

How to Be Both Rich and Happy: Combining Quantitative and Qualitative Strategic Reasoning

(Extended abstract of preliminary report¹)

Nils Bulling and Valentin Goranko

Abstract

We propose a logical framework combining a game-theoretic study of abilities of agents to achieve *quantitative objectives* in multi-player games by optimizing payoffs or preferences on outcomes with a logical analysis of the abilities of players for achieving *qualitative objectives* of players, i.e., reaching or maintaining game states with desired properties. We enrich concurrent game models with payoffs for the normal form games associated with the states of the model and propose a quantitative extension of the logic ATL* enabling the combination of quantitative and qualitative reasoning.

1 Introduction

Amongst the different traditions in studying strategic abilities of agents in multi-player games we can distinguish two rather characteristic ones:

Quantitative reasoning about strategic abilities to achieve *quantitative objectives*, such as optimizing payoffs (e.g., maximizing rewards or minimizing cost). This tradition comes from game theory and economics and usually studies one-shot normal form games, their (finitely or infinitely) repeated versions, and extensive form games.

Qualitative reasoning deals with strategic abilities of players for achieving *qualitative objectives*: reaching or maintaining outcome states with desired properties, e.g., winning states, or safe states, etc. It comes from logic and computer science.

Put as a slogan, the former tradition is concerned with *how a player can become maximally rich, or how to pay as little cost as possible*, while the latter tradition – with *how a player can achieve a state of ‘happiness’, e.g. winning, or to avoid reaching a state of ‘unhappiness’ (losing) in the game*.

The most essential technical difference between qualitative and quantitative players’ objectives is that the former typically refer to (a temporal pattern over) Boolean properties of game states on a given play and can be monitored locally whereas the latter are determined by the entire history of the play (accumulated payoffs) or even the whole play (its value, being a limit of average payoffs, or of discounted accumulated payoffs). It is therefore generally computationally more demanding and costly to design strategies satisfying quantitative objectives or to verify their satisfaction under a given strategy of a player or coalition.

Remark. More generally, decision theory and game theory study rational behaviour of players aiming at optimizing their *preferences* between outcomes. Preferences can be regarded as both qualitative and quantitative objectives and, if equipped with a suitable mechanism for preference aggregation over a series of outcomes accumulated in the course of the play, then our work presented here, based on quantitative payoffs in naturally ordered domains, can be suitably generalized to that setting.

¹This is a slightly revised version of [10].

The two traditions above have followed rather separate developments, with generally quite different agendas, methods and results, including, inter alia:

- on the purely qualitative side, *logics of games and multiagent systems*, such as the Coalition logic CL [29], the Alternating time temporal logic ATL [4], and variations of it, see e.g. [23], [32], [26], etc., formalizing and studying qualitative reasoning in concurrent game models;
- some *single-agent and multi-agent bounded resource logics* [9, 3, 27] extending or modifying concurrent game models with some quantitative aspects by considering cost of agents' actions and reasoning about what players with bounded resources can achieve.
- extensions of qualitative reasoning (e.g., reachability and Büchi objectives) in multi-player concurrent games with 'semi-quantitative' aspects by considering a preference preorder on the set of qualitative objectives, see e.g., [6], [7], thereby adding payoff-maximizing objectives and thus creating a setting where traditional game-theoretic issues such as game value problems and Nash equilibria become relevant.
- deterministic or stochastic *infinite games on graphs*, with qualitative objectives: typically, reachability, and more generally – parity objectives or objectives specified as ω -regular languages over the set of plays, see e.g. [19], [17], [18], [20], [11], [15].
- on the purely quantitative side, first to mention *repeated games*, extensively studied in game theory (see e.g., [28]), which can be naturally treated as simple, one-state concurrent game models with accumulating payoffs paid to each player after every round and no qualitative objectives;
- from a more computational perspective, various stochastic games with quantitative objectives, such as energy games and discounted and mean-payoff games have been studied, see e.g. [30].
- the conceptually different but technically quite relevant study of *counter automata, Petri nets, vector addition systems*, etc. – essentially a study of the purely quantitative single-agent case of concurrent game models (see e.g. [22]), where only accumulated payoffs but no qualitative objectives are taken into account and a typical problem is to decide reachability of payoff configurations satisfying formally specified arithmetic constraints from a given initial payoff configuration.

Besides all these, there have been several recent threads of research proposing the combination of qualitative and quantitative game analysis, and thus coming closer in spirit to the present work:

- Zelonka [33] considers infinite 2-player turn-based games where every move is associated with a 'reward' (e.g., priority in parity games) after every move and eventually player 2 pays player 1 a payoff determined by the resulting infinite sequence of rewards. The objectives of the players are to optimize payoff (for player 1), respectively cost (for player 2).
- *mean-payoff parity games* [16], [14], [13], [5] study 2-player games combining parity objectives with quantitative requirements for the mean payoff;
- *parity energy games* [12] study 2-player turn-based games combining parity objectives with requirements for maintenance of non-negative energy.

