Negotiating Intelligently

John Debenham & Simeon Simoff

Faculty of IT, University of Technology, Sydney, Australia.
{debenham,simeon}@it.uts.edu.au
WWW home page: http://e-markets.org.au/

Abstract. The predominant approaches to automating competitive interaction appeal to the central notion of a utility function that represents an agent's preferences. Agent's are then endowed with machinery that enables them to perform actions that are intended to optimise their expected utility. Despite the extent of this work, the deployment of automatic negotiating agents in real world scenarios is rare. We propose that utility functions, or preference orderings, are often not known with certainty; further, the uncertainty that underpins them is typically in a state of flux. We propose that the key to building intelligent negotiating agents is to take an agent's historic observations as primitive, to model that agent's changing uncertainty in that information, and to use that model as the foundation for the agent's reasoning. We describe an agent architecture, with an attendant theory, that is based on that model. In this approach, the utility of contracts, and the trust and reliability of a trading partner are intermediate concepts that an agent may estimate from its information model. This enables us to describe intelligent agents that are not necessarily utility optimisers, that value information as a commodity, and that build relationships with other agents through the trusted exchange of information as well as contracts.

1 Introduction

The potential value of the e-business market — including e-procurement — is enormous. Given that motivation and the current state of technical development it is surprising that a comparatively small amount of automated negotiation is presently deployed.[1] Technologies that support automated negotiation include multiagent systems and virtual institutions, game theory and decision theory.

Multiagent systems technology [1] — autonomous, intelligent and flexible software agents — provides the basis for constructing automated negotiation agents. Of particular significance are deliberative architectures that incorporate proactive planning systems [2]. *Virtual Institutions* are software systems composed of autonomous agents, that interact according to predefined conventions on language and protocol, and that guarantee that certain norms of behaviour are enforced. Norms may be specified that regulate the interactions, and so protect the integrity of the commitments exchanged. A Virtual Institution is in

[1] Auction bots such as those on eBay, and automated auction houses do a useful job, but do not automate negotiation in the sense described here.

a sense a natural extension of the social concept of institutions as regulatory systems that shape human interactions [3].

Game theory tells an agent what to do, and what outcome to expect, in many well-known negotiation situations [4], but these strategies and expectations are derived from assumptions about the agent's preferences and about the preferences of the opponent. One-to-one negotiation is generally known as *bargaining* [5] — it is the natural negotiation form when the negotiation object comprises a number of issues. For example, in bargaining over the supply of steel issues could include: the quantity of the steel, the quality of the steel, the delivery schedule, the settlement schedule and, of course, the price. Beyond bargaining there is a wealth of material on the theory of *auctions* [6] for one-to-many negotiation, and *exchanges* for many-to-many negotiation. Fundamental to this analysis is the central role of the utility function, and the notion of rational behaviour by which an agent aims to optimise its utility, when it is able to do so, and to optimise its *expected* utility otherwise.

We propose that utility functions, or preference orderings, are often not known with certainty; further, the uncertainty that underpins them is typically in a state of flux. We propose that the key to building intelligent negotiating agents is to take an agent's historic observations as primitive, to model that agent's changing uncertainty in that information, and to use that model as the foundation for the agent's reasoning. We call such agents *information-based agents*. In Sec. 2 we describe the ideas behind these agents, and formalise these ideas in Sec. 3. Sec. 4 relates commitments made to their eventual execution, this leads to a formalisation of trust. Strategies for information-based agents are discussed in Sec. 5.

2 The foundation of information-based agents

We discuss the issues that an intelligent negotiating agent should be designed to address, and give informal descriptions of how information-based agents address those issues. This section provides the rationale for the formal work that follows.

Percepts, the content of messages, are all that an agent has to inform it about the world and other agents. The validity of percepts will always be uncertain due to the reliability of the sender of the message, and to the period that has elapsed since the message arrived. Further, the belief that an agent has in the validity of a percept will be determined by the agent's level of individual *caution*. The information-based agent's *world model* is deduced from the percepts using *inference rules* that transform percepts into statements in probabilistic logic. These rules are peculiar to the agent. These world models are expressed in first-order logic where the validity of statements is expressed as a probability distribution over some valuation space that is not necessarily true/false and may have no natural ordering. The agent's personal *caution* is a component of these epistemic probabilities.

The integrity of percepts decreases in time. The way in which it decreases will be determined by the type of the percept, as well as by the issues listed in the previous paragraph including 'caution'. An agent may have background knowledge concerning the expected integrity of a percept as $t \to \infty$. Information-based agents represent this background knowledge as a *decay limit distribution*. If the background knowledge is incomplete then one possibility for an agent is to assume that the decay limit distribution has maximum entropy [7] whilst being consistent with the data.

All messages are valueless unless their integrity can be verified to some degree at a later time, perhaps for a cost. To deal with this issue we employ an *institution agent* that always reports promptly and honestly on the execution of all commitments, forecasts, promises and obligations. This provides a conveniently simple solution to the integrity verification issue. The institution agent also takes care of "who owns what". This enables the agents to negotiate and to evaluate the execution of commitments by simple message passing.

An agent's percepts generally constitute a sparse data set whose elements have differing integrity. An agent may wish to induce tentative conclusions from this sparse and uncertain data of changing integrity. Percepts are transformed by inference rules into statements in probabilistic logic as described above. Information-based agents may employ entropy-based logic [8] to induce complete probability distributions from those statements. This logic is consistent with the laws of probability, but the results derived assume that the data is 'all that there is to know' — Watt's Assumption [9].

