

An Agent for Ecological Deliberation

John Debenham¹ and Carles Sierra²

¹ University of Technology, Sydney, Australia
debenham@it.uts.edu.au

² Institut d'Investigació en Intel·ligència Artificial - IIIA,
Spanish Scientific Research Council, CSIC
08193 Bellaterra, Catalonia, Spain
sierra@iia.csic.es

Abstract. An agent architecture supports the two forms of deliberation used by human agents. 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 agent community to the predominant form of deliberation used by mankind.

1 Introduction

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 [1] and Vernon Smith [2] as being within 'two worlds'. Hayek and Smith identify; *constructivist rationality* that underpins rational predictive models of decision making; and, *ecological rationality* that refers to social institutions and practices that *emerge* from the history of an agent's interactions and are *not* pre-designed.

For intelligent agency we interpret Hayek and Smith's two rationalities as:

- Constructivist. An agent's actions are determined by a theory that may be independent of the particular environment in which the agent is situated, and typically requires access to data.
- Ecological. An agent's actions are the product of prior agents' actions only — deliberation that uses past experience and contextual triggers to build action sequences from experiential memory.

This paper is concerned with the issue generally known as *bounded rationality* that dates back to David Hume [3] 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 [4], and is discussed in all good books on artificial intelligence, e.g. [5].

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 [6]. Despite these advances, work in multiagent systems has been heavily influenced by

game theory, decision theory and logic [7]; this is in contrast to an original motivation for investigating ‘distributed artificial intelligence’ in the mid 1970s where intelligence *emerged* from the interactions between systems.

Why would an agent be motivated to deliberate in a non-constructivist way? First, it may not be aware of a constructivist theory that addresses its goals. Second, it may have difficulty articulating its needs and its context completely and accurately in the theory. Third, the data required by the theory to determine its actions may not be readily available. Fourth, it may not have sufficient time for all this to happen. Fifth, it may favour ecological deliberation simply because it leads to a superior outcome. For example, 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.

The main contribution of this paper is to describe a single agent that exhibits ecological deliberation, we show how it evolves as its experience grows. Various preliminaries are described in Section 2. Section 3 introduces the essential features of the agent architecture including the world model, and a ‘social model’ that is essential to ecological deliberation. Section 4 describes expectations of the effect of actions in the experiential memory— these expectations include measures of trust. Section 5 describes the ecological deliberative process, and Section 6 concludes.

2 Preliminaries

A multiagent system $\{\alpha, \beta_1, \dots, \beta_o, \xi, \theta_1, \dots, \theta_t\}$, 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 , [8]. 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 [9].

An agent’s *in-coming messages* 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 3.

Some messages trigger the agent’s *reactive logic* that overrides other activities. The agent aims to satisfy its *needs* using one of two forms of *deliberation*: *constructivist* (described in [10]) that is based on theories that call on *plans*, and *ecological* that uses past experience and contextual triggers to retrieve or build action sequences from experiential memory.

An *ecologically rational* agent α with need ν in context Θ^t will act using the lottery $E_\alpha(\mathcal{H}_\alpha^t, \nu, \Theta^t) \in \Delta(\mathcal{M} \times \mathcal{B})$ where:

$$E_\alpha : \mathcal{H} \times \mathcal{N} \times \mathcal{I} \rightarrow \Delta(\mathcal{M} \times \mathcal{B}) \quad (1)$$

where E_α is a function that is *not* founded on an abstraction or theory that models, explains, constrains, describes or prescribes the behaviour of agents or the environment; it encapsulates α 's particular ecological rationality. As above, if α has an open dialogue then Equation 1 determines (non-deterministically) the next utterance that α makes in that dialogue. Ecologically rational agents are non-deterministic. The action performed by a *deterministic ecologically rational agent* is determined by: $E_\alpha : \mathcal{H} \times \mathcal{N} \times \mathcal{I} \rightarrow \mathcal{M} \times \mathcal{B}$.

