# A Framework for Resolving Multiagent Collaboration Dilemmas

John Chia-chin Wang
*National Taiwan University, Taipei, Taiwan, Republic of China*
*jcwang@csie.ntu.edu.tw*

## Abstract

*Intelligent agents that participate in free-forming collaborations choose strategies to maximize their effectiveness in achieving goals. In a heterogeneous agent society where each has no knowledge of the logic behind other's motives or real intentions, it is difficult to conceive how an algorithm would provide an effective collaboration strategy. We propose a set of general, quantitative criteria for detecting collaboration situations leading to the duoagent collaboration dilemma (DCD). We show that DCD is a widely applicable collaboration scenario and it can take advantage of a class of problems social scientists have studied extensively. Lastly, we provide steps to developing practical computable strategies for agents to avoid and resolve problems in a subclass of DCD.*

## 1. Introduction

Imagine the following situation: your intelligent agent Alice finds agent Bob over the Internet, thinking that they needed each other to achieve their goals. Knowing that Bob may be intending to conceal facts or manipulate the communication so that he can be more effective in arriving at his own goals (at the cost of compromising Alice's competency), how best should Alice devote her resources and carry through the collaboration?

The obvious step towards solving the problem is probably to design an agent architecture to completely eliminate the possibility of "cheating." Work in distributed AI often takes this approach and designs frameworks to contain the behavior of the intelligent agents. However, *practical* results have been very limited [4] and it seems worthwhile to approach the problem from a different direction.

### 1.1. Fundamental difficulties

Direct analysis on the constituents of an intelligent agent community faces at least three major problems. Traditional computer science approaches may be fundamentally ineffective in dealing with collaboration decisions [6]. We identify three observations that complicate the situation.

**Computational imprecision** The control logic behind each intelligent agent can be completely opaque to one another. This situation is unlike traditional studies of computer systems where each component is fully described by a clear set of algorithms. Agents in general, especially when having conflicting goals, have no interest in exposing their own "bottomlines," hence we cannot compute the situations with the assumption that we always have accurate descriptions of them.

**Membership imprecision** Due to the asynchronous nature of the Real World (where communication inevitably incurs a loss of undetermined time), there is no way we can exactly label each agent to be in or out of a set. Try to determine the number of web pages connected to the Internet at a given instant, or the number of agents that will eventually receive your next broadcast message, and you should never be able to have an exact answer [3].

Thus even if each node can be specified with mathematical logic, we can hardly have an exact picture of all the entities participating in a collaboration to carry out a complete, overall analysis.

**Goal conflict** Each intelligent agent has its own set of *goals* to achieve at any given moment. There need not be any agreement among the agents on the goals, and agents can be stepping on each other's toes in the process of carrying out their tasks [5]. If the environment tries to set a "public standard" that interferes with an agent's goals (i.e., data encryption standard), we should expect the agent to find ways to invalidate the constraints to achieve its aims.

These three inevitable characteristics of intelligent agents research—computational imprecision, membership imprecision, and goal conflict—make our studies depart fundamentally from traditional computer science.

### 1.2. Implementation difficulties

In addition to the fundamental differences, there are several real problems when we choose to study agents that work over

1060-3425/98 $10.00 (c) 1998 IEEE

the Internet. The Internet has several peculiar qualities that used to receive very little attention but can be a real threat to enforcing multiagent policies or to protecting innocent agents. Three of which are discussed below.

**Intermittence** The Internet is based on the TCP/IP protocol, a reliable, asynchronous message-passing mechanism. When a network fault occurs, data packet transmission is retried until target responds with a positive acknowledgement. When an expected message is delayed, there is no way for the waiting agent to correctly deduce whether it was because the sender went down, the network caused transmission delay, or data became corrupt during delivery and is being retransmitted.

Under this condition, when a request goes without a response, we cannot decide whether the agent at the other end should be held responsible for not providing service. This makes it hard to enforce agreements over the network.

**Amorphism** On the Internet, entities are named by their *IP addresses*, a string of digits that allows parties to be located, functionally resembling telephone numbers. Imagine trying to determine the real identity of a stranger behind a telephone handset. It is not an easy task, if feasible at all. Worse yet, two voices coming from two different telephone numbers cannot be determined to have come from the same person or two different entities.