It may be necessary to establish the *bona fide* of a commitment by including a 'default' obligation that an agent is forced to perform by an authority external to the multiagent system in the event that a commitment is not executed correctly. We do not address this issue here, and assume that defaults can be specified, and that some legal mechanism can be invoked if necessary.

In considering a proposed commitment or promise, an agent evaluates that commitment in some way, perhaps in utilitarian terms. Information-based agents express evaluations as probability distributions over some evaluation space that may not be totally ordered. An *evaluation space* is a complete disjoint set whose elements may be qualitative. An evaluation space may have a lattice ordering that represents an agent's known preferences. An agent may not have sufficient information to determine its preferences. If subjective judgement is required then an agent may be ambivalent. If there is a temporal consideration, such as in selecting a trading partner for a long-term relationship, then an agent may have no clear preference. If a negotiation object involves multiple issues then an agent may know its preference along each dimension, but not across the entire space. If an agent's goals are expressed at a high-level, such as 'the goal of this organisation is to achieve world domination', then such a goal may not translate into preferences that enable an agent to choose between two different types of pencil to purchase, for example. Further, preferences, like everything else, are derived from the agent's information and so may change in time — this may mean that an attempt to be a 'utility optimiser' is flawed by the fact that the utility function is uncertain, and its uncertainty is in a state of flux. To place utility optimisation in perspective — it aims to optimise an inherently uncertain function of permanently changing integrity that takes no account of the value of information exchange.

Information is strategic. It has value as information. With the exception of highly cautious messages such as "make me an offer" in simple bargaining, *everything* that an agent says gives away information. This is a central issue here — the exchange of information is a strategic component of competitive interaction. Information-based agents evaluate information received as the reduction in entropy of the agent's current world model. Information transmitted is evaluated as the agent's expectation of the reduction of entropy in the recipient's world model (by assuming that the recipient's reasoning apparatus mirrors its own).

In deciding how to act, an agent takes account of the value of the inflow and outflow of information; it takes account of its personal preferences, and, if these preferences may be totally ordered then they may be represented as a utility function. It also takes account of functions that evaluate trust and other parameters. From its information an information-based agent extracts: information measures, trust measures, measures of preference, utility or monetary value, all being expressed as probability distributions that represent the changing uncertainty in an agent's belief in the integrity of all of its information. Measures of the value of information, and utilitarian measures of the value of contracts, should not be unified into one overall measure. They are different and orthogonal concepts. Information-based agents exploit this orthogonality.

An agent will form expectations of other agents' behaviours by observing the difference between commitments encapsulated in contracts, promises and statements of intent, and their subsequent execution. These observations may take account of whether those differences, if any, are 'good' or 'bad' for the agent. In Sec. 4 we describe various measures for doing this. Some of these measures involve ideas from information theory.

An agent acts in response to some *need* or needs. A need may be exogenous such as the agent's 'owner' needs a bottle of wine, or a message from another agent offering to trade may trigger a latent need to trade profitably. A need may also be endogenous such as the agent deciding that it owns more wine that it requires. An agent may be attempting to satisfy a number of needs at any time, and may have expectations of its future needs. A need may involve acquiring some physical object, it may also be a need to acquire information or to develop some on-going relationship. Information-based agents have a planning system that is invoked to aim to satisfy needs — the planning system is flexible and explores a number of (possibly inter-related) options at the same time, and devotes resources (if necessary) to the most promising. One component of the world model will necessarily be an estimate of the expectation that some contract will subsequently satisfy a need in an acceptable way — this will be in the form of a probability distribution over some evaluation space that measures 'the expectation of the acceptability of the execution of a contract with respect to the satisfaction of that need'.

In general there will be many 'acceptable' ways of satisfying an agent's needs. An agent has strategies that aim to reach acceptable outcomes that are in some sense ‘good’. In a one-one negotiation, strategies determine the dynamic sequence in which an agent makes acceptable proposals to another agent, and a mechanism for accepting proposals from that agent. Information-based agents’ strategies exploit the orthogonality between information in proposals and preferences (possibly the utility) of proposals. For example, if our agents wish to make an offer at some ‘preference level’ then they may select from the available equi-preferable offers on the basis of their information revelation (an agent may wish this to be high, low) or selected from some section of the ontology. Alternatively, an agent may wish to make an offer at some ‘information revelation level’ — for example, the equitable information revelation strategy [10]. That strategy responds to a message with a response that gives the recipient expected information gain similar to that which the message gave to the agent — in single-issue bargaining if the opponent works with a fixed increment/decrement then the equitable information revelation strategy will respond likewise.

Strategies also address the issue of choosing which agent to attempt to negotiate with, and how needs may be met in the future. We have described trust and honour models for information-based agents [11]. This leads on to the development of business relationships that we believe are founded on the trust that grows both from an expectation of ‘good’ contracts and contract executions, and from an expectation of valuable and reliable information exchange.

Not withstanding the sequence of the discussion above, the strategies will determine the agent’s world model. The agent’s dynamic world model provides the ‘data’ for the strategies. The communication language determines the sorts of percept that an agent may receive, and the strategies determine the probability distributions in the world model. So the communication language and a strategy will together determine the requirements for the inference rules.

