# Metastrategies in the Colored Trails Game

Tracking Number: 127


## ABSTRACT

This paper presents a novel method to describe and analyze strategic interactions in settings that include multiple actors, many possible actions and relationships among goals, tasks and resources. It shows how to reduce these large interactions to a set of bilateral normal-form games in which the strategy space is significantly smaller than the original setting, while still preserving many of its strategic characteristics. We demonstrate this technique on the Colored Trails (CT) framework, which encompasses a broad family of games defining multi-agent interactions and has been used in many past studies. We define a set of representative heuristic metastrategies in a three-player CT setting. When players' strategies are chosen from this set, the original CT setting decomposes into smaller, bilateral games that correspond to the well-known Prisoners' Dilemma, Stag Hunt and Ultimatum games. We formalize a set of criteria that need to hold for a general CT setting in order to guarantee the decomposition. Using sampling and simulation, these criteria are shown to work in practice for many different instances of the CT game. Our results have significance for multi-agent systems researchers in mapping large multi-player task settings to well-known bilateral normal-form games in a way that facilitates the analysis of the original setting.



## Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]; J.4 [**Social and Behavioral Sciences**]

## General Terms

Design, Experimentation

## Keywords

Colored Trails, Metastrategies


## 1. INTRODUCTION

Computer systems are increasingly being deployed in task settings where multiple participants make decisions together—whether collaboratively, competitively or in between—in order to accomplish individual and group goals. Recently a new testbed has been introduced to enable evaluation and comparison between computational strategies for a wide variety of task settings, i.e. the Colored Trails (CT) framework [3].[1] CT has spawned 30 publications in diverse multi-agent settings, such as repeated negotiation, interruption management, team formation and space research [4, 12, 10]. CT is particularly attractive because it is grounded in a situated task domain and is rich enough to reflect features of real-life interactions. The CT framework encompasses a family of different games that provide an analogue to the ways in which goals, tasks and resources interact in real-world settings. CT is parametrized to allow for increasing complexity along a number of dimensions, such as task complexity, the availability of and access to information, the dependency relationships that hold between players, and the communication protocol. In abstracting from particular domains, CT provides a general framework to analyze multi-agent interactions.

A fundamental problem when performing (game-theoretic) analysis of complex games is dealing with large action spaces. Once we go beyond typical two-player two-action normal form games, the curse of dimensionality occurs in terms of finding equilibria and analyzing dynamics. For example, when analyzing the evolutionary dynamics of auctions or Poker, we need to abstract over atomic actions by introducing metastrategies [18, 16, 14], thus reducing large-scale interactions to smaller games. In a similar vein, this paper suggests a way of reducing multi-player interactions in the CT framework to a set of smaller games. It provides a mapping between a particular CT task setting and normal-form games in a way that preserves much of the strategic qualities of the original setting. To this end, it defines a set of heuristic metastrategies for each player that are domain-independent and make minimal assumptions about the way other players make decisions. Moreover, it lays down a set of criteria that need to hold such that the generated CT game instances are "interesting", implying that they allow a reduction to canonical social dilemma games taking place between (pairs of) players, i.e., Stag Hunt, Prisoners' Dilemma, and Ultimatum games. In these games, the metastrategies correspond to Nash equilibria and/or Pareto-optimal strategies. The mapping from CT game instances to well-known social dilemmas allows to compare participants' behavior in CT with prior results from these smaller, more traditional settings. Given the mapping of the CT game to social dilemmas, our analysis is extended by assessing the effect of adding social factors to participants' decision-making process.

Results from simulation experiments that sample thousands of CT game instances confirm that participants' outcomes from playing metastrategies in the original CT game correspond to the outcomes from playing the same strategies in the reduced Prisoners' Dilemma, Stag Hunt, and Ultimatum games. The results in this paper have significance for agent-designers in that they facilitate the


**Cite as:** Metastrategies in the Colored Trails Game, Paper No. XXX, *Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, Yolum, Tumer, Stone and Sonenberg (eds.), May, 2–6, 2011, Taipei, Taiwan, pp. XXX–XXX.



[1]Colored Trails is free software and is available for download a http://www.eecs.harvard.edu/ai/ct. A complete list of publications can also be found at this link.

comparison of computational strategies in different task settings with results obtained in more traditional idealized settings. Moreover they allow to generate new types of interactions in task settings that meet scientifically interesting criteria.

## 2. RELATED WORK

The idea to consider aggregate or metastrategies for facilitating (game-theoretic) analysis of a complex game is not new. In related work, strategies are often aggregated using heuristics, allowing the construction of e.g. *heuristic payoff tables* [18, 19]. Generally, a normal-form-game payoff matrix is replaced by a heuristic payoff table, since assembling all possible actions into a matrix is impractical for complex games (the resulting matrix would have too many dimensions). A heuristic allows to define metastrategies over the atomic actions, reducing the number of actions that have to be explicitly taken into account. A metastrategy typically represents a philosophy, style of play, or a rule of thumb.

Recent domains in which the heuristic approach has been followed include auctions [17, 11] and Poker [16]. In these domains, expert knowledge is available to assist in the establishment of suitable heuristics. For instance, in auctions, there are many well-known automated trading strategies such as Gjerstad-Dickhaut, Roth-Erev, and Zero Intelligence Plus [15, 14]. In Poker, experts describe metastrategies based on only a few features, such as players' willingness to participate in a game, and players' aggression-factor once they do participate. Examples of metastrategies in Poker, based on these features, are the tight-passive (a.k.a. Rock), tight-aggressive (a.k.a. Shark), loose-passive (a.k.a. Fish) and loose-aggressive (a.k.a. Gambler) metastrategies. Depending on the actions taken by a player over a series of games, it may be categorized as belonging to a specific type of player, i.e., as using a certain metastrategy. This allows researchers to analyze real-world Poker games, in which the metastrategy employed by each player in a particular series of games can be identified. Subsequently, obtained payoffs in this series of games may be used to compute heuristic payoff tables for each metastrategy [16]. These tables then allow to study the evolutionary dynamics of Poker.