The unlimited forms an entity can assume on the Internet make it a greater impossibility to enforce agreements. Whenever it deems profitable to, an entity can instantly disappear and re-emerge under a different identifier.

**Suffocation** Nothing in the world today prevents an agent from blocking out another on the Internet. If anyone wants to paralyze an information search service, he can write a simple script to continuously feed the service with dummy queries until the server becomes too busy to usefully serve anyone else.

When threats are possible on the Internet, an agent can effectively shut down another's services or claim that others are inhibiting it from completing its tasks. Again responsibility for failing a contract becomes hard to determine.

To be applicable in reality, we need new solutions that can handle or avoid all these problems, and this is exactly what we have set out to study. We start by a formulation of the domain we wish to investigate, followed by specific strategies agents can make use of to strive in collaborations on such a domain. At all times we keep in mind the list of problems we have raised in this section and formulate our solutions to avoid building our foundation on quick sand.

## 2. Problem formulation

We start by defining the process of multiagent collaboration. We characterize a collaboration process with resources put into it and payoffs derived as a result.

### 2.1. Native characteristics

**Definition 1 (Collaboration payoff)** *Given* $R_1, \cdots, R_n$ *over a scalar field $\mathcal{F}$ (e.g., real numbers) as resources invested in a collaboration by agents $\alpha_1, \cdots, \alpha_n$, respectively, the* payoffs *to agent $\alpha_i$, $i = 1, \cdots, n$, of the collaboration is a function $p_i(R_1, \cdots, R_n) \in \mathcal{F}$.*

In order to compute the *exact* payoff from a collaboration, the above model would require we know the *exact* resources invested by each agent in an *n*-agent collaboration. This can be troublesome in situations involving agents with the quality of computational imprecision. Such agents may consume resources from channels that escape monitoring, effectively prohibiting an outsider to make a precise measurement of the qualities. Human agents are of a particular class that consume resources implicitly or in units difficult to measure (i.e., environmental effects, emotional effects, etc.).

The model we wish to arrive at must be more tolerant on the resource issue. As we will see, our formulation only requires a relative ordering among the $R_i$'s which is in most cases easier to identify than the actual values.

In collaborations, we wish to address the fact that two agents working together may produce more or less than the sum of each working individually, hence yielding a profit or deficit (it is for this profit that agents collaborate). For our model we will discuss only those payoff functions that are linear combinations of the resources invested by each participant.

**Definition 2 (Linear payoff)** *A* linear payoff *to agent $\alpha_i$ in an* n*-agent collaboration has the form $p_i(R_1, \cdots, R_n) = \sigma_i(R_1, \cdots, R_n) \sum_{j=1}^{n} c_j R_j$, where $\sigma_i(R_1, \cdots, R_n) \in \mathcal{F}$ is the profit share function with $\sum_{i=1}^{n} \sigma_i(R_1, \cdots, R_n) \leq 1.0$ and $c_j \in \mathcal{F}$ the* collaboration effectiveness factor *on agent $\alpha_j$'s contribution.*

Linear payoff offers an arithmetic model that is easy to compute. Perhaps more important, error ranges for linear payoffs are bounded proportional to the magnitude of uncertainty in $R_i$'s (recall computational and membership imprecisions). In our formulation, a $c_j$ greater than $1.0$ shows that the effort spent by agent $\alpha_j$ is indeed *amplified* through collaboration. A value less than $1.0$ for $c_j$ would indicate otherwise (probably due to incompatible working attitude, habit, style, etc.).

The $\sigma_i$'s are fractions of the total payoff that goes to $\alpha_i$. If $\sigma_i = 1/n$ for all *n* agents, we say the payoff shares are *completely external*. That is, the participating agents have absolutely no influence over the distribution of collaboration results. Fame and reputation are close examples to a completely external payoff in a collaboration. If $\sigma_i = R_i / \sum_j R_j$, then the payoffs are *completely internal*. Distribution of payment and gains in experience are examples of completely internal payoffs. There is a continuum of share ratios among the two extremes as we formalize below.