Other relevant references discussing the interaction between qualitative and quantitative reasoning in multi-player games, include [31], [24],

Among the most studied models in the qualitative reasoning tradition are *concurrent game models* [4, 29]. On the one hand they are richer than normal form games, as they incorporate a whole family of such games, each associated with a state of a transition system; but on the other hand, they are somewhat poorer because the outcomes of each of these normal form games, associated with a given state, are simply the successor states with their associated games, etc. whereas no payoffs, or even preferences on outcomes, are assigned. Thus, plays in concurrent game models involve a sequence of possibly different one-shot normal form games played in succession, and all that is taken into account in the purely logical framework are the properties – expressed by formulae of a logical language – of the states occurring in the play. Concurrent game models can also be viewed as generalization of (possibly infinite) extensive form games where cycles and simultaneous moves of different players are allowed, but no payoffs are assigned.

This project purports to combine the two agendas in a common logical framework, by enriching concurrent game models with payoffs for the one-shot normal form games associated with the states, and thus enabling the combination of quantitative game-theoretic reasoning with the qualitative logical reasoning. Again, put as a slogan, our framework allows reasoning about whether/how a player can reach or maintain a state of ‘happiness’ while becoming, or remaining, as rich as (rationally) possible, or paying the least possible price on the way. Here we introduce and discuss a general framework of models and logics for combined quantitative and qualitative reasoning that would naturally cover each of the topics listed above, and to initiate a long term study on it.

2 Preliminaries

A **concurrent game model** [4] (CGM) $\mathcal{S} = (\text{Ag}, \text{St}, \{\text{Act}_a\}_{a \in \text{Ag}}, \{\text{act}_a\}_{a \in \text{Ag}}, \text{out}, \text{Prop}, \text{L})$ comprises:

- a non-empty, fixed set of players $\text{Ag} = \{1, \dots, k\}$ and a set of actions $\text{Act}_a \neq \emptyset$ for each $a \in \text{Ag}$.
For any $A \subseteq \text{Ag}$ we will denote $\text{Act}_A := \prod_{a \in A} \text{Act}_a$ and will use $\vec{\alpha}_A$ to denote a tuple from Act_A . In particular, Act_{Ag} is the set of all possible action profiles in \mathcal{S} .
- a non-empty set of game states St .
- for each $a \in \text{Ag}$ a map $\text{act}_a : \text{St} \rightarrow \mathcal{P}(\text{Act}_a)$ setting for each state s the actions available to a at s .
- a transition function $\text{out} : \text{St} \times \text{Act}_{\text{Ag}} \rightarrow \text{St}$ that assigns the (deterministic) *successor (outcome) state* $\text{out}(q, \vec{\alpha}_{\text{Ag}})$ to every state q and action profile $\vec{\alpha}_{\text{Ag}} = \langle \alpha_1, \dots, \alpha_k \rangle$ such that $\alpha_a \in \text{act}_a(q)$ for every $a \in \text{Ag}$ (i.e., every α_a that can be executed by player a in state q).
- a set of atomic propositions Prop and a labelling function $\text{L} : \text{St} \rightarrow \mathcal{P}(\text{Prop})$.

Thus, all players in a CGM execute their actions synchronously and the combination of these actions, together with the current state, determines the transition to a (unique) successor state in the CGM.

The **logic of strategic abilities** ATL^* (*Alternating-Time Temporal Logic*), introduced and studied in [4], is a logical system, suitable for specifying and verifying qualitative objectives of players and coalitions in concurrent game models. The main syntactic construct of ATL^* is a formula of type $\langle\langle C \rangle\rangle \gamma$, intuitively meaning: “*The coalition C has a collective strategy to guarantee the satisfaction of the objective γ on every play enabled by that strategy.*” Formally, ATL^* is a multi-agent extension of the branching time logic CTL^* , i.e., multimodal logic extending the linear-time temporal logic LTL —comprising the temporal operators **X** (“at the next state”), **G** (“always from now on”) and **U** (“until”)—with

strategic path quantifiers $\langle\langle C \rangle\rangle$ indexed with coalitions C of players. There are two types of formulae of ATL^* , *state formulae*, which constitute the logic and that are evaluated at game states, and *path formulae*, that are evaluated on game plays. These are defined by mutual recursion with the following grammars, where $C \subseteq \text{Ag}$, $p \in \text{Prop}$: state formulae are defined by $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle C \rangle\rangle\gamma$, and path formulae by $\gamma ::= \varphi \mid \neg\gamma \mid \gamma \wedge \gamma \mid \mathbf{X}\gamma \mid \mathbf{G}\gamma \mid \gamma \mathbf{U}\gamma$.

The logic ATL^* is very expressive and that comes at a high computational price: satisfiability and model checking are **2ExpTime**-complete. A computationally better behaved fragment is the logic ATL , which is the multi-agent analogue of CTL, only involving state formulae defined by the following grammar, for $C \subseteq \text{Ag}$, $p \in \text{Prop}$: $\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle C \rangle\rangle\mathbf{X}\varphi \mid \langle\langle C \rangle\rangle\mathbf{G}\varphi \mid \langle\langle C \rangle\rangle(\varphi \mathbf{U}\varphi)$. For this logic satisfiability and model checking are **ExpTime**-complete and **P**-complete, respectively. We will, however, build our extended logical formalism on the richer ATL^* because we will essentially need the path-based semantics for it.

