2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology

An agent supports constructivist and ecological rationality

John Debenham
Centre for Quantum Computation & Intelligent Systems
University of Technology, Sydney
NSW 2007 Australia
debenham@it.uts.edu.au

Carles Sierra
Institut d'Investigació en Intel·ligència Artificial, IIIA
Spanish Scientific Research Council, CSIC
08193 Bellaterra, Catalonia, Spain
sierra@iiia.csic.es

***Abstract*—An agent architecture supports the two forms of deliberation used by human agents. The work is founded on the two forms of rationality described by the two Nobel Laureates Friedrich Hayek and Vernon Smith. Cartesian, constructivist rationalism leads to game theory, decision theory and logical models. Ecological rationalism leads to deliberative actions that are derived from agents' prior interactions and are *not* designed; i.e., they are strictly *emergent*. This paper aims to address the scant attention paid by the multiagent systems community to the predominant form of deliberation used by mankind.**

***Keywords*-intelligent agents; rationality; deliberation;**

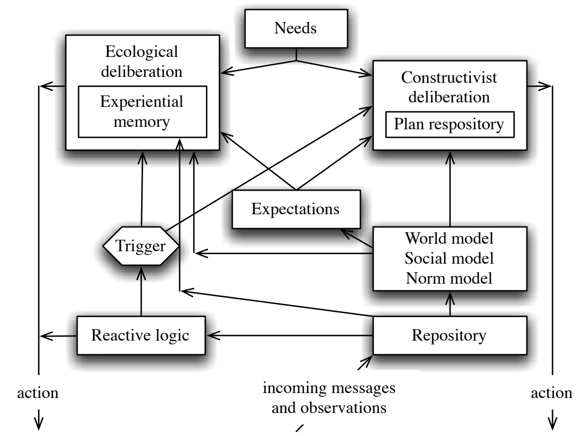
I. Introduction

This paper is concerned with the issue generally known as *bounded rationality* that dates back to David Hume [1] and more recently to the early work of Herbert Simon. Bounded rationality refers to systems that are not founded on Cartesian rationalism; it has been widely addressed in economics [2], and is discussed in all good books on artificial intelligence, e.g. [3].

For over fifty years artificial intelligence research has spawned countless theories and systems that are *not* founded on Cartesian rationalism; one classic contribution being Rodney Brooks' work reported in his 'Computers and Thought' award-winning paper [4]. Despite these advances, work in multiagent systems has been heavily influenced by game theory, decision theory and logic [5]. This paper describes a form of agency that enables rational agents to move beyond Cartesian rationalism. The work is founded on the two forms of rationality described by the two Nobel Laureates Friedrich Hayek [6] and Vernon Smith [7] as being within 'two worlds'.

Hayek and Smith identify; *constructivist rationality* that underpins rational predictive models of decision making; and, *ecological rationality* founded on the concept of "spontaneous order" that refers to social institutions and practices that *emerge* from the history of an agent's interactions and are *not* pre-designed. Past experience is a precursor to ecological rationality. For example, as we have described them previously, trust and honour [8] and reputation [9], are purely ecological concepts. An agent's *experiential memory* contains a record of: what occurred when each prior individual action was made, and a record of what occurred when each goal-directed sequence of actions was completed.

Figure 1. The agent framework.



This work is based on the intelligent agent framework illustrated in Figure 1. An agent's *in-coming messages* (the actions of other agents) and *observations* of the effect of its own actions are tagged with the identity of the sending agent and the time received, and are stored in a *repository*. A *world model* contains beliefs of the state of the other agents and the environment, and a *social model* contains beliefs of the state of the agent's *relationships* with the other agents. The agent's *experiential memory* contains complete historic information concerning prior actions and sequences of actions — this is detailed in Section II.

This paper is organised as follows. Various preliminaries are described in Section II as well as the essential features of the agent architecture including the 'social model' that is essential to ecological deliberation. Section III describes expectations of the effect of actions in the experiential memory — these expectations include measures of trust. Section IV describes the ecological deliberative process, and Section V concludes.