**Definition 3 (Quality of share)** *In a linear payoff, let the payoff share function $\sigma_i(R_1, \cdots, R_n)$ be of the form*

$$\frac{1}{n}(1 + q\frac{(n-1)R_i - \sum_{j \neq i} R_j}{\sum_{j=1}^{n} R_j})$$

*then $q \in \mathcal{F}$ is the* quality of share *factor.*

It is easy to verify that $q = 0.0$ denotes the condition of a completely external payoff share, and $q = 1.0$ denotes a completely internal payoff share.

**Definition 4 (Collaboration profit)** *Given collaboration payoff function $p_i$ for agent $\alpha_i$, the* collaboration profit *for the agent is defined as $\pi_i = p_i - R_i$. If $p_i$ is a linear payoff, then the profit function $\pi_i$ is called a* linear profit.

Again for simplicity, we will analyze only linear profits. Having defined profit for an agent in a collaboraion, we define the generic agent collaboration problem (ACP) as the following.

**Definition 5 (Agent collaboration problem)** *Given a set of agents about to participate in a (possibly infinite) sequence of collaborations, a solution to the* agent collaboration problem (ACP) *for each participating agent is an algorithm that decides a sequence of resource contribution decisions to maximize the agent's accumulated profit during the collaborations.*

## 2.2. Modeling duoagent collaborations with linear profits

The solution to an ACP depends on the nature of the interaction among the agents. We need to determine a class of interactions for which interesting solutions exist and can be found.

The case we investigate in this study is a class that includes exactly two agents, $\alpha_1$ and $\alpha_2$ where agent $\alpha_i$ has the ability to make a choice between investing resources of magnitude $R_i$ and $r_i$, and that $R_i > r_i$, then for agent $\alpha_1$, equations (1) through (4) follow directly from definitions in the previous section.

$$\begin{aligned}
\pi_1^T &= \pi_1(r_1, R_2) \\
&= \frac{1}{2}(1 + q\frac{r_1 - R_2}{r_1 + R_2})(c_1 r_1 + c_2 R_2) - r_1 \qquad (1) \\
\pi_1^R &= \pi_1(R_1, R_2) \\
&= \frac{1}{2}(1 + q\frac{R_1 - R_2}{R_1 + R_2})(c_1 R_1 + c_2 R_2) - R_1 \qquad (2) \\
\pi_1^P &= \pi_1(r_1, r_2) \\
&= \frac{1}{2}(1 + q\frac{r_1 - r_2}{r_1 + r_2})(c_1 r_1 + c_2 r_2) - r_1 \qquad (3) \\
\pi_1^S &= \pi_1(R_1, r_2) \\
&= \frac{1}{2}(1 + q\frac{R_1 - r_2}{R_1 + r_2})(c_1 R_1 + c_2 r_2) - R_1 \qquad (4)
\end{aligned}$$

Payoffs for $\alpha_2$ are symmetrically defined.

For each agent, $\pi^T$ can be interpreted as the *temptation to defect*, or the profit for cheating on a cooperating partner. $\pi^R$ is the *reward for cooperation*, the profit when both agents cooperate. $\pi^P$ is the *penalty for defection* when both agents defect. Lastly $\pi^S$ is the *sucker's payoff* for the agent that is being cheated for cooperating.

Four quantities form six inequalities. We are to determine which are the conditions that describe the case we wish to investigate.

**Reward for cooperation vs. penalty for defection** This is probably the easiest inequality to decide. Unless the situation requires $\pi^R > \pi^P$, both agents will look forward to cheating on each other rather than to trying to achieve mutual cooperation. Usually it is not the case that both agents devoting less would guarantee better a return.

**Temptation to defect vs. reward for cooperation** We discuss the case of $\pi^T > \pi^R$, where the agent is tempted to do less for a greater profit (by making the other party do the work). In a society where cheating off a well-intended partner *does not* yield any extra profit, it is easy to reach an agreement to cooperate where neither agent has the motive to defect. More interesting is the case $\pi^T > \pi^R$ observed in many situations where we have a dilemma.

**Penalty for defection vs. sucker's payoff** With its partner not putting effort into a collaboration, we should agree to have an agent be better off stand clear from the collaboration. Usually an agent who conserves its resources in this situation should do better than one that spends them unconditionally. That is to say, $\pi^P > \pi^S$.

These three inequalities are actually all that's needed to decide the remaining three conditions, namely $\pi^T > \pi^S$, $\pi^T > \pi^P$, and $\pi^R > \pi^S$.

