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.**

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 is in contrast to an original motivation for investigating 'distributed artificial intelligence' in the mid 1970s where intelligence *emerged* from the interactions between systems.

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'.

The work of Hayek and Smith is concerned with human agents and with economic agents, and not directly with computerised, intelligent agents where the predominant logical distinction is between *deliberative* and *reactive* logic. Hayek and Smith's two rationalities lead to two distinct forms of deliberation, and has little directly to do with autonomic reactivity that typically overrides other processes in both the human neuropsychological system and in intelligent agents; for example, when we touch something that is very hot.

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.

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.
- Ecological. An agent's actions are the product of prior agents' actions only — this includes observations that an agent has made of its environment.

As the name suggests, ecological rationality is concerned with a richer form of bounded rationality than simplifying the calculation of a theoretically 'optimal' action by: rules for simplifying search, rules for terminating search or heuristic decision rules to select actions from an incomplete set of options [8]. Ecological rationality is taken in the context of the Hayekian view [6] in which agents evolve themselves together with the norms[1] of the systems they inhabit[2] whilst their environment changes. Norms, that may well be incompatible with an agent's goals, enable agents to interact in an orderly way, and to achieve goals through collective deliberation that they could not achieve individually. This evolutionary collective deliberation may override personal instinctual behaviours such as indiscriminate greed driven by a desire to survive; it enables agents to become, and to remain, civilised. In this context, *ecological rationality* is deliberation that uses past experience and contextual triggers

---

[1]Hayek uses the term 'order' that refers to: traditions, customs, norms, rules and guidelines. An agent may belong to a number of normative systems (or, electronic institutions [9]) whose norms may be shared with, or in conflict with, those of other systems. The 'extended order' includes the whole show. If a multiagent system interacts with human society then its norms will respect the rules and laws that apply to society as a whole.

[2]The evolution of individual agents and component systems are not considered in isolation — the whole ensemble evolves in response to itself and to the environment — they are *complex* systems. For example, in Hayek's extensive writing there is little mention of ethics as it too evolves.

to build action sequences from experiential memory.

Past experience is a precursor to ecological rationality. For example, as we have described them previously, trust and honour [10] and reputation [11], 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. The extent of an agent's experiential memory will be limited by the extent to which it can observe actions performed by other agents and can gauge their effect.

Building action sequences from experiential memory involves more than just retrieval. An agent has: to learn to imitate the actions that it believes that others do, to form expectations of the effect of actions, to select actions from a set of candidates, to adapt actions to suit the current norms and state of the environment, and when things don't work out to learn to experiment with untested actions.

The co-evolution of agents and systems is central to Hayek's view of how things should be [6]. The main driver for this co-evolutionary process is competition. This does not mean that co-operation is bad. Creating or joining co-operative sub-systems are the means by which agents achieve goals beyond their individual capacity, and then these co-operative sub-systems attain their goals in the competitive world in which they are situated. Such competition requires that the agents that comprise them, and the norms that bind them, must co-evolve for such a co-operative sub-system to survive. This all sounds rather Darwinian, but Hayek is careful to distinguish between genetic evolution and cultural evolution [op. cit. page 23].

Hayek places constructivist theories in their context — they are produced through a cultural evolutionary process by human agents and by the norms by which they abide [6]. He exposes the fallacy in using theories to *control* agents and systems by: calling "on human reason to seize the reigns and control future development" [op. cit. page 22].

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[3]. 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,

[3]For example, the agent may desire to act so as to strengthen, or weaken, a relationship with a particular agent, perhaps to discharge or generate some social obligation, or it may desire to act so that it is seen to be behaving a particular way, perhaps by apparently behaving altruistically — there may not be a theory that satisfactorily balances these desires with more mundane desires concerning the effect of the actions that it can take.