In this paper, we pursue a similar approach, although a lack of heuristic expertise implies that we need to first perform an in-depth study of the game and possible means of aggregating strategies. Since expert knowledge on heuristics within the CT framework is not available, we cannot readily label a certain chip exchange as being, e.g., an egocentric or a social one. We aim to provide an analysis that does allow us to label chip exchanges in this manner.

We discuss three distinct levels within the CT framework. On the highest level, we have the complete *framework* itself, i.e., all possible CT games. The intermediate level identifies a certain *game* within the framework, e.g., the three-player variant we study in this paper. The lowest level is a *game instance*, e.g., one specific board configuration with a certain allocation of chips and a certain position for each of the three players and the goal. Going up from the lowest level, we see that players can perform certain *actions* in a CT game instance, can adhere to certain *strategies* in a CT game, and can use certain *metastrategies* in the CT framework.

While we restrict our analysis to one CT game (the three-player variant discussed below), the same analysis also applies to other games within the framework. Therefore, the analysis indeed leads to the identification of metastrategies. These metastrategies may be used as a solid basis to come up with heuristic payoff tables.

## 3. COLORED TRAILS

We focus on a variant of a three-player negotiation variant [2] of CT that includes a board of 4x4 squares, colored in one of five colors. Each player possesses a piece located on the board and a set of colored chips. A colored chip can be used to move a player's piece to an adjacent square (diagonal movement is not allowed) of the same color. The general goal is to position pieces onto or as close as possible to a goal location indicated by a flag. Each player receives points purely based on its own performance. There are three distinct players in the game: two proposers ($P1$ and $P2$) and a responder ($R$). Figures 1(a) and 1(c) show two examples of game instances. These instances include game boards with goal and player locations, as well as the chip sets that have been allocated to each player. The two instances will be used as running examples throughout the paper. For example, in Figure 1(a), proposer $P1$ is missing a green chip to get to the goal (by moving left-up-up), proposer $P2$ is missing a gray or green chip (moving up-up-right or up-right-up) to get to the goal, and responder $R$ is missing a gray chip and a blue chip to get to the goal (moving right-3up-right).

Proposers can offer a chip exchange to the responder. The responder can accept exactly one or no proposal at all. Although all players can observe the board, proposers cannot observe each other's chips, but can observe the chips of the responder. The responder can observe the chips of all of the players. Both proposers make offers to the responder simultaneously; they cannot observe each other's offer. The responder can only accept or reject proposals and is not allowed to make a counter-proposal.

The CT game is divided into a sequence of three phases and ends with an automatic evaluation.

**Initial phase.** The game board and the chip sets are allocated to the players. This initial phase allows participants to locate their own piece on the board and reason about the game.

**Proposal phase.** The two proposers can make chip exchange offers to the responder.

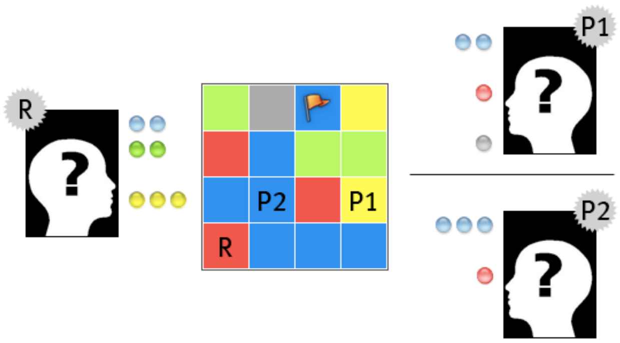
**Reaction phase.** The responder is presented with the two proposals. It can accept one of the proposals or reject both.

**Termination and Scoring Phase.** In this phase, players automatically exchange chips if they have reached agreement, and the icon of each player is advanced as close as possible towards its goal (using the Manhattan path with the shortest distance) given the result of the negotiation. The game ends and scores are automatically computed for each player: for every step between the goal and the player's position, 25 penalty points are subtracted. For every chip the player has not used, it receives 10 extra points. All players receive a 100 point bonus to guarantee that scores are not negative.

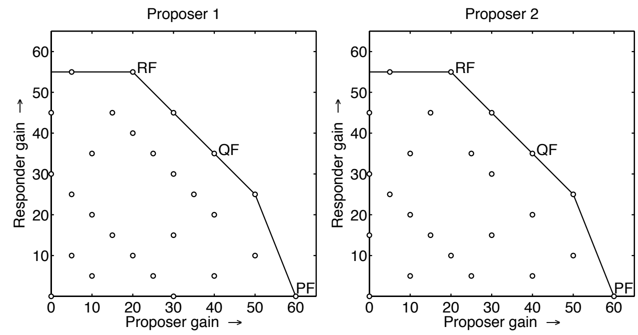
In the current paper, we use the following terminology associated with scores. First, the *base score* for a player $p \in \{R, P1, P2\}$ is the score the player receives when there is no agreement.[2] Second, the *gain* for a player $p$ and a chip exchange proposal $s$ denotes the difference in the score in the game (given that $s$ is realized) and the base score, and is denoted as $G_p(s)$.The base score for $p$, i.e., the gain when there is no agreement, is denoted as $G_p(\emptyset)$.

For example, in Figure 1(a), $G_{P1}(\emptyset) = 80$. This is because if there is no agreement, the player can only move one square to the left by using its red chip. It is still two squares away from the goal, yielding $2 \times 25 = 50$ penalty points. It has 3 remaining chips, yielding $3 \times 10 = 30$ points. With the 100-point bonus, the total base score becomes 80. In this particular game, $G_{P2}(\emptyset) = 80$ as well, with the optimal move being one to the right, using one red chip. The responder has a base score of 75; it can spend two blue

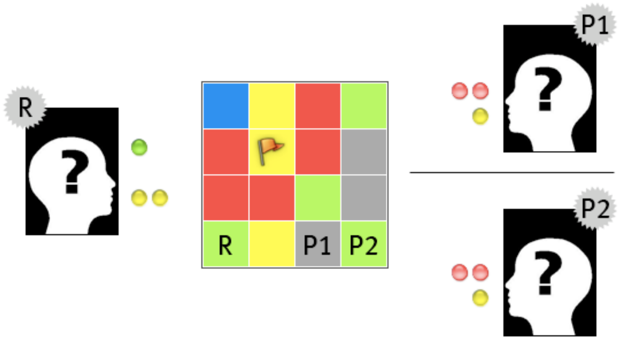
[2]Whenever we are not referring to one specific proposer, we will use the general notation '$P$' when we imply '$P1$ or $P2$'.