The “ecological rationality” in E_α is based on the belief that the wisdom in \mathcal{H}_α^t can somehow tell α how to act rationally. This belief is not an abstraction or “theory” (as described above); it is simply a belief that the wisdom embedded in prior observations are a basis for rational action. In a simple form, this belief may be that prior agent behaviour reliably indicates future behaviour¹. That is, ecological rationality may be founded on *a sense of trust* that α has in agents, i.e. that they will continue to behave with no less integrity than that which they have displayed in the past. Ecological rationality may also be founded on the reputation that another agent has, or on trust in the institution (i.e. a normative multiagent system to which all the agents belong) to ensure that agents behave in some way [11]. In addition, Ecological rationality may be founded on subtle observations mined from \mathcal{H}_α^t . As a simple example, “Whenever John is running a marketing promotion Carles invariably gives excellent service”. In general, ecologically rational behaviour will be derived from \mathcal{H}_α^t using data mining techniques. Ecological rationality, like trust, is an experience-based phenomenon — it can not exist without experience. At the “top end”, ecological rationality embodies all the models that have been described from information-based agents including the important notion of integrity².

3 Agent Architecture

α acts to satisfy a *need*. Needs either trigger α 's constructivist, goal/plan deliberative reasoning described in [10], or ecological deliberation described in Section 5.

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: (i) the *action*, a_{act} , i.e. the utterance, the sending and receiving agents, and the time at which the action was taken, (ii) the *trigger*, or precondition, that signalled when the action was to be performed, a_{trig} , and (iii) any observed *effect(s)*, a_{effect} , i.e. any identifiable responses that are an effect of a_{act} — see Section 4.

¹ Such a belief may incorporate context. For example, “John is most reliable *except* on Mondays”.

² The extent to which a partner agent will take advantage of his private information when enacting his commitments. E.g. “I haven’t got the strawberries you ordered because they were all rain damaged.”

Then a *sequence experience*, s , consists of: (i) the *goal* of the sequence, s_{goal} , that may have been to satisfy a need, (ii) a sequence of *action experiences*, $s_{\mathbf{a}} = (a_i)_{i=1}^n$, where each action experience a_i is described as above, (iii) 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* (see Section 3.2) between the agents, s_{social} , i.e. $s_{\text{env}} = \{s_{\text{norm}}, s_{\text{agents}}, s_{\text{social}}\}$, (iv) a *rating* of the outcome of the action sequence, s_{rate} , that enables an ecologically rational agent to develop its repertoire of actions.

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.

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 integrity of beliefs derived from observations decreases in time. α may have background knowledge concerning the expected integrity of a belief as $t \rightarrow \infty$. Such background knowledge is represented as a *decay limit distribution*. If the background knowledge is incomplete then one possibility is for α to assume that the decay limit distribution has maximum entropy whilst being consistent with the data. Given an uncertain belief represented as the distribution, $\mathbb{P}(X_i)$, and a decay limit distribution $\mathbb{D}(X_i)$, $\mathbb{P}(X_i)$ decays by:

$$\mathbb{P}^{t+1}(X_i) = \Delta_i(\mathbb{D}(X_i), \mathbb{P}^t(X_i)) \quad (2)$$

where Δ_i is the *decay function* for the X_i satisfying the property that $\lim_{t \rightarrow \infty} \mathbb{P}^t(X_i) = \mathbb{D}(X_i)$. For example, Δ_i could be linear: $\mathbb{P}^{t+1}(X_i) = (1 - \nu_i) \times \mathbb{D}(X_i) + \nu_i \times \mathbb{P}^t(X_i)$, where $\nu_i < 1$ is the decay rate for the i 'th distribution. Either the decay function or the decay limit distribution could also be a function of time: Δ_i^t and $\mathbb{D}^t(X_i)$.

3.1 World Model

In the absence of in-coming messages the integrity of \mathcal{M}^t decays by Equation 2. The following procedure updates \mathcal{M}^t for all utterances expressed in \mathcal{C} . Suppose that α receives a message μ from agent β at time t . Suppose that this message states that something is so with probability z , and suppose that α attaches an epistemic belief $\mathbb{R}^t(\alpha, \beta, \mu)$ to μ — this probability reflects α 's level of personal *caution*. Each of α 's active plans, s , contains constructors for a set of distributions $\{X_i\} \in \mathcal{M}^t$ together with associated *update functions*, $J_s(\cdot)$, such that $J_s^{X_i}(\mu)$ is a set of linear constraints on the posterior distribution for X_i . Examples of these update functions are given in [12]. Denote the prior distribution $\mathbb{P}^t(X_i)$ by \mathbf{p} , and let $\mathbf{p}(\mu)$ be the distribution with

