Expressiveness and Nash Equilibrium in Iterated Boolean Games

Julian Gutierrez
Dept. of Computer Science
University of Oxford
Oxford, United Kingdom
julian.gutierrez@cs.ox.ac.uk

Paul Harrenstein
Dept. of Computer Science
University of Oxford
Oxford, United Kingdom
paulhar@cs.ox.ac.uk

Giuseppe Perelli
Dept. of Computer Science
University of Oxford
Oxford, United Kingdom
giuseppe.perelli@cs.ox.ac.uk

Michael Wooldridge
Dept. of Computer Science
University of Oxford
Oxford, United Kingdom
mjw@cs.ox.ac.uk

ABSTRACT

We introduce and investigate a novel notion of expressiveness for temporal logics that is based on game theoretic properties of multi-agent systems. We focus on *iterated Boolean games*, where each agent i has a goal γ_i, represented using (a fragment of) Linear Temporal Logic (LTL). The goal γ_i captures agent i's preferences: the models of γ_i represent system behaviours that would satisfy i, and each player is assumed to act strategically, taking into account the goals of other players, in order to bring about computations satisfying their goal. In this setting, we apply the standard game-theoretic concept of Nash equilibria: the Nash equilibria of an iterated Boolean game can be understood as a (possibly empty) set of computations, each computation representing one way the system could evolve if players chose strategies in Nash equilibrium. Such an equilibrium set of computations can be understood as expressing a temporal property—which may or may not be expressible within a particular LTL fragment. The new notion of expressiveness that we study is then as follows: *what* LTL *properties are characterised by the Nash equilibria of games in which agent goals are expressed in fragments of* LTL? We formally define and investigate this notion of expressiveness and some related issues, for a range of LTL fragments.

Keywords

Iterated Boolean games; LTL; Expressiveness; Nash equilibrium

1. INTRODUCTION

Temporal logics are probably the most successful and widely used class of formalisms for the specification and verification of computer systems [10]. In particular, temporal logics have proven to be enormously valuable in model checking, where a standard question is whether all computations of a given system satisfy a particular temporal logic property φ [7]. A natural question relating to temporal logics is that of their *expressive power*: what system properties is it possible to express within a particular temporal logic or temporal logic fragment? For example, the relative expressiveness of linear *versus* branching time temporal logics was a major research topic in theoretical computer science for more than a decade [11, 31], and still generates debate to the present day.

In this article, we are interested in the use of temporal logic for reasoning about multi-agent systems, and in particular, we are interested in questions relating to expressiveness that arise in such settings. We use *iterated Boolean games* as our abstract model of multi-agent systems [15]. In this model, each agent i exercises exclusive control over a subset of Boolean variables, and the game is played over an infinite number of rounds, where at each round each player chooses a valuation for their variables. The result of play is an infinite computation, which can be understood as a model for Linear Temporal Logic (LTL) [26, 10]. To represent agent preferences in iterated Boolean games, each player i is assumed to have a goal γ_i, expressed using (a fragment of) LTL: the models of γ_i represent computations/plays that would satisfy i. Each player is assumed to act strategically, taking into account the goals of other players, in order to try to bring about computations that will satisfy their goal. We can then apply the standard game-theoretic concept of Nash equilibrium: the Nash equilibria of an iterated Boolean game can be understood as a (possibly empty) set of computations, with each computation representing one way the system could evolve if players in the game chose strategies in equilibrium.

Our main interest in the present paper is as follows. The Nash equilibria of an iterated Boolean game are a set of computations, and such a set of computations can be understood as expressing a temporal property. Now, suppose we have a game G in which each player i has a goal γ_i expressed in a fragment L of LTL: then, *what temporal properties can be expressed by the equilibria of* G? In particular, it is very natural to ask whether the equilibria of an L-game can be characterised within L itself. We formally define and investigate this novel notion of expressiveness as well as some related issues for a range of known fragments of LTL.

The problem of expressing Nash equilibria of concurrent games has, of course, been considered elsewhere. For instance, a popular approach is to develop new formalisms for representing temporal properties of Nash equilibria, and similar game-theoretic solution concepts, *in the object language* by adding new operators to existing temporal logics [5, 14]. We believe that our approach—focussing on the temporal properties that Nash equilibrium can distinguish in logic-based (Boolean) games—is markedly different.

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

As a motivating example, consider the following temporal variation on the *Battle of the Sexes* game (see, for instance, [21]).

