Rationalising decisions using multiple values

Luis Antunes and Helder Coelho
Faculdade de Ciências, Universidade de Lisboa
Campo Grande, 1700-162 Lisboa, Portugal
Ph: +351-1-7500087, Fax: +351-1-7500084
{xarax, hcoelho}@di.fc.ul.pt

ABSTRACT.
Even a simple artificial agent will have to make decisions and consider their consequences. An agent can achieve this by evaluating situations along several dimensions. The BVG (Beliefs, Values and Goals) architecture is a means to build up such an evaluation machinery. The motivational part of the agent's mind is captured by the concept of value, able to capture the preferences of the agent, which cannot be reduced to either beliefs or goals, or a combination of the two. We sketch a calculus for agent decision, and look into the problem of feedbacking the consequences of actions into the agent's decision mechanism, in order to adjust its behaviour for future occasions.

Keywords.
Agent model and architecture, decision machinery, multi-agent systems, autonomy.

INTRODUCTION.

Let us consider a setting in which autonomous agents are inserted into a shared environment and have to cope with scarcity of resources, complexity, dynamism and unpredictability. By agents we mean entities that will act either in behalf of someone (say, their user) or simply representing themselves. By autonomous we mean that those agents have motivations of their own, and can and will decide what to do based upon their own self interest. The simple fact that we expect our environment to have more than one agent, together with the notion of autonomy just given, ensures that our world is unpredictable and complex, and cannot be dealt with by computing in advance all possible courses of action, and so be prepared for every circumstance. Our agents will have to be smarter than that.

Decisions and rationality: utility.

The problem we will be addressing is that of decision, that is, how can an agent *choose* from a given set of alternative candidate actions. What justifies this choice, what are the expected consequences of that choice, how do these expectations help illuminate the choice process (deliberation).

In classical decision theory the concept of utility emerges as a possible answer to all these questions. Utility is a way of representing the preferences of the agent. Utility theory says that every state of the world has a degree of usefulness, or utility, to an agent, and that the agent will prefer states with higher utility. Utility theory can be combined with probability theory to build the Principle of Maximum Expected Utility (PMEU) that states that an agent is rational if and only if it chooses the action that yields the higher expected utility, averaged over all the possible outcomes of the action [Russell and Norvig 1995].

One of the constraints Russell and Norvig propose as central to utility theory is orderability. Orderability requires the agents to either prefer one of any two given states, or be indifferent between the two. It may well be the case that an agent does not relate at all two given states, and any attempt of comparing the two would be prejudicial, and proven wrong when new information arrives. Moreover, in a complex, dynamic environment, it is impossible to have strong beliefs about almost anything.

Further criticism on utility-based decision is summarised in [Nwana et al 1996]. In particular, they criticise that all the agents have knowledge of the payoff matrix, and therefore full knowledge of the other agent's preferences. Of course, assuming rationality is defined as utility maximising, this would amount to an agent being fully capable of predicting all the other agent's choices. The world is necessarily closed, and becomes predictable, given enough computational power.

The case here is not one of rationality versus irrationality. We are discussing the basic requirements an agent should fulfil to use utility theory for decision. And those requirements are too demanding on the capabilities of the agent. Surely if an agent fulfilled those, utility theory would be a rather strong candidate to inform a decision theory. The problem is that agents cannot raise themselves to be so perfect. As [Simon 1996] notes, "a large body of evidence shows that human choices are not consistent and transitive as they would be if a utility function existed."

BDI agents.

The Belief-Desire-Intention (BDI) agent model [Rao and Georgeff 1991, Bratman 1987, Wooldridge and Parsons 1998, Cohen and Levesque 1987] is a different tentative at solving the decision-action problem. Basically, agents commit to the choices they have made (and that have turned desires into intentions), and do not change this commitment until some significant change occurs (the intention is fulfilled or it becomes impossible to fulfil).

In [Wooldridge and Parsons 1998] a clear distinction is made between deliberation (the process of deciding what to do) and action selection (deciding how to do it). They are especially interested in the reconsideration of intentions, and so keep desires out of their agent architecture. It is still an issue to see where intentions come from. In [Pollack and Ringuette 1990], it is the deliberation mechanism that provides a serialisation of intentions, through the use of payoffs- and of utilities-based functions. Utilities are also used in [Wooldridge and Parsons 1998] to define an optimal action selection function, but nothing is said about how to build it.