Arithmetic Constraints. We define a simple language of arithmetic constraints to express conditions about the accumulated payoffs of players on a given play. For this purpose, we use a set $V_{\text{Ag}} = \{v_a \mid a \in \text{Ag}\}$ of special variables to refer to the accumulated payoffs of the players at a given state and denote by V_A the restriction of V_{Ag} to any group $A \subseteq \text{Ag}$. The payoffs can be integers, rationals², or any reals. We denote the domain of possible values of the payoffs, assumed to be a subset of the reals \mathbb{R} , by \mathbb{D} and use a set of constants symbols X , with $0 \in X$, for names of special real values (see further) to which we want to refer in the logical language.

For fixed sets X and $A \subseteq \text{Ag}$ we build the set $T(X, A)$ of *terms over X and A* from $X \cup V_A$ by applying addition, e.g. $v_a + v_b$. An evaluation of a term $t \in T(X, A)$ is a mapping $\eta : X \cup V_A \rightarrow \mathbb{D}$. We write $\eta \models t$ to denote that t is *satisfied* under the evaluation η . Moreover, if some order of the elements $X \cup V_A$ is clear from context, we also represent an evaluation as a tuple from $\mathbb{D}^{|X|+|V_A|}$ and often assume that elements from X have their canonic interpretation. The set $\text{AC}(X, A)$ of *arithmetic constraints* over X and A consists of all expressions of the form $t_1 * t_2$ where $*$ $\in \{<, \leq, =, \geq, >\}$ and $t_1, t_2 \in T(X, A)$. We use $\text{ACF}(X, A)$ to refer to the set of Boolean formulae over $\text{AC}(X, A)$; e.g. $(t_1 < t_2) \wedge (t_2 \geq t_3) \in \text{ACF}(X, A)$ for $t_1, t_2, t_3 \in T(X, A)$. We note that the language $\text{ACF}(X, A)$ is strictly weaker than Presburger arithmetic, as it involves neither quantifiers nor congruence relations.

We also consider the set $\text{APC}(X, A)$ of *arithmetic path constraints* being expressions of the type $w_a * c$ where $a \in \text{Ag}$, $*$ $\in \{<, \leq, =, \geq, >\}$ and $c \in X$. The meaning of w_a is to represent the value of the current play for the player a . That value can be defined differently, typically by computing the accumulated payoff over the entire play, by using a future discounting factor, or by taking the limit – if it exists – of the mean (average) accumulated payoff (cf. [28]). We note that the discounted, accumulated, mean or limit payoffs may take real values beyond the original domain of payoffs \mathbb{D} ; so, we consider the domain for X to be a suitable closure of \mathbb{D} .

3 Concurrent Game Models with Payoffs and Guards

Definition 1 (Guards) *Let $a \in \text{Ag}$. An (individual) a -guard is an arithmetic constraint formula α over $X \cup V_{\{a\}}$.*

We now extend concurrent game models with utility values for every action profile applied at every state and with guards that determine which actions are available to a player at a given configuration, consisting of a state and a utility vector, in terms of arithmetic constraints on the utility of that player.

²Note that models with rational payoffs behave essentially like models with integer payoffs, after once-off initial re-scaling.

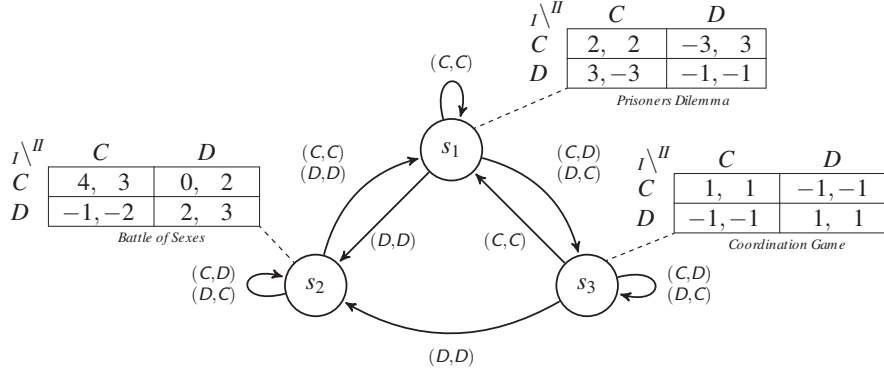


Figure 1: A simple CGM^{GP} .

Definition 2 (Guarded CGM with payoffs) A guarded CGM with payoffs (CGM^{GP}) is a tuple

$$\mathfrak{M} = (\mathcal{S}, \text{payoff}, \{g_a\}_{a \in \text{Ag}})$$

where $\mathcal{S} = (\text{Ag}, \text{St}, \{\text{Act}_a\}_{a \in \text{Ag}}, \{\text{act}_a\}_{a \in \text{Ag}}, \text{out}, \text{Prop}, L)$ is a CGM and:

- $\text{payoff} : \text{Ag} \times \text{St} \times \text{Act}_{\text{Ag}} \rightarrow \mathbb{D}$ is a payoff function assigning for every state s and action profile applied at s a payoff to every agent. We will write $\text{payoff}_a(s, \vec{\alpha})$ for $\text{payoff}(a, s, \vec{\alpha})$.
- $g_a : \text{St} \times \text{Act}_a \rightarrow \text{ACF}(X, a)$, for each $a \in \text{Ag}$, is a guard function that assigns for each state $s \in \text{St}$ and action $\alpha \in \text{Act}_a$ an arithmetic constraint formula $g_a(s, \alpha)$ that determines whether α is enabled for a at the state s given the value of a 's current accumulated payoff in the play.