EXAMPLE 1 (BOOLEAN BALLET). *Alice and Bob go out every weekend, either alone or together. They have to decide repeatedly whether to go to the prize fight or to the ballet, but are unable to correlate their actions. Their preferences, however, are slightly divergent. Alice wishes always to go out together, whereas Bob wants to go at least once to the ballet together with Alice but also at least once wishes to see the prize fight alone. The situation can be modelled as an iterated Boolean game where Alice controls variable p and Bob variable q. Setting p to true indicates Alice going to the ballet and similarly for q and Bob. Setting p, (respectively, q), to false means going to the prize fight. The preferences of the two agents can then be represented as follows, where $\mathsf{G}\,\varphi$ stands for "always φ," $\mathsf{F}\,\varphi$ for "eventually φ," and $\bar{q}$ for "not q":*

$$\gamma_{Alice} = \mathsf{G}(p \leftrightarrow q) \qquad \gamma_{Bob} = \mathsf{F}(p \wedge q) \wedge \mathsf{F}(p \wedge \bar{q})$$

Thus, the preferences of the players are phrased using only the temporal operators F and G. Observe that the set of equilibrium runs includes the run $pq, \bar{p}\bar{q}, \bar{p}\bar{q}, \bar{p}\bar{q}, \ldots$: Alice and Bob go out to the ballet together the first time and meet at the prize fight forever after. Alice will thus be satisfied. Not so Bob, but there is little Bob can do apart from going to the ballet alone at some point and be miserable as well, or miss out on the ballet altogether. The set of equilibrium runs, however, does not include $pq, pq, \bar{p}\bar{q}, \bar{p}\bar{q}, \bar{p}\bar{q}, \ldots$ as Bob would dash off to the prize fight after having been to the ballet first. The runs that are sustained by a Nash equilibrium are precisely those in which always either pq or $\bar{p}\bar{q}$ is the case and in which either pq or $\bar{p}\bar{q}$ occurs at most once. This, however, is not a property that can be expressed in the (stutter-invariant) fragment of LTL *with F and G as the only temporal operators.*[1] *Moreover, it raises the question as to which properties can be characterised as the set of runs sustained by a Nash equilibrium of some iterated Boolean game with the players' preferences formulated in this and other fragments.*

2. PRELIMINARIES

Our analysis uses *Linear Temporal Logic* and the iterated Boolean games based on it. In this section, we present the core concepts of these frameworks along with a number of auxiliary notions.

*Linear Temporal Logic (*LTL*).*

We use the well-known framework of **Linear Temporal Logic** (LTL) [26, 10, 2]. The formulae of LTL are constructed in the usual fashion from a non-empty and finite set Φ of propositional variables using the Boolean connectives *negation* ($\neg\varphi$) and *disjunction* ($\varphi \vee \psi$), as well as the temporal operators *next* ($\mathsf{X}\,\varphi$), *eventually* ($\mathsf{F}\,\varphi$), *always* ($\mathsf{G}\,\varphi$), and *until* ($\varphi\,\mathsf{U}\,\psi$). *truth* ($\top$), *falsity* ($\bot$), *conjunction* ($\varphi \wedge \psi$), *implication* ($\varphi \rightarrow \psi$), and *bi-implication* ($\varphi \leftrightarrow \psi$), are introduced as the usual abbreviations of $p \vee \neg p$, $\neg\top$, $\neg(\neg\varphi \vee \neg\psi)$, $\neg\varphi \vee \psi$, and $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$, respectively. For p a propositional variable we will write $\bar{p}$ for $\neg p$. We also omit conjunctions in conjunctive clauses and, for instance, denote $p \wedge \neg q \wedge r$ by $p\bar{q}r$.

By a **valuation** v we understand a subset of propositional variables, that is, $v \subseteq \Phi$. Thus the set of valuations over Φ is given by 2^Φ. Intuitively, a propositional variable p is set to true at valuation v whenever $p \in v$. For a valuation $v \subseteq \Phi$, we have χ_v^Φ denote the **characteristic clause for** v given by $\chi_v^\Phi = \bigwedge_{p \in v} p \wedge \bigwedge_{q \in \Phi \setminus v} \bar{q}$. Thus, for $w, v \subseteq \Phi$, we have $w \models \chi_v^\Phi$ if and only if $v = w$. We will also identify valuations and their characterising clauses, that is, write $p\bar{q}r$ for valuation $\{p, r\}$ if $\Phi = \{p, q, r\}$.

The formulae of LTL are interpreted with respect to **runs** $\rho = v_1, v_2, \ldots$, which we define as infinite sequences (or ω-*words*) over valuations in Φ, that is, $\rho \in (2^\Phi)^\omega$. We denote the set of runs over valuations in 2^Φ by $runs(\Phi)$, again omitting the reference to Φ when clear from the context.

