An Architecture for Purely Probabilistic Negotiating Agents: Pessimism, Punishment, and Laissez-Faire Paths

Ronald P. Loui
Department of Computer Science and Engineering and
Washington University in St. Louis
St. Louis, Missouri, USA 63130
loui@cs.wustl.edu

October 15, 2005

1 Acknowledgements

This paper was presented at the Probability as a Guide to Life Symposium at the University of Rochester, Department of Philosophy, October 2004, in honor of my advisor, Henry E. Kyburg, Jr.

These ideas date from 1998 and result from collaboration on negotiation games with Anne Jump.

2 Introduction

Instead of viewing negotiation as a game with a solution, this paper conceives of negotiation as a process. This view of negotiation is useful both for the design of automated agents, and for the mathematical description of the social interaction. The view may be adopted as a psychological claim about the measurable forces within human agents, but that claim is not made here.

The process of negotiation is regulated by two principal kinds of linguistic act: the extension of an offer and the notification of unilateral breakdown. Other linguistic acts such as questioning, answering, focusing, and threatening, are not included here for the sake of simplicity.

At any time, an objective probability can be attached to each potential settlement, representing the probability that it will be the settlement if there is agreement. Meanwhile, an objective probability can be attached to the possibility of non-agreement, or breakdown. This objective probability is based on past observations of similarly situated agents in similar strategic circumstances, having made similar progress toward settlement. The exact ontological basis for induction and inductive method are beyond the scope of this paper, but any inductive statistical approach will suffice. It is crucial that this probability not be a subjective probability: it cannot be altered by will or desire, strategy or design, bias or dogma, or any other psychological mechanism. There is interpersonal agreement on what the probability is, once the data and the method are fixed. Furthermore, objective probability is binding upon the rational agent as a guide for action.

The process of negotiation is driven by the dynamics of objective probability assessments, conditioned on the amount of time that has elapsed since last nontrivial progress. Specifically, as time since last progress grows, the probability of breakdown rises, and the expected utility of potential settlement falls. It is an empirical hypothesis that there is a granularity of time with respect to which probability of breakdown rises monotonically as progress is not made. It is hard to imagine situations where the weight of empirical evidence does not warrant this hypothesis. The expected utility of settlement will fall as pessimism rises until expectation reaches either the best offer extended by the other agent, or else the security level of the agent.

Any offer that is slightly above security would be accepted if pessimism had no compensatory force. Of course, this behavior is not witnessed in practice, nor is it acceptable behavior for designed agents.

Whenever the other agent fails to make progress, the utility of unilateral breakdown rises. That is, the act of breaking down on the other agent becomes more valuable, the more uncooperative that agent's behavior is perceived to be. A nonstandard procedural utility is attached to the time since last progress by the other agent. The effect of this procedural utility is to create a race between the falling expectation, due to pessimism, and the rising value of breakdown, based on resentment and the prospect of punishment. It is the principal claim of this paper, as a matter of design, that these two forces suffice to govern the basic negotiating agent and produce a variety of acceptable and desirable behaviors.

The one-sided rational agent is given deadlines at which time an offer must be extended, an offer must be accepted, or unilateral breakdown must occur. There are many situations in which a one-sided bound is appropriate. For example, when a negotiating broker has the power to act on behalf of another, but does not have the full proxy to suppress independent acts by the person or persons represented by the broker. A trade representative, for example, may have the power to extend an offer, but not the exclusive power to extend offers. Similarly, the single agent may be governed by a society of minds, a large part of which will behave according to expected utility, but some of which may yet behave with independent initiative.

A two-sided rational agent is not precluded by the approach here. Such an agent would have both an objectively determined earliest time to act and an objectively determined latest time to act. However, a point-rational agent, where an exact earliest and latest time to act are prescribed, and the times coincide, violates the constructive philosophy of this approach. One-sided rational constraints provide normative guides for action, but there may be other grounds for acting "ahead of schedule."