978-0-7695-3801-3/09 $26.00 © 2009 IEEE
DOI 10.1109/WI-IAT.2009.158

II. Agent Architecture

We assume that a multiagent system $\{\alpha, \beta_1, \ldots, \beta_o, \xi, \theta_1, \ldots, \theta_t\}$, contains an agent α that interacts with negotiating agents, β_i, and information providing agents, θ_j. We assume that each dialogical interaction takes place within a particular institution that is represented by an *institutional agent*, ξ, [10]. Institutions, or normative systems, play a central role in this work. We will describe an *ontology* that will permit us both to structure the dialogues and to structure the processing of the information gathered by agents. Our agent α has two languages: $\mathcal{C}$ is an illocutionary-based language for communication, and $\mathcal{L}$ is a probabilistic first-order language for internal representation including the representation of its *world model* $\mathcal{M}^t$. $\mathcal{C}$ is described in [11].

Agent α receives all messages expressed in $\mathcal{C}$, they are time-stamped and sourced-stamped, qualified with a subjective belief function $\mathbb{R}^t(\alpha, \beta, \mu)$ that normally decays with time (see below), and are stored in a *repository* $\mathcal{Y}^t$ that contains information concerning every action that α observes — presumably this will include those actions that α takes.

α's *experiential memory* contains a history of what happened when any goal-directed sequence of actions was triggered or when any individual action was observed. First an individual *action experience*, a, consists of:

- the action, a_{act}, i.e. the utterance, the sending and receiving agents, and the time at which the action was taken,
- the trigger, or precondition, that signalled when the action was to be performed, a_{trig},
- any observed effect(s), a_{effect}, i.e. any identifiable responses that are an effect of a_{act} — see Section III.

Then a *sequence experience*, s, consists of:

- the *goal* of the sequence, s_{goal}, that may have been to satisfy a need,
- a sequence of action experiences, $s_{\vec{a}} = (a_i)_{i=1}^n$, where each action experience a_i is described as above,
- beliefs of the prevailing *environment*, s_{env}, that includes: the institutional norms that apply at the time, s_{norm}, the agents involved in the interaction, s_{agents}, and the state of the *social model* between the agents, s_{social}, i.e. $s_{\text{env}} = \{s_{\text{norm}}, s_{\text{agents}}, s_{\text{social}}\}$,
- a *rating*[1] of the outcome of the action sequence, s_{rate}, that enables an ecologically rational agent to develop its repertoire of actions.

α uses the contents of its experiential memory to: reuse successful action sequences, build new sequences from individual actions, and improve prior sequences by using its knowledge of individual action experiences.

The *social model* contains beliefs of the state of α's relationships with other agents — it consists of two components. First, an *intimacy model* that for each agent β consists of α's model of β's private information, *and*, α's model of the private information that β has about α. Second, a *balance model* of the extent of reciprocity between pairs of agents.

Intimacy and balance were first reported in [11] to support argumentative negotiation where they were based on five illocutionary categories. Our requirements here are more general, and the models are quite different but we retain the same names. The spirit of them remains the same: *intimacy* — degree of closeness, and *balance* — degree of fairness. Intimacy is defined in terms of information gain, and balance in terms of ratings.

Private information is categorised first by the type of statement, using a set of illocutionary particles $\mathcal{F}$, and second by the contents of the statement, using the ontology $\mathcal{O}$. A categorising function $\kappa : U \to \mathcal{P}(\mathcal{F})$, where U is the set of utterances, allocates utterances to one or more category in the framework. The power set, $\mathcal{P}(\mathcal{F})$, is required as some utterances belong to multiple categories.

$I^t_{\alpha/\beta}$ is α's model of β's private information; it is represented as real numeric values over $\mathcal{F} \times \mathcal{O}$. Suppose α receives utterance u from β and that category $f \in \kappa(u)$ then: $I^t_{\alpha/\beta(f,c)} = I^{t-1}_{\alpha/\beta(f,c)} + \lambda \times \mathbb{I}(u) \times \text{Sim}(u, c)$ for any $c \in \mathcal{O}$, where λ is the learning rate, $I^t_{\alpha/\beta(f,c)}$ is the intimacy value in the (f, c) position in $\mathcal{F} \times \mathcal{O}$, $\mathbb{I}(u)$ is the Shannon information gain in $\mathcal{M}^t$ due to receiving u, and Sim is a measure of semantic similarity [12]. Additionally, the intimacy model decays in time in any case by $I^t_{\alpha/\beta} = \delta \times I^{t-1}_{\alpha/\beta}$ where $\delta < 1$ and very close to 1 is the decay rate.