The semantics of LTL then interprets LTL-formulae with respect to a run $\rho = v_0, v_1, v_2, \ldots$ and time index $t \in \mathbb{N}$ as follows.

$$\begin{array}{lll}
\rho, t \models p & \text{iff} & p \in v_t \quad (\text{for } p \in \Phi) \\
\rho, t \models \neg\varphi & \text{iff} & \rho, t \not\models \varphi \\
\rho, t \models \varphi \vee \psi & \text{iff} & \rho, t \models \varphi \text{ or } \rho, t \models \psi \\
\rho, t \models \mathsf{X}\,\varphi & \text{iff} & \rho, t+1 \models \varphi \\
\rho, t \models \mathsf{F}\,\varphi & \text{iff} & \rho, t' \models \varphi \text{ for some } t' \geq t \\
\rho, t \models \mathsf{G}\,\varphi & \text{iff} & \rho, t' \models \varphi \text{ for all } t' \geq t \\
\rho, t \models \varphi\,\mathsf{U}\,\psi & \text{iff} & \text{for some } t' \geq t \text{ both } \rho, t' \models \psi \text{ and} \\
 & & \rho, t'' \models \varphi \text{ for all } t \leq t'' < t'
\end{array}$$

We say that a run ρ **satisfies** a formula φ if $\rho, 0 \models \varphi$. The set of runs that satisfy formula φ we denote by $[\![\varphi]\!]$. A formula φ is **satisfiable** if some run satisfies φ. Observe that $\neg\,\mathsf{X}\,\varphi$ is equivalent to $\mathsf{X}\,\neg\varphi$, and $\mathsf{F}\,\varphi$ to $\neg\,\mathsf{G}\,\neg\varphi$.

We also employ a number of auxiliary concepts. For $\Psi \subseteq \Phi$ and $\rho \in runs(\Phi)$, we write $\rho|_\Psi$ for the **restriction** (or **projection**) of ρ to Ψ, that is, if $\rho = v_1, v_2, \ldots$ then $\rho|_\Psi = w_1, w_2, \ldots$ where $w_t = v_t \cap \Psi$ for each $t \geq 1$. For $X \subseteq runs(\Phi)$ and $\Psi \subseteq \Phi$, we denote by $X|_\Psi$ the set $\{\rho|_\Psi : \rho \in X\}$.

By a **history** we understand a *finite* and possibly empty sequence $\pi = v_0, \ldots, v_k$ in V^*. We let $length(\pi)$ denote the **length** of π. We furthermore use the usual operations on histories. For π and π', we have $\pi; \pi'$ denote the *concatenation* of π and π', and $\pi^\omega = \pi; \pi; \pi; \ldots$ the *infinite iteration* of π. We also specify subsets of runs in $runs(\Phi)$ by ω-*regular expressions* over the alphabet $\Sigma = 2^\Phi$ using the operations *concatenation* ($X; Y$), *union* ($X + Y$), *finite iteration* (X^*), and *infinite iteration* (X^ω).

A history $\pi = w_0, \ldots, w_k$ is a **prefix** of a run $\rho = v_0, v_1, v_2, \ldots$ (in symbols $\pi \leq \rho$) if $v_t = w_t$ for all $t \leq k$. The set of prefixes of a run ρ we denote by $prefix(\rho)$. We say that a history π is the (unique) **maximal common prefix** of two (finite or infinite) runs ρ and ρ' if π is a prefix of both ρ and ρ' and no prefix of greater length has this property. For instance, $\pi = pq, p\bar{q}$ and $\pi' = pq, p\bar{q}, \bar{p}q$ are both a common prefix of the runs $\rho = pq, p\bar{q}, \bar{p}q, pq, pq, \ldots$ and $\rho' = pq, p\bar{q}, \bar{p}q, \bar{p}\bar{q}, \bar{p}\bar{q}, \ldots$, but only π' is maximal.

We will invoke frequently the notion of an ultimately periodic run. A run $\rho = v_0, v_1, v_2, \ldots$, is said to be **ultimately periodic** if there are $s, p \geq 0$, called the **starting index** and **period**, respectively, such that for every $t \geq s$, we have that $v_t = v_{t+p}$. The *Ultimately Periodic Model Theorem* in [29] establishes that every satisfiable LTL-formula is satisfied by an ultimately periodic run.

THEOREM 2 (SISTLA AND CLARKE [29]). *Every satisfiable* LTL-*formula is satisfied by an ultimately periodic run.*

Moreover, every ultimately periodic run $\rho = v_0, v_1, \ldots$ over 2^Φ with starting index s and period p is characterised by the formula

$$\chi_\rho = \bigwedge_{0 \leq t \leq s+p} \mathsf{X}^t \chi_{v_t} \wedge \mathsf{X}^s\,\mathsf{G}(\bigwedge_{v \in 2^\Phi} (\chi_v \rightarrow \mathsf{X}^p \chi_v)),$$

where $\mathsf{X}^k \varphi = \underbrace{\mathsf{X} \cdots \mathsf{X}}_{k \text{ times}} \varphi$.