The procedural utility here that causes rejection of an 11-cent offer in favor of a 10-cent security level at breakdown in, for example, a zero-sum split of a dollar, is not the same as a substantive utility attached to the unfairness of a split. It has become famously popular in recent years to model the outcome of negotiation in such a way that objects of value take into account perceptions of property rights or distributive justice. Fair-split effects are easily modeled prior to the process of negotiation by transforming the payoff functions. Here we are modeling the procedural effect of altering objects of value based on how they are reached through negotiation, not just what they represent to an agent in relation to other potential settlements. Here we are interested in the possibility that a unilateral 10-cent outcome has more value if the path to that outcome is torturous. Prior authors have considered the possibility that a 10-cent outcome has less value if there were an aspiration for a dollar.

In either case, whether the interest is distributional or procedural justice, the transformation of payoffs models an internal value of dignity or pride, and a willingness to invest in a social mechanism that produces external pressures on other agents to behave better.

3 Pairs of Agents with Payoffs

Two agents A and B are conjoined in action when the payoff to each agent depends on the choice pair (a, b) from the space of joint choices, Ω_A and Ω_B for A and B respectively. In strategic form, the former is the set of rows and the latter is the set of columns. In the absence of coordination, A chooses from Ω_A and B chooses from Ω_B simultaneously. In this model, the space of joint choices is also the set of potential settlements. $\Omega = \Omega_A \times \Omega_B$, but we often carry the cross-product into the expressions to emphasize the pairing of action.

The payoffs in this model are u_A and u_B.

$$u_A, u_B : \Omega_A \times \Omega_B \to \Re.$$

4 Dialogue

At any time, $\Upsilon_A(t) \subseteq (\Omega_A \times \Omega_B)$ is the offer set extended by A at t, the set of offers A has extended to B. Similarly, $\Upsilon_B(t)$ is the offer set extended by B at t. If at any time the intersection is non-null, we declare agreement and an end of dialogue, even if the intersection is non-unique.

Suppose that offer sets are monotonic in time;

$$\forall t' > t : \Upsilon(t) \subseteq \Upsilon(t')$$

i.e., offers cannot be rescinded.

Dialogue can end with the acceptance of an offer, which might as well be the reciprocal extension of an offer that has been extended by the other agent. Dialogue can also end in either agent's unilateral breakdown. This is a simple and final notification of termination, either bd_A or bd_B, depending on who declares breakdown first. Breakdown implies the absence of an agreement.

5 Security and Breakdown

In the absence of agreement, we can assume that each agent has a method for selecting an uncoordinated choice. These are a_{bd} and b_{bd},

$$bd = (a_{bd}, b_{bd}) \in (\Omega_A \times \Omega_B)$$

where the security level of A is thus $u_A(bd)$, and for B, it is $u_B(bd)$.

For monotonic payoff functions on ordered choice sets, where u_A is decreasing as choice from Ω_A or Ω_B increases, and u_B is increasing as choice from Ω_A or Ω_B increases, the security levels may as well be determined by the unique Nash equilibrium, $a_{bd} = min_{\prec}\Omega_A$, $b_{bd} = max_{\prec}\Omega_B$. In terms of strategic form, this means that the breakdown act is in the upper right corner, $bd = (1, |\Omega_B|)$, if the lowest ordered choice pair, e.g., (1,1), is in the upper left.

For all t the agent has an objective probability of breakdown, $Prob^{bd}_A(t)$, which at the moment does not distinguish between breakdown of A on B, or B on A, or simultaneous breakdown. There is a probability of settlement, distributed over all the potential settlements Ω,

$$Prob^s_A(t) = \Sigma_{\omega\in\Omega} Prob^s_A(t)(\omega) = 1 - Prob^{bd}_A(t).$$

6 Admissible and Probable Settlements

At t, the value of the best offer to A is

$$\nu_{\cdot A}(t) = max_{\Upsilon_B(t)} u_A$$

which might also be called A's value of A's best offer from B. This is what A would get if A were to settle immediately on B's best terms at time t.