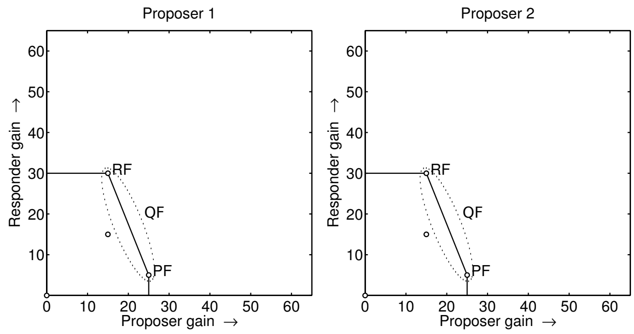

(a) First example Colored Trails game instance.




(b) Gain graph of the game instance presented in (a).




(c) Second example Colored Trails game instance.




(d) Gain graph of the game instance presented in (c).

**Figure 1: Example Colored Trails game instances and gain graphs. In (a) and (c), the three players ($R$, $P1$ and $P2$) are shown, along with their chip sets. The two proposers cannot observe each others' chip sets. All players can see the board, on which their locations are indicated, as well as the goal state (a yellow flag). In (b) and (d), we show the gain graphs for both proposers. These graphs plot proposer gain versus responder gain for each possible proposal with non-zero benefit. The convex hull in this graph denotes the Pareto-front. The meta-strategies PF, RF and QF are located on this front, as indicated. In (b), QF is a pure meta-strategy; in (d), it is a mixed meta-strategy (of PF and RF), since there is no proposal on the convex hull between PF and RF.**

chips to go right and upward, yielding a distance of 3 to the goal (i.e., 75 penalty points) and 5 remaining chips (50 points), plus 100 bonus points. One possible proposal for $P1$ is to offer a red and a grey chip for a blue chip, a green chip, and three yellow chips from the responder. In this case the proposer can get to the goal, and receives a gain of 60. Meanwhile, the responder can use this exchange to get one square away from the goal, but it uses all of its chips then. The gain from this exchange to the responder is zero.

# 4. DEFINING METASTRATEGIES

Although the rules of the CT game are simple, it is not trivial to analyze. Both proposers need to reason about the tradeoff between making beneficial offers to the responder and offers that are beneficial for themselves, especially because they compete with each other for making the best offer to the responder. Moreover, the number of possible strategies is large. In the example instance presented in Figure 1(a), the number of unique proposals for $P1$ is 240, while $P2$ can choose from 144 unique proposals.[3] The responder can choose to accept or reject any of these offers, so the size of the strategy space for the responder is $240 \times 144 \times 2$. The size of the combined strategy space makes it difficult to analyze this game in a principled way. In this section we show how to reduce this large setting to smaller interactions in a way that preserves the strategic flavor of the original CT scenario.

[3]Two proposals are unique if they do not use the same chips.

## 4.1 Initial Assumptions

We first describe two assumptions we make about the various players in the game. We will relax the first assumption later.

**Rational responder.** The responder $R$ has three possible actions, i.e., to accept the proposal of $P1$, to accept the proposal of $P2$, or to accept neither of them. For the responder, the game is thus similar to an Ultimatum game with proposer competition [9]. Initially, in our analysis, we assume that the responder plays according to a rational strategy. If both proposals do not provide it with a positive gain, it rejects both; if both proposals yield an equal gain, it accepts one of them with equal probability for both; if one proposal is strictly better, it accepts this proposal.

**Semi-rational proposers.** In order to select a strategy, i.e., a proposal to offer to the responder, proposers have to take into account the gain resulting from each proposal for themselves as well as the responder. For our analysis, we assume that proposer $P$ limits the set of possible proposals to those that (1) lead to a non-negative personal gain, i.e., $G_P(s) \geq 0$, and (2) have a chance of being accepted by the responder, i.e., $G_R(s) \geq 0$. For example, in Figure 1(a), $P1$ ($P2$) has 79 (50) valid proposals given this limitation.

## 4.2 Analysis of Scenario

In the canonical Ultimatum game, the optimal strategy $s$ for the proposer $P$ against a rational responder maximizes its gain while providing a non-negative gain for the responder (i.e., the optimal

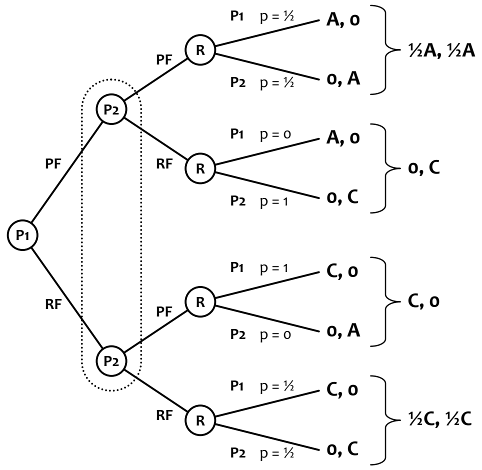



**Figure 2: Extensive-form representation of the three-player negotiation variant of CT with two proposer metastrategies. The payoff for the rational responder R is not shown.**

strategy is $\arg\max_s G_P(s)$). However, in the two-proposer setting we consider, proposers compete with each other, which means proposers have to take into account the gain of the responder. To facilitate analysis, we plot the gains $G_R(s)$ against $G_P(s)$ for each possible proposal $s$ in a *gain graph*. Gain graphs for the two example games are given in Figures 1(b) and 1(d). In the interaction between a proposer and the responder, the Pareto-dominant proposals are located on the convex hull, as indicated in the figures.