A guard $g_a(s, \alpha)$ is called state-based if $g_a(s, \alpha) \in \text{ACF}(X)$.

Every guard g_a must satisfy consistency conditions that enable at least one action for a at s . Formally, for each $s \in \text{St}$, the formula $\bigvee_{\alpha \in \text{Act}_a} g_a(s, \alpha)$ must be valid.

An **initialized CGM^{GP} ($ICGM^{GP}$)** is a pair $(\mathfrak{M}, \text{init})$ where \mathfrak{M} is a CGM^{GP} and $\text{init} = (s_0, u_1^0, \dots, u_k^0)$ is an initial configuration, where $s_0 \in \text{St}$ is an initial state and $\vec{u}_0^0 = (u_1^0, \dots, u_k^0)$ is the vector of initial utilities of all players.

The guard g_a refines the function act_a from the definition of a CGM, which can be regarded as a guard function assigning to every state and action a constant arithmetic constraint **true** or **false**. In our definition the guards assigned by g_a only depend on the current state and the current accumulated payoff of a . The idea is that when the payoffs are interpreted as costs, penalties or, more generally, consumption of resources the possible actions of a player would depend on her current availability of utility/resources.

Example 1 Consider the CGM^{GP} shown in Figure 1 with 2 players, I and II, and 3 states, where in every state each player has 2 possible actions, C (cooperate) and D (defect). The transition function is depicted in the figure. The normal form games associated with the states are respectively versions of the Prisoners Dilemma at state s_1 , Battle of the Sexes at state s_2 and Coordination Game at state s_3 .

The guards for both players are defined at each state so that the player can apply any action if she has a positive current accumulated payoff, may only apply action C if she

has accumulated payoff 0; and must play an action maximizing her minimum payoff in the current game otherwise. The initial payoffs of both players are 0.

Configurations, plays, and histories. Let \mathfrak{M} be a CGM^{GP} defined as above. A *configuration* (in \mathfrak{M}) is a pair (s, \vec{u}) consisting of a state s and a vector $\vec{u} = (u_1, \dots, u_k)$ of currently accumulated payoffs, one for each agent, at that state. Hereafter we refer to *accumulated payoffs* as *utility*, at a given state. We define the set of possible configurations as $\text{Con}(\mathfrak{M}) = \text{St} \times \mathbb{D}^{|\text{Ag}|}$. The (*partial*) *configuration transition function* is defined as

$$\widehat{\text{out}} : \text{Con}(\mathfrak{M}) \times \text{Act}_{\text{Ag}} \rightarrow \text{Con}(\mathfrak{M}),$$

such that $\widehat{\text{out}}((s, \vec{u}), \vec{\alpha}) = (s', \vec{u}')$ iff:

- (i) $\text{out}(s, \vec{\alpha}) = s'$ (s' is a successor of s if $\vec{\alpha}$ is executed).
- (ii) assigning the value u_a to v_a satisfies the guard $g_a(s, \alpha_a)$ for each $a \in \text{Ag}$, i.e. $u_a \models g_a(s, \alpha_a)$ (each agent's move α_a is enabled at s by the respective guard g_a applied to the current accumulated utility value u_a).
- (iii) $u'_a = u_a + \text{payoff}_a(s, \vec{\alpha})$, for each $a \in \text{Ag}$.

An initialized CGM^{GP} with a designated initial configuration (s_0, \vec{u}_0) gives rise to a *configuration graph on \mathfrak{M}* consisting of all configurations in \mathfrak{M} reachable from (s_0, \vec{u}_0) by the configuration transition function. A *play* in a CGM^{GP} \mathfrak{M} is an infinite sequence $\pi = c_0 \vec{\alpha}_0, c_1 \vec{\alpha}_1, \dots$ from $(\text{Con}(\mathfrak{M}) \times \text{Act})^\omega$ such that $c_n \in \widehat{\text{out}}(c_{n-1}, \vec{\alpha}_{n-1})$ for all $n > 0$. The set of all plays in \mathfrak{M} is denoted by $\text{Plays}_{\mathfrak{M}}$. Given a play π we use $\pi[i]$ and $\pi[i, \infty]$ to refer to the i th element and to the subplay starting in position i of π , respectively. A *history* is any finite initial sequence $h = c_0 \vec{\alpha}_0, c_1 \vec{\alpha}_1, \dots, c_n \in (\text{Con}(\mathfrak{M}) \times \text{Act})^* \text{Con}(\mathfrak{M})$ of a play in $\text{Plays}_{\mathfrak{M}}$. The set of all histories is denoted by $\text{Hist}_{\mathfrak{M}}$. For any history h we also define $h[i]$ as for plays and additionally $h[\text{last}]$ and $h[i, j]$ to refer to the last state on h and to the sub-history between i and j , respectively. Finally, we introduce functions \cdot^c , \cdot^u , and \cdot^s which denote the projection of a given play or history to the sequence of its configurations, utility vectors, and states, respectively. For illustration, let us consider the play $\pi = c_0 \vec{\alpha}_0, c_1 \vec{\alpha}_1, \dots$. We have that $\pi[i, \infty] = c_i \vec{\alpha}_i, c_{i+1} \vec{\alpha}_{i+1}, \dots$; $\pi[i] = c_i \vec{\alpha}_i$; $\pi^c[i, \infty] = c_i, c_{i+1}, \dots$; $\pi^c[i] = c_i$; $\pi^a[i] = \vec{\alpha}_i$; $\pi^u[i] = v_i$; and $\pi^s[i] = s_i$ where $c_i = (s_i, \vec{u}_i)$.