In either case, the question of choice remains interesting and unexplored. In most BDI models, the deliberation process filters through the desires to provide intentions. But little is said about such a process, except the logical features the resulting intentions must possess: they must be consistent with the previously existing intentions, they must be achievable in accordance with the agent's beliefs, and so on. It is too clean of a choice process, in the sense that all we have is technical reasons to choose or not. We lack real world, 'dirty' reasons for choice, the preferences of the agent. The truth is that in real world our choices aren't that clean: we choose some things apparently for no reason, just because we want to give it a try; we pursue in activities although we're not sure whether we will ever be successful; we are not rational. Or are we? This is the question we will be trying to answer in the rest of the paper.

In the next section we provide an answer to this question by postulating values as the mental objects that allow choice to be made. We evaluate situations and possible courses of action against several dimensions, and use this multiple evaluation to finally choose the action to perform. Moreover, the outcome of that action is feedbacked into the agent in order to enhance future decisions.

MULTIPLE VALUES.

The move away from utility-based decision we make here has been done previously in different manners. In [Sichman 1995], the attempt was to base social interaction in the theory of social dependency [Castelfranchi 1990, Castelfranchi et al 1992]. Some agents' abilities complement another agents' needs. Keeping in mind that an agent's behaviour is fundamentally determined by its own motivations, social interactions (such as goal adoption, social exchange, etc.) are determined by the agent's place in the network of social dependences involving it. So, the whole notion of sociality is based on this relational concept of dependence, allowing for such interesting forms of social situations as manipulation and exploitation, as well as exchange and cooperation.

In the present work, the emphasis is on the individual mental components and mechanisms an agent must possess in order to give meaning to the expression "pursue of self interest."

In a previous paper, [Antunes 1997] has proposed the use of multiple values as a new kind of mental objects that has a fundamental role in decision-making, and so in defining the character of the agents. A value is a dimension against which a situation can be evaluated. A dimension is a non empty set endowed with an order relation. Most interesting situations from the decision standpoint will have several such dimensions, and so most decisions are based on multiple evaluations of the situation and alternative courses of action [Miceli and Castelfranchi 1992].

Values for choice.

In a similar line of reasoning, Simon has proposed the notion of aspiration level: "Aspirations have many dimensions: one can have aspirations for pleasant work, love, good food, travel, and many other things. For each dimension, expectations of the attainable define an aspiration level that is compared with the current level of achievement. (…) There is no simple mechanism for comparison *between* dimensions. In general, a large gain along one dimension is required to compensate for a small loss along another — hence the system's net satisfactions are history-dependent, and it is difficult for people to balance compensatory effects."

Values can be said to constitute the motivational part of the agent's mind. Different agents decide differently in the same situation because they base those decisions in different value systems. The same agent can decide differently in the same situation because its value system has evolved as a result of the interaction between the agent and the world.

It is possible that the several evaluations do not all agree in selecting the same candidate for execution. Belonging to different dimensions it may not be easy to collapse these evaluations into a function that serialise the candidates. However, the agent must decide, so such a function is needed. It could then be argued that this function *is* the decision function, and is equivalent to a utility function. Apparently, in this case there would be no need for the multiple evaluations setting, they would all amount to components of the decision function. On the contrary, we defend the independent existence for the evaluations along the multiple dimensions, and for a number of reasons: new dimensions can be added or subtracted, or are not relevant, or are especially relevant in some situations; agents can communicate about them, exchanging experiences, adopting new measures or even dimensions from other agents; the evaluation functions along each dimension can evolve over time. The existence of such objects in mind allows for a greater flexibility in a number of situations than would be possible if all were collapsed into some kind of generalised utility function. There is also the argument that such a function does exist, as a kind of 'general currency' in mind, one dimension that would serve as a conversion basis for all others. Cabanac's paper in [Meyer and Wilson 1991] proposes pleasure (and pain) for this role, which agrees with the behavioural theories of reinforcement and punishment. This proposal is questionable, but worth further thought.

Adaptation over time.

When an agent makes a decision, it will monitor the features of the world that are related with the consequences of that decision. In particular, it will check whether the action was successful or not. The agent updates its belief and goal structure according to what has happened as a result of its action. And the information concerning that decision and its consequences will be used to enhance the decision system for future situations.

The agent has to feedback the relevant information and use it to readjust its value system. This means that the agent must evaluate the quality of its previous evaluation, the one that lead to the decision. We can put the question of what information to feedback and how to use it.