Figure 1.   The agent framework.
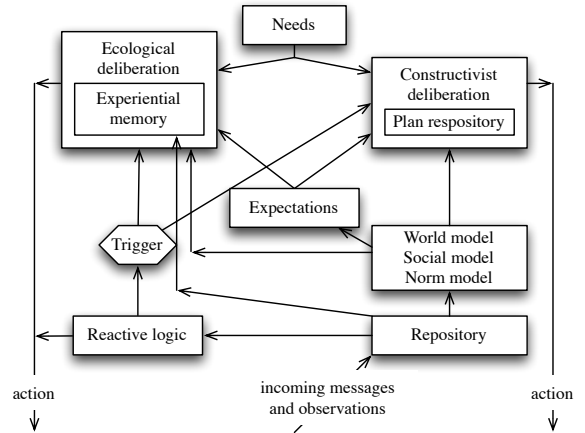
whereas others rely on their merchant to make a choice for them — this choice is purely ecological, its 'rationality' is in the trust that has been built through repeated interaction.

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

II. PRELIMINARIES

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 III.

Some messages trigger the agent's *reactive logic* that overrides other activities and may cause an *action* to be performed or may *trigger* further deliberative processes. Summarising techniques are used to distil the large number of incoming messages into summary *expectations* of the effect of actions including: trust, honour and reliability. These expectations may be used by the agent's constructivist deliberation, and are vital to its ecological deliberation. The agent aims to satisfy its *needs* using one of two forms of *deliberation*: *constructivist* (described in [12]) 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.

This paper draws from our work on information-based agency [13] that is well-suited to this purpose. It supports rich models of inter-agent relationships [14] that are a quintessential feature of emergent behaviour between agents, it supports rich models of trust, honour and reliability [10] that provide the rationale for ecologically rational behaviour, it includes a generate and test approach to planning [12], additionally it uses tools from information theory to manage uncertainty in a nice way. 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. The evolution of norms, institutions and systems will be the subject of future work.

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, [9]. 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 [14].

In order to structure agent dialogues we need an ontology that includes a (minimum) repertoire of elements: a set of *concepts* (e.g. quantity, quality, material) organised in a is-a hierarchy (e.g. platypus is a mammal, australian-dollar is a currency), and a set of relations over these concepts (e.g. price(beer,Australian-dollar)).[4] We model ontologies following an algebraic approach [15].

An ontology is a tuple $\mathcal{O} = (C, R, \leq, \sigma)$ where:

1) C is a finite set of concept symbols (including basic data types);
2) R is a finite set of relation symbols;
3) $\leq$ is a reflexive, transitive and anti-symmetric relation on C (a partial order);
4) $\sigma : R \to C^+$ is the function assigning to each relation symbol its arity.

where $\leq$ is a traditional *is-a* hierarchy, and R contains relations between the concepts in the hierarchy.

The concepts within an ontology are closer, semantically speaking, depending on how far away are they in the structure defined by the $\leq$ relation. Semantic distance plays a fundamental role in this work. A measure [16] bases the *semantic similarity* between two concepts on the path length induced by $\leq$ (more distance in the $\leq$ graph means less semantic similarity), and the *depth* of the subsumer concept

(common ancestor) in the shortest path between the two concepts (the deeper in the hierarchy, the closer the meaning of the concepts). Semantic similarity is then defined as:

$$\text{Sim}(c, c') = e^{-\kappa_1 l} \cdot \frac{e^{\kappa_2 h} - e^{-\kappa_2 h}}{e^{\kappa_2 h} + e^{-\kappa_2 h}}$$

where l is the length (i.e. number of hops) of the shortest path between the concepts, h is the depth of the deepest concept subsuming both concepts, and κ_1 and κ_2 are parameters scaling the contribution of shortest path length and depth respectively.

Given a formula $\varphi \in \mathcal{C}$ in the communication language we define the vocabulary or *ontological context* of the formula, $O(\varphi)$, as the set of concepts in the ontology used in it. Thus, we extend the previous definition of similarity to sets of concepts in the following way:

$$\text{Sim}(\varphi, \psi) = \max_{c_i \in O(\varphi)} \min_{c_j \in O(\psi)} \{\text{Sim}(c_i, c_j)\} \qquad (1)$$

These definitions of semantic similarity are based only on the structure of the ontology, and are a first approximation to 'semantic distance' in a rich sense.