$I^t_{\alpha\backslash\beta}$ is α's model of the private information that β has about α. Assuming that confidential information is treated in confidence[2] α will know what β knows about α. This means that the same method can be used to model $I^t_{\alpha\backslash\beta}$ as $I^t_{\alpha/\beta}$ with the exception of estimating $\mathbb{I}(u)$ as it is most unlikely that α will know the precise state of β's world model — for this we resort to the assumption that β's world model mirrors α's and 'estimate' the information gain. Then the *intimacy model* is $I^t_{\alpha\beta} = (I^t_{\alpha/\beta}, I^t_{\alpha\backslash\beta})$. In [11] balance was defined as the element by element numeric difference of $I^t_{\alpha/\beta}$ and $I^t_{\alpha\backslash\beta}$. That definition is not suitable here.

$R^t_{\alpha/\beta}$ is a model of α's aggregated rating of β's actions in assisting α to achieve her goals and satisfy her needs. α will have a variety of goals including the acquisition of goods, information, offering and receiving advice, gossip, and so

[1]This rating is not simply in terms of the extent to which the sequence outcome met the original need, but in a sense that includes the possibility that the other agents involved may have adapted their actions to take account of changes in circumstance that occur during the sequence itself, or even that they went "over the odds" and gave more than was expected of them in some sense. These ratings are on a fuzzy scale from -5 to $+5$ where 0 means "is perfectly acceptable", -5 means "ghastly, completely unacceptable" and $+5$ means "better than I could have dreamed of". Ratings are not a 'utility function' in any sense — they are a subjective assessment of outcomes that is totally dependent on the prevailing state of the environment.

[2]See [13] for a discussion on measuring *confidentiality* i.e. 'information leakage'.

on. These goals are categorised using a set of illocutionary particles $\mathcal{G}$ and the ontology $\mathcal{O}$. Suppose α triggers an action sequence s with goal $g = (k, d)$ when the state of the environment is e and on completion of the sequence rates the outcome as $\rho(\alpha, s, e)$ then:

$$R^t_{\alpha/\beta(k,c)} = R^{t-1}_{\alpha/\beta(k,c)} + \lambda \times \rho(\alpha, s, e) \times \text{Sim}(d, c)$$

for any $c \in \mathcal{O}$, where $\rho(\alpha, s, e)$ is the fuzzy rating of the outcome of s as an integer in the range $[-5, +5]$, λ is the learning rate, $R^t_{\alpha/\beta(k,c)}$ is the aggregated rating in the (k, c) position in $\mathcal{G} \times \mathcal{O}$, and Sim is a semantic similarity function. Additionally, the model decays[3] in time in any case by $R^t_{\alpha/\beta} = \delta \times R^{t-1}_{\alpha/\beta}$ where $\delta < 1$ and very close to 1 is the decay rate.

α should have "a pretty good idea" of how β rates α's actions in assisting β to achieve her goals, and $R^t_{\alpha\backslash\beta}$ models α's estimates of β's rating of α's performance. Then the *balance model* is the pair $R^t_{\alpha\beta} = (R^t_{\alpha/\beta}, R^t_{\alpha\backslash\beta})$. This structure is a historical summary of how α believes it has "done the right thing", or otherwise, by other agents. It also exposes social debts, obligations and opportunities.

III. Expectations

An ecologically rational agent's rationality lies only in its past experience. To behave rationally it will require some expectation, based on that experience, of what other agents will do. Experiential memory records each of the agent's individual experiences; it does not address expectation. We now derive expectations from this historic data. Expectations are considered for the two classes of experience in experiential memory. First, expectations concerning the effect of making a single action (i.e. utterance), second, expectations of the effect of triggering an action sequence.