minimum relative entropy³ with respect to \mathbf{p} : $\mathbf{p}_{(\mu)} = \arg \min_{\mathbf{r}} \sum_j r_j \log \frac{r_j}{p_j}$ that satisfies the constraints $J_s^{X_i}(\mu)$. Then let $\mathbf{q}_{(\mu)}$ be the distribution:

$$\mathbf{q}_{(\mu)} = \mathbb{R}^t(\alpha, \beta, \mu) \times \mathbf{p}_{(\mu)} + (1 - \mathbb{R}^t(\alpha, \beta, \mu)) \times \mathbf{p} \tag{3}$$

and then let:

$$\mathbb{P}^t(X_{i(\mu)}) = \begin{cases} \mathbf{q}_{(\mu)} & \mathbf{q}_{(\mu)} \text{ is more interesting than } \mathbf{p} \\ \mathbf{p} & \text{otherwise} \end{cases} \tag{4}$$

A general measure of whether $\mathbf{q}_{(\mu)}$ is ‘more interesting than’ \mathbf{p} is: $\mathbb{K}(\mathbf{q}_{(\mu)} \parallel \mathbb{D}(X_i)) > \mathbb{K}(\mathbf{p} \parallel \mathbb{D}(X_i))$, where $\mathbb{K}(\mathbf{x} \parallel \mathbf{y}) = \sum_j x_j \ln \frac{x_j}{y_j}$ is the Kullback-Leibler distance between two probability distributions \mathbf{x} and \mathbf{y} .

Finally merging Equations 4 and 2 we obtain the method for updating a distribution X_i on receipt of a message μ :

$$\mathbb{P}^{t+1}(X_i) = \Delta_i(\mathbb{D}(X_i), \mathbb{P}^t(X_{i(\mu)})) \tag{5}$$

This procedure deals with integrity decay, and with two probabilities: first, any probability z in the message μ , and second the belief $\mathbb{R}^t(\alpha, \beta, \mu)$ that α attached to μ .

$\mathbb{R}^t(\alpha, \beta, \mu)$ is estimated by measuring the ‘difference’ between μ and its subsequent verification. Suppose that μ is received from agent β at time u and is verified by ξ as μ' at some later time t . Denote the prior $\mathbb{P}^u(X_i)$ by \mathbf{p} . Let $\mathbf{p}_{(\mu)}$ be the posterior minimum relative entropy distribution subject to the constraints $J_s^{X_i}(\mu)$, and let $\mathbf{p}_{(\mu')}$ be that distribution subject to $J_s^{X_i}(\mu')$. We now estimate what $\mathbb{R}^u(\alpha, \beta, \mu)$ should have been in the light of knowing *now*, at time t , that μ should have been μ' .

The idea of Equation 3, is that $\mathbb{R}^t(\alpha, \beta, \mu)$ should be such that, *on average* across \mathcal{M}^t , $\mathbf{q}_{(\mu)}$ will predict $\mathbf{p}_{(\mu')}$ — no matter whether or not μ was used to update the distribution for X_i , as determined by the condition in Equation 4 at time u . The *observed reliability* for μ and distribution X_i , $\mathbb{R}_{X_i}^t(\alpha, \beta, \mu) \mid \mu'$, on the basis of the verification of μ with μ' , is the value of k that minimises the Kullback-Leibler distance:

$$\mathbb{R}_{X_i}^t(\alpha, \beta, \mu) \mid \mu' = \arg \min_k \mathbb{K}(k \cdot \mathbf{p}_{(\mu)} + (1 - k) \cdot \mathbf{p} \parallel \mathbf{p}_{(\mu')})$$

3.2 Social Model

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 [9] 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

³ Entropy-based inference is a form of Bayesian inference that is convenient when the data is sparse [13] and encapsulates common-sense reasoning.

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 $\kappa : U \rightarrow \mathcal{P}(\mathcal{F})$, 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.

$I_{\alpha/\beta}^t$ 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: $I_{\alpha/\beta(f,c)}^t = I_{\alpha/\beta(f,c)}^{t-1} + \lambda \times \mathbb{I}(u) \times \text{Sim}(u, c)$ for any $c \in \mathcal{O}$, where $\text{Sim}(\cdot)$ is a semantic similarity function [14], λ is the learning rate, $I_{\alpha/\beta(f,c)}^t$ 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 using Equation 5, and Sim is as above. Additionally, the intimacy model decays in time in any case by $I_{\alpha/\beta}^t = \delta \times I_{\alpha/\beta}^{t-1}$ where $\delta < 1$ and very close to 1 is the decay rate.