Before we propose alternative ways to tackle this issue, let us note that again the evaluation of the decisions quality could serve as a candidate for the general decision function. That is, if this function constitutes the ultimate assessment of the quality of the decisions, why not use it for decision itself? The answer is that none of these functions is expected to be optimal, and so it is a better option to let all of them operate freely to enhance its contribution to the general behaviour. It could even be the case that the optimal behaviour does not exist, and there is no sense in looking for the perfect decision function. The world is too open and unpredictable for that to be possible to obtain. Furthermore, as we shall see, it could be just as difficult to use this feedbacked information as is to solve the decision problem itself.

We propose, as a way out, three alternative ways of assessing the quality of the decision. The first is to consider some measure of 'goodness' of the decision, that will assess to what extent the action was successful. The simplest case is to test for success or failure; in a more promising approach, the action is evaluated and some (say, numerical, a real number belonging to [0, 1]) grade is assigned to the decision and corresponding action. This was the option used in [Antunes et al 1993], where a game strategy was designed which used the assessment of our opponent's move to dynamically adapt our estimation of her strength, in order to allow for exploitation of her possible weaknesses.

The second alternative is to observe in the world the consequences of the selected action in terms of the same dimensions that the agent used for decision. For instance, if the agent used time and money as the key dimensions upon which the decision was based, then it should observe what the consequences were of that decision in terms of time and money themselves. This should ease up the problem of updating the system of values, since there are no mediators to consider between the evaluation and its assessment.

Finally, we propose to make the assessment of the decision by measuring its results in terms of the set of dimensions that the designer of the agent is interested in. When we design an agent, we are interested in obtaining some behaviour, and we can postulate some dimensions and corresponding measures to evaluate the quality of the behaviour we obtain to compare it with the desired one. But it is not necessarily the case that when designing the agent's decision machinery, we use those exact dimensions and measures we had in mind. To evaluate the agents performance with our own measures (that they did not have access to) amounts to look for emergent behaviour.

It is not obvious how to use this feedbacked information to enhance decision-making as it is not obvious how to design the decision mechanism in the first place. We have conducted experiments with the mechanisms we have just described, to gain insight into these problems. We concluded that a lot of *ad hoc* design options must be made in each concrete decision setting. However, some structural features can be obtained.

THE BVG AGENT ARCHITECTURE.

The BVG (Beliefs, Values and Goals) architecture puts in evidence the role of choice in the behaviour of the agents. The agent is kept as simple as possible in what concerns its basic blocks. A structure of goals and beliefs is completed by values, which will be used by the choice mechanism to obtain a decision. These values are not rigid, they can and will be changed during the interaction.

Basic mental objects.

The intention here was to simplify as much as possible the architecture where we wanted to introduce the concept of value. So we started by taking beliefs, and grouped together the pro-active mental objects under the umbrella designation of goal. Of course there are several types of goals, and intentions are certainly among the most important and useful, whereas desires are usually allowed the flexibility and freedom of representing anything the agent might want, without further concerns. We aim at skipping these technical details, so that we concentrate on what will be new: the use of values to inform choice (see for example [Wooldridge and Parsons 1998] and their inability to face this issue). Beliefs represent what the agent *knows*; goals represent what the agent *wants*; values will represent what the agent *likes*.

Once this introduction of values accomplished, we can expand our architecture. We will take the general classes of informational and motivational attitudes (represented by beliefs and goals), and leave their specialisation until later. We can subscribe almost entirely the architectural options in [Wooldridge and Parsons 1998]. However, we kept desires out of the picture for different reasons. We believe that the sources of the agent's goals can come from suitably informed beliefs. It will be a further role of values to help select among the believed things which ones to pursue. But this role of values in deliberation will be left for further research as we examine the choice process.

The use of values.
The agent operates as follows. In each cycle of operation, the agent's beliefs and evaluations are updated. This is similar to the function 'next state' in [Wooldridge and Parsons 1998], which is thought of as a belief revision function. It determines the new set of beliefs of the agent, taking as basis the present set of beliefs and also the state of the world. This new set of beliefs includes any new information that origins from the agent's perceptions.

Afterwards, deliberation occurs, during which the agent computes the decision to be taken and the corresponding action to be executed. The selected action is then executed, and the world is updated accordingly. The agent notes these changes in the world in the immediate cycle. A more asynchronous scheme would certainly be more realistic, allowing for pre-emptive processes of reasoning, but for the moment, this scheme suffices to gain insight into the problem at hand.