We consider expectations concerning the effect of triggering an action sequence. Suppose that α triggers an action sequence, s with goal g where the state of the environment is e then we are interested in the rating of the outcome r. Given the rich meaning of the environment it is reasonable to consider:

$$\mathbb{P}(\text{Observe}^{t'}(r) \mid \text{Enact}^t(s), e) \qquad (1)$$

If $\Omega \in e$ is the set of agents in e, then the aggregated rating[4] of their responsive actions leading to the sequence outcome is a subjective measure of their collective *trust*, *honour* or *reliability* — a fuller account of these estimates is given in [8].

We first consider a special case of the expected rating of a diminutive action sequence consisting of a single agent, $\Omega = \{\beta\}$, and a single action — as is observed in the case of "commitment followed by subsequent enactment". In this case if we use the method in [11] to estimate $\mathbb{P}^t_\beta(v|u)$ where u is the commitment and v the enactment then:

$$T_\alpha(\beta, u, e) = \sum_v \rho(\alpha, v, e) \times \mathbb{P}^t_\beta(v|u)$$

Then α's estimate of the *trust*, *honour* or *reliability* of β with respect to a class of utterances U will be:

$$T_\alpha(\beta, U, e) = \sum_{u \in U} T_\alpha(\beta, u, e) \times \mathbb{P}^t_\alpha(u)$$

where $\mathbb{P}^t_\alpha(u)$ is as above.

For action sequences in general we abbreviate the expectation of Equation 1 to $\mathbb{P}^t(r|s, e)$ that we may estimate directly using the same reasoning for estimating $\mathbb{P}^t_\beta(v|u)$ in [11] as r is over a discrete space. Then $T_\alpha(\Omega, s, e) = \mathbb{E}^t_\Omega(r|s, e)$ and $T_\alpha(\Omega, S, e) = \sum_{s\in S} T_\alpha(\Omega, s, e) \times \mathbb{P}^t_\alpha(s)$.

IV. Ecological deliberation

Human agents employ ecological deliberation for all but a very small proportion of the decisions that they make. The neurological processes that enable human non-Cartesian deliberative processes are not well understood. It appears that given a need, contextual triggers somehow retrieve appropriate action sequences from experiential memory. The retrieval process does not require a complete match and operates tentatively when the perceived environment is new, possibly by adapting the action sequence. This is reminiscent of the work of Roger Schank on dynamic memory.

α has the following assets at its disposal to support ecological deliberation:

- an *experiential memory* — Section II
- *expectations* — Section III
- a *world model* — described in [11]
- a *social model* — Section II

Together experiential memory and expectations make a potent pair. Experiential memory contains details of action sequences, and expectations tell us what to expect if those sequences are reused. The world and social models describe the states of affairs that α desires to change.

What is needed is an evolutionary method of some sort — that may well be how humans operate. A problem with evolutionary methods is that we may not be prepared to accept poor performance while the method evolves, although permitting a method to explore and make mistakes may also enable it to discover. A strategy is reported in [14] on how to place all of one's wealth as win-bets indefinitely on successive horse races so as to maximise the rate of growth; this is achieved by proportional gambling, i.e. by betting a proportion of one's wealth on each horse equal to the probability that that horse will win. This result is interesting as the strategy is independent of the betting odds. The situation that we have is not equivalent to the horse

[3] This form of decay means that in the limit all values in the model decay to 0 meaning "is perfectly acceptable". This may appear to be odd, but the model is used only to gauge divergence from the norm; it is *not* used to select a trading partner — that is a job for the trust model.

[4] See Footnote 1.

race, but it is tempting to consider the strategy that selects sequence s_i with probability q_i:

$$q_i = \frac{r(s_i)^c}{\sum_k r(s_k)^c} \tag{2}$$

where $c > 0$ is a real constant that moderates the degree of exploration. This strategy will favour those sequences whose expected performance moderated by a 'super-Sim' function is greater.