We impose one more set of conditions to limit our scope of study, and that is the requirement of $\pi_1^R + \pi_2^R > \pi_1^T + \pi_2^S$ and $\pi_1^R + \pi_2^R > \pi_2^T + \pi_1^S$. This requires that the two agents would wish to develop mutual cooperation rather than victimizing each other in alternation. This is often favored in real situations where if for nothing else, we would like to promote a peaceful atmosphere rather than a strictly competitive one.

**Definition 6 (Duoagent collaboration dilemma)** *An agent collaboration problem with exactly two agents for whom the profit functions obey the inequalities $\pi_i^T > \pi_i^R > \pi_i^P > \pi_i^S$ for $i = 1, 2$, $\pi_1^R + \pi_2^R > \pi_1^T + \pi_2^S$ and $\pi_1^R + \pi_2^R > \pi_2^T + \pi_1^S$, is a* duoagent collaboration dilemma (DCD).

Once these inequalities hold, our problem domain covers a well-studied social problem: the prisoner's dilemma game (PDG) [2]. In one interpretation of PDG, two criminals are arrested for a crime. If they both cooperate and conceal the evidence, they would both be acquitted (both receive the score of $\pi^R$). If exactly one of them defects and discloses the evidence, he is acquitted and given an reward (score of $\pi^T$), while the one who conceals the evidence gets an especially heavy punishment (score of $\pi^S$). If both disclose in a mutual defection, they would both be convicted with a typical sentence (score of $\pi^P$). The scores are required to follow the same inequality relationships we derived for a DCD.

**Table 1: Payoff structure for a PDG**

| | | Player B: Cooperate | Player B: Defect |
|---|---|---|---|
| Player A | Cooperate | $\pi_1^R, \pi_2^R$ | $\pi_1^S, \pi_2^T$ |
| | Defect | $\pi_1^T, \pi_2^S$ | $\pi_1^P, \pi_2^P$ |

If we examine Table 1 with the inequalities in mind, we realize that for each agent, choosing defection (choosing $r_i$ over $R_i$ where $R_i > r_i$) is the "safest" strategy irrespective of what the other does (in game theory lingo, the *dominant strategy*). Since we rule out the possibility of agents reaching an enforceable contract before making its choice (due to intermittence, amorphism, etc.), choosing cooperation runs the risk of receiving the lower score ($\pi^R$ or $\pi^S$) than that defection yields ($\pi^T$ or $\pi^P$, respectively). Consequently it makes sense to defect unconditionally, in which case there is a possibility of receiving the highest score $\pi^T$ and at the same time avoiding getting the worst outcome $\pi^S$. Of course, the dilemma is that by doing so, the agents miss the opportunity of scoring the sum of $\pi_1^R + \pi_2^R$ which is better than either $\pi_1^T + \pi_2^S$ or $\pi_2^T + \pi_1^S$, and is also greater than $\pi_1^P + \pi_2^P$ that mutual defection yields.

Axelrod [1] examines this intriguing situation in repeated PDGs and gathers contests for experts in game theory, economics, etc., to design software subroutines to compute the most favorable strategy during encounters of this type. The results convey great insights: while there is no one strategy that outperforms all its opponents, those that are *nice* (cooperate until partner defects), *retaliatory* (punish by defection if partner defects voluntarily), *forgiving* (cooperative as soon as partner bends and starts cooperating), and *clear* (easy for partner to anticipate) have an advantage. In fact the most successful strategy "TIT FOR TAT" does nothing but repeat what its partner did in the previous encounter, whereas some of the less successful went as far as developing statistical models trying to find opportunities to take advantage of their partner.

PDGs are interesting but rarely applicable in a computationally imprecise domain. PDGs require the precise values of the payoffs be written in a table and each agent's decision be truthfully announced to the public. This is by far too much to ask for among intelligent agents that we can at best treat as black boxes. Later on we will analyze strategic implications of DCDs and see how we can make results of the PDG useful in our domain.