A's concession level is A's least value among A's offers to B:

$$\nu^{\cdot}_A(t) = min_{\Upsilon_A(t)} u_A \ .$$

In this discussion, A's objective probability of settlement at t should be supported only for outcomes A values at least as much as $\nu_{\cdot A}(t)$ and no more than $\nu^{\cdot}_A(t)$. Furthermore, if $\nu_{\cdot B}(t)$ and $\nu^{\cdot}_B(t)$ are known to A, then A's probability of settlement is restricted to outcomes with u_B in that range as well:

for private payoffs (u_A only known to A, u_B only known to B),

$\forall t : \forall \omega \in (\Omega_A \times \Omega_B)$:
if $(u_A(\omega) < \nu_{.A}(t)$ or $u_A(\omega) > \nu^{\cdot}_{A}(t))$
then $Prob^s_A(t)(\omega) = 0$.

for public payoffs (both u_A and u_B known to both A and B),

$\forall t : \forall \omega \in (\Omega_A \times \Omega_B)$:
if $(u_A(\omega) < \nu_{.A}(t)$, or $u_A(\omega) > \nu^{\cdot}_{A}(t)$,
or $u_B(\omega) < \nu_{.B}(t)$, or $u_B(\omega) > \nu^{\cdot}_{B}(t))$
then $Prob^s_A(t)(\omega) = 0$.

The practical effect is to prune the set of admissible settlements as proposals are made. The joint pruning effect is remarkably effective in practice. For the agent guided by objective probability, this means that even a short exchange of reasonable offers can quickly identify the likely outcomes of negotiation if there is settlement, and tight bounds on resultant utilities if there is settlement.

The concession level is equal to A's value of A's best offer to B, when payoffs are monotonic:

$$\nu'^{\cdot}_{A} = u_A(argmax_{\Upsilon_A(t)} u_B) \ .$$

For payoffs that are not monotonic, it may be better to use the latter expression as an upper bound on the support. Actual experience may include data where the values of settlement lay outside even this enlarged best-offer-to/best-offer-from range, but perhaps such data represents negotiation with agents who are not fully rational, or negotiation situations that are peculiar.

7 Expectation

As soon as there is a probability distribution over settlements,

$$Prob^s_A(t) : \Omega \rightarrow [0, 1] \ ,$$

there is an expected value of settlement,

$$Eu^s_A(t) = \Sigma_{\omega \in \Omega} Prob^s_A(t)(\omega) u_A(s)$$

which weighs each settlement's value with its likelihood in the usual way. Absent an actual distribution, $Prob^s_A(t)$, a convenient way to think about this is with a Hurwicz parameter, α. (We resist calling this an optimism-pessimism parameter as it is traditionally called because we are using pessimism differently below.) Presuming α_A and α_B, or just α when the distinction is not important,

$$Eu^s_A(t) = (1 - \alpha)\nu^{\cdot}_{A}(t) + \alpha \nu_{.A}(t) \ .$$

Expected utility takes the expected utility given settlement and the utility given breakdown, and mixes them by the probability of breakdown:

$$Eu_A(t) = Prob^{bd}_A(t) u_A(bd) + Prob^s_A(t) Eu^s_A(t) \ .$$

8 One-Sided Rational Acts

Rationality requires latest times for certain acts:

1. If at t, $\exists \omega \in \Upsilon_B(t) : Eu_A(t) \geq u_A(\omega)$, then A accepts an offer of B; else

2. If at t, $\exists \omega \in \Omega, \omega \notin \Upsilon_A(t) : Eu_A(t) \leq u_A(\omega)$, then A extends an offer to B; else

3. If at t, $Eu_A(t) \leq u_A(bd)$, then A unilaterally breaks down on B.

There is no problem ordering these so that one rule defeats another if two rules prescribe conflicting action. This is especially important with discrete time.

9 Pessimism

The probability of breakdown, $Prob^{bd}_A(t)$, naturally exhibits a dynamics we refer to as pessimism: the longer the time since non-trivial progress was made, the greater the probability of breakdown. This is an empirical hypothesis for our agents, since their assessment of probability is induced by their actual experience. Furthermore, this hypothesis seems plausible only after an initial reactionary period, post progress, during which the probability of breakdown might actually fall. Basically, if one waits long enough and nothing has happened, the probability of an agreement becomes more remote.

It is sufficient here to theorize about a granularity of time, δ, and to pay attention only to times since progess was made by the other agent. Non-trivial progress here is any offer that improves $\nu_{.A}$, though others could define non-triviality more stringently. Others might also drive agent behavior using the probabilities conditioned on time since either agent makes progress, but this choice is not important here.