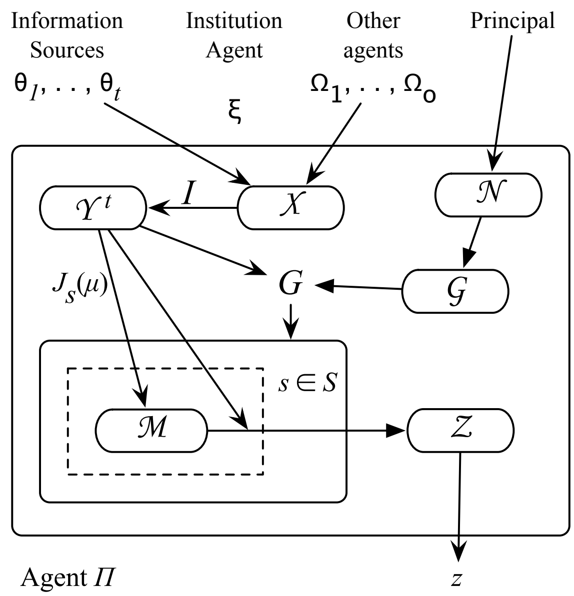
3 The architecture of information-based agents

The architecture of our information-based agent is shown in Fig. 1. The agent begins to function in response to a percept (received message) that expresses a need $N \in \mathcal{N}$. A *need* can be either exogenous (typically, the agent receives a message from another agent $\{\Omega_1, \ldots, \Omega_o\}$), or endogenous. The agent has a set of pre-specified *goals* (or *desires*), $G \in \mathcal{G}$, from which one or more is selected to satisfy its perceived needs. Each of these goals is associated with one or more plans, $s \in S$. This is consistent with the BDI model [1], and we do not detail these aspects here. The agent in Fig. 1 also interacts with information sources $\{\theta_1, \ldots, \theta_t\}$ that in our experiments[2] include unstructured data mining and text mining ‘bots’ that retrieve information from the agent market-place and from general news sources.

Finally the agent in Fig. 1 interacts with an ‘Institution Agent’, ξ, that reports honestly and promptly on the fulfilment of contracts. The Institution Agent is a conceptual device to prevent the requirement for agents to have ‘eyes’

[2] `http://e-markets.org.au`

Fig. 1. Agent architecture.



and effectors. For example, if agent Π wishes to give an object Y to agent Ω_k then this is achieved by Π sending a message to ξ requesting the transfer the ownership of Y from Π to Ω_k, once this this is done, ξ sends a message to Ω_k advising him that he now owns Y. Given such an Institution Agent, agents can negotiate and evaluate trade by simply sending and receiving messages.

Π has two languages: $\mathcal{C}$ and $\mathcal{L}$. $\mathcal{L}$ is a first-order language for internal representation — precisely it is a first-order language with sentence probabilities optionally attached to each sentence representing Π's epistemic belief in the validity of that sentence. $\mathcal{C}$ is an illocutionary-based language for communication [12]. Messages expressed in $\mathcal{C}$ from $\{\theta_i\}$ and $\{\Omega_i\}$ are received, time-stamped, source-stamped and placed in an *in-box* $\mathcal{X}$. The illocutionary particles in $\mathcal{C}$ are:

– Offer(Π, Ω_k, δ). Agent Π offers agent Ω_k a contract $\delta = (\pi, \varphi)$ with action commitments $\pi \in \mathcal{L}$ for Π and $\varphi \in \mathcal{L}$ for Ω_k.

– Accept(Π, Ω_k, δ). Agent Π accepts agent Ω_k's previously offered contract δ.

– Reject$(\Pi, \Omega_k, \delta[, \mathit{info}])$. Agent Π rejects agent Ω_k's previously offered contract δ. Optionally, information $\mathit{info} \in \mathcal{L}$ explaining the reason for the rejection can be given.

– Withdraw$(\Pi, \Omega_k[, \mathit{info}])$. Agent Π breaks down negotiation with Ω_k. Extra $\mathit{info} \in \mathcal{L}$ justifying the withdrawal may be given.

– Inform$(\Pi, \Omega_k, \mathit{info})$. Agent Π informs Ω_k about $\mathit{info} \in \mathcal{L}$ and commits to the truth of info.

– Reward($\Pi, \Omega_k, \delta, \phi$[, *info*]). Intended to make the opponent accept a proposal with the promise of a future compensation. Agent Π offers agent Ω_k a contract δ. In case Ω_k accepts the proposal, Π commits to make $\phi \in \mathcal{L}$ true. The intended meaning is that Π believes that worlds in which ϕ is true are somehow desired by Ω_k. Optionally, additional information in support of the contract can be given.

– Threat($\Pi, \Omega_k, \delta, \phi$, [*info*]) Intended to make the opponent accept a proposal with the menace of some sort of retaliation. Agent Π offers agent Ω_k a contract δ. In case Ω_k does not accept the proposal, Π commits to make $\phi \in \mathcal{L}$ true. The intended meaning is that Π believes that worlds in which ϕ is true are somehow *not* desired by Ω_k. Optionally, additional information in support of the contract can be given.

– Appeal(Π, Ω_k, δ, *info*) Intended to make the opponent accept a proposal as a consequence of the belief update that the accompanying information might bring about. Agent Π offers agent Ω_k a contract δ, and passes information in support of the contract.

The accompanying information, *info*, can be of two basic types: (i) referring to the process (plan) used by an agent to solve a problem, or (ii) data (beliefs) of the agent including preferences. When building relationships, agents will therefore try to influence the opponent by changing their processes (plans) or by providing new data.

3.1 World model

Everything that Π has at its disposal is derived from the messages in the inbox $\mathcal{X}$. As messages age, the degree of belief that Π associates with them will decrease. We call this *information integrity decay.* A factor in the integrity of a message will be the reliability of the source. This subjective decay is a feature of the agent, and agents will differ in their subjective estimates.