PROPOSITION 3. *Let $\rho = v_0, v_1, \ldots$ be an ultimately periodic run over 2^Φ. Then, ρ is the unique run over 2^Φ that satisfies χ_ρ.*

[1] Interestingly, had Bob's goal been to go both to the ballet with Alice and to the prize fight alone infinitely often, that is, if Bob's goal had been $\mathsf{G}(\mathsf{F}(p \wedge q) \wedge \mathsf{F}(p \wedge \bar{q}))$, the set of Nash equilibria of the resulting iterated Boolean game would be expressible in this fragment, namely by $\mathsf{G}(p \leftrightarrow q) \wedge (\mathsf{F}\,\mathsf{G}\,pq \vee \mathsf{F}\,\mathsf{G}\,\bar{p}\bar{q})$.

Fragments of Linear Temporal Logic.

We study the expressive power of the most natural, and therefore most widely known, fragments of LTL. Such fragments are the "stutter-invariant" fragment (technically, the X-free fragment), denoted by L_{U}, as well as other fragments where the use of the "until" operator is restricted to simply being G or F, leading to the following sublogics: $L_{\mathsf{G,F,X}}$, where only G and F and X are allowed, and with similar interpretations, the sublogics $L_{\mathsf{G,F}}$, $L_{\mathsf{X,G^+}}$, and $L_{\mathsf{X,F^+}}$, where the "+" notation indicates that negations are allowed only in front of propositional variables (otherwise, for instance, the L_{G} fragment would be the same as the L_{F} fragment). The fragment $L_{\mathsf{G,F}}$ was briefly discussed in Example 1. We will not study the extremely weak sublogics $L_{\mathsf{G^+}}$, $L_{\mathsf{F^+}}$, L_{X}, and $L_{\emptyset}$ (the latter referring to propositional logic). On the one hand, the two latter sublogics cannot express interesting properties of infinite runs—all their models are finite words; on the other hand, the first two sublogics can express only very limited classes of temporal properties: only "safety" properties in the case of $L_{\mathsf{G^+}}$ and only "reachability" properties in the case of $L_{\mathsf{F^+}}$. Indeed, our study covers the most relevant LTL fragments in the literature. Finally, by $L_{\omega\text{-reg}}$ we refer to the set of ω-regular expressions, which are seen as an *extension* of LTL, rather than a fragment [33]. Sometimes we refer explicitly to the set Ψ of variables over which L is defined, and write $L(\Psi)$ for L.

Iterated Boolean Games.

Boolean games were introduced by Harrenstein et al. [17] and further popularised by, among others, Bonzon, Lang, and Wooldridge (see, for instance, [4, 12]). In this paper, we adopt the framework of *iterated Boolean games* as proposed by Gutierrez et al. [15], where players play a Boolean game over an infinite number of rounds and where the players's goals are given by an LTL-formula. For L a fragment of LTL over Φ, an L-**iterated Boolean game** is a tuple

$$G = (N, \Phi, \Phi_1, \dots, \Phi_n, \gamma_1, \dots, \gamma_n),$$

where N is a set of players, each $\Phi_i \subseteq \Phi$ is a subset of propositional variables under control of player i, and γ_i is a formula in L representing player i's preferences over $runs(\Phi)$. We assume $\Phi_1, \dots, \Phi_n$ to partition Φ, that is, $\Phi_1 \cup \dots \cup \Phi_n = \Phi$ and $i \neq j$ implies $\Phi_i \cap \Phi_j = \emptyset$. Henceforth, we refer to an L-iterated Boolean game as an L-**game**. We say that G is an L-game over Φ if we make explicit reference to the set Φ.

An iterated Boolean game takes place in an infinite number of rounds and in every round each player i simultaneously makes a **choice** $v_i \subseteq \Phi_i$ of values for the propositional variables under its control based on the values chosen by *all* players in previous rounds. Formally, a **strategy** for a player i is a function $f_i \colon (2^{\Phi})^* \to 2^{\Phi_i}$ which associates with every history $\pi \in V^*$ a choice $f_i(\pi) \in 2^{\Phi_i}$. A **strategy profile** is a tuple $f = (f_1, \dots, f_n)$ that associates with each player i a strategy f_i and induces an infinite run $\rho(f) = v_0, v_1, v_2, \dots$ defined as follows:

$$\begin{aligned} v_0 &= f_1(\epsilon) \cup \dots \cup f_n(\epsilon) \\ v_{t+1} &= f_1(v_0, \dots, v_t) \cup \dots \cup f_n(v_0, \dots, v_t) \end{aligned}$$

A player i strictly prefers runs that satisfy γ_i to runs that do not and is indifferent otherwise, that is, i **strictly prefers** run ρ to run ρ' if and only if $\rho \not\models \gamma_i$ and $\rho' \models \gamma_i$. Thus, each player's preferences in iterated Boolean games are dichotomous, dividing the set of runs into those that are preferred and those that are not preferred.