We need to look further and see what are required of the individual agents to fall into a DCD when they join. For example, you might insist that, even if partner sits around and does nothing towards the collaboration, there are still situations where an agent's increased effort can lead to positive profit gains. That is to say, $\pi^S > \pi^P$ is also a plausible case. This can happen if an agent's effort is at least doubled during a collaboration (with a collaboration effectiveness factor over 2.0), in which case after giving half of the share (when $q$ is 0.0) to the partner and taking away its original resource costs, the agent still makes a profit. In the next section, we will do detailed analysis to see how the collaboration parameters, namely the $R_i$'s, $r_i$'s, $c_i$'s and $q$'s, will force the situation to be an instance of the DCD.

## 2.3. Characteristics of agents in the DCD

It is delightful to find that we have located a sub-problem of the ACP where we are able to take advantage of results from game theory. In this section we continue to look at this particular situation and see how it maps to the optimal strategy for the agents.

The characteristic inequalities $\pi^T > \pi^R > \pi^P > \pi^S$, $\pi_1^R + \pi_2^R > \pi_1^T + \pi_2^S$ and $\pi_1^R + \pi_2^R > \pi_2^T + \pi_1^S$ yield relationships among the most basic collaboration variables, namely the $R_i$'s, $r_i$'s, $c_i$'s and $q$'s. We first define a number of shorthand notations to simplify the conditional expressions and help develop our intuition towards the inequalities.

**Definition 7 (Net impact, gross impact)** *Given an agent $\alpha_i$ participating in a collaboration with an effectiveness factor $c_i$ and the freedom to choose between investing resources $R_i$ and $r_i$ with $R_i > r_i$, the* net impact *of the agent is $net_i = (1-(c_i-1))(R_i - r_i) = (2-c_i)(R_i - r_i)$ and the* gross impact $grs_i = (1+(c_i-1))(R_i - r_i) = c_i(R_i - r_i)$.

The impact of an agent is intuitively its range of performance dynamics, or the difference between the most and least possible resource devotions. Impact indicates the ability of an agent in dictating the outcome of a collaboration. The gross impact takes collaboration effectiveness into account, hence is the real impact perceived by the agents during collaboration. The net impact on the other hand indicates an agent's dynamic range taken away effects due to collaboration. Gross impact and net impact together gives an agent a sense of the real inner capability (the true range of resource devotion) it possesses.

Another expression that frequently appears among the inequalities is shorthanded below as the "differential share." It indicates the perceived difference between the payoff that goes to the agent and that to its partner. Dependence on differential share indicates the agent's need to look at "if I am doing better/worse than my partner."

**Definition 8 (Differential share)** *In a duoagent collaboration problem where agents $\alpha_1$ and $\alpha_2$ contribute $R_1$ and $R_2$, respectively, the* differential share *to agent $\alpha_i$ is the function $dsh_i(R_1, R_2) = \frac{R_i - R_j}{R_1 + R_2}(c_1 R_1 + c_2 R_2)$, $i \neq j$, where $c_1$ and $c_2$ are the corresponding collaboration effectiveness factors for $\alpha_1$ and $\alpha_2$.*

Figure 1 shows the shape of the coefficient $\frac{R_1 - R_2}{R_1 + R_2}$ of a differential share function for $R_1$ and $R_2$ in the range of 0.0 to 10.0. It can be seen as a bias of share towards the partner with greater contributions.

Solving inequalities for $\alpha_1$ in a duoagent collaboration dilemma (DCD) and we obtain the following simplified inequalities.

$$\pi_1^T > \pi_1^R : \quad net_1 > q(dsh_1(R_1, R_2) - dsh_1(r_1, R_2)) \quad (5)$$

$$\pi_1^P > \pi_1^S : \quad net_1 > q(dsh_1(R_1, r_2) - dsh_1(r_1, r_2)) \quad (6)$$