Now we focus on our choice mechanism. Values are used to relate goals with beliefs. Each goal is characterised by a set of values that work as standards (something like Simon's aspiration levels) against which candidate sub-goals will be matched.

$\text{Goal}\ ((V_1, v_1), \ldots, (V_n, v_n))\ g$ (1)

Plans are represented as beliefs. The values that inform these beliefs represent the characterisations the agent associates with each of the alternative plans. For instance, we could have:

$\text{Bel}\ ((V_1,v_{11}), \ldots, (V_n, v_{1n})\ s_1 \text{->} g$ (2)

$\text{Bel}\ ((V_1,v_{21}), \ldots, (V_n, v_{2n})\ s_2 \text{->} g$ (3)

We have now the conditions to make our choice. For that purpose, we need a calculus that collapses the several measures over the values into a function that picks up the best option. For instance, a calculus based on [Simon 1996] would pick s_i such that i is the minimum that satisfies $\forall j,\ v_{ij} \geq v_j$.

In general, we will have a choice function F that will take g and s_i as arguments and assign each s_i a real number that will allow for the serialisation of the candidates. The chosen candidate will be proposed as a new goal for execution, and so it needs to be assigned its own standards.

$\text{Goal}\ ((V_1=F_0(v_1, v_{i1})), \ldots)\ s_i$ (4)

As an example, consider that our value V_i is a deadline for the achievement of goal g. v_{i1} would be the amount of time the alternative s_i is believed to spend. So we should take $F_0(x, y)= x-y$, so that the sub-goal s_i has v_1-v_{i1} as deadline.

The insertion of a BVG agent into its world.
Now we insert the agent in its world according to this previous scheme. The feedback of the previous choice and corresponding action is made by looking at the appropriate features in the world. In a more elaborate architecture, this mechanism would be driven by expectations, by now we will use the appropriately chosen beliefs, and handle this information in the decision process.

When the decision is made, it gives rise to a new sub-goal to be fulfilled. The decision mechanism assigns the corresponding values for the sub-goal standards, and keeps track of the performance of its execution. When the goal is tentatively fulfilled, the quality of this execution is assessed and the standards are updated in the super-goal. This means that when some goal is executed, the quality of this execution reflects not on itself, but on the choice that led to it. If the subordinate fails, we blame the one that wrongly chose this candidate over the others.

The values that characterise our plans can be updated as well. If some alternative did not perform as we expected, the features that lead to it being chosen should be revised. It may seem that one of the two revisions could be enough to reflect the information observed into the system of values. But a short example will show otherwise. Imagine we have

Goal (…, (Time, 100), …) g (5)

And that the alternative goal g1 was chosen among the candidates:

$\text{Bel}\ (\ldots(\text{Prob. succ.}, 0.8), (\text{Time}, 10), \ldots)\ g_1 \text{->} g$ (6)

$\text{Bel}\ (\ldots, (\text{Time}, 20), \ldots)\ g_2 \text{->} g$ (7)

If g_1 fails we should not only update its own description as a plan (e.g., Bel (…(Probability of success, 0.7), (Time, 10), …) g_1 -> g) but we should also update g: Goal (…, (Time, 90), …) g, since g_1 has spent 10 time units.

So we have $r(s_i) \in$ Results (the set where the assessment of the decisions takes values), the evaluation of the execution of the selected plan. to update our system of values, we need a function G that takes $v_{i1}, \ldots, v_{in}$ and $r(s_i)$ as arguments and returns $v_{i1}', \ldots, v_{in}'$, the new characterisation of s_i. G also acts over g, the goal 'responsible' for choosing its sub-goal s_i.

CONCLUDING REMARKS.

We addressed the problem of choice at the light shed by the introduction of multiple values to assess a decision situation. The results of some experiments confirm that the agent's choice machinery becomes more clear, as agents express their preferences through the use of this multiple value framework. Choice is performed by collapsing the various assessments into a choice function, that cannot be considered equivalent to a utility function, since it is computed in execution time. Moreover, by feedbacking assessments of the quality of the previous decision into the agent's decision process, we gain in adaptation capabilities. Our agents' decisions no longer depend solely on the past events as known at design time. Instead, events are incorporated into the decision machinery as time passes, and the components of those processes evolve continuously to be aggregated just when a decision is needed. This is a step towards real autonomy. Future research will address the increase of the agent's adaptability by enhancing the use of the feedbacked information. Another issue is the expansion of an agent's system of values as a result of interaction with other agents.