It can easily be seen that with the players, strategies and preferences defined in this way, each iterated Boolean game defines a **strategic game** in the game-theoretic sense of the word (see, e.g., [24, 22, 28]). Accordingly, the usual game theoretic solution concepts are available for the analysis of iterated Boolean games. This in particular holds for **Nash equilibrium**, which in our present setting is a strategy profile $f^* = (f_1^*, \dots, f_n^*)$ such that for all players i and all of i's strategies g_i, we have that $\rho(f^*_{-i}, g_i) \models \gamma_i$ implies $\rho(f^*) \models \gamma_i$, where (f^*_{-i}, g_i) denotes the profile $(f_1^*, \dots, f_{i-1}^*, g_i, f_{i+1}^*, \dots, f_n^*)$. We furthermore say that a run $\rho \in runs(\Phi)$ is *sustained* by a Nash equilibrium f^* in a game G whenever $\rho(f^*) = \rho$. We then refer to ρ as an **equilibrium run**. The set of equilibrium runs of G—rather than the set of equilibria itself—we denote by $NE(G)$.

Expressiveness.

Given a suitable model-theoretic semantics, the expressive power of a logic can be measured in terms of the sets of models it can characterise. In the case of LTL, the models are runs of valuations and by a **linear time property** given a set Φ of propositional variables, we understand any subset $X \subseteq runs(\Phi)$ of runs.

We say that an LTL-fragment L can **express** a property $X \subseteq runs(\Phi)$ if there is some formula $\varphi \in L$ such that $X = [\![\varphi]\!]$. A weaker notion of expressiveness applies when, given a temporal property, the fragment can express a stronger property. Thus, L is said to **weakly express** the non-empty property X whenever there is a *satisfiable* formula $\varphi \in L$ with $[\![\varphi]\!] \subseteq X$.

The concept of weak expressiveness is much weaker than standard expressiveness, in the sense that fragments are likely to be able to express considerably more properties in the weak sense than they can under the standard notion. Indeed, for an LTL-fragment L to weakly express *every* LTL-property it suffices to be able to express $\emptyset$ and $\{\rho\}$ for every ultimately periodic run $\rho \in runs(\Phi)$. By virtue of Proposition 3, this holds, for instance, for the fragment $L_{\mathsf{X,F,G}}$. In settings where a designer uses temporal logic to specify the desired behaviour of a system, however, the concept of weak expressiveness seems to be quite natural. If the designer wants a system, for instance a multi-agent system, to behave accordingly to the specification and she manages to design it so that it behaves in equilibrium according to a *stronger* but consistent specification, she should still be satisfied.

A third notion of expressiveness abstracts away from the propositional variables available in the language. Rather than requiring that a temporal property X on 2^{Φ} *coincide* with the set of runs satisfying some formula φ in a fragment $L(\Phi)$, it demands that X be the set of *projections to* Φ of the runs satisfying some formula φ of L in *an extended set of propositional variables*. Formally, we say that LTL-fragment $L(\Phi)$ can **projectively express** property $X \subseteq runs(\Phi)$ if there is some finite set Ψ of auxiliary variables and some formula $\varphi \in L(\Phi \cup \Psi)$ such that $X = \{\rho|_{\Phi} \in runs(\Phi) \colon \rho \in [\![\varphi]\!]\}$.

Finally, a fragment $L(\Phi)$ is said to be able to **weakly projectively express** a non-empty property $X \subseteq runs(\Phi)$ if there is some finite set Ψ of variables and some formula $\varphi \in L(\Phi \cup \Psi)$ such that for every run $\rho \in runs(\Phi \cup \Psi)$ we have that $\rho \in [\![\varphi]\!]$ implies $\rho|_{\Phi} \in X$.

On the basis of the above concepts we also introduce concepts of *relative* expressiveness. Thus, we say that fragment L_1 is **at least as expressive as** another fragment L_2, in symbols $L_1 \geq L_2$, if every property that can be expressed by L_2 can also be expressed by L_1. Furthermore, $L_1 \geq^{\mathsf{w}} L_2$ denotes that every property that L_2 can express (not weakly express!), can also be weakly expressed by L_1. The notations $L_1 \geq^{\mathsf{p}} L_2$ and $L_1 \geq^{\mathsf{wp}} L_2$ are defined analogously for projective and weak projective expressiveness. Note that the relations $\geq^{\mathsf{p}}$ and $\geq^{\mathsf{wp}}$ need not be reflexive. Finally, $L_1 >^{\mathsf{w}} L_2$ denotes that $L_1 \geq^{\mathsf{w}} L_2$ and there is a property $X \in runs(\Phi)$ that is not expressible in L_2 but that L_1 can weakly express. The notations $L_1 > L_2$, $L_1 >^{\mathsf{p}} L_2$, and $L >^{\mathsf{wp}} L_2$ are introduced analogously.[2]

[2] Observe that it is *not* generally the case that $L_1 >^{\mathsf{w}} L_2$ if and only if $L_1 \geq^{\mathsf{w}} L_2$ and not $L_2 \geq^{\mathsf{w}} L_1$. Similarly for $>^{\mathsf{p}}$ and $>^{\mathsf{wp}}$.