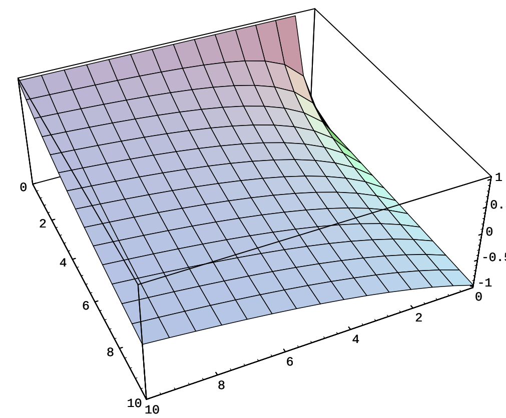


**Figure 1: A plot of the coefficient $\frac{R_1 - R_2}{R_1 + R_2}$ (with discontinuity at $R_1 = R_2 = 0.0$).**

Inequalities (5) and (6) explain that in a DCD, an agent's net impact on the collaboration must be greater than the fair share gained by contributing more. For completely external payoff shares where $q = 0.0$, this leads to the above-mentioned requirement that $c_1 < 2.0$. The intuition can be obtained by plugging in various values for $c_1$ back into inequalities (1) through (4) with $q = 0.0$. We see that for $c_1$ greater than 2.0, the agent unconditionally receives more as it devotes greater investment. In that case the agent never feels being cheated or has a temptation to defect, and we do not have a dilemma to solve.

In reality, most likely we are not in a situation where our efforts double just by participating in a collaboration, regardless partner's actions. That is to say that we usually find equations (5) and (6) automatically satisfied in interesting cases. Situations where $q \neq 0.0$ are slightly more tricky and we will leave the interpretation to the interested reader (possibly with the aid of Figure 1).

$$\pi_1^R > \pi_1^P : \quad grs_2 - net_1 > q(dsh_1(R_1, R_2) - dsh_1(r_1, r_2)) \quad (7)$$

Inequality (7) basically says that the difference between the real impact partner has and my own net impact (the latter being my impact taken away the benefits of collaboration), should be proportional to the differential share resulted from turning into mutual cooperation. Roughly, if partner does not have more power than I do in improving the situation, then there is no incentive for me to expect to benefit from turning mutual defection into mutual cooperation. Solving (7) with (5) and (6) actually yields $c_i > 0.0$ if $q = 0.0$.

$$\pi_1^R + \pi_2^R > \pi_1^T + \pi_2^S : \quad grs_1 > net_1 \quad (8)$$

$$\pi_1^R + \pi_2^R > \pi_2^T + \pi_1^S : \qquad grs_2 > net_2 \tag{9}$$

Inequalities (8) and (9) require that the gross impact be greater than the net impact. This leads to $c_i > 1.0$ regardless what value $q$ takes. It can be interpreted that the collaboration *must have* an incentive for each agent to devote more effort, since $c_i > 1.0$ indicates that the collaboration in nature expands the value of the resources an agent puts into it.

Given agents whose range of resource dedication can be identified, and a collaboration where the effectiveness factors are in a plausible range (greater than 1.0 and under 2.0 when the share is completely external), we have determined the quality of share necessary for the collaboration to be a DCD.

## 3. Collaboration strategies

In the previous section we looked at the DCD from an omniscient eye to get a clear view of the situation. However, an agent can only determine the nature of the situation from its profits and deficits. In this section we will study strategies that allow an agent to perform "well" in DCD scenarios.

### 3.1. Rounds in collaboration sequences

**Proposition 1 (Destined defection)** *An agent participating in a DCD equipped with prior knowledge about the (finite) number of rounds to collaborate is guaranteed the best accumulated profit if it always defects.*

The proposition can be proved by induction on the number of rounds. If there is exactly one round of collaboration (that the agents will never meet again afterwards), an agent choosing defection is in a position where another agent cannot make it "lose." That is, for any given choice of the partner, an agent can only score lower if it chooses cooperation (since $\pi^R < \pi^T$ and $\pi^S < \pi^P$). Assuming destined defection holds for fixed $k$-round DCDs, then for a $(k+1)$-round game, all except the first move are bound to be defections. Knowing it cannot change the subsequent choices of the partner, an agent would defect also on the first round following the same logic for exactly one round.

The result from mathematical induction does not apply when the number of rounds is not fixed (e.g. when a dice toss decides whether there will be a next round), with the possibility that the sequence of encounters is infinite. The fact that ($\pi^R > \pi^P$) in a DCD makes it enticing for agents to finding mutual cooperation possibilities rather than falling into the less desirable state of mutual defection if none of of the two converts. The prospects of mutual cooperation is non-existent in the case of fixed rounds (see the inductive step). Since we have already ruled out the possibility of having an enforceable contract among intelligent agents (due to intermittence and amorphism), the only way an agent can find the signs to establish mutual cooperation is by knowing the results of its partner's previous collaborations. Actually for some interesting cases (in fact, those that coincide with the PDG) this requirement can be relaxed as we shall see.