**Two proposer metastrategies.** We note the following proposals located on the convex hull.

1. **Proposer focus (PF)**. PF is the strategy in which the proposer first maximizes its own gain, and then finds the maximum gain for the responder.

$$PF_P = \arg\max_{s'} G_R(s'), s' \in \arg\max_{s} G_P(s), s \in \mathcal{S}.$$

2. **Responder focus (RF)**. RF is the strategy in which the proposer first maximizes the responder's gain, and then finds the maximum gain for itself.

$$RF_P = \arg\max_{s'} G_P(s'), s' \in \arg\max_{s} G_R(s), s \in \mathcal{S}.$$

We call these proposals *metastrategies*, as their definition does not depend on the actual CT setting. The proposals corresponding to the metastrategies PF and RF for the example CT games appear in Figures 1(b) and (d). In the example instance of Figure 1(b), the strategy PF for P1 corresponds to the chip exchange we mentioned before, in which P1 offers one red chip and one gray chip in exchange for a blue, a green, and three yellow chips, leading to a gain, if accepted, of 60 for P1 and 0 for R), while RF corresponds to giving a blue chip, a red chip, and a gray chip in exchange for two green chips (leading to a gain of 20 for P1 and 55 for R here).

**Interactions between metastrategies.** Suppose that proposers play only the metastrategies PF and RF. We show an extensive-form representation of the resulting CT scenario in Figure 2 (for proposer 2). We do not list the payoff for the responder from playing its rational strategy. In the figure, the two decision nodes of $P2$ are grouped into into one information set, because the players make their proposals simultaneously. Once $P2$ has chosen, the static and rational strategy of the responder (which is indicated in the figure) leads to certain expected gains.[4] Here, $A$ denotes the gain that a proposer receives when playing PF and being accepted; $C$ denotes the gain for RF being accepted. Clearly, this extensive-form game can be represented in a 2x2 matrix game which omits the responder's strategy. The gain matrix of the symmetrical game between the two proposers is given below.

| | PF | RF |
|---|---|---|
| **PF** | $\frac{1}{2}A, \frac{1}{2}A$ | $0, C$ |
| **RF** | $C, 0$ | $\frac{1}{2}C, \frac{1}{2}C$ |

Since the game between the two proposers is a 2x2 matrix game, it is straightforward to analyze. The game depends on the relationship between $A$ and $C$, as follows. For $A < 2C$, the game is a Prisoners' Dilemma, with one Nash Equilibrium at (RF, RF) and a Pareto-optimal outcome at (PF, PF). For $A \geq 2C$, we obtain a Stag Hunt game, with two Nash Equilibria; the RF-equilibrium (shorthand notation) is risk-dominant, while the PF-equilibrium is payoff-dominant. Both the Prisoners' Dilemma and the Stag Hunt game are well-known *social dilemmas* [13].

The strategic qualities of the original CT game are preserved in the 2x2 matrix game played between metastrategies. In the original CT game, the RF metastrategy corresponds to offering the best possible offer to the responder. RF is therefore also the risk-dominant proposal in the original game, because it guarantees a positive gain to the proposer. Even if both proposers play RF, the expected gain for each proposer will be positive. In contrast, the PF metastrategy is payoff-dominant but risky, because it provides a low (or even zero) gain to the responder. It will yield the most positive possible gain (payoff) for the proposer if the other proposer also plays PF, but will yield no gain at all otherwise. The proposers' dilemma in the original game (favoring themselves or the responder) is thus accurately reflected in the reduced 2x2 matrix game.

In the example CT instance shown in Figure 1(a) and (b), we find that the PF strategy yields a gain to the proposer of 60 and a gain of 0 to the responder if accepted, while the RF strategy yields a gain of 20 to the proposer and 55 to the responder. Hence, $A = 60$ and $C = 20$ here, and $A > 2C$; the game played between the two proposers is thus a Stag Hunt. In a similar manner, we can conclude that the CT game of Figure 1(c) and (d) is a Prisoners' Dilemma, because $A = 25$ and $C = 15$ yields $A < 2C$.

**Introducing a third metastrategy.** While we distinguish only two metastrategies thus far, a proposer's actual strategy $s$ may be mixed, yielding (in theory) infinitely many possible (mixed) strategies $s$ based on the two metastrategies.

We now demonstrate that a proposer can benefit from employing a metastrategy other than such a mixed strategy $s$. This is illustrated in Figure 3 (left and right). In the gain graph, all mixed strategies of PF and RF are located on the straight line connecting PF and RF. From the proposer's perspective, any mixed strategy $s$ is strictly dominated by a strategy $s^*$ for which $G_P(s^*) > G_P(s)$. This constraint is met by all points to the right of $s$ in the plot. Given that the responder behaves rationally, we say that $s^*$ strictly dominates $s$ iff $G_R(s^*) > G_R(s)$. All points above $s$ in the plot meet this constraint. Thus, strategies that lie in the white area of the graphs in Figure 3 strictly dominate the mixed strategy $s$.

[4] When calculating these expected gains, we assume that metastrategy pairs (e.g., $PF_{P1}$ and $PF_{P2}$) yield the same gain for the responder, so the responder is indifferent to the two metastrategies. However, a CT game may generally not be (fully) symmetrical, but as we will see, games are symmetrical in expectation. We will come back to this when discussing our experiments.

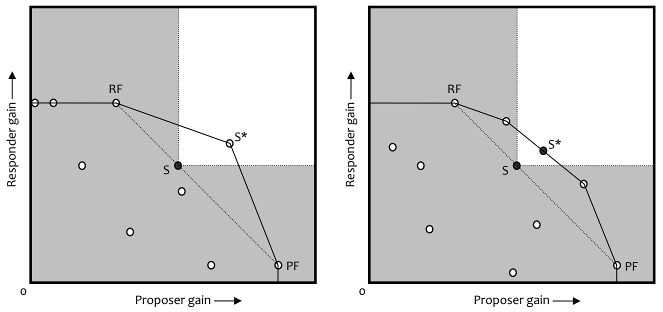



**Figure 3: Pure or mixed strategies on the convex hull may strictly dominate a mixed strategy of PF and RF. In the example on the left, we find a pure strategy $s^*$ on the convex hull that strictly dominates a mixed strategy $s$ of PF and RF. On the right, a mixed strategy $s^*$ strictly dominates $s$.**

In some cases, the convex hull may lie on the straight line between PF and RF, as for instance in the second example game (Figure 1(d)); then, there is no strategy that strictly dominates $s$. In other cases however, as in Figures 3 and 1(b), the convex hull may be located above the line PF-RF. In these cases, we can *always* find a strategy $s^*$ that strictly dominates $s$, except if $s$ is a pure strategy itself, i.e., if $s$ assigns a probability of 1 to a certain metastrategy. For instance, in Figure 3 (left), we find a pure strategy $s^*$ on the convex hull that dominates $s$. In Figure 3 (right), a mixed strategy $s^*$ on the convex hull dominates $s$. In Figure 1(b), both proposers have three pure strategies (and an infinite number of mixed strategies involving one or more of these pure strategies) that dominate a mixed strategy of PF and RF.