Each plan is driven by its expectations of the state of the world, and by the states of the other agents. These states will generally be quite numerous, and so we assume that at any time the agent's active plans will form expectations of certain *features* only, where each feature will be in one of a finite number of states[3]. Suppose that there are m such features, introduce m random variables, $\{X_i\}_{i=1}^{m}$. Each value, $x_{i,j}$, of the i'th random variable, X_i, denotes that the i'th feature is in the j'th perceivable state, or *possible world*, of that feature.

The messages in $\mathcal{X}$ are then translated using an *import function* I into sentences expressed in $\mathcal{L}$ that have integrity decay functions (usually of time) attached to each sentence, they are stored in a *repository* $\mathcal{Y}^t$. And that is all that happens until Π triggers a goal.

In general Π will be uncertain of the current state of each feature. Π's *world model*, $M \in \mathcal{M}$, consists of probability distributions over each of these random variables. If these m features are independent then the overall uncertainty, or

[3] We thus exclude the possibility of continuous variables.

entropy, of Π's world model is:

$$\mathbb{H}^t(M) = -\sum_{i=1}^{m} \mathbb{E}(\ln \mathbb{P}^t(X_i)) = -\sum_{i=1}^{m}\sum_{j=1}^{n_i} \mathbb{P}^t(X_i = x_{i,j}) \ln \mathbb{P}^t(X_i = x_{i,j})$$

The general idea is that if Π receives new information then the overall uncertainty of the world model is expected to decrease, but if Π receives no new information then it is expected to increase.

3.2 Reasoning

Consider first what happens if Π receives no new information. Each distribution, $\mathbb{P}^t(X_i)$, is associated with a *decay limit distribution*, $\mathbb{D}(X_i)$, that represents the expected limit state of the i'th feature in the absence of any observations of the state of that feature: $\lim_{t\to\infty} \mathbb{P}^t(X_i) = \mathbb{D}(X_i)$. For example, if the i'th feature is whether it is raining in Sydney, and $x_{i,1}$ means "it is raining in Sydney" and $x_{i,2}$ means "it is not raining in Sydney" — then if Π believes that it rains in Sydney 5% of the time: $\mathbb{D}(X_i) = (0.05, 0.95)$. If Π has no background knowledge about $\mathbb{D}(X_i)$ then the decay limit distribution is the maximum entropy, "flat", distribution. In the absence of incoming information, $\mathbb{P}(X_i)$ decays by:

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

where Δ_i is the *decay function* for the i'th feature satisfying the property that $\lim_{t\to\infty} \mathbb{P}^t(X_i) = \mathbb{D}(X_i)$. For example, Δ_i could be linear:

$$\mathbb{P}^{t+1}(X_i) = (1 - \nu_i)\mathbb{D}(X_i) + \nu_i \times \mathbb{P}^t(X_i) \tag{1}$$

where $\nu_i < 1$ is the decay rate for the i'th feature. Either the decay function or the decay limit distribution could also be a function of time: Δ_i^t and $\mathbb{D}^t(X_i)$.

If Π receives a message expressed in $\mathcal{C}$ then it will be transformed by inference rules into statements expressed in $\mathcal{L}$. We introduce this procedure with an example. Preference information is a statement by an agent that it prefers one class of contracts to another where contracts may be multi-issue. Preference illocutions may refer to particular issues within contracts — e.g. "I prefer red to yellow", or to combinations of issues — e.g. "I prefer a car with a five year warranty to the same car with a two year warranty than costs 15% less". An agent will find it useful to estimate which contract under consideration is favoured most by the opponent. Preference information can assist with this estimation as the following example shows. Suppose Π receives preference information from Ω_k through an Inform$(\Omega_k, \Pi, \mathit{info})$ illocution: *info* = "for contracts with property Q_1 or property Q_2, the probability that the contract Ω_k prefers most will have property Q_1 is z" — the ontology in $\mathcal{C}$ is assumed to contain an illocutionary particle that can express this statement. What happens next will depend on Π's plans. Suppose that Π has an active plan $s \in S$ that calls for the probability distribution $\mathbb{P}^t(\text{Favour}(\Omega_k, \Pi, \delta)) \in M$ over all δ, where $\text{Favour}(\Omega_k, \Pi, \delta)$

means that "δ is the contract that Ω_k prefers most from Π". Suppose Π has a prior distribution $\boldsymbol{q} = (q_1, \dots)$ for $\mathbb{P}^t(\text{Favour}(\cdot))$. Then s will require an inference rule: $J_s^{\text{Favour}}(\mathit{info})$ that is the following linear constraint on the posterior $\mathbb{P}^t(\text{Favour}(\Omega_k, \Pi, \delta))$ distribution:

$$z = \frac{\sum_{\delta:Q_1(\delta)} p_\delta}{\left(\sum_{\delta:Q_1(\delta)} p_\delta\right) + \left(\sum_{\delta:Q_2(\delta)} p_\delta\right) - \left(\sum_{\delta:Q_1 \wedge Q_2(\delta)} p_\delta\right)} \tag{2}$$

and is determined by the *principle of minimum relative entropy* — a form of Bayesian inference is that is convenient when the data is sparse [13] — as described generally below. The inference rule $J_s^{Favour}(\cdot)$ infers a constraint on a distribution in M from an illocution expressed in $\mathcal{C}$. Inferences of this sort are necessary for Π to operate, but their validity is a personal matter for Π to assume.