Ecological deliberation is by no means the poor relation of its Cartesian brother. When selecting a bottle of wine, some human agents refer to books of ratings and prices and make a constructivist choice, whereas others rely on their merchant to make a choice for them — this choice is purely ecological, its 'rationality' is in the trust that has been built through repeated interaction. It may be that the recommendations of the wine merchant are better in all respects than those that the agent could derive from the data available. If this is so then a rational agent should surely prefer ecological deliberation.

A committed constructivist might respond by saying that clearly the agent should learn as much about wine as the merchant and then everything becomes Cartesian again. Building an agent into a "Mr Know-It-All" is dangerous if it means that the agent believes his knowledge will remain superior in a competitive world to that of other agents, he may then live and decay in a state of sublime ignorance.

A rational agent builds an experiential memory and maintains an open mind on whether to choose constructivist or ecological deliberation. It reinforces the choices it makes by forming a view on which performs better by using its subjective ability to evaluate outcomes.

V. Discussion

The full realisation of the Hayekian vision of self-evolving agents situated in a world of self-evolving institutions is an extensive research agenda that is the subject of ongoing research. For example, there is no clear means of achieving an orderly self-evolution of normative systems in a multi-system context. The contribution of this paper is to describe how a single agent can engage in ecological deliberation in addition to well-understood constructivist deliberation. This enables agents to evolve and adapt their deliberative processes as their environment and their fellow agents evolve.

The social model contains beliefs of the strength of agents' relationships, enables agents to form desires of how those relationships could be, and to form intentions of how to make them so. A possible next step is to experiment with a norm model in a similar fashion. If this can be achieved through ecological deliberation then we will be close to understanding self-evolving electronic institutions that will take multiagent systems technology to a new level.

References

[1] D. Hume, *An Enquiry concerning Human Understanding*, 3rd ed., P. N. Nidditch, Ed. Clarendon Press, Oxford., 1777.

[2] A. Rubinstein, *Modeling Bounded Rationality*. MIT Press, Cambridge, MA, 1998.

[3] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed. Prentice Hall, 2002.

[4] R. A. Brooks, "Intelligence without reason," in *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, R. Myopoulos and J. Reiter, Eds. Sydney, Australia: Morgan Kaufmann, August 1991, pp. 569–595.

[5] S. Russell, "Rationality and intelligence," *Artificial Intelligence*, vol. 94, no. 1-2, pp. 57–77, July 1997.

[6] F. A. Hayek, *The Fatal Conceit: The Errors of Socialism*, W. W. Bartley, Ed. University Of Chicago Press, 1991.

[7] V. L. Smith, *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge University Press, 2007.

[8] C. Sierra and J. Debenham, "Trust and honour in information-based agency," in *Proceedings Fifth International Conference on Autonomous Agents and Multi Agent Systems AAMAS-2006*, P. Stone and G. Weiss, Eds. Hakodate, Japan: ACM Press, New York, May 2006, pp. 1225 – 1232.

[9] ——, "Information-based reputation," in *First International Conference on Reputation: Theory and Technology (ICORE'09)*, M. Paolucci, Ed., Gargonza, Italy, 2009, pp. 5–19.

[10] J. L. Arcos, M. Esteva, P. Noriega, J. A. Rodr´guez, and C. Sierra, "Environment engineering for multiagent systems," *Journal on Engineering Applications of Artificial Intelligence*, vol. 18, 2005.

[11] C. Sierra and J. Debenham, "The LOGIC Negotiation Model," in *Proceedings Sixth International Conference on Autonomous Agents and Multi Agent Systems AAMAS-2007*, Honolulu, Hawai'i, May 2007, pp. 1026–1033.

[12] Y. Li, Z. A. Bandar, and D. McLean, "An approach for measuring semantic similarity between words using multiple information sources," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 4, pp. 871 – 882, July / August 2003.

[13] C. Sierra and J. Debenham, "Information-based deliberation," in *Proceedings Seventh International Conference on Autonomous Agents and Multi Agent Systems AAMAS-2008*, L. Padgham, D. Parkes, J. Müller, and S. Parsons, Eds. Estoril, Portugal: ACM Press, New York, May 2008.

[14] J. J. Kelly, "A new interpretation of information rate," *IEEE Transactions on Information Theory*, vol. 2, no. 3, pp. 185–189, September 1956.