III. Agent Architecture

α acts by delivering utterances, and observes by receiving them. α acts to satisfy a *need* that may be exogenous such as a need to trade profitably, triggered by another agent's actions, or endogenous such as α deciding that it owns more wine than it requires. Needs either trigger α's constructivist, goal/plan deliberative reasoning described in [12], or ecological deliberation described in Section V.

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[5] 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}[6], i.e. any identifiable responses that are an effect of a_{act} — see Section IV.

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

---

[4]Usually, a set of axioms defined over the concepts and relations is also required. We will omit this here.

[5]Practicality is not a concern here.

[6]These may be difficult to identify precisely, but recording effects is considerably more economical than recording posterior world states.

- 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* (see Section III-B) between the agents, s_{social}, i.e. $s_{\text{env}} = \{s_{\text{norm}}, s_{\text{agents}}, s_{\text{social}}\}$,
- a *rating*[7] 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 integrity of beliefs derived from observations decreases in time. α may have background knowledge concerning the expected integrity of a belief as $t \to \infty$. 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:

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

where Δ_i is the *decay function* for the X_i satisfying the property that $\lim_{t \to \infty} \mathbb{P}^t(X_i) = \mathbb{D}(X_i)$. 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 $\mathbb{D}^t(X_i)$.

A. 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*.

[7]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.

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 Section IV-A. Denote the prior distribution $\mathbb{P}^t(X_i)$ by $\vec{p}$, and let $\vec{p}_{(\mu)}$ be the distribution with minimum relative entropy[8] with respect to $\vec{p}$: $\vec{p}_{(\mu)} = \arg\min_{\vec{r}} \sum_j r_j \log \frac{r_j}{p_j}$ that satisfies the constraints $J_s^{X_i}(\mu)$. Then let $\vec{q}_{(\mu)}$ be the distribution:

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

and then let:

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

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

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

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

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 $\vec{p}$. Let $\vec{p}_{(\mu)}$ be the posterior minimum relative entropy distribution subject to the constraints $J_s^{X_i}(\mu)$, and let $\vec{p}_{(\mu')}$ 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 $\mathcal{M}^t$, $\vec{q}_{(\mu)}$ will predict $\vec{p}_{(\mu')}$ — 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, $\mathbb{R}_{X_i}^t(\alpha, \beta, \mu) | \mu'$, on the basis of the verification of μ with μ', is the value of k that minimises the Kullback-Leibler distance:

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

B. 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

[8]Entropy-based inference is a form of Bayesian inference that is convenient when the data is sparse [17] and encapsulates common-sense reasoning [18].

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 [14] 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 $\kappa : U \rightarrow \mathcal{P}(\mathcal{F})$, 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: $I_{\alpha/\beta(f,c)}^t = I_{\alpha/\beta(f,c)}^{t-1} + \lambda \times \mathbb{I}(u) \times \mathrm{Sim}(u,c)$ for any $c \in \mathcal{O}$, where 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 in Equation 1. 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.

$I_{\alpha\backslash\beta}^t$ is α's model of the private information that β has about α. Assuming that confidential information is treated in confidence[9] α will know what β knows about α. This means that the same method can be used to model $I_{\alpha\backslash\beta}^t$ 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 $I_{\alpha\beta}^t = (I_{\alpha/\beta}^t, I_{\alpha\backslash\beta}^t)$. In [14] balance was defined as the element by element numeric difference of $I_{\alpha/\beta}^t$ and $I_{\alpha\backslash\beta}^t$. 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:

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

for any $c \in \mathcal{O}$, where $\rho(\alpha, s, e)$ is the fuzzy rating of the

[9]See [12] for a discussion on measuring *confidentiality* i.e. 'information leakage'.

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 in Equation 1. Additionally, the model decays[10] 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.