We need a *public, neutral and unmistakable service* that faithfully records the profits for each agent in every collaboration (in a publicly known measurement such as market currency). Monthly bank statements and the U.S. government's Internal Revenue Service are examples of such services in human terms. Such a service would prevent an agent from running away from its previous encounters since all future partners will see the agent's past. Once an agent commits to its first collaboration endeavor, it has effectively started interacting with all other agents. The only way an agent can quit from its history is by applying amorphism, trashing all its past relationships with its decision.

### 3.2. DDCD strategies with binary history

One way an agent can look for signs of cooperation possibilities is by looking at partner's actions during previous encounters. In PDG, this is done by having the participants publicly state their choice of C (for cooperation) or D (for defection). However, in DCDs usually this is not possible. Computational imprecision again kicks in and prevents others from determining the actual amount of resources the partner invested during encounters, especially when the difference between $R_i$ and $r_i$ is small.

The discernable DCD (DDCD) is a subclass of the DCD domain we can prove correspondence with the PDG in applying collaboration strategies. In a DDCD, the two agents have *sufficiently equal* ability in investing resources and attract sufficiently equal collaboration effectiveness factors, denoted as $R_1 \approx R_2$, $r_1 \approx r_2$ and $c_1 \approx c_2$, although the exact numbers may be undetermined. In addition, the $R_i$'s must be *discernably greater* than 0.0 which is in turn discernably greater than the $r_i$'s, denoted as $R_i \gg 0.0$ ($R_i$ is *discernably positive*) and $0.0 \gg r_i$ ($r_i$ is *discernably negative*). We say $A \gg B$ if the fact $A > B$ is recognized by all participating agents based on their perception of their profits earned. $A \approx B$ if and only if neither $A \gg B$ nor $B \gg A$. The $\gg$ relation satisfies transitivity. Consequently $R_i \gg r_i$ in a DDCD.

The DDCD is usually the domain that human beings prefer to work on. People who collaborate on one task usually have comparable capabilities and expectations towards the collaboration. When professional skills are unmatched, for example, monetary compensation (in the form of payments) or psychological rewards (as social recognition or fulfilled sympathy) is induced to make the resource investments as much matched as possible. In addition, when either one decides to cheat in collaboration, usually the partner can detect the situation. It is under these condition that humans feel comfortable working since means are available to measure the contribution the partner has committed. Now our intelligent agents are enabled to base their decisions on similar terms.

**Proposition 2 (DDCD-PDG correspondence)** *Given agents $\alpha_1^D$, $\alpha_2^D$ in DDCD and agents $\alpha_1^P$, $\alpha_2^P$ in repeated PDG, where $\alpha_i^D$ invests $R_i$ for every $\alpha_i^P$'s choice of C and $r_i$ for the choice of D. Define payoff structure for $\alpha_1^P$, $\alpha_2^P$ as the following*

$$\begin{aligned}
\pi_1^T &= \pi_2^T = \frac{1}{2}(cR - (2-c)r - q\frac{R-r}{r+R}(cr+cR)) \\
\pi_1^R &= \pi_2^R = (c-1)R \\
\pi_1^P &= \pi_2^P = (c-1)r \\
\pi_1^S &= \pi_2^S = \frac{1}{2}(cr - (2-c)R + q\frac{R-r}{r+R}(cr+cR))
\end{aligned}$$

*then for every strategy of $\alpha_1^P$ and $\alpha_2^P$ that depends on partner's choice between C and D, there exist a corresponding decision rule for $\alpha_1^D$ and $\alpha_2^D$ that depends on a discernably positive or negative collaboration profit such that $\alpha_1^P$ scores higher than $\alpha_2^P$ if and only if $\alpha_1^D$ is more profitable than $\alpha_2^D$.*

For every strategy in PDG that depends on partner choosing C in the previous round, the corresponding strategy in DDCD depends on a profit gain; for every strategy in PDG that depends on D, that in DDCD depends on a profit loss, or a deficit.