$I_{\alpha \setminus \beta}^t$ is α 's model of the private information that β has about α . Assuming that confidential information is treated in confidence⁴ α will know what β knows about α . This means that the same method can be used to model $I_{\alpha \setminus \beta}^t$ as $I_{\alpha/\beta}^t$ 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 $I_{\alpha\beta}^t = (I_{\alpha/\beta}^t, I_{\alpha \setminus \beta}^t)$. In [9] balance was defined as the element by element numeric difference of $I_{\alpha/\beta}^t$ and $I_{\alpha \setminus \beta}^t$. That definition is not suitable here.

$R_{\alpha/\beta}^t$ 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 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:

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

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, $R_{\alpha/\beta(k,c)}^t$ is the aggregated rating in the (k, c) position in $\mathcal{G} \times \mathcal{O}$, and Sim is as above. Additionally, the model decays in time in any case by $R_{\alpha/\beta}^t = \delta \times R_{\alpha/\beta}^{t-1}$ where $\delta < 1$ and very close to 1 is the decay rate. 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.

α should have “a pretty good idea” of how β rates α 's actions in assisting β to achieve her goals, and $R_{\alpha \setminus \beta}^t$ models α 's estimates of β 's rating of α 's performance. Then the *balance model* is the pair $R_{\alpha\beta}^t = (R_{\alpha/\beta}^t, R_{\alpha \setminus \beta}^t)$. This structure is a historical summary

⁴ See [10] for a discussion on measuring *confidentiality* i.e. ‘information leakage’.

of how α believes it has “done the right thing”, or otherwise, by other agents. It also exposes social debts, obligations and opportunities.

4 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.

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, as described in Section 3, it is reasonable to consider:

$$\mathbb{P}(\text{Observe}^{t'}(r) \mid \text{Enact}^t(s), e) \quad (6)$$

If $\Omega \in e$ is the set of agents in e , then the aggregated rating 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 [12].

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 estimate $\mathbb{P}_\beta^t(v|u)$ where u is the commitment and v the enactment then:

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

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

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

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

For action sequences in general we abbreviate the expectation of Equation 6 to $\mathbb{P}^t(r|s, e)$ that we may estimate directly using the same reasoning for estimating $\mathbb{P}_\beta^t(v|u)$ because r is over a discrete space. Then $T_\alpha(\Omega, s, e) = \mathbb{E}_\Omega^t(r|s, e)$ and $T_\alpha(\Omega, S, e) = \sum_{s \in S} T_\alpha(\Omega, s, e) \times \mathbb{P}_\alpha^t(s)$.

5 Ecological Deliberation

Human agents employ ecological deliberation for all but a very small proportion of the decisions that they make [15]. 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.

When an agent is ‘born’ it will have no experiential memory, and its only rational basis for deliberative action will be either through pre-programmed constructivist deliberation or by imitating a ‘parent’ or ‘teacher’. Thereafter, whenever it acts it will observe the effects of its actions and its experiential memory will expand.

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

- an *experiential memory* — Section 3
- *expectations* — Section 4
- a *world model* — Section 3.1
- a *social model* — Section 3.2

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.

An agent acts to satisfy its needs. An ecological agent’s rationality lies in the trust that it has of others. This means that an ecological agent’s desires should address its social needs as well as its material needs — these may not be compatible. And this means that the actions that an ecological agent takes should attempt to shape its social model as well as its world model⁵. An agent’s social structures, and the structures of the institutions that it inhabits, are its means to transcend its individual ability.

Rather than give a tedious description of how each of the above operations may be performed we simply assume that they all have been, and that we are confronted with an enormous selection of previous, improved, adapted, simplified and created action sequences.

Our problem then is: given a current need, the current norm state, and the current states of the world and social models, to select one sequence. We deal with the complexity of matching the current goal and environment to those of previously observed sequences with a ‘super-Sim’ function that moderates the expected rating (Section 4) of each previously recorded sequence, s , to give expectations of the rating, $r(s) \in [0, 1]$, of how that sequence would perform if it was reused now for the current need.