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

IV. 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.

A. Expected effect of a single action

We consider expectations concerning the effect of making a single action; that is, the expected a_{effect} given a_{act}. To make this problem tractable we consider only utterances for which a particular *form* of response is expected. For example, "what is the time?" or "send me a bottle of Protos[11]". For these utterances α utters u and expects to observe utterances, v, from a particular set of agents, Ω, and of a form from the set F. α's expectations are that:

$$\forall u \in U \cdot \mathrm{Enact}_\alpha^t(u) \rightarrow \forall \beta \in \Omega \cdot \exists v \in U \cdot \exists w \in F$$
$$(\mathrm{Observe}_\alpha^{t_\beta}(\mathrm{Enact}_\beta(v)) \wedge \mathrm{In}(v, w) \wedge \mathrm{Form}(u, \beta, w)) \quad (6)$$

where $\mathrm{Form}(u, \beta, w)$ means that w is a form of response that α expects having uttered u, $\mathrm{In}(v, w)$ means that v is an instantiation of w, and $t_\beta > t$. For example, u could be "what is the price of Protos", w could be "the price of Protos is x", and v could be "the price of Protos is €40". [19] describes a neat way of dealing with instantiation using a refinement relation.

For each agent $\beta \in \Omega$ we abbreviate the expectation of Equation 6 to $\mathbb{P}_\beta^t(v|u)$. In the absence of in-coming messages the conditional probabilities, $\mathbb{P}_\beta^t(v|u)$, should tend to ignorance as represented by the *decay limit distribution* and Equation 2. We now show how Equation 5 may be used

[10]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.

[11]A fine wine from the 'Ribera del Duero' region, Spain.

to revise $\mathbb{P}^t(v|u)$ as observations are made. Let the set of possible utterances be $\Phi = \{v_1, v_2, \ldots, v_m\}$ with prior distribution $\vec{p} = \mathbb{P}^t_\beta(v|u)$. Suppose that message w is received, we estimate the posterior $\vec{p}_{(w)} = (p_{(w)i})^m_{i=1} = \mathbb{P}^{t+1}(v|u)$.

First, given the expectation $\mathbb{P}^t_\beta(v|u)$, if α observes that β utters v_k then α may use this observation to estimate $p_{(v_k)k}$ as some value d at time $t+1$. We estimate the distribution $\vec{p}_{(v_k)}$ by applying the principle of minimum relative entropy as in Equation 5 with prior $\vec{p}$, and the posterior $\vec{p}_{(v_k)} = (p_{(v_k)j})^m_{j=1}$ satisfying the single constraint: $J^{(v|u)}(v_k) = \{p_{(v_k)k} = d\}$.

Second, α may use the above observation to revise $\mathbb{P}^t_\beta(v'|u')$ when u and u' are semantically close in the sense of Equation 1. For example, u could be "please send me a chicken on Tuesday" and u' could be "please send me a duck on Thursday". Following the notation above this is achieved by: $J^{(v'|u')}(v_k) = \{p_{(v_k)k} = d \times g(\mathrm{Sim}(u, u'))\}$ *provided that:* $d \times g(\mathrm{Sim}(u, u')) > p_k$, where g is a function that moderates the values of the Sim function, and p_k is the prior value. Equation 4 will ensure that this update process only applies when $d \times g(\mathrm{Sim}(u, u'))$ is sufficiently large to deliver positive information gain to $\mathbb{P}^{t+1}_\beta(v'|u')$.

The entropy $\mathbb{H}^t_\beta(v|u)$ estimates α's uncertainty in β's response given that α has uttered u. α may interact with more than one agent. Suppose that agent γ is an ideal agent who always responds impeccably then:

$$T_\alpha(\beta, \gamma, u) = 1 - \sum_v \mathbb{P}^t_\gamma(v|u) \log \frac{\mathbb{P}^t_\gamma(v|u)}{\mathbb{P}^t_\beta(v|u)}$$

measures the relative entropy between this ideal distribution, $\mathbb{P}^t_\gamma(v|u)$, and the distribution of β's expected actions, $\mathbb{P}^t_\beta(v|u)$, where the "1" is an arbitrarily chosen constant being the maximum value that this measure may have. This estimate is with respect to a single u. It makes sense to aggregate these values over a class of utterances, say over those u that are in the ontological context o, that is $u \leq o$:

