $\sum p_i = 1$. Choosing the prospect with highest expected utility will maximize the agent's utility in the long term and is therefore considered as rational.

On the other side experiments with humans show that human decision making deviates from the above. Kahneman and Tversky [1979], for example, describe a series of experiments that led to the formulation of prospect theory, a theory of human decision making under risk. In this section we describe a selection of their experiments that are used in section 4 to be repeated by extended behavior networks.

2.1 Weighting Function

In their first experiment students had the choice between winning 2500 Israeli pounds[1] with a probability of 0.33, winning 2400 with probability 0.66 and nothing with probability 0.01 (A) compared to winning 2400 for sure (B). The expected utility of decision theory for A and B are $eu_A = 0.33 \times 2500 + 0.66 \times 2400 + 0.01 \times 0 = 2409$ and $eu_B = 1.0 \times 2400 = 2400$. So a rational agent should prefer A over B. However, in the experiment 82% (significant*) of the students chose B, the certain outcome.

Name	Prospect	Expected Utility	Human choices
A	(2500,0.33; 2400,0.66; 0,0.01)	2409	18
B	(2400, 1.0)	2400	82*

Table 1. Problem 1: One uncertain one certain prospect

[1] The average monthly income of a family was 3000 Israeli pounds at that time

The second experiment repeats the first but eliminates a chance of winning 2400 with probability 0.66 from both prospects. Table 2 shows the results of the experiments. Now with both prospects being uncertain the majority of students prefer C over D. 62% of the students took combination B and C.

Name	Prospect	Expected Utility	Human choices
C	(2500,0.33; 0,0.67)	825	83*
D	(2400, 0.34; 0,0.66)	816	17

Table 2. Problem 2: Two uncertain prospects

This and more experiments showed that humans overestimate low probabilities (except the impossible outcome) and underestimate high probabilities (except the certain outcome). Prospect theory therefore introduces a non-linear weighting function mapping probabilities to decision weights. Figure 1 shows a qualitative sketch of the weighting function [Kahnemann and Tversky, 1979].

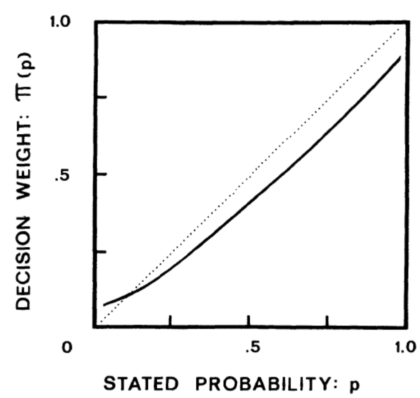

Fig. 1. A hypothetical weighting function as proposed by Kahneman and Tversky.

2.2 Value Function

In another experiment Tversky and Kahneman [1981] have shown that students preferred a certain win of 240$ (A) compared to a 25% chance of winning 1000$ (B) despite the fact that expected utility of (A) is less. In the same experiment a 75% chance for a loss of 1000$ (D) was prefered over a certain loss of 750$ (C) despite the fact that both have the same expected utility (see table 3). 73% of the students chose the combination of (A) and (C), 3% chose the combination (B) and (D). In another experiment students had the choice between a 25% chance of winning 240$ and a 75% chance of loosing 760$ (E) or a 25% chance of winning 250$ and a 75% chance of loosing 750$ (F). Not surprisingly all students choose option (F) (see table 4).

Name	Prospect	Expected Utility	Human choices
A	(240,1.0)	240	84
B	(1000, 0.25; 0,0.75)	250	16
C	(-750, 1.0)	-750	13
D	(-1000,0.75; 0,0.25)	-750	87

Table 3. Problem 3: decision under gains and losses

Name	Prospect	Expected Utility	Human choices
E	(240,0.25; -760,0.75)	-510	0
F	(250,0.25; -750,0.75)	-500	100

Table 4. Problem 4: intransitive decision with respect to problem 3

This experiment is particularly interesting since the combination (A) and (C) chosen by most students in the first experiment is equivalent with respect to decision theory to option (E) of the second experiment while the combination (B) and (D) is equivalent to option (F). So decision making of the majority was intransitive.

Prospect theory suggests that gains and losses are not linearly mapped to the subjective value of human decision makers. The value function is rather "generally concave for gains and commonly convex for losses and steeper for losses than for gains" [Kahnemann and Tversky, 1979]. A qualitative value function with this properties is displayed in figure 2.