Given that we now face the problem of devising a method that selects an action sequence it is worth considering first what we expect of that method. What it should *not* do is to say “That one is the best choice” that is pure constructivism — it says “Carles and John have greater knowledge than can ever emerge from the environment”. Worse still it would mean that by determining the agent’s actions it would then pervert the agent’s experiential memory for ever more⁶.

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.

⁵ In future work we propose to address how it should also attempt to shape the norms of the institutions that it inhabits.

⁶ Unfortunately this complication also applies to the definition of ‘super-Sim’.

The point of this digression is to excuse ourselves for presenting only a ‘quasi-ecological’ method that permits the agent to explore whilst guiding it in an apparently sensible direction. A strategy is reported in [16] 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 race, but it is tempting to consider the strategy that selects sequence s_i with probability q_i :

$$q_i = \frac{r(s_i)^c}{\sum_k r(s_k)^c} \quad (7)$$

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

Finally we consider how an agent combines constructivist and ecological deliberation.

Ecological deliberation is by no means the poor relation of its Cartesian brother. Referring back to the ‘wine merchant’ example in Section 1, it may simply 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. Creating a “Mr Know-It-All” agent 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.

6 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 on-going 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. Hayek, F.A.: *The Fatal Conceit: The Errors of Socialism*. University Of Chicago Press (1991)
2. Smith, V.L.: *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge University Press, Cambridge (2007)
3. Hume, D.: *An Enquiry concerning Human Understanding*, 3rd edn, 1975 edn. Clarendon Press, Oxford (1977)
4. Rubinstein, A.: *Modeling Bounded Rationality*. MIT Press, Cambridge (1998)
5. Russell, S., Norvig, P.: *Artificial Intelligence: A Modern Approach*, 2nd edn. Prentice-Hall, Englewood Cliffs (2002)
6. Brooks, R.A.: Intelligence without reason. In: Myopoulos, R., Reiter, J. (eds.) *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, Sydney, Australia, pp. 569–595. Morgan Kaufmann, San Francisco (1991)
7. Russell, S.: Rationality and intelligence. *Artificial Intelligence* 94, 57–77 (1997)
8. Arcos, J.L., Esteva, M., Noriega, P., Rodríguez, J.A., Sierra, C.: Environment engineering for multiagent systems. *Journal on Engineering Applications of Artificial Intelligence* 18 (2005)
9. Sierra, C., Debenham, J.: The LOGIC Negotiation Model. In: *Proceedings Sixth International Conference on Autonomous Agents and Multi Agent Systems AAMAS-2007*, Honolulu, Hawai'i, pp. 1026–1033 (2007)
10. Sierra, C., Debenham, J.: Information-based deliberation. In: Padgham, L., Parkes, D., Müller, J., Parsons, S. (eds.) *Proceedings Seventh International Conference on Autonomous Agents and Multi Agent Systems AAMAS-2008*, Estoril, Portugal. ACM Press, New York (2008)
11. Sierra, C., Debenham, J.: Information-based reputation. In: Paolucci, M. (ed.) *First International Conference on Reputation: Theory and Technology (ICORE 2009)*, Gargonza, Italy, pp. 5–19 (2009)
12. Sierra, C., Debenham, J.: Trust and honour in information-based agency. In: Stone, P., Weiss, G. (eds.) *Proceedings Fifth International Conference on Autonomous Agents and Multi Agent Systems AAMAS-2006*. pp. 1225–1232. ACM Press, New York (2006)
13. Cheeseman, P., Stutz, J.: On The Relationship between Bayesian and Maximum Entropy Inference. In: *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, pp. 445–461. American Institute of Physics, Melville (2004)
14. Li, Y., Bandar, Z.A., McLean, D.: An approach for measuring semantic similarity between words using multiple information sources. *IEEE Transactions on Knowledge and Data Engineering* 15, 871–882 (2003)
15. Levine, D.S.: Brain mechanisms for making, breaking, and changing rules. In: Huang II, D.S., Levine, D.C.W., Levine, D.S., Jo, K.H. (eds.) *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques*, pp. 345–355. Springer, Heidelberg (2008)
16. Kelly, J.J.: A new interpretation of information rate. *IEEE Transactions on Information Theory* 2, 185–189 (1956)