Pessimism. (Empirical Hypothesis) Agents are likely to discover (for some $\delta > 0$) that

$$Prob^{bd}_A(t^{lp} + k + \delta) > Prob^{bd}_A(t^{lp} + k) \text{ for all } k > 0,$$

where t^{lp} is the time of the last non-trivial progress by the other agent, i.e., t^{lp} is the earliest time at which $\nu_{.A}$ took on its current value at time t

$$t^{lp} = argmin_z \nu_{.A}(z) = \nu_{.A}(t) \text{ and}$$

A linear pessimism is one in which $Prob^{bd}(t^{lp} + k + \delta) = (k + \delta)\pi$ for all $k > 0$, for a fixed parameter π, so long as this value is less than or equal to one. Note that linear pessimism implies a resetting of $Prob^{bd}$ to zero whenever there is non-trivial progress by the other agent.

Agents who experience unbounded pessimism are manipulable by agents who simply wait to make concessions. By withholding progress, the mendacious agent can cause the pessimism-driven agent to have a falling expectation until, eventually, $Eu_A(t)$ becomes not $Eu_A(t) = Prob_A^{bd}(t)u_A(bd) + Prob_A^s(t)Eu_A^s(t)$, but simply $u_A(bd)$.

If the manipulating agent has made an offer worth any ϵ greater than $u_A(bd)$, then the unbounded pessimistic agent will eventually accept that offer rather than break down. This is simply the "rational" behavior of taking a $u_A(bd) + \epsilon$ "sure thing" over an expecation of $u_A(bd)$. This behavior is undesirable even if there is no intention to manipulate by the agent who witholds offers while the other agent's pessimism rises.

At this point, a game theorist might try to release the agent from even the weak, one-sided rationality constraints and make room for a counter-strategy of adversarial waiting. We take a different tack. The agent who waits for offers is also the agent who experiences a rise in the procedural utility attached to unilateral breakdown. Put simply, if one agent offends the other by making the other agent wait, the second agent will find breakdown sweeter.

10 Resentment and Punishment

A procedural utility $u_A^{bd}(t)$ is attached to the act of unilateral breakdown, so that the actual value of breakdown, when initiated by A, is not $u_A(bd)$ but is instead $u_A(bd) + u_A^{bd}(t)$. For simplicity, assume the procedural effect resets, $u_A^{bd}(t) = 0$, whenever B makes non-trivial progress. That is, there is no memory of prior offense. It is just as easy to imagine a memory parameter, μ, such that $u^{bd}(t') = \mu u^{bd}(t)$ when B makes non-trivial progress at t.

u^{bd} is a source of utility which presupposes that value can be created by process.

There are many ways to understand u^{bd}. It can be a model of dignity, pride, or social investment. We were probably all taught to attach a u^{bd} value to unilateral breakdown, and would instruct our pupils to do the same; we can do no differently for our artificial agents. There may be a negative u^{bd} reflecting a penalty for breakdown, when agents have been especially cooperative. We do not theorize about negative u^{bd} here. u^{bd} is introduced solely as a force to compensate for pessimism.

There is the possibility that the agent does not actually feel the value u^{bd} in the same way that the agent feels u. There are two responses. First, artificial agents don't feel either, so they can presumably be designed with arbitrary choices for both u and u^{bd}. Second, the agent who does not actually feel u^{bd} should either not use it in calculating the value of breakdown, or else should seek more self-awareness.

The presence of this measure amends one of the rationality constraints,

3'. If at t, $Eu_A(t) \leq u_A(bd) + u_A^{bd}(t)$, then A unilaterally breaks down on B.

There are also subtle changes the equation for $Eu_A(t)$ which do not change the statement of rationality constraints 1 and 2, but which change their dynamics:

$$Eu_A(t) = Prob_A^{bd}(t)(u_A(bd) + u_A^{bd}(t)) + Prob_A^s(t)Eu_A^s(t)$$

This equation permits the value of the negotiation to rise based on speculation about the value of eventual breakdown. If one charts the Eu curve against time, it can sometimes turn upwards when the agent who is close to breaking down decides to postpone breakdown (as well as capitulation) until u^{bd} has risen.