Now, more generally, suppose that Π receives a percept μ from agent Ω_k at time t. Suppose that this percept states that something is so with probability z, and suppose that Π attaches an epistemic belief probability $\mathbb{R}^t(\Pi, \Omega_k, \mu)$ to μ. Π's set of active plans will have a set of model building functions, $J_s(\cdot)$, such that $J_s^{X_i}(\mu)$ is a set of linear constraints on the posterior distribution for X_i where the prior distribution is $\mathbb{P}^t(X_i) = \boldsymbol{q}$. Let $\boldsymbol{p} = (p_1, \dots)$ be the distribution with minimum relative entropy with respect to $\boldsymbol{q}$: $\boldsymbol{p} = \arg\min_{\boldsymbol{p}} \sum_j p_j \log \frac{p_j}{q_j}$ that satisfies the constraints $J_s^{X_i}(\mu)$. Then let $\boldsymbol{r}$ be the distribution:

$$\boldsymbol{r} = \mathbb{R}^t(\Pi, \Omega_k, \mu) \times \boldsymbol{p} + (1 - \mathbb{R}^t(\Pi, \Omega_k, \mu)) \times \boldsymbol{q}$$

and then for a small time step δt let:

$$\mathbb{P}^{t+\delta t}(X_i) = \begin{cases} \boldsymbol{r} & \text{if } \mathbb{K}(\boldsymbol{r} \| \mathbb{D}(X_i)) > \mathbb{K}(\mathbb{P}^t(X_i) \| \mathbb{D}(X_i)) \\ \boldsymbol{q} & \text{otherwise} \end{cases} \tag{3}$$

where $\mathbb{K}(\boldsymbol{x} \| \boldsymbol{y}) = \sum_j x_j \ln \frac{x_j}{y_j}$ is the Kullback-Leibler distance between two probability distributions $\boldsymbol{x}$ and $\boldsymbol{y}$. The idea in Eqn. 3 is that the vector $\boldsymbol{r}$ will only update $\mathbb{P}^t(X_i)$ if it contains more information with respect to the decay limit distribution than the prior $\boldsymbol{q}$. Then combining Eqn. 3 with Eqn. 1 let:

$$\mathbb{P}^{t+1}(X_i) = (1 - \nu_i)\mathbb{D}(X_i) + \nu_i \times \mathbb{P}^{t+\delta t}(X_i) \tag{4}$$

and note that this procedure has dealt with integrity decay, and with two probabilities: first, the probability z in the percept μ, and second the epistemic belief probability $\mathbb{R}^t(\Pi, \Omega_k, \mu)$ that Π attached to μ. Given a probability distribution $\boldsymbol{q}$, the *minimum relative entropy distribution* $\boldsymbol{p} = (p_1, \dots, p_I)$ subject to a set of J linear constraints $\boldsymbol{g} = \{g_j(\boldsymbol{p}) = \boldsymbol{a_j} \cdot \boldsymbol{p} - c_j = 0\}, j = 1, \dots, J$ (that must include the constraint $\sum_i p_i - 1 = 0$) is: $\boldsymbol{p} = \arg\min_{\boldsymbol{p}} \sum_j p_j \log \frac{p_j}{q_j}$. This may be calculated by introducing Lagrange multipliers $\boldsymbol{\lambda}$: $L(\boldsymbol{p}, \boldsymbol{\lambda}) = \sum_j p_j \log \frac{p_j}{q_j} + \boldsymbol{\lambda} \cdot \boldsymbol{g}$. Minimising L, $\{\frac{\partial L}{\partial \lambda_j} = g_j(\boldsymbol{p}) = 0\}, j = 1, \dots, J$ is the set of given constraints $\boldsymbol{g}$, and a solution to $\frac{\partial L}{\partial p_i} = 0, i = 1, \dots, I$ leads eventually to $\boldsymbol{p}$.

3.3 Estimating $\mathbb{R}^t(\Pi, \Omega_k, \mu)$

Π attaches an epistemic belief probability $\mathbb{R}^t(\Pi, \Omega_k, \mu)$ to each message μ. A historic estimate of $\mathbb{R}^t(\Pi, \Omega_k, \mu)$ may be obtained by measuring the 'difference' between commitment and execution. Π's plans will have constructed a set of distributions. We measure this 'difference' as the error in the effect that μ has on each of Π's distributions. Suppose that μ is received from agent Ω_k at time u and is verified at some later time t. For example, μ could be a chunk of information: "the interest rate will rise by 0.5% next week", and suppose that the interest rate actually rises by 0.25% — represent what the message should have been μ'. What does all this tell agent Π about agent Ω_k's reliability? Consider one of Π's distributions for X that is $\boldsymbol{q}^u$ at time u. Let $\boldsymbol{p}^u_\mu$ be the posterior minimum relative entropy distribution subject to the constraint $J^X_s(\mu)$, and let $\boldsymbol{p}^u_{\mu'}$ be that distribution subject to $J^X_s(\mu')$. We now estimate what $\mathbb{R}^u(\Pi, \Omega_k, \mu)$ should have been in the light of knowing *now*, at time t, that μ should have been μ'.