Thus, proposers indeed may benefit from employing additional metastrategies, since these additional metastrategies may dominate (mixed strategies of) the two metastrategies we already defined. We therefore introduce a third metastrategy, named QF (where Q is chosen simply because it is between P and R), which is to play the median proposal on the convex hull. Note that the median may be defined for any number of proposals on the convex hull; for an even number, we probabilistically select a proposal from the two median ones. Thus, proposals corresponding to the third metastrategy may again be found in any CT game.

Figures 1(b) and 1(d) show the mixed metastrategy QF for the two example CT instances. We see that the first instance has a pure QF metastrategy which dominates a mixed strategy of PF and RF. The gain graph shows that QF yields a proposer (responder) gain of $40$ ($35$) here. The second instance has no strategies on the convex hull that dominate a mixed strategy of PF and RF; still, QF may be defined by choosing probabilistically from PF and RF. The (expected) gain for QF is then the average of the gains for PF and RF, i.e., 20 for the proposers and $17.5$ for the responder.

**Interactions between three metastrategies.** With three metastrategies, the game between the two proposers becomes a 3x3 matrix game as follows.

| | **PF** | **QF** | **RF** |
|---|---|---|---|
| **PF** | $\frac{1}{2}A,\frac{1}{2}A$ | $0,B$ | $0,C$ |
| **QF** | $B,0$ | $\frac{1}{2}B,\frac{1}{2}B$ | $0,C$ |
| **RF** | $C,0$ | $C,0$ | $\frac{1}{2}C,\frac{1}{2}C$ |

Here, $A \geq B \geq C$. As with the two-strategy game, we can find different types of game depending on the relation between $A$, $B$, and $C$. It is easy to see that potential equilibria are located on the diagonal of the matrix. Moreover, as in the two-strategy game, (RF, RF) is an equilibrium. Depending on the values of $A$, $B$, and $C$, we may distinguish four different games. For all games in which $A < 2B < 4C$, (RF, RF) is the sole equilibrium. For $A \geq 2B \geq 4C$, all three diagonal strategies are equilibria. For $A < 2B$ and $B \geq 2C$, the equilibria are (RF, RF) and (QF, QF). For $A \geq 2B$ and $B < 2C$, we find equilibria at (RF, RF) and (PF, PF).

In the example of Figure 1(a), we find $B = 40$ ($A = 60$ and $C = 20$ still holds); thus, $A < 2B$ and $B = 2C$, meaning the 3x3 matrix game has two equilibria, i.e., the RF- and the QF-equilibrium. In Figure 1(c), we find $A = 25$, $B = 20$ and $C = 15$, so $A < 2B$ and $B < 2C$, yielding a single equilibrium at (RF, RF).

**Adding social factors to the responder model.** Thus far we have assumed the responder to be rational. Empirical evidence (in Ultimatum game settings) suggests that human responders are actually not fully rational [5]. One of the most well-known alternative models for Ultimatum-game responder behavior is *inequity aversion* [1]. The responder does not act directly on its gain $G_R(s)$, but instead on a utility function $U_R(s)$, which depends on its own gain, but also on how it compares to the gain of the proposer, $G_P(s)$. The original model distinguishes two components in the utility function, namely greed and compassion, both of which decrease the responder's utility in comparison to the actual gain. The greed component is generally far stronger (in humans); we do not consider the compassion component here. Translated to our settings, the responder's utility function may be then defined as follows.

$$U_R(s) = \begin{cases} G_R(s) & G_R(s) \geq G_P(s) \\ G_R(s) - \alpha_R\,(G_P(s) - G_R(s)) & \text{otherwise} \end{cases}$$

There is one parameter, $\alpha_R$, which determines how strongly the responder dislikes a proposal which gives a proposer more gain than the responder.

To illustrate the effect of inequity aversion, we apply it to the gain graph of proposer 1 in the first example game (Figure 1(a) and (b)). The effect for $\alpha_R = 0.5$ is visualized in Figure 4. For proposals that give the proposer more gain than the responder (i.e., below the diagonal), the utility (perceived gain) for the responder is lower than the actual gain. As a result, some proposals that may be accepted by a rational responder are not accepted by an inequity-averse responder. As is visible from the figure, the convex hull changes, as does the location of the PF metastrategy. If the re-

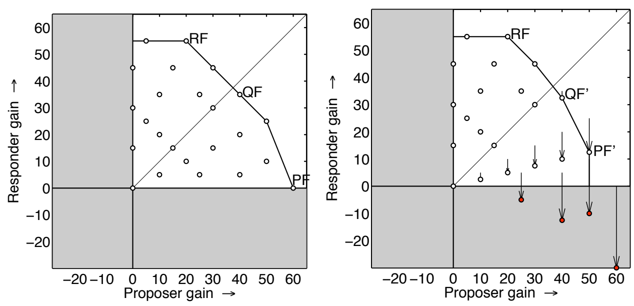



**Figure 4: Effect of inequity aversion on the utility (perceived gain) of the responder: the convex hull and the metastrategies change. PF will no longer be accepted by the responder, which means proposers need to offer PF' instead.**

sponder is inequity-averse and the proposers do not take this into account, they may coordinate (without communication) to offer the PF proposal, expecting that one of them will be accepted, while in reality, the responder will reject both proposals.