The roots of *projective expressiveness* go back to the work of Beth [3] and Craig [8] in the 1950s on definability in first order logic. The concept has numerous applications in model theory [6, 18] and has recently also been studied in the context of modal and temporal logics [13, 16]. In our setting, projective expressiveness defines relatively weak constraints on a fragment and should be carefully distinguished from standard expressiveness. Thus, it is well known that the properties that LTL can express are *non-counting* and cannot, for instance, characterise temporal property $even(p)$, the set of runs on 2^Φ in which p is set to true at every even state (see, for instance [33]). Still, $q \wedge \mathsf{G}(q \leftrightarrow \mathsf{X}\,\neg q) \wedge \mathsf{G}(q \to p)$ projectively expresses exactly this property. It is known from the literature that every *ω-regular* property can be projectively expressed by LTL if you can use an *unbounded* number of additional propositional variables [13, 30]. We reformulate this result for our setting.

PROPOSITION 4. LTL $\geq L_{\mathsf{X},\mathsf{F},\mathsf{G}} \geq^{\mathsf{p}} L_{\omega\text{-reg}}$.

Proof: Let $A = (Q, 2^\Phi, \delta, Q_0, F)$ be a nondeterministic Büchi automaton [2]. We construct a formula φ_A in $L_{\mathsf{F},\mathsf{G},\mathsf{X}}(\Phi \cup Q)$, where Φ and Q are disjoint, as follows.

$$\begin{aligned}
\varphi^A_{init} &= \textstyle\bigvee_{q \in Q_0} q\\
\varphi^A_{trans} &= \mathsf{G}(\textstyle\bigwedge_{q \in Q}(q \to \bigvee_{\{(q',v)\,:\; q' \in \delta(q,v)\}}(\chi^\Phi_v \wedge \mathsf{X}\, q')))\\
\varphi^A_{accept} &= \mathsf{G}\,\mathsf{F} \textstyle\bigvee_{q \in F} q\\
\varphi^A_{invar} &= \mathsf{G}(\textstyle\bigvee_{q \in Q}(q \wedge \bigwedge_{q' \neq q} \bar{q}'))
\end{aligned}$$

Then set $\varphi^A = \varphi^A_{init} \wedge \varphi^A_{trans} \wedge \varphi^A_{accept} \wedge \varphi^A_{invar}$. By an inductive argument it then follows that $\mathscr{L}_A = \{\rho|_\Phi : \rho \models \varphi_A\}$, where $\mathscr{L}_A$ is the language accepted by A. It follows that $L_{\mathsf{X},\mathsf{F},\mathsf{G}}$ can projectively express all ω-regular temporal logic properties. Recalling that $L_{\omega\text{-reg}}$ is the class of languages over 2^Φ accepted by nondeterministic Büchi automata and that LTL $\geq L_{\mathsf{X},\mathsf{F},\mathsf{G}}$, we obtain the result. $\square$

It is interesting to note that allowing for additional variables along with projection has a similar effect as, for instance, extending LTL to *Extended Temporal Logic* (ETL) by including to the logical language suitable *grammar-operators* as proposed in [33].

Expressiveness in Equilibrium.

Nash equilibrium fundamentally pertains to strategy profiles rather than to runs, and the former have a much richer structure than the latter. A strategy profile $f = (f_1, \dots, f_n)$ not only prescribes which actions each player has to perform at every time during the run $\rho(f)$, it also prescribes actions off this path, that is, for histories that are *not* a prefix of ρ. Thus the following two things can hold in an L-game G:

(*i*) different strategy profiles f and g may induce the same run, that is, $\rho(f) = \rho(g)$, even if f is an equilibrium and g is not;

(*ii*) different runs ρ and ρ' may satisfy the same players' goals, even if one is sustained by a Nash equilibrium and the other one is not.

Possibility (*ii*) makes that the set of equilibrium runs of a game may seem quite unrelated to the goals of the players. To see this, consider the following example.

EXAMPLE 5. *Let G be the two-player game where players 1 and 2 control p and q, respectively, and whose goals are given by $\gamma_1 = \mathsf{F}(p \wedge q)$ and $\gamma_2 = \top$, respectively. Then, run $(\bar{p}\bar{q})^\omega$ is sustained by the Nash equilibrium where the two players always set p and q to false, respectively. Thus, γ_2 is satisfied, but γ_1 is not. This is also true of run $(\bar{p}q)^\omega$, which, however, is not sustained by any Nash equilibrium: no matter what strategies are played, player 1 would deviate to a strategy that sets p to true at some point.*

This last phenomenon illustrates that for fragments L of LTL it is quite possible that the Nash equilibrium runs of L-games are distinct from the sets of runs satisfying L-formulae.