The idea of Eqn. 3, is that the current value of $\mathbb{R}^t(\Pi, \Omega_k, \mu)$ should be such that, *on average*, $\boldsymbol{p}^u_\mu$ will be "close to" $\boldsymbol{p}^u_{\mu'}$ when we eventually discover μ' — no matter whether or not μ was used to update the distribution for X, as determined by the acceptability test in Eqn. 3 at time u. The *observed reliability* for μ and distribution X, $\mathbb{R}^t_X(\Pi, \Omega_k, \mu)|\mu'$, on the basis of the verification of μ with μ', is the value of r that minimises the Kullback-Leibler distance:

$$\mathbb{R}^t_X(\Pi, \Omega_k, \mu)|\mu' = \arg\min_r \mathbb{K}(r \cdot \boldsymbol{p}^u_\mu + (1-r) \cdot \boldsymbol{q}^u \| \boldsymbol{p}^u_{\mu'})$$

If $\mathbf{X}(\mu)$ is the set of distributions that μ affects, then the overall *observed reliability* on the basis of the verification of μ with μ' is:

$$\mathbb{R}^t(\Pi, \Omega_k, \mu)|\mu' = 1 - (\max_{X \in \mathbf{X}(\mu)} |1 - \mathbb{R}^t_X(\Pi, \Omega_k, \mu)|\mu'|)$$

Then for each ontological context o_j, at time t when μ has been verified with μ':

$$\mathbb{R}^{t+1}(\Pi, \Omega_k, o_j) = (1-\nu) \times \mathbb{R}^t(\Pi, \Omega_k, o_j) + \nu \times \mathbb{R}^t(\Pi, \Omega_k, \mu)|\mu' \times \text{Sim}(o_j, O(\mu))$$

where Sim measures the semantic distance between two sections of the ontology, and ν is the learning rate. Over time, Π notes the ontological context of the various μ received from Ω_k, and over the various ontological contexts calculates the relative frequency, $\mathbb{P}^t(o_j)$, of these contexts, $o_j = O(\mu)$. This leads to an overall expectation of the *reliability* that agent Π has for agent Ω_k:

$$\mathbb{R}^t(\Pi, \Omega_k) = \sum_j \mathbb{P}^t(o_j) \times \mathbb{R}^t(\Pi, \Omega_k, o_j)$$

4 Commitment and Execution

The interaction between agents Π and Ω_k will eventually lead to some sort of *contract*: $\delta = (\pi, \varphi)$ where π is Π's commitment and φ is Ω_k's commitment.

No matter what these commitments are, Π will be interested in any variation between Ω_k's commitment, φ, and what actually happens, the execution, φ'. The form of this commitment could be a promise to deliver goods, or abide by certain trading terms that extend over a period of time, or that some information that may, or may not, prove to be correct. We denote the relationship between commitment and execution, $\mathbb{P}^t(\text{Execute}(\varphi')|\text{Commit}(\varphi))$ simply as $\mathbb{P}^t(\varphi'|\varphi)$. In general we assume that such commitment and execution takes place in the context of a *relationship* ρ between Π and Ω_k.

Beliefs 'evaporate' as time goes by. If we don't keep an ongoing relationship, we become unsure how *trustworthy* a trading partner is. This decay is what justifies a continuous relationship between agents. The conditional probabilities, $\mathbb{P}^t(\varphi'|\varphi)$, should tend to ignorance as represented by the *decay limit distribution* $\boldsymbol{d} = \{d_i\}$. If we have the set of observations $\Phi = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$ then complete ignorance of the opponent's expected behaviour means that given the opponent commits to φ, the conditional probability for each observable outcome φ' becomes $d_i = \frac{1}{n}$, but Π may have background beliefs about Ω_k's decay limit distribution. This natural decay of belief is offset by new observations. We define the evolution of the probability distribution as: $\mathbb{P}^{t+1}(\varphi'|\varphi) = \big((1-\nu)\cdot \boldsymbol{d} + \nu \cdot \mathbb{P}^t_+(\varphi'|\varphi)\big)$, where $\nu \in [0,1]$ is the learning rate, and $\mathbb{P}^t_+(\varphi'|\varphi)$ represents the posterior distribution for $(\varphi'|\varphi)$ given an observed contract execution as the following shows.

Suppose that Π has a business relationship ρ with agent Ω_k, that Ω_k commits to φ, and this commitment is sound. The material value of φ to ρ will depend on the future use that Π makes of it, and that is unlikely to be known. So Π estimates the value of φ to the relationship ρ he has with Ω_k using a probability distribution $(p_1, \ldots, p_n)$ over a *relationship evaluation space* $E = (e_1, \ldots, e_n)$ that could range from "that is what I expect from the perfect trading partner" to "it is totally useless" — E may contain hard or fuzzy values. $p_i = w_i(\rho, \varphi)$ is the probability that e_i is the correct evaluation of the enactment φ in the context of relationship ρ, and $\boldsymbol{w} : \mathcal{L} \times \mathcal{L} \to [0,1]^n$ is the *evaluation function.*

Let $(\varphi_1, \ldots, \varphi_m)$ be the set of possible contract executions in some order. Then for a given φ_k, $(\mathbb{P}^t(\varphi_1|\varphi_k), \ \ldots, \mathbb{P}^t(\varphi_m|\varphi_k))$ is the prior distribution of Π's estimate of what will actually occur if Ω_k committed to φ_k occurring and $\boldsymbol{w}(\rho, \varphi_k) = (w_1(\rho, \varphi_k), \ldots, w_n(\rho, \varphi_k))$ is Π's evaluation over E with respect to the relationship ρ of Ω_k's commitment φ_k. Π's expected evaluation of what will occur given that Ω_k has committed to φ_k occurring is:

$$\boldsymbol{w}^{\text{exp}}(\rho, \varphi_k) = \left(\sum_{j=1}^{m} \mathbb{P}^t(\varphi_j|\varphi_k) \cdot w_1(\rho, \varphi_j), \ldots, \sum_{j=1}^{m} \mathbb{P}^t(\varphi_j|\varphi_k) \cdot w_n(\rho, \varphi_j) \right).$$