Instead of offering PF, proposers should offer PF', yielding a gain of 50 instead of 60 in the example. They are able to find PF' if they are aware of the inequity aversion present in the responder (i.e., the value of $\alpha_R$), which implies they can calculate the modified gain graph. We deal with both unaware as well as aware proposers in our empirical evaluation.

## 4.3 Generalizing the analysis

In this section we show that our analysis may be generalized to more than three metastrategies and other variants of the CT game.

**An arbitrary number of metastrategies.** We may introduce additional metastrategies in a similar manner as the third one; e.g., we could have five metastrategies, corresponding to the minimal, first quartile, median, third quartile, and maximal proposal on the convex hull. Generally speaking, for $n$ metastrategies, we obtain an $n$x$n$ matrix game. The $n$ diagonal strategies may each be equilibria or not, except the ever-present RF-equilibrium. Depending on the gains for each metastrategy pair, we may thus distinguish $2^{n-1}$ different possibilities for equilibria in the $n$x$n$ game.

**Generalizing to other CT variants.** The analysis above is specifically performed on the three-player negotiation variant of CT. However, results generalize to other games within the CT framework, since chip exchanges are a vital part of the framework [7]. We will provide a few pointers here.

A common different variant is a two-player negotiation game (i.e., one proposer and one responder), potentially with multiple phases and/or alternating roles [8]. The one-shot two-player game also allows us to construct the gain graph and identify the metastrategies. Since the dilemma (and the associated competition) between the proposers is missing, the single proposer may get away with offering PF every time.[5]

More generally, the concept of the gain graph naturally extends to negotiation games with any number of proposers and responders. For instance, a game with three proposers and one responder leads to a three-player social dilemma between the proposers, which may be modeled for instance as a Public Goods game, and interactions similar to the Ultimatum game between each proposer and the responder. In a game with multiple responders we can still construct gain graphs between pairs of proposer(s) and responders, with an Ultimatum Game with responder competition [6] taking place between these pairs. Analytical and experimental studies in the Ultimatum Game clearly indicate that players benefit from an increased number of opponents in the opposite role (e.g., responders fare well with more proposers) [6, 9].

# 5. EMPIRICAL EVALUATION

In this section, we outline how balanced and challenging instances of the CT game may be generated. We then discuss how players that perform actions according to the metastrategies may be heuristically implemented. Finally, we generate a large number of games, have our heuristic players play them, and evaluate the empirical payoff tables, which can be compared with analytical results.

## 5.1 Criteria for balanced games

This section outlines three criteria that are essential to create well-balanced and challenging instances of the the CT game.

[5]Repeated games fall outside the scope of this paper.

**Baseline scores.** The initial board state (positions and chip sets) should yield baseline scores that are comparable for all three players. We generate games that limit the difference in baseline scores to be less than $\epsilon$.

$$\max\left\{G_{P1}\left(\emptyset\right), G_{P2}\left(\emptyset\right), G_{R}\left(\emptyset\right)\right\} \\ -\min\left\{G_{P1}\left(\emptyset\right), G_{P2}\left(\emptyset\right), G_{R}\left(\emptyset\right)\right\} < \epsilon$$

**Negotiation requirement.** No player should be able to reach the goal location on its own without engaging in a chip trade. We define $isSolution(P, C) = true$ iff player $P$ can reach the goal given a chip set $C$. The initial chip set of a player $P$ is given by $chips(P)$.

$$\neg isSolution(P1, chips(P1)) \wedge \\ \neg isSolution(P2, chips(P2)) \wedge \\ \neg isSolution(R, chips(R))$$

**Mutual dependence.** Due to the negotiation requirement, both proposers depend on a *subset* of the responder's chip set. In turn, the responder relies on a subset of either proposer 1 or proposer 2. A one-sided proposal (i.e. asking for all chips or dispensing of all chips) may not lead to a chip set allowing both the proposer and the responder to reach the goal.

$$\exists C_{P1}, C_R \in chips(P1) \cap chips(R)\ s.t. \\ C_{P1} \cap C_R = \emptyset \wedge isSolution(P1, C_{P1}) \wedge isSolution(R, C_R) \\ \exists C_{P2}, C_R \in chips(P2) \cap chips(R)\ s.t. \\ C_{P2} \cap C_R = \emptyset \wedge isSolution(P2, C_{P1}) \wedge isSolution(R, C_R)$$

We implement these three criteria by generating many pseudo-random games and checking them against the criteria, keeping only those games that match.

## 5.2 Experimental setup

For the empirical evaluation of the proposed metastrategies we generate a database of $10K$ games that adhere to the criteria listed above (we chose $\epsilon = 20$). Below, we discuss how the metastrategies are implemented in heuristic players and how empirical payoffs are computed from games played between these players.

**Heuristic players.** We implemented three heuristic players, each following one of the three metastrategies, i.e. PF, QF and RF. All three heuristic players start by enumerating all possible chip exchange proposals. Proposals that yield negative gains for either the proposer or the responder are neglected. Heuristic players following metastrategies PF and RF are a straightforward implementation of the definitions given earlier. Metastrategy QF requires to compute the PF and RF strategy points in the gain graph, as well as the convex hull connecting both.[6] The median proposal on the convex hull is then selected. For an even number, the heuristic player probabilistically selects a proposal from the two median ones.

**Computing empirical payoffs.** A single entry of the empirical payoff matrix is computed as follows. The row determines the metastrategy played by P1, while the column determines the metastrategy for P2. For each game in the database, chip exchanges proposed by the players are evaluated by the responder and if a proposal is excepted, chips are exchanged and scores evaluated. The resulting payoff is the difference between final and baseline scores (i.e. gain) averaged over all $10K$ games. This process leads to a full empirical payoff table for the game as a whole.

## 5.3 Results

In this section, empirical payoff tables obtained by the metastrategies are presented and compared to the predicted payoff tables.

[6]While any convex hull algorithm is adequate, our implementation is based on the time-efficient Graham scam algorithm.

**Two-strategy game.** With two metastrategies PF and RF, we obtain an empirical payoff table as follows.

| | PF | RF |
|---|---|---|
| PF | 21.0, 20.6 | 2.9, 11.7 |
| RF | 11.8, 2.6 | 6.5, 6.2 |

The empirical payoffs yield a Stag Hunt game, with $A \approx 42$ and $C \approx 12$. When we compare the empirical payoff table to the analytical one, we notice two things.