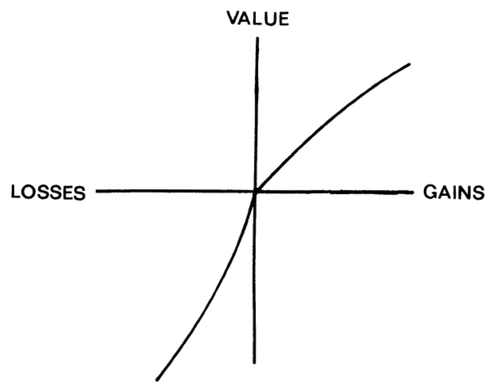

Fig. 2. A hypothetical value function as proposed by Kahneman and Tversky.

3 Extended Behavior Networks

Behavior Networks [Maes, 1989] use a mechanism of activation spreading to decide between a couple of executable behaviors combining reactive and deliberative decision making. Extended Behavior Networks (EBNs) [Dorer, 2004; 1999a] changed the mechanism of activation spreading so that activation is a measure of expected utility of a behavior according to decision theory. They have been successfully used, for example, as decision mechanism for the magmaFreiburg team scoring 2^{nd} in RoboCup 1999 simulation league competition [Dorer, 1999b]. In this section we give a short overview of the relevant activation spreading mechanism in EBNs before we describe how this mechanism needs to be changed to model human decision making according to prospect theory.

3.1 Network Definition

Extended behavior networks consist of goals, resource nodes and so called competence modules that are linked into a network.

Definition 1. *A* goal *consists of a tuple (*GCon*,* ι*,* RCon*) with*

- GCon *the* goal condition *(conjunction of propositions, i.e. possibly negated atoms), the situation in which the goal is satisfied,*
- $\iota \in [0..1]$ *the (static)* importance *of the goal,*
- RCon *the* relevance condition *(conjunction and disjunction of propositions), i.e. the situation-dependent (dynamic) importance of the goal.*

Definition 2. *A* resource node *is a tuple (*res*,* g*,* θ_{Res}*) with*

- $res \in \mathcal{R}$ *the* resource *represented by the node,*
- $g \in \mathbb{R}^+$ *the amount of* bound resource units*, i.e. units that are bound by a currently executing competence module and*

- $\theta_{Res} \in]0..\theta]$ *the resource specific* activation threshold *(where θ is the global activation threshold of the network).*

Definition 3. *A* competence module k *consists of a tuple (*Pre*, b,* Post*,* Res*, a) with*

- Pre *the* precondition *and* $e = \tau_P(Pre, s)$ *the* executability *of the competence module in situation s where* $\tau_P(Pre, s)$ *is the (fuzzy) truth value of the precondition in situation s;*
- b *the* behavior *that is performed once the module is selected for execution;*
- Post *a set of tuples (Eff, ex), where Eff is an expected effect (a proposition) and* $ex = P(Eff|Pre)$ *is the* probability *of Eff getting true after execution of behavior b,*
- a *the* activation $\in \mathbb{R}$*, representing a notion of the expected utility of the behavior (see below).*
- Res *is a set of* resources $res \in \mathcal{R}$ *used by behavior b.* $\tau_U(k, res, s)$ *is the situation-dependent amount of resource units expected to be used by behavior b.*

Definition 4. *An* extended behavior network *EBN consists of a tuple* $(\mathcal{G}, \mathcal{M}, \mathcal{U}, \Pi)$*, where* $\mathcal{G}$ *is a set of goals,* $\mathcal{M}$ *a set of competence modules,* $\mathcal{U}$ *a set of* resource nodes *and* Π *is a set of* parameters *that control activation spreading (see below)*

- $\gamma \in [0..1[$ *controls the influence of activation of modules,*
- $\delta \in [0..1[$ *controls the influence of inhibition of modules,*
- $\beta \in [0..1[$ *the inertia of activation across activation cycles,*
- $\theta \in [0..\hat{a}]$ *the activation threshold that a module has to exceed to be selected for execution, with* $\hat{a}$ *the upper bound for a module's activation,*
- $\Delta\theta \in]0..\theta]$ *the threshold decay.*

3.2 Activation Spreading

The decision of which behavior to adopt should be based on the the expected utility of executing such behavior. In EBNs, the expected utility of a behavior is approximated by a mechanism called *activation spreading.* The competence modules are connected to the goals and other competence modules of the network. Across those links activation is spread from the goals to the competence modules and among competence modules.