Example 2 *Some possible plays starting from s_1 in Example 1 are given in the following where we assume that the initial accumulated payoff is 0 for both agents. We note that this implies that the first action taken by any agent is always C.*

1. *Both players cooperate forever:* $(s_1, 0, 0), (s_1, 2, 2), (s_1, 4, 4), \dots$
2. *After the first round both players defect and the play moves to s_2 , where player I chooses to defect whereas II cooperates. Then I must cooperate while II must defect but at the next round can choose any action, so a possible play is:* $(s_1, 0, 0), (s_1, 2, 2), (s_2, 1, 1), (s_2, 0, -1), (s_2, 0, 1), (s_2, 0, 3), (s_2, 0, 5), \dots$
3. *After the first round player I defects while II cooperates and the play moves to s_3 , where they can get stuck indefinitely, until – if ever – they happen to coordinate, so a possible play is:* $(s_1, 0, 0), (s_1, 2, 2), (s_3, 5, -2), (s_3, 4, -3), (s_3, 3, -4), \dots (s_3, 0, -7), (s_3, -1, -8), \dots$

Note, however, that once player I reaches accumulated payoff 0 he may only apply C at that round, so if player II has enough memory or can observe the accumulated payoffs of I he can use the opportunity to coordinate with I at that round by cooperating, thus escaping the trap at s_3 and making a sure transition to s_2 .

4. If, however, the guards did not force the players to play C when reaching accumulated payoffs 0, then both players could plunge into an endless misery if the play reaches s_3 .

Strategies. A *strategy* of a player a is a function $s_a : \text{Hist} \rightarrow \text{Act}$ such that if $s_a(h) = \alpha$ then $h^u[\text{last}]_a \models g_a(h^s[\text{last}], \alpha)$; that is, actions prescribed by a strategy must be enabled by the guard. Our definition of strategy is based on histories of configurations and actions, so it extends the notion of strategy from [4] where it is defined on histories of states, and includes strategies, typically considered e.g. in the study of repeated games, where often strategies prescribe to the player an action dependent on the previous action, or history of actions, of the other player(s). Such are, for instance, TIT-FOR-TAT or GRIM-TRIGGER in repeated Prisoners Dilemma; likewise for various card games, etc. Since our notion of strategy is very general, it easily leads to undecidable model checking problems. So, we also consider some natural restrictions, such as: *state-based*, *action-based* or *configuration-based*, *memoryless*, *bounded memory*, of *perfect recall* strategies³. Here we adopt a generic approach and assume that two classes of strategies \mathcal{S}^p and \mathcal{S}^o are fixed as parameters, with respect to which the proponents and opponents select their strategies, respectively. The proponent coalition A selects a \mathcal{S}^p -strategy s_A (i.e. one agreeing with the class \mathcal{S}^p) while the opponent coalition $\text{Ag} \setminus A$ selects a \mathcal{S}^o -strategy $s_{\text{Ag} \setminus A}$. The outcome play $\text{outcome_play}_{\mathfrak{M}}(c, (s_A, s_{\text{Ag} \setminus A}), l)$ in a given CGM^{GP} \mathfrak{M} determines the play emerging from the execution of the (complete) strategy profile $(s_A, s_{\text{Ag} \setminus A})$ from configuration c in \mathfrak{M} .

4 The Logic: Quantitative ATL*

We now extend the logic ATL* to the logic QATL* with atomic quantitative objectives being state or path arithmetic constraints over the players' accumulated payoffs. The semantics of QATL* naturally extends the semantics of ATL* over CGM^{GP} s, but parameterised with the two classes of strategies \mathcal{S}^p and \mathcal{S}^o .

Definition 3 (The logic QATL*) *The language of QATL* consists of state formulae φ , which constitute the logic, and path formulae γ , generated as follows, where $A \subseteq \text{Ag}$, $\text{ac} \in \text{AC}$, $\text{apc} \in \text{APC}$, and $p \in \text{Prop}$:*

$$\varphi ::= p \mid \text{ac} \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\langle A \rangle\rangle\gamma \text{ and } \gamma ::= \varphi \mid \text{apc} \mid \neg\gamma \mid \gamma \wedge \gamma \mid \mathbf{X}\gamma \mid \mathbf{G}\gamma \mid \gamma\mathbf{U}\gamma.$$

Let \mathfrak{M} be a CGM^{GP} , c a configuration, $\varphi, \varphi_1, \varphi_2$ state-formulae and $\gamma, \gamma_1, \gamma_2$ path formulae. Further, let \mathcal{S}^p and \mathcal{S}^o be two classes of strategies as described above. The semantics of the path constraints is specified according to the limit-averaging or discounting mechanism adopted for computing the value of a play for a player. Then the truth of a QATL formula at a configuration in \mathfrak{M} is defined by mutual recursion on state and path formulae as follows:*

$\mathfrak{M}, c \models p$ for $p \in \text{Prop}$ iff $p \in L(c^s)$;

$\mathfrak{M}, c \models \text{ac}$ for $\text{ac} \in \text{AC}$ iff $c^u \models \text{ac}$,

$\mathfrak{M}, c \models \varphi_1 \wedge \varphi_2$ iff $\mathfrak{M}, c \models \varphi_1$ and $\mathfrak{M}, c \models \varphi_2$,

$\mathfrak{M}, c \models \neg\varphi$ iff $\mathfrak{M}, c \not\models \varphi$,

$\mathfrak{M}, c \models \langle\langle A \rangle\rangle\gamma$ iff there is a collective \mathcal{S}^p -strategy s_A for A such that for all collective \mathcal{S}^o -strategies $s_{\text{Ag} \setminus A}$ for $\text{Ag} \setminus A$ we have that $\mathfrak{M}, \text{outcome_play}_{\mathfrak{M}}(c, (s_A, s_{\text{Ag} \setminus A})) \models \gamma$.