Now suppose that Π observes the event $(\phi'|\phi)$ in another relationship ρ' also with agent Ω_k. Eg: Π may buy wine and cheese from the same supplier. Π may wish to revise the prior estimate $\boldsymbol{w}^{\text{exp}}(\rho, \varphi_k)$ in the light of the observation $(\phi'|\phi)$ to:

$$(\boldsymbol{w}^{\text{rev}}(\rho, \varphi_k) \mid (\varphi'|\varphi)) = \boldsymbol{g}(\boldsymbol{w}^{\text{exp}}(\rho, \varphi_k), \boldsymbol{w}(\rho', \phi), \boldsymbol{w}(\rho', \phi'), \rho, \rho', \varphi, \phi, \phi'),$$

for some function $\boldsymbol{g}$ — the idea being, for example, that if the commitment, ϕ, concerning the purchase of cheese, ρ', was not kept then Π's expectation that the commitment, φ, concerning the purchase of wine, ρ, will not be kept should increase. We estimate the posterior $\mathbb{P}^t_+(\varphi'|\varphi)$ by applying the principle of minimum relative entropy: $\left(\mathbb{P}^t_+(\varphi_j|\varphi)\right)_{j=1}^m = \arg\min_{\boldsymbol{p}} \sum_{i=1}^m p_i \log \frac{p_i}{\mathbb{P}^t(\varphi_i|\varphi)}$ where $\boldsymbol{p} = (p_j)_{j=1}^m$, satisfies the n constraints:

$$\sum_{j=1}^{m} p_j \cdot w_i(\rho, \varphi_j) = g_i(\boldsymbol{w}^{\exp}(\rho, \varphi_k), \boldsymbol{w}(\rho', \phi), \boldsymbol{w}(\rho', \phi'), \rho, \rho', \varphi, \phi, \phi')$$

for $i = 1, \ldots, n$. This is a set of n linear equations in m unknowns, and so the calculation of the minimum relative entropy distribution may be impossible if $n > m$. In this case, we take only the m equations for which the change from the prior to the posterior value is greatest. That is, we attempt to select the most significant factors.

Consider a distribution of expected fulfilment of commitments that represent Π's "ideal" for a relationship with Ω_k, in the sense that it is the best that Π could reasonably expect Ω_k to do. This distribution will be a function of Ω_k, Π's history with Ω_k, anything else that Π believes about Ω_k, and general environmental information including time — denote all of this by e, then we have $\mathbb{P}^t_I(\varphi'|\varphi, e)$. For example, if Π considers that it is unacceptable for the execution φ' to be less preferred than the commitment φ then $\mathbb{P}^t_I(\varphi'|\varphi, e)$ will only be non-zero for those φ' that Π prefers to φ. The distribution $\mathbb{P}^t_I(\cdot)$ represents what Π expects, or hopes, Ω_k will do. *Trust* is the relative entropy between this ideal distribution, $\mathbb{P}^t_I(\varphi'|\varphi, e)$, and the distribution of the observation of fulfilled commitments, $\mathbb{P}^t(\varphi'|\varphi)$. That is:

$$\text{Trust}(\Pi, \Omega_k, \varphi) = 1 - \sum_{\varphi'} \mathbb{P}^t_I(\varphi'|\varphi, e) \log \frac{\mathbb{P}^t_I(\varphi'|\varphi, e)}{\mathbb{P}^t(\varphi'|\varphi)} \tag{5}$$

where the "1" is an arbitrarily chosen constant being the maximum value that trust may have. This equation defines trust for one, single commitment φ — for example, my trust in my butcher if he commits to provide me with a 10% discount for the rest of the year. It makes sense to aggregate these values over a class of commitments, say over those φ that are subtypes of a particular relationship ρ, that is $\varphi \leq \rho$. In this way we measure the trust that I have in my butcher in relation to the commitments he makes for red meat generally:

$$\text{Trust}(\Pi, \Omega_k, \rho) = 1 - \frac{\sum_{\varphi:\varphi\leq\rho} \mathbb{P}^t(\varphi) \left[\sum_{\varphi'} \mathbb{P}^t_I(\varphi'|\varphi, e) \log \frac{\mathbb{P}^t_I(\varphi'|\varphi, e)}{\mathbb{P}^t(\varphi'|\varphi)}\right]}{\sum_{\varphi:\varphi\leq\rho} \mathbb{P}^t(\varphi)}$$

where $\mathbb{P}^t(\varphi)$ is a probability distribution over the space of commitments that the next commitment Ω_k will make to Π is φ. Similarly, for an overall estimate of Π's trust in Ω_k:

$$\text{Trust}(\Pi, \Omega_k) = 1 - \sum_{\varphi} \mathbb{P}^t(\varphi) \left[\sum_{\varphi'} \mathbb{P}^t_I(\varphi'|\varphi, e) \log \frac{\mathbb{P}^t_I(\varphi'|\varphi, e)}{\mathbb{P}^t(\varphi'|\varphi)}\right]$$