A competence module receives *activation* directly from a goal if the module has an effect that is equal to a proposition of the goal condition of that goal.

$$a_{kg_i}^{t\,\prime} = \gamma \cdot u(\iota_{g_i}, r_{g_i}^t) \cdot \nu_\gamma(p_j) \cdot ex_j \;, \qquad (2)$$

$u(\iota_{g_i}, r_{g_i}^t)$ is the utility function mapping importance ι_{g_i} and relevance $r_{g_i}^t$ to a utility value. In section 4 we will show how this utility function has to be changed to reproduce human decision making described in section 2.2. ν_γ determines how activation is distributed to multiple propositions of the goal condition. ex_j is the probability of the corresponding effect to come true. In section 4 we will introduce a weighting function for this probability corresponding to section 2.1. The amount of activation depends on the probability of an effect to come true and the utility of the proposition in the goal condition as described in equation 1.

A competence module is *inhibited* by a goal if it has an effect proposition that is equal to a proposition of the goal condition and one of the two propositions is negated. Inhibition represents negative expected utility and is used to avoid the execution of behaviors that would lead to undesired effects.

$$a_{kg_i}^{t\,\prime\prime} = -\delta \cdot u(\iota_{g_i}, r_{g_i}^t) \cdot \nu_\delta(p_j) \cdot ex_j \ , \tag{3}$$

A competence module x is linked to another competence module y if x has an effect that is equal to a proposition of the precondition of y. y is called a *successor* module of x. Module x gets activation from the successor the amount of which depends on the utility of the precondition and the probability of the effect to come true. The utility of propositions that are not part of a goal condition is not available directly. It can be determined indirectly using the activation of the containing module and the truth value of the proposition. In this way, unsatisfied preconditions get implicit sub-goals of the network. Their utility directly depends on the utility of the competence module itself.

Finally a competence module x is linked to another competence module y if it has an effect that is equal to a proposition of the precondition of y and one of the two propositions is negated. y is called a *conflictor* of x, because it has an effect that destroys an already satisfied precondition of x. Again, a conflictor link from x to y is inhibiting (negative activation) to avoid undesired effects.

The activation of a module k at time t is then the sum of all incoming activation and the previous activation of the module decayed by β (defined in the set of parameters Π):

$$a_k^t = \beta a_k^{t-1} + \sum_i a_{kg_i}^t, \tag{4}$$

where $a_{kg_i}^t$ is the maximal activation module k receives at time t from goal g_i to which the module is linked directly or indirectly across incoming successor and conflictor links of other competence modules. For more details on activation spreading see [Dorer, 1999a].

Behavior selection is done locally in each competence module in a cycle containing the following steps. The details of behavior selection are not relevant in this context.

1. Calculate the activation a of the module.
2. Calculate the executability e of the module.
3. Calculate the execution-value $h(a, e)$ as the product of both.
4. Choose those competence modules for execution that have an execution-value above that of each resource node linked to. For each resource there have to be enough units available.
5. Reduce θ of each resource node not used $\Delta\theta$ and go to 1.

4 EBNs and Prospect Theory

In this section we show how the calculation of activation in EBNs has to be changed to correspond to findings reported in section 2. Experiments with both versions of EBNs reproduce the discrepancy between decision theoretic and human decision making reported in that section.

4.1 Theory

As described in section 2, prospect theory introduces non-linear value and weighting functions to explain results of experiments on human decision making.

The very same can be done for EBNs. The already existing value function u (see equation 2, 3) needs to be changed according to prospect theory. The utility function was chosen to correspond to the measure of risk taking defined by Arrow and Pratt [Eisenführ and Weber, 1999]:

$$r(x) = \frac{u''(x)}{u'(x)} \tag{5}$$

Using US-American tax data, Friend and Blume [1975] showed that investors exhibited decreasing absolute and constant relative risk taking behavior $x \cdot r(x)$. A utility function

$$u(x) = \begin{cases} x^{2\rho} & : \quad x \geq 0 \\ -x^{2\rho} & : \quad x < 0 \end{cases} \tag{6}$$

corresponds to this observation if x is a normed value and $\rho \in]0..1]$ is used as risk parameter. Using $\rho = \frac{1}{2}$ the utility function is linear and corresponds to decision theory exhibiting risk neutral behavior. $\rho < \frac{1}{2}$ corresponds to a risk-aversive behavior in case of gains ($x \geq 0$) and risk-taking behavior in case of losses ($x < 0$) as was observed in section 2. For $\rho > \frac{1}{2}$ it is vice versa.