$\mathfrak{M}, \pi \models \varphi$ iff $\mathfrak{M}, \pi[0] \models \varphi$;

$\mathfrak{M}, \pi \models \text{apc}$ for $\text{apc} \in \text{APC}$ iff $\pi^u \models \text{apc}$.

$\mathfrak{M}, \pi \models \mathbf{X}\gamma$ iff $\mathfrak{M}, \pi[1] \models \gamma$,

³We note that all strategies need to be consistent with the guards, so state-based strategies are only applicable in models where the guards only take into account the current state, but not the accumulated payoffs.

$\mathfrak{M}, \pi \models \mathbf{G}\gamma$ iff $\mathfrak{M}, \pi[i] \models \gamma$ for all $i \in \mathbb{N}_0$,

$\mathfrak{M}, \pi \models \gamma_1 \mathbf{U}\gamma_2$ iff there is $j \in \mathbb{N}_0$ such that $\mathfrak{M}, \pi[j] \models \gamma_2$ and $\mathfrak{M}, \pi[i] \models \gamma_1$ for all $0 \leq i < j$.

Ultimately, we define $\mathfrak{M}, c \models \varphi$ as $\mathfrak{M}, c, 1 \models \varphi$.

If not clear from context, we assume $\models_{(S^v, S^o)}$ for \models .

The semantics presented above extends the standard semantics for ATL^* and is amenable to various refinements and restrictions, to be studied further. For instance, if appropriate, an alternative semantics can be adopted, based on irrevocable strategies [1] or, more generally, on strategy contexts [8] or other mechanisms for strategy commitment and release [2]. Also, the nested operators as defined here access the accumulated utility values and require plays to be infinite. Similarly to [9], one can consider variants of these settings which may yield decidable model checking and better complexity results.

As the logic QATL^* extends ATL^* , it allows expressing all purely qualitative ATL^* properties. It can also express purely quantitative properties, e.g.: $\langle\langle \mathbf{a} \rangle\rangle \mathbf{G}(v_a > 0)$ meaning “Player \mathbf{a} has a strategy to maintain his accumulated payoff to be always positive”, or $\langle\langle A \rangle\rangle (w_a \geq 3)$ meaning “The coalition A has a strategy that guarantees the value of the play for player \mathbf{a} to be at least 3”. Moreover, QATL^* can naturally express combined qualitative and quantitative properties, e.g. $\langle\langle \mathbf{a}, \mathbf{b} \rangle\rangle ((v_a + v_b \geq 1) \mathbf{U}p)$, etc.

Example 3 The following QATL^* state formulae are true at state s_1 of the CGM^{GP} in Example 1, where p_i is an atomic proposition true only at state s_i , for each $i = 1, 2, 3$:

- (i) $\langle\langle I, II \rangle\rangle \mathbf{F}(p_1 \wedge v_I > 100 \wedge v_{II} > 100) \wedge \langle\langle I, II \rangle\rangle \mathbf{XXX} \langle\langle II \rangle\rangle (\mathbf{G}(p_2 \wedge v_I = 0) \wedge \mathbf{F} v_{II} > 100)$.
- (ii) $\neg \langle\langle I \rangle\rangle \mathbf{G}(p_1 \vee v_I > 0) \wedge \neg \langle\langle I, II \rangle\rangle \mathbf{F}(p_3 \wedge \mathbf{G}(p_3 \wedge (v_I + v_{II} > 0)))$.

5 (Un)Decidability: Some Preliminary Results

Generally, the CGM^{GP} models are too rich and the language of QATL^* is too expressive to expect computational efficiency, or even decidability, of either model checking or satisfiability testing. Some preliminary results and related work show that model checking of QATL^* in CGM^{GP} s is undecidable under rather weak assumptions, e.g. if the proponents or the opponents can use memory-based strategies. These undecidability results are not surprising as CGM^{GP} s are closely related to Petri nets and vector addition systems and it is known that model checking over them is generally undecidable. In [21], for example, this is shown for fragments of CTL and (state-based) LTL over Petri nets. Essentially, the reason is that the logics allow to encode a “test for zero”; for Petri nets this means to check whether a place contains a token or not. In our setting undecidability follows for the same reason, and we will sketch some results below.

Undecidability results. The logic QATL restricts QATL^* in the same way as ATL restricts ATL^* , due to lack of space we skip the formal definition. As a first result we show that model checking QATL is undecidable even if only the proponents are permitted to use perfect recall strategies and the opponents are bound to memoryless strategies. More formally, let S^{pr} denote the class of perfect recall state-based strategies and S^m the class of memoryless state-based strategies. That is, strategies of the former class are functions of type $\text{St}^* \rightarrow \text{Act}$ and of the latter class functions of type $\text{St} \rightarrow \text{Act}$.