5 Strategies

An agent requires *strategies* for deciding who to interact with, and for deciding how to manage interaction using the language $\mathcal{C}$. In $\mathcal{C}$ as defined in Sec. 3, contracts may be for a single trade, or may encapsulate an on-going trading relationship. Interaction is normally bound by an *interaction protocol* that moderates the interaction sequence, and so may limit the range of model building functions, $J_s(\cdot)$. Consider the protocol in which statements in $\mathcal{C}$ are exchanged between pairs of agents, and Offer($\cdot$) statements are binding until countered, or until one of the pair issues a Quit($\cdot$). That is, an agent would only enter into a negotiation — ie: offer exchange — if it were prepared to commit. To manage this protocol, agent Π requires the following probability estimates in M where Π is bargaining with opponent Ω_k, in satisfaction of some need $N \in \mathcal{N}$:

1 $\mathbb{P}^t(val(\Pi, \Omega_k, N, \delta) = v_i)$ — for any deal, δ, the probability distribution over some valuation space $\{v_i\}$ that measures how "good" the deal δ is to Π.
2 $\mathbb{P}^t(acc(\Pi, \Omega_k, \delta))$ — for any deal, δ, the probability that Ω_k would accept δ.
3 $\mathbb{P}^t(conv(\Pi, \Omega_k, \Delta))$ — for any sequence of offer-exchanges, Δ, the probability that that sequence will converge to an acceptable deal.
4 $\mathbb{P}^t(trade(\Pi, \Omega_k, o_j) = u_i)$ — for an ontological context o_j, the probability distribution over some valuation space $\{u_i\}$ that measures how "good" Ω_k is as a trading partner to Π for deals in ontological context o_j.

The estimation of these distributions has been described previously [10]. Π's strategy determines how it uses these distributions. An approach to issue-tradeoffs is described in [14]. That strategy attempts to make an acceptable offer by "walking round" the iso-curve of Π's previous offer (that has, say, an acceptability of α) towards Ω_k's subsequent counter offer. In terms of the machinery described here: $\arg\max_\delta\{\ \mathbb{P}^t(acc(\Pi, \Omega_k, \delta)) \mid \mathbb{E}^t(val(\Pi, \Omega_k, N, \delta)) \approx \alpha\ \}$. By including the "information dimension" Π can implement strategies that go beyond utilitarian thinking. Π evaluates every illocution for its utilitarian value, and for its value as information. For example, the *equitable information revelation* strategy [10] responds to a message μ with a message that gives the recipient expected information gain similar to that which μ gave to Π; these responses are also "reasonable" from a utilitarian point of view. An information-based agent evaluates all exchanges in terms of both their estimated utilitarian value, and their information value.

Estimations of trust — Sec. 4 — may be used to select a trading partner. One interesting question is to determine a set of partners to maintain for deals from a particular section of the ontology — this is a question of risk management. Having identified such a set, the agent then has to decide which one of these partners to use for the next negotiation. A nice strategy is to choose the partner with a probability equal to the probability that they are the best choice — as determined by trust, or some other means.

An information-based agent additionally requires strategies to manage the exchange of information, and to be strategic in their information acquisition. This includes strategies for dealing with the information sources $\{\theta_1, \ldots, \theta_t\}$, which becomes interesting if those sources are not always available, charge a fee, or take some time to deliver. This also includes strategies for the acquisition of information by both covert and overt strategic interaction with other agents $\{\Omega_1, \ldots, \Omega_o\}$. These information strategies are the subject of current research.

6 Conclusion

We do not claim that this is the end of the matter in deploying automated negotiators, and the approach described here has yet to be trialed extensively. But we do maintain the strategic apparatus of intelligent negotiating agents should include the intelligent use of information. We have proposed a theoretical basis for managing information in the context of competitive interaction, and have shown how that theory may be computed by an intelligent agent. Information theory provides the theoretical underpinning that enables such an informed agent to value, manage and exchange her information intelligently.

References

1. Wooldridge, M.: Multiagent Systems. Wiley (2002)
2. Ghallab, M., Nau, D., Traverso, P.: Automated Planning : Theory and Practice. Morgan Kaufmann (2004)
3. 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)
4. Rosenschein, J., Zlotkin, G.: Rules of Encounter. MIT Press (1998)
5. Muthoo, A.: Bargaining Theory with Applications. Cambridge UP (1999)
6. Klemperer, P.: The Economic Theory of Auctions : Vols I and II. Edward Elgar (2000)
7. Paris, J.: Common sense and maximum entropy. Synthese **117** (1999) 75 – 93
8. Jaynes, E.: Probability Theory — The Logic of Science. Cambridge University Press (2003)
9. Jaeger, M.: Representation independence of nonmonotonic inference relations. In: Proceedings of KR'96, Morgan Kaufmann (1996) 461–472
10. Debenham, J., Simoff, S.: An agent establishes trust with equitable information revelation. In Subrahmanian, V., Regli, W., eds.: Proceedings of the 2005 IEEE 2nd Symposium on Multi-Agent Security and Survivability, Drexel University, Philadelphia, USA, IEEE (2005) 66 – 74
11. Sierra, C., Debenham, J.: Trust and honour in information-based agency. In: Proceedings Fifth International Conference on Autonomous Agents and Multi Agent Systems AAMAS-2006, Hakodate, Japan, ACM Press, New York (2006)
12. Sierra, C., Jennings, N., Noriega, P., Parsons, S.: A framework for argumentation-based negotiation. In: Intelligent Agents IV: Agent Theories, Architectures, and Languages (ATAL-97), Springer-Verlag: Heidelberg, Germany (1998) 177–192
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. American Institute of Physics, Melville, NY, USA (2004) 445 – 461
14. Faratin, P., Sierra, C., Jennings, N.: Using similarity criteria to make issue trade-offs in automated negotiation. Journal of Artificial Intelligence **142** (2003) 205–237