For this reason we now also introduce the following notions of expressiveness, which relate to the temporal properties that are characterised by the sets of equilibria runs of iterated Boolean games in a given fragment L. Thus, we say that LTL-fragment L can **express in equilibrium** property $X \subseteq runs(\Phi)$ if there is an L-game G with $X = NE(G)$. Example 1 shows how the standard notion of expressiveness can be different from expressiveness in equilibrium. The concepts of **weak expressiveness in equilibrium**, **projective expressiveness**, and **weak projective expressiveness** are then defined analogously to weak expressiveness, projective expressiveness, and weak projective expressiveness, respectively, with the role of the extensions $[\![\varphi]\!]$ of L-formulae being taken over by the Nash equilibrium runs $NE(G)$ of L-games G.

The corresponding relative expressiveness notions are similarly defined as expected. Here, we have L^{NE} refer to the temporal language over 2^Φ that is defined by the sets of equilibrium runs of L-games. For fragments L_1 and L_2 we then write $L_1^{NE} \geq L_2$ if every property that can be expressed by L_2, can also be expressed in equilibrium by L_1, that is, if for every $\varphi \in L_2$, there is an L_1-game such that $NE(G) = [\![\varphi]\!]$. The notations $L_1 \geq L_2^{NE}$, $L_1^{NE} \geq L_2^{NE}$ have an analogous interpretation and, moreover, extend naturally to $\geq^{\mathsf{w}}$, $\geq^{\mathsf{p}}$, and $\geq^{\mathsf{wp}}$. Clearly, $L_1 \geq L_2$ implies both $L_1 \geq^{\mathsf{w}} L_2$ and $L_1 \geq^{\mathsf{p}} L_2$. Moreover, each of the following expressiveness relations $\geq$, $\geq^{\mathsf{w}}$, $\geq^{\mathsf{p}}$, and $\geq^{\mathsf{wp}}$ is easily shown to be transitive.

One would expect that, if a fragment L can express a property X, it can also express X in equilibrium. Although generally true, there is one notable exception: if there is only a single propositional variable p and $X = \emptyset$. For L, any unsatisfiable formula will do. For L^{NE}, however, one has to observe that in the setting of iterated Boolean games, control over a single variable p can be assigned to a single player i only, who is then also the only player who could possibly deviate from a profile. If X is the empty property, i should want to deviate from *every* given run ρ and, consequently, i does not have its goal satisfied at any of them. Then, there is no run that i would possibly want to deviate to and the (non-empty) set of all runs coincides with the set of equilibrium runs.

If L expresses X by φ and there are at least two variables p and q, one can construct the **L-game** G^{mp}_φ, which bears a resemblance to the well-known game of *matching pennies* [24], as follows. Let there be two players, i and j, such that i controls p and j all other variables, including q, and whose goals are given by:

$$\gamma_i = \varphi \vee (p \leftrightarrow q) \qquad\qquad \gamma_j = \varphi \vee (p \leftrightarrow \bar{q}).$$

To see that $[\![\varphi]\!]$ is exactly the set of equilibrium runs of G^{mp}_φ, observe that every run that satisfies φ, also satisfies both players' goals and, hence, is sustained by an equilibrium. If, by contrast, a run ρ does not satisfy φ, the players will be entangled in a matching pennies game on p and q. Consequently, ρ is not sustained by an equilibrium. Thus we obtain that

$$L_1 \geq L_2 \text{ implies } L_1^{NE} \geq L_2, \text{ provided that } |\Phi| \geq 2.$$

Observe that G^{mp}_φ could also contain additional variables. Thus, for projective expressiveness in equilibrium we immediately have the following statement.

$$L_1 \geq L_2 \text{ implies } L_1^{NE} \geq^{\mathsf{p}} L_2.$$

Using a construction that is similar to G_φ^{mp}, it can also be shown that, if a fragment L can express a property X, it can also projectively express in equilibrium the complement of X. This result is especially relevant for fragments that are not closed under negation.

LEMMA 6. *Let φ be a formula in fragment L. Then, L can projectively express $[\![\neg\varphi]\!]$ in equilibrium.*

Proof: Let $p, q \notin \Phi$. Consider the two-player L-game G with $\Phi_1 = \{p\}$, $\Phi_2 = \Phi \cup \{q\}$ and the players' goals being given by:

$$\gamma_1 = \varphi \wedge (p \leftrightarrow q) \qquad\qquad \gamma_2 = \bar{p} \leftrightarrow q.$$

First observe that no equilibrium run in G satisfies φ. To see this, let f be an arbitrary profile with $\rho(f) \models \varphi$. If $\rho(f) \not\models \gamma_1$, player 1 would deviate by choosing the opposite value for p in the first round. If, on the other hand, $\rho(f) \models \gamma_1$, player 2 would deviate by choosing the opposite value for q in the first round. Hence, $NE(G)|_\Phi \subseteq [\![\neg\varphi]\!]$. To see that this inclusion also holds in the opposite direction, assume that $\rho \not\models \varphi$ and that profile f induces ρ. Since, neither p nor q occurs in φ, we may assume without loss of generality that $\rho \models \bar{p} \leftrightarrow q$. Therefore, player 2 has its goal achieved and will not deviate. As, moreover, player 1 controls no variables occurring in φ, no deviation from f will satisfy its goal. It follows that f is a Nash equilibrium and $\rho \in NE(G)$, as desired. $\square$

3. EQUILIBRIA AND FRAGMENTS

We now will proceed to investigate expressiveness in equilibrium of a number of well-known fragments of LTL.