The payoff structure follows naturally from the profit functions in a DCD where $R_1 = R_2$, $r_1 = r_2$ and $c_1 = c_2$. In DDCD, this is exactly how agents perceive their partner's resource investments. There cannot be any profit gain or loss due to unequal collaboration parameters or else the required sufficient equality $R_1 \approx R_2$, $r_1 \approx r_2$, and $c_1 \approx c_2$ would fail. Therefore the payoffs *truthfully reflect the profitability of the agents*.

Given $2.0 > c > 1.0$ and $R_i \gg 0.0$ and $0.0 \gg r_i$, we easily see that $\pi_i^T \gg 0.0$, $\pi_i^R \gg 0.0$ (a discernably positive profit), and $0.0 \gg \pi_i^P$, $0.0 \gg \pi_i^S$ (a discernably negative profit). The former corresponds to partner's choice of $R$ (C in PDG) and the latter corresponds to $r$ (D in PDG). This is all it is required to establish DDCD-PDG correspondence.

There is one additional requirement before coding a PDG strategy into a decision program in DDCD. There must be the aforementioned global information service that supplies a *binary history* of an agent's collaborations: if a previous partner has made a profit with this agent, then agent receives a '1' bit; otherwise it is stamped a '0' bit. Recall "TIT-FOR-TAT" which Axelrod found effective in his tournaments. An agent in DDCD implementing TIT-FOR-TAT would simply look at the last bit in its partner's history and devote $R$ if it finds an '1', $r$ if it finds a '0'.

It follows from here that a DDCD agent with access to global binary history can implement all game theoretic strategies for solving PDGs that are computable, be it finite, regular (for agents based on finite-state automata), or context-free (for those based on the Turing machines).

## 4. Conclusion

We have identified the requirement for multiagent collaboration scenarios where strategies for the prisoner's dilemma game (PDG) can apply. For this domain, called the discernable duoagent collaboration dilemma (DDCD), we have developed mechanisms for programming intelligent agents to implement any computable game theoretic strategy found for the PDG, all without violating the assumption of computational imprecision. The agents base their decisions purely on profits or deficits for other agents, which are recorded in binary history by some neutral referral agency. This is in contrast against game theoretic results which depend on the *decisions* of the agents that are difficult to recognize by other agents in practical terms.

We have also, on a general basis, introduced the formal notion of collaboration efforts, effectiveness, profits, and profit share. Based on these structures we can define payoffs that are linear in nature, which are simple to evaluate arithmetically and have error ranges bound within a multiple of the sampling precision. This may serve as a promising framework for future investigations.

## 5. Future research

We have dug into a small pit in the universe of all agent collaboration problems, even though we must admit this pit contains a great deal of interesting intelligent agent interactions. This is not to say that other situations are not worthy of further investigation.

For DDCDs we have devised a scheme for leveraging the best PDG strategies social scientists have found or will find in the future. We would be eager to relax DDCDs to perhaps categorical DCDs: those where agents can perceive multiple categories of profits rather than just the binary profit/deficit. It would be curious to see whether the ability to discern the degree of profitability allows an agent be more *efficient* in a search for profit, or if the two DCDs are actually equivalent.

A vast portion of the problem space we have not touched on is that with more than two participating agents during each collaboration encounter. New problems surface when we consider the possibility of intelligent agent alliances and voting processes. We might again find game theoretical results, especially those involving multiple players, to be insightful in leading us towards practical solutions in interesting multi-agent ACPs.

## References


[1] Robert Axelrod, *The Evolution of Cooperation*, Basic Books, New York, 1984.

[2] Andrew Colman, *Game Theory and Experimental Games*, Pergamon Press, Oxford, 1982.

[3] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi, "Coordinated Attack", *Reasoning About Knowledge*, MIT Press, Massachusetts, 1995, pp. 176-183.

[4] Les Gasser and Michael N. Huhns, "Themes in Distributed Artificial Intelligence Research", *Distributed Artificial Intelligence*, vol. 2, Pitman, London, 1989.

[5] Jeffery S. Rosenschein and Gilad Zlotkin, "Designing Conventions for Automated Negotiation", *AI Magazine*, AAAI Press, Fall 1994.

[6] Dave West, "Hermeneutic Computer Science", *Communications of the ACM*, vol. 40 no. 4, ACM Press, New York, April 1997.