Undecidability can be shown using ideas from e.g. [9, 21]. Here, we make use of the construction of [9] to illustrate the undecidability by simulating a two-counter machine (TCM). A TCM [25] can be considered as a transition system equipped with two integer counters that enable/disable transitions. Each step of the machine depends on the current state, symbol on the tape, and the counters, whether they are zero or not. After each step

the counters can be incremented (+1), or decremented (-1), the latter only if the respective counter is not zero. A TCM is essentially a (nondeterministic) push-down automaton with two stacks and exactly two stack symbols (one of them is the initial stack symbol) and has the same computation power as a Turing machine (cf. [25]). A *configuration* is a triple (s, w_1, w_2) describing the current state (s), the value of counter 1 (w_1) and of counter 2 (w_2). A *computation* δ is a sequence of subsequent configurations effected by transitions.

For the simulation, we associate each counter with a player. The player's accumulated payoff encodes the counter value; actions model the increment/decrement of the counters; guards ensure that the actions respect the state of the counters. The accepting states of the two-counter machine are encoded by a special proposition *halt*. Now, the following lemma stating the soundness of the simulation can be proved:

Lemma 1 (Reduction) *For any two-counter machine A we can construct a finite CGM^{GP} \mathfrak{M}^A with two players and proposition *halt* such that the following holds: A halts on the empty input iff \mathfrak{M}^A contains a play π with $\pi^c = (s^0, (v_1^0, v_2^0))(s^1, (v_1^1, v_2^1)) \dots$ such that there exists $j \in \mathbb{N}$ with $\text{halt} \in L(s^j)$.*

The next theorem gives two cases for which the model checking problem is undecidable. By the previous Lemma we have to ensure that the halting state is reached which can be expressed by $\langle\langle 1 \rangle\rangle \mathbf{Fhalt}$. We can also use purely state-based guards and encode the consistency checks in the formula as follows: $\langle\langle 1 \rangle\rangle (v_1 \geq 0 \wedge v_2 \geq 0 \wedge \text{test} \rightarrow (e_1 \leftrightarrow v_1 = 0 \wedge e_2 \leftrightarrow v_2 = 0)) \mathbf{Uhalt}$ where the proposition e_i is added to the model to indicate that the value of counter i is zero. This has to be tested at distinguished “test states” that are labelled with proposition *test*. Note that this information is static and obtained from the transition relation of the automaton.

Proposition 1 *Model checking the logic QATL is undecidable, even for the 2 agent case and no nested cooperation modalities, where $S^p = S^{pr}$ and $S^o = S^m$.*

Restoring decidability. There are some natural semantic restrictions of QATL* where decidability may be restored; these include for instance, the enabling of only memoryless strategies, imposing non-negative payoffs, constraints on the transition graph of the model, bounds on players utilities etc. Combined with restrictions on the language, esp. on the arithmetic constraints on payoffs, these can enable reductions to known decidable problems, e.g., reachability and coverability problems in Petri nets or to solving energy parity games. Some such results will be reported in the full paper.

6 Concluding Remarks

This paper proposes a long-term research agenda bringing together issues, techniques and results from several research fields. It aims at bridging the two important aspects of reasoning about objectives and abilities of players in multi-player games: quantitative and qualitative, and eventually providing a uniform framework for strategic reasoning in multi-agent systems.

References

- [1] T. Ågotnes, V. Goranko, and W. Jamroga. Alternating-time temporal logics with irrevocable strategies. In D. Samet, editor, *Proceedings of TARK XI*, pages 15–24, 2007.