ACKNOWLEDGEMENTS.

We deeply thank our colleague João Faria for his help in working the case study, and for his support throughout the research. Thanks are also due to João Balsa, José Castro Caldas, José Cascalho, Cristiano Castelfranchi, Rosaria Conte, Maria Miceli, Luis Moniz, Leonel Nóbrega, Pedro Rodrigues and Jorge Louçã. This research has been partially supported by project Praxis 2/2.1/TIT/1662/95 (SARA).

REFERENCES.

[Antunes 1997] Luis Antunes, Towards a model for value-based motivated agents, in proceedings of MASTA'97 (EPIA '97 workshop on Multi-agent Systems: Theory and Applications), Coimbra, October 1997.

[Antunes et al 1993] Luis Antunes, Luis Moniz and Carlos Azevedo, RB+: The Dynamic Estimation of the Opponent's Strength, in proceedings of the AISB Conference, IOS Press, Birmingham 1993 (11 pages).

[Axelrod 1995] Robert Axelrod, A model of the emergence of new political actors, in Artificial Societies — The computer simulation of social life (proceedings of SimSoc'93), Nigel Gilbert and Rosaria Conte (eds.), University College London Press, 1995.

[Bratman 1987] Michael E. Bratman, Intentions, Plans and Practical Reasoning, Harvard University Press, London, 1987.

[Castelfranchi 1990] C. Castelfranchi, Social Power. A point missed in Multi-Agent, DAI and HCI., in Decentralized AI - Proceedings of MAAMAW'90, Yves Demazeau and J.P. Müller (Eds.), Elsevier Science Publishers B. V., Amesterdam, 1990.

[Castelfranchi et al 1992] Cristiano Castelfranchi, Maria Miceli, and Amedeo Cesta, Dependence relations among autonomous agents, in Decentralized AI 3, Eric Werner and Yves Demazeau (eds.), Elsevier Science Publishers B. V., Amesterdam, 1992.

[Cohen and Levesque 1987]] Philip R. Cohen and Hector J. Levesque, Intention = Choice + Commitment, Proceedings of AAAI'87, 1987.

[Kinny and Georgeff 1991] David Kinny and Michael Georgeff, Commitment and Effectiveness of Situated Agents, Proceedings of the Twelfth International Joint Conference on Artificial Intelligence, IJCAI-91), Sydney, Australia, 1991.

[Meyer and Wilson 1991] Jean-Arcady Meyer and Stewart W. Wilson (eds.), From Animals to Animats: Proceedings of the First International Conference on Simulation of Adaptive Behaviour (SAB'90), The MIT Press, Cambridge MA, 1991.

[Miceli and Castelfranchi 1992] Maria Miceli and Cristiano Castelfranchi, The Cognition of Value — A cognitive theory of the evaluation mechanisms and processes, FrancoAngeli, Milano, 1992 (in Italian).

[Nwana et al 1996] Hyacinth Nwana, Lyndon Lee and Nick Jennings, Coordination in Software Agent Systems, BT Technology Journal, 14 (4), 1996.

[Pollack and Ringuette 1990] Martha E. Pollack and Marc Ringuette, Introducing the tileworld: experimentally evaluating agent architectures, in Proceedings of the Conference of the American Association for Artificial Intelligence (AAAI'90), 1990.

[Rao and Georgeff 1991]] Anand S. Rao and Michael P. Georgeff, Modeling Rational Agents within a BDI-Architecture, in proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning, Morgan Kaufmann, 1991.

[Russell and Norvig 1995] Stuart Russell and Peter Norvig, Artificial intelligence: a modern approach, Prentice Hall, 1995.

[Sichman 1995] Jaime Sichman, Du raisonnement social chez les agents, une approche fondée sur la théorie de la dépendance, Thèse de Doctorat, Institut National Polytechnique de Grenoble, 1995.

[Simon 1996] Herbert Simon, The Sciences of the Artificial (third edition), the MIT Press, Cambridge, 1996.

[Wooldridge and Parsons 1998] Michael Wooldridge and Simon Parsons, Intention Reconsideration Reconsidered, Pre-Proceedings of the 5th International Workshop on Agent Theories, Architectures and Languages (ATAL'98), 4-7 July, Paris, 1998.