First, the game is nearly, but not completely symmetrical. This may be explained by the relatively small size of the board, which leads to relatively large differences (i.e. possible disbalances) between the two proposers. On the small board we use, symmetry arises from repeated play. The game is guaranteed to be symmetrical in expectation, since proposers' positions are randomized.

Second, there are (small) positive values where we expected values of 0. In some instances, a certain proposer's PF proposal is preferred by the responder over the other proposer's RF proposal. Once again, this issue may be dealt with by using larger boards, which would reduce the probability that PF 'wins' from RF. However, larger boards are (even) more difficult for human players.

**Three-strategy game.** The empirical payoff table for the three-strategy game is given below. The values in the corners of the table are identical to those in the two-strategy game.

| | PF | QF | RF |
|---|---|---|---|
| PF | 21.0, 20.6 | 5.7, 24.3 | 2.9, 11.7 |
| QF | 24.5, 5.6 | 14.8, 14.2 | 6.0, 9.8 |
| RF | 11.8, 2.6 | 10.0, 5.7 | 6.5, 6.2 |

We see that $B \approx 25$. It is interesting to consider the interactions between the 'neighboring' metastrategies. Looking at the interaction between PF and QF, we find a Prisoners' Dilemma. The QF metastrategy is very strong against PF, giving proposers a strong incentive to defect. Between QF and RF, we find a Stag Hunt. The payoff table thus yields a game with two equilibria, namely the QF- and the RF-equilibrium.

**Inequity aversion (unaware proposers).** In our next experiment, we determine the effect of introducing social considerations (inequity aversion) in the responder's decision-making, without the proposers being aware of this. We provide the empirical payoff matrices for two reasonable values of $\alpha_R$, restricting ourselves to the two-strategy game.

$\alpha_R = 0.5$

| | PF | RF |
|---|---|---|
| PF | 9.6, 9.9 | 1.2, 12.0 |
| RF | 12.0, 1.2 | 6.4, 6.3 |

$\alpha_R = 1.0$

| | PF | RF |
|---|---|---|
| PF | 5.3, 5.3 | 0.7, 11.9 |
| RF | 12.0, 0.7 | 6.5, 6.1 |

The second equilibrium (PF, PF) disappears, because proposers expect their PF proposal to be accepted more than it actually is. The game thus turns into a Prisoners' Dilemma with one Pareto-dominated equilibrium at (RF, RF). The higher the value of $\alpha_R$, the stronger this effect.

**Inequity aversion (aware proposers).** We also investigate what happens if the proposers *do* know that the responder is inequity-averse. The payoff matrices for the same values of $\alpha_R$ are:

$\alpha_R = 0.5$

| | PF | RF |
|---|---|---|
| PF | 17.0, 17.2 | 3.4, 11.3 |
| RF | 11.3, 3.2 | 6.5, 6.1 |

$\alpha_R = 1.0$

| | PF | RF |
|---|---|---|
| PF | 15.3, 15.1 | 4.6, 10.4 |
| RF | 10.5, 4.4 | 6.4, 6.1 |

The second equilibrium is back again; proposers appropriately adjust their PF proposals to the expectations of the responder. The payoff for PF is (sensibly) lower against itself than in the original game with a rational responder. PF does increasingly well against RF, simply because PF is (slightly) more similar to RF when proposers take into account the responder's expectation.

# 6. DISCUSSION

The previous sections have provided conditions under which CT games can generally be decomposed into a set of multiple normal-form games that are characterized by social dilemmas (i.e., Prisoners' Dilemma, Stag Hunt and Ultimatum games), as visualized in Figure 5, using a number of metastrategies defined on the possible chip exchange proposals in the game.

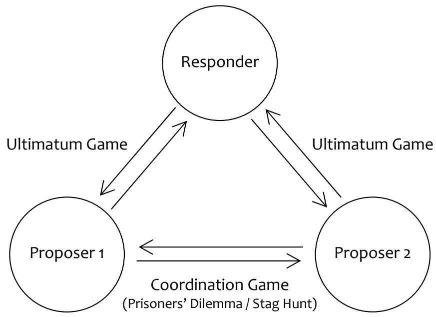



**Figure 5: Decomposing the three-player negotiation variant of the Colored Trails Game.**

We show that the metastrategy that favors the responder (RF) is always an equilibrium in the reduced normal form game, regardless of the number of metastrategies. This is because the responder has an advantage in the CT setting we consider, in that no player receives a gain if it does not accept an offer. An empirical analysis of a large set of game instances illustrates that, in expectation, two metastrategies yield two equilibria, reducing it to a Stag Hunt game. We may also find game instances that are Prisoners' Dilemmas, i.e., with only the equilibrium that favors the responder. Using three metastrategies in the same set of game instances also yields two equilibria in expectation, namely those two metastrategies that are most favorable for the responder (QF and RF).

Adding social factors to the responder allows this player to enforce a higher payoff—essentially, the proposers are driven to defection in the Stag Hunt or Prisoners' Dilemma game they play, because the responder is better at exploiting the proposer competition in the Ultimatum game component. This increased power for the responder may be countered by introducing multiple responders, as in an Ultimatum game with responder competition [6].

As noted, our analysis of a single game instance assumes that responders are indifferent between the gains from the two proposals resulting from any pair of metastrategies, i.e., for the metastrategies $PF_{P1}$ and $PF_{P2}$, we have that $G_R(PF_{P1}) = G_R(PF_{P2})$ (and similar for QF and RF). If this condition does not hold, the responder will favor one of the metastrategy proposals over the other, which means the actual game instance does not reduce to a Stag Hunt or Prisoners' Dilemma. We observed that approximately $25\%$ of the $10K$ games we generated (and that met our three criteria) were actually games in which the responder is indifferent between metastrategy pairs. Of these $25\%$, approximately one-fifth are Prisoners' Dilemmas, and four-fifth are Stag Hunts. The remaining games (i.e., $75\%$) were not symmetrical, meaning that one proposer has a strategic advantage over the other proposer. Even though our symmetry assumption thus does not hold for a majority of generated game instances, our empirical results show that, even for games in which the assumption does not hold, the *expected* gains to proposers from playing metastrategies do in fact correspond to Stag Hunt and Prisoners' Dilemma games.