- [2] T. Ågotnes, V. Goranko, and W. Jamroga. Strategic commitment and release in logics for multi-agent systems (extended abstract). Technical Report IfI-08-01, Clausthal University of Technology, 2008.
- [3] Natasha Alechina, Brian Logan, Nguyen Hoang Nga, and Abdur Rakib. Logic for coalitions with bounded resources. *J. Log. Comput.*, 21(6):907–937, 2011.
- [4] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time Temporal Logic. *Journal of the ACM*, 49:672–713, 2002.
- [5] Aaron Bohy, Véronique Bruyère, Emmanuel Filiot, and Jean-François Raskin. Synthesis from LTL specifications with mean-payoff objectives. In Nir Piterman and Scott A. Smolka, editors, *TACAS*, volume 7795 of *Lecture Notes in Computer Science*, pages 169–184. Springer, 2013.
- [6] Patricia Bouyer, Romain Brenguier, Nicolas Markey, and Michael Ummels. Nash equilibria in concurrent games with Büchi objectives. In S. Chakraborty and A. Kumar, editors, *FSTTCS'2011 LIPIcs*, pages 375–386. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011.
- [7] Patricia Bouyer, Romain Brenguier, Nicolas Markey, and Michael Ummels. Concurrent games with ordered objectives. In L. Birkedal, editor, *Proc. of FoSSaCS'2012, Springer LNCS, vol. 7213*, pages 301–315, 2012.
- [8] T. Brihaye, A. Da Costa, F. Laroussinie, and N. Markey. ATL with strategy contexts and bounded memory. Technical Report LSV-08-14, ENS Cachan, 2008.
- [9] Nils Bulling and Berndt Farwer. On the (Un-)Decidability of Model-Checking Resource-Bounded Agents. In H. Coelho and M. Wooldridge, editors, *Proc. of ECAI 2010*, pages 567–572. IOS Press, Amsterdam, 2010.
- [10] Nils Bulling and Valentin Goranko. How to be both rich and happy: Combining quantitative and qualitative strategic reasoning about multi-player games (extended abstract). In *Proceedings of the 1st International Workshop on Strategic Reasoning*, Electronic Proceedings in Theoretical Computer Science, pages 33–41, Rome, Italy, March 2013.
- [11] Krishnendu Chatterjee, Luca de Alfaro, and Thomas A. Henzinger. Qualitative concurrent parity games. *ACM Trans. Comput. Log.*, 12(4):28, 2011.
- [12] Krishnendu Chatterjee and Laurent Doyen. Energy parity games. *Theor. Comput. Sci.*, 458:49–60, 2012.
- [13] Krishnendu Chatterjee, Laurent Doyen, Thomas A. Henzinger, and Jean-François Raskin. Generalized mean-payoff and energy games. In K. Lodaya and M. Mahajan, editors, *Proc. of FSTTCS'2010*, volume 8 of *LIPIcs*, pages 505–516. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2010.
- [14] Krishnendu Chatterjee and Thomas A. Henzinger. Reduction of stochastic parity to stochastic mean-payoff games. *Inf. Process. Lett.*, 106(1):1–7, 2008.
- [15] Krishnendu Chatterjee and Thomas A. Henzinger. A survey of stochastic ω -regular games. *J. Comput. Syst. Sci.*, 78(2):394–413, 2012.
- [16] Krishnendu Chatterjee, Thomas A. Henzinger, and Marcin Jurdzinski. Mean-payoff parity games. In *Proc. of LICS'2005*, pages 178–187. IEEE Computer Society, 2005.

- [17] Krishnendu Chatterjee, Marcin Jurdzinski, and Thomas A. Henzinger. Simple stochastic parity games. In M. Baaz and J. Makowsky, editors, *Proc. of CSL'2003*, volume 2803 of *LNCS*, pages 100–113. Springer, 2003.
- [18] Krishnendu Chatterjee, Marcin Jurdzinski, and Thomas A. Henzinger. Quantitative stochastic parity games. In J. Ian Munro, editor, *Proc. of SODA'2004*, pages 121–130. SIAM, 2004.
- [19] Luca de Alfaro and Thomas A. Henzinger. Concurrent omega-regular games. In *Proc. of LICS 2000*, pages 141–154, 2000.
- [20] Luca de Alfaro, Thomas A. Henzinger, and Orna Kupferman. Concurrent reachability games. *Theor. Comput. Sci.*, 386(3):188–217, 2007.
- [21] Javier Esparza. Decidability of model checking for infinite-state concurrent systems. *Acta Informatica*, 34:85–107.
- [22] Javier Esparza. Decidability and complexity of petri net problems - an introduction. In *In Lectures on Petri Nets I: Basic Models*, pages 374–428. Springer-Verlag, 1998.
- [23] Valentin Goranko and Wojciech Jamroga. Comparing semantics of logics for multi-agent systems. *Synthese*, 139(2):241–280, 2004.
- [24] Erich Grädel and Michael Ummels. Solution Concepts and Algorithms for Infinite Multiplayer Games. In Krzysztof Apt and Robert van Rooij, editors, *New Perspectives on Games and Interaction*, volume 4 of *Texts in Logic and Games*, pages 151–178. Amsterdam University Press, 2008.
- [25] JE Hopcroft and JD Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, Massachusetts, 1979.
- [26] W. Jamroga and T. Ågotnes. Constructive knowledge: What agents can achieve under incomplete information. *Journal of Applied Non-Classical Logics*, 17(4):423–475, 2007.
- [27] Dario Della Monica, Margherita Napoli, and Mimmo Parente. On a logic for coalitional games with priced-resource agents. *Electr. Notes Theor. Comput. Sci.*, 278:215–228, 2011.
- [28] M. Osborne and A. Rubinstein. *A Course in Game Theory*. MIT Press, 1994.
- [29] M. Pauly. A modal logic for coalitional power in games. *J. of Logic and Computation*, 12(1):149–166, 2002.
- [30] H.J.M. Peters and O.J. Vrieze, editors. *Surveys in Game Theory and Related Topics*. CWI, Amsterdam, 1987.
- [31] Sophie Pinchinat. A generic constructive solution for concurrent games with expressive constraints on strategies. In K. Namjoshi et al, editor, *Proc. of ATVA'2007*, Springer LNCS, vol. 4762, pages 253–267, 2007.
- [32] W. van der Hoek and M. Wooldridge. Cooperation, knowledge and time: Alternating-time Temporal Epistemic Logic and its applications. *Studia Logica*, 75(1):125–157, 2003.
- [33] Wiesław Zielonka. An invitation to play. In J. Jędrzejowicz and A. Szepietowski, editors, *Proc. of MFCS'2005*, volume 3618 of *LNCS*, pages 58–70. Springer, 2005.

Nils Bulling
Clausthal University of Technology, Germany
Email: bulling@in.tu-clausthal.de

Valentin Goranko
Technical University of Denmark
Email: vfgo@dtu.dk