Even more subtly, for agents who distinguish between unilaterally breaking down and being unilaterally broken down upon, the expression requires yet another parameter which weighs the probability of the former, e.g.

$$Eu_A(t) = Prob_A^{bd}(t)(u_A(bd) + (.5)u_A^{bd}(t)) + Prob_A^s(t)Eu_A^s(t)$$

when each is equally likely.

The agent who witholds an offer now creates a race between falling expectation due to pessimism and rising value of breakdown, due to resentment. Since acting on resentment is a kind of punishment for the agent who does not make progress, we refer to these agents as pessimism-punishment agents (PP agents).

A linear resentment (with no memory) is one in which $u^{bd}(t^{lp} + k) = \rho k$. This implies that $u^{bd} = 0$ when there is progress, and it implies there is no bound on how much u^{bd} can grow.

A linear PP agent is thus specified by two parameters, π and ρ.

11 Laissez-Faire Paths

For two PP agents, each agent is governed by a $Prob^{bd}$ and a u^{bd}. In the case of linear PP agents, there are four parameters to note: π_A, π_B, ρ_A, and ρ_B.

A closed offer set is one in which all settlements $s \in \Omega(t)$ that A values no more than $\nu^{\cdot}_A(t)$ are in $\Upsilon_A(t)$, and all settlements that B values no more than $\nu^{\cdot}_B(t)$ are in $\Upsilon_B(t)$.

Closed offer sets can therefore be characterized uniquely by the concession levels of each agent, $(\nu^{\cdot}_A(t), \nu^{\cdot}_B(t))$, from which one can infer $(\nu_{.A}(t), \nu_{.B}(t))$. For agents who close their offer sets, this point is like a state of the negotiation at time t. There are actually two more values to add to the state, the time since A's last progress, and the time since B's last progress. The state of negotiation for two PP agents is thus

$$\sigma(t) = (\nu_{.A}(t), \nu_{.B}(t), t_A^{lp}, t_B^{lp}) \ .$$

A successor state to $\sigma(t)$ is either a change in $\nu_{.A}$, because B makes a substantive new proposal, which resets t_B^{lp} to zero, or a change in $\nu_{.B}$, because A makes a new proposal, resetting t_A^{lp} to zero. In discrete time, A and B can both make progress in the same time step. There are two other possibilities: agreement and breakdown. Agreement can happen if A's expectation falls to $\nu_{.A}$, or if B's expectation falls to $\nu_{.B}$. Breakdown can happen if A's expectation falls below $u_A(bd) + \rho_A t_B^{lp}$, or if B's expectation falls below $u_B(bd) + \rho_B t_A^{bd}$.

A series of successor states is a laissez-faire path. All laissez faire paths end in an agreement or a breakdown. Given initial levels of concession, public valuations u_A and u_B, and public announcement of linear PP parameters, the trajectory and final state of a laissez-faire path is determinable.

The path of negotiation can be controlled by accelerating offers so that the agents move from one laissez-faire path to another.

12 Discussion

In this paper we have attempted three things. First, an expectation-driven negotiating agent is contemplated. The insistence on probabilities that are objective and the estimation of a probability of breakdown made it possible to force the agent to make rational concessions, so long as the expectation is falling due to empirically-based pessimism. The compensatory force was the willingness to punish based on resentment born of the procedural utility attached to unilateral breakdown on noncooperative negotiating partners.

Second, the resulting PP agents are decoupled from speculation about each other's mental states or solution strategies. PP agents can even announce that they are playing according to their π and ρ and are only slightly manipulable since they are essentially treating the game as a game against nature. An adversary can postpone offers until the latest time before breakdown, thus receiving the maximal offer set from the PP agent; but eventually the PP agent will break down unless concessions are matched. Memory parameters further protect the PP agent from manipulation.

Third, the picture of negotiation becomes a picture of path-selection among various laissez-faire paths, at least for a pair of PP agents negotiating with each other. The problem is not one-shot coordination within the set of possible settlements, e.g., the identification of a unique equilibrium. The problem is to identify and assess the acceptability of the path that the process of negotiation is taking, and to decide whether it is worth controlling that path with alternately motivated or alternately justified action.