In case a certain experiment requires all games to be Stag Hunts or Prisoners' Dilemmas (i.e., not only in expectation), the assumption of responder indifference can be enforced during game generation. We note that, for the case in which responders are assumed to be rational, we do not need to make assumptions about the gains to proposers from pairs of metastrategies (a rational responder does not consider those gains), while for inequity-averse responders, the gains to proposers for every metastrategy pair must also be equal, i.e., $G_{P1}(PF_{P1}) = G_{P2}(PF_{P2})$ (and similar for QF and RF).

# 7. CONCLUSION AND FUTURE WORK

In this paper, we show how to reduce a large multi-agent task-setting, i.e., a game in the often-used Colored Trails (CT) framework, to a set of smaller, bilateral normal-form games, while still preserving most of the strategic charachteristics of the original setting. We provide a set of criteria and assumptions that need to hold in order for the normal-form games to correspond to canonical social dilemma games from the literature. We show how to define representative heuristic metastrategies in the CT setting that make minimal assumptions about the other players. The games taking place between metastrategies are shown to correspond to Prisoners' Dilemma, Stag Hunt and Ultimatum games. We demonstrate that the metastrategies' analytical payoff tables, which we generated on the basis of assumptions that are not always met, nonetheless correspond to empirical payoff tables by sampling from thousands of CT game instances and showing that the outcome to players from using the metastrategies corresponds *on average* to the outcomes from the social dilemma games.

Although our analysis and examples are based on a particular CT scenario (a three-player take-it-or-leave-it game), they demonstrate the possibility of using metastrategies to reduce other CT games (e.g., games with a different number of players in each role, or even games that are further removed from the game under consideration here), and multi-agent interactions in general, to (social dilemma) normal-form games. More precisely, the techniques we present apply to general multi-agent interactions in which optimal actions can be computed, given that other players are using specified strategies.

We are currently extending our approach in two ways. First, we are developing metastrategies that consider other social factors that affect people's behavior in task settings, such as altruism and generosity, also from the perspective of the proposers. Second, we are considering more complex CT scenarios that include repeated negotiation, in which metastrategies will need to account for players' trust and reciprocity relationships.

# 8. REFERENCES


[1] E. Fehr and K. Schmidt. A Theory of Fairness, Competition and Cooperation. *Quart. J. of Economics*, 114:817–868, 1999.

[2] S. G. Ficici and A. Pfeffer. Modeling how Humans Reason about Others with Partial Information. In *Proceedings of the Seventh International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2008.

[3] Y. Gal, B. Grosz, S. Kraus, A. Pfeffer, and S. Shieber. Agent decision-making in open-mixed networks. *Artificial Intelligence*, 2010.

[4] Y. Gal and A. Pfeffer. Modeling reciprocity in human bilateral negotiation. In *National Conference on Artificial Intelligence (AAAI)*, 2007.

[5] H. Gintis. *Game theory evolving*. Princeton University Press Princeton:, 2000.

[6] B. Grosskopf. Reinforcement and directional learning in the ultimatum game with responder competition. *Experimental Economics*, 6(2):141–158, 2003.

[7] B. J. Grosz, S. Kraus, S. Talman, B. Stossel, and M. Havlin. The Influence of Social Dependencies on Decision-Making. Initial investigations with a new game. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2004.

[8] G. Haim, Y. Gal, S. Kraus, and Y. Blumberg. Learning Human Negotiation Behavior Across Cultures. In *HuCom10 - Second International Working Conference on Human Factors and Computational Models in Negotiation*, 2010.

[9] C. Hauert. Ultimatum game with proposer competition. GameLab experiments, consulted 10 August 2010.

[10] D. Hennes, K. Tuyls, M. Neerincx, and M. Rauterberg. Micro-scale social network analysis for ultra-long space flights. In *The IJCAI-09 Workshop on Artificial Intelligence in Space*, 2009.

[11] M. Kaisers, K. Tuyls, F. Thuijsman, and S. Parsons. Auction analysis by normal form game approximation. In *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT'08. IEEE/WIC/ACM International Conference on*, volume 2, pages 447–450. IEEE, 2009.

[12] E. Kamar, Y. Gal, and B. Grosz. Modeling user perception of interaction opportunities for effective teamwork. In *Symposium on Social Intelligence and Networking, IEEE Conference on Social Computing*, 2009.

[13] D. Messick and M. Brewer. Solving social dilemmas: A review. *Review of personality and social psychology*, 4:11–44, 1983.

[14] S. Phelps, K. Cai, P. McBurney, J. Niu, S. Parsons, and E. Sklar. Auctions, evolution, and multi-agent learning. In *Proceedings of the Symposium on Adaptive Learning Agents and Multi-Agent Systems*, 2007.

[15] S. Phelps, M. Marcinkiewicz, and S. Parsons. A novel method for automatic strategy acquisition in n-player non-zero-sum games. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 705–712. Springer, 2001.

[16] M. Ponsen, K. Tuyls, M. Kaisers, and J. Ramon. An evolutionary game-theoretic analysis of poker strategies. *Entertainment Computing*, 1:39–45, 2009.

[17] P. Vytelingum, D. Cliff, and N. Jennings. Evolutionary stability of behavioural types in the continuous double auction. In *Agent-Mediated Electronic Commerce. Automated Negotiation and Strategy Design for Electronic Markets*, pages 103–117. Springer, 2007.

[18] W. Walsh, R. Das, G. Tesauro, and J. Kephart. Analyzing complex strategic interactions in multi-agent systems. In *AAAI-02 Workshop on Game-Theoretic and Decision-Theoretic Agents*, pages 109–118, 2002.

[19] W. Walsh, D. Parkes, and R. Das. Choosing samples to compute heuristic-strategy Nash equilibrium. *Agent-Mediated Electronic Commerce*, V:109–123, 2004.