A weighting function was not envisaged in original EBNs, but can easily be introduced. In equations 2 and 3 the probability ex_j is replaced with $\pi(ex_j)$ where the weighting function $\pi(x)$ is defined as follows:

$$\pi(x) = \begin{cases} 0 & : \quad x = 0 \\ 1 & : \quad x = 1 \\ e^{x-1} - \frac{1}{4} & : \quad 0 < x < 1 \end{cases} \tag{7}$$

This weighting function shows the properties described in section 2.1. For the impossible ($x = 0$) and certain ($x = 1$) outcome weighting function and probability match ($\pi(x) = x$). Low probabilities are overestimated ($\pi(x) > x$) while high probabilities are underestimated ($\pi(x) < x$).

The simplicity with which EBNs can be adjusted to model human decision making and prospect theory is also underlined by the amount of code changes necessary for implementation. Changing the existing value function required adding 4 lines of code. Adding the weighting function required changes in 4 lines of code (the calculation of activation for each type of connection) and adding another 6 lines of code.

4.2 Experiments

The changes above have been applied to the problems described in section 2. In the following figures, the upper level nodes of the networks are the goals with corresponding value. The lower level nodes are the competence modules representing the alternatives to choose from. The set of connections from a competence module represent the prospects with corresponding probabilities. The parameters for all the networks were chosen in order to have activation values matching expected utility of decision theory. As values we used $\gamma = 1.0, \delta = 1.0$[2] and $\beta = 0.0$. In the prospect theoretic cases we used as risk parameter $\rho = 0.4$.

Figure 3 shows the results of running an original EBN with decision theoretic activation calculation on problem 1 and problem 2. The activation of the modules correspond to the expected utility of decision theory. The network prefers a over b and c over d accordingly.

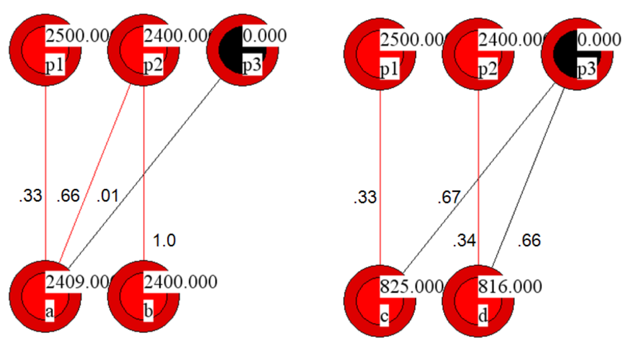

Fig. 3. Decision Theoretic EBN for problem 1 (left) and problem 2 (right) of section 2.

Figure 4 shows the results of running an EBN with new value and weighting function on the same problems. Now the network prefers b over a and c over d as the majority of students did.

In the same way, decision theoretic EBNs and EBNs according to prospect theory have been applied to problems 3 and 4. Figure 5 shows the results of an EBN with decision theoretic activation calculation. Again the activation of the modules correspond to the expected utility of decision theory. The network prefers b over a, is indifferent with respect to c and d and prefers f over e.

Figure 6 shows the results of an EBN according to prospect theory on the same problems. Here the network prefers a over b and d over c as the majority of students did. Also it prefers f over e showing the same intransitive decision taken by a significant amount of students.

[2] A value of 1.0 for γ and δ is outside the definition area that guarantees convergence of activation. In our case this is no problem since no activation spreading between competence modules is done

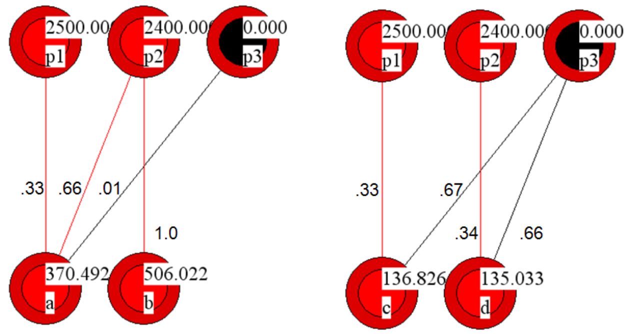

Fig. 4. EBN for problem 1 (left) and problem 2 (right) of section 2.