$$T_\alpha(\beta, \gamma, o) = 1 - \frac{\sum_{u:u \leq o} \mathbb{P}^t_\alpha(u)\,[1 - M_\alpha(\beta, \gamma, u)]}{\sum_{u:u \leq o} \mathbb{P}^t_\beta(u)}$$

where $\mathbb{P}^t_\alpha(u)$ is a probability distribution over the space of utterances that α's next utterance to β is u.

B. Expected rating of 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, as described in Section III, it is reasonable to consider:

$$\mathbb{P}(\mathrm{Observe}^{t'}(r) \mid \mathrm{Enact}^t(s), e) \qquad (7)$$

If $\Omega \in e$ is the set of agents in e, then the aggregated rating[12] of their responsive actions leading to the sequence outcome

[12]See Footnote 7.

is a subjective measure of their collective *trust*, *honour* or *reliability* — a fuller account of these estimates is given in [10].

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 of Section IV-A 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 7 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 Section IV-A as r is over a discrete space. Then $T_\alpha(\Omega, s, e) = \mathbb{E}^t_\Omega(r|s, e)$ and $T_\alpha(\Omega, S, e) = \sum_{s \in S} T_\alpha(\Omega, s, e) \times \mathbb{P}^t_\alpha(s)$.

V. 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.

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 III
- *expectations* — Section IV
- a *world model* — Section III-A
- a *social model* — Section III-B

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 model[13]. An agent's social structures, and the structures of the institutions that it inhabits, are its means to transcend its individual ability.

An agent will make an ecologically rational deliberative action by:

- reusing an existing action sequence[14],
- improving an existing action sequence by referring to prior action experiences,
- adapting an existing action sequence when its norms or initial state are dissimilar to those prevailing,
- simplifying an existing action sequence by delegating responsibility for part of it to another (trusted) agent,
- experimenting — possibly by attempting to second-guess the rationale behind other agents' actions

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 IV-B) 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 more[15].

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.

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 [20] 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:

$$q_i = \frac{r(s_i)^c}{\sum_k r(s_k)^c} \qquad (8)$$

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

A. Overall Strategy

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 I, 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. 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.

VI. 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.

---

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

[14]In case this appears to be a simple application of case-based-reasoning-style case retrieval, note the complexity of the all important environment. The devil is in the environment.

[15]Unfortunately this complication also applies to the definition of 'super-Sim'.

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] G. Gigerenzer and R. Selten, Eds., *Bounded Rationality —- The Adaptive Toolbox*. MIT Press, Cambridge, MA, 2002.

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

[10] 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.

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

[12] ——, "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.

[13] ——, "Information-based agency," in *Proceedings of Twentieth International Joint Conference on Artificial Intelligence IJCAI-07*, Hyderabad, India, January 2007, pp. 1513–1518.

[14] ——, "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.

[15] Y. Kalfoglou and M. Schorlemmer, "IF-Map: An ontology-mapping method based on information-flow theory," in *Journal on Data Semantics I*, ser. Lecture Notes in Computer Science, S. Spaccapietra, S. March, and K. Aberer, Eds. Springer-Verlag: Heidelberg, Germany, 2003, vol. 2800, pp. 98–127.

[16] 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.

[17] P. Cheeseman and J. Stutz, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Melville, NY, USA: American Institute of Physics, 2004, ch. On The Relationship between Bayesian and Maximum Entropy Inference, pp. 445 – 461.

[18] J. Paris, "Common sense and maximum entropy," *Synthese*, vol. 117, no. 1, pp. 75 – 93, 1999.

[19] A. Fàbregues, J. Madrenas, C. Sierra, and J. Debenham, "Supplier performance in a digital ecosystem," in *Proceedings of the IEEE International Conference on Digital Ecosystems and Technologies (IEEE-DEST 2009)*, Istanbul, Turkey, 2009.

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