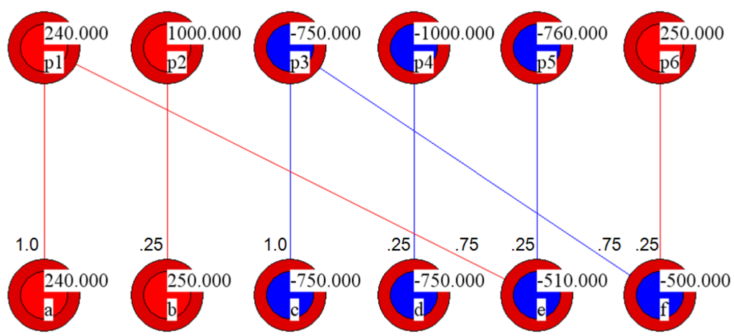

Fig. 5. EBN for problem 3 and problem 4 using decision theoretic activation spreading.

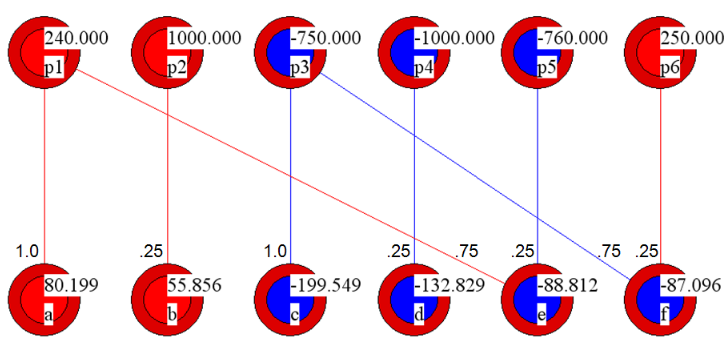

Fig. 6. EBN for problem 3 and 4 using activation spreading based on prospect theory.

5 Discussion and Future Work

In this paper we showed that EBNs can be used to reproduce human decision making deviating from rational decision making with respect to decision theory. However, the experiments of Kahneman and Tversky required only relatively small EBNs and no activation spreading between competence modules. Future work should investigate if the results can be used for bigger EBNs using activation spreading. The RoboCup domain is particularly interesting since EBNs have already been successfully applied to it.

The next steps should then be:

1. Play a number of soccer games against agents of an opponent team using EBNs based on prospect theory for decision making. This is already possible and first results indicate that such a team shows different behavior. Investigate if and how the first team of agents can improve their performance by modeling their opponents using EBNs based on prospect theory compared to a team of agents using EBNs for opponent modeling based on decision theory.
2. Replace the opponent team by a real soccer team and see if it shows improved performance when switching between EBNs using decision theory or prospect theory. If step 1 is successful then the results of this paper indicate that also this step should be.

References

[1999a] K. Dorer. Behavior networks for continuous domains using situation–dependent motivations. *Proceedings of the Sixteenth International Conference of Artificial Intelligence*, pages 1233–1238, 1999.

[1999b] K. Dorer. Extended behavior networks for the magmafreiburg team. In S. Coradeschi, T. Balch, G. Kraetzschmar, and P. Stone, editors, *Team Descriptions Simulation League*, pages 79–83. Linköping University Electronic Press, Stockholm, 1999.

[2004] K. Dorer. Extended behavior networks for behavior selection in dynamic and continuous domains. In U. Visser, H.-D. Burkhard, P. Doherty, and G. Lakemeyer, editors, *Proceedings of the ECAI workshop Agents in dynamic domains*. ECAI, Valenzia, 2004.

[1999] F. Eisenführ and M. Weber. *Rationales Entscheiden*. Springer Verlag, Berlin; Heidelberg, 1999.

[1975] I. Friend and M. Blume. The demand for risky assets. *American Economic Review*, 64:900–922, 1975.

[1979] D. Kahnemann and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

[1989] P. Maes. The dynamics of action selection. *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 991–997, 1989.

[1981] A. Tversky and D. Kahneman. The framing of decisions and the psychology of choice. *Science*, 22:453–458, 1981.