3.1 The Full and Empty Fragments

It is not necessarily the case that a fragment can express or projectively express more temporal properties in equilibrium than it normally can. A prime example is *propositional calculus*, that is, the LTL-fragment $L_\emptyset$. A fundamental observation in the literature on Boolean games is that the formula

$$\textstyle\bigwedge_{1\le i\le N}(\bigvee_{\theta\colon \Phi_i\to\{\top,\bot\}} \theta(\gamma_i) \to \gamma_i),$$

where $\theta(\gamma_i)$ results from γ_i by replacing every occurrence of $p \in \Phi_i$ by $\theta(p)$, characterises the set of equilibrium runs of the $L_\emptyset$-game $(N, \Phi, \Phi_1, \ldots, \Phi_n, \gamma_1, \ldots, \gamma_n)$ (see, e.g., [4]). Thus, $L_\emptyset$ can express every property that it can express in equilibria, and vice versa—provided that Φ contains at least two propositional variables.

PROPOSITION 7. *$L_\emptyset > L_\emptyset^{NE}$, if there is at most one propositional variable. Otherwise, both $L_\emptyset \geq L_\emptyset^{NE}$ and $L_\emptyset^{NE} \geq L_\emptyset$.*

The full fragment LTL is at the other end of the syntactic spectrum of the linear temporal logics considered in this paper. Barring the borderline case in which $|\Phi| \le 1$, we have

$$\mathrm{LTL}^{NE} \geq L$$

for all fragments L we consider. Thus, the question remains how expressive LTL is w.r.t. LTL^{NE}, that is, to what extent the temporal properties defined by the Nash equilibria of LTL-games can be expressed by LTL itself. As a first step in this direction we prove game-theoretic pendant of Theorem 2: if an LTL-game has an *equilibrium run*, it also has an ultimately periodic *equilibrium run*.

PROPOSITION 8. *Let G be an* LTL*-game with the players' goals given by $\gamma_1, \ldots, \gamma_n$. Then, for every run $\rho \in NE(G)$, there is an ultimately periodic run $\rho^* \in NE(G)$ such that $\rho \models \gamma_i$ if and only if $\rho^* \models \gamma_i$ for every player i.*

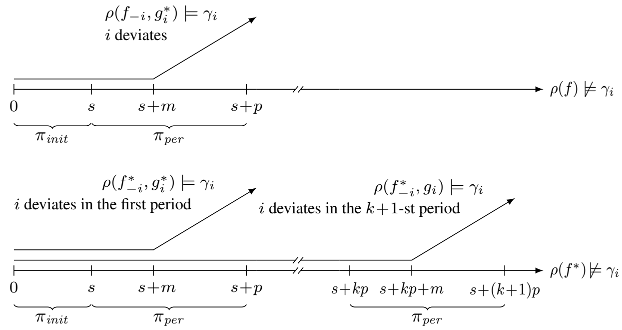

Figure 1: Proof of Proposition 8. For equilibrium f, construct profile f^* such that $\rho(f^*)$ is an ultimately periodic equilibrium run. If we assume that i deviates from f^* in the $k+1$-st period, it follows that i would deviate in the first period as well. As f and f^* coincide on any history with a maximal common prefix in $\pi_{init};\pi_{per}$, it follows that i also deviates from f.

Proof: Let ρ be an equilibrium run of G and let f be a Nash equilibrium sustaining ρ. Let $W = \{i \in N\colon \rho \models \gamma_i\}$ be the set of *winners* at ρ and consider the formula

$$\gamma_\rho = \textstyle\bigwedge_{i\in W}\gamma_i \wedge \bigwedge_{i\in N\setminus W}\neg\gamma_i.$$

Clearly, $\rho \models \gamma_\rho$. By virtue of Theorem 2, there is also an ultimately periodic run $\rho^* = \pi_{init};\pi_{per}^\omega$ with start index s and period p such that $\rho^* \models \gamma_\rho$. Moreover, due to Lemma 4.5 in [29], we may assume for every $t \ge s$ that we have $\rho^*, t \models \gamma_i$ if and only if $\rho^*, t{+}p \models \gamma_i$ for all players i in the game.

We define a strategy profile f^* that induces ρ^* and show that it is a Nash equilibrium of G. Hence, $\rho^* \in NE(G)$. First observe that every history π has a unique maximal common prefix π_{com} with ρ^*, which is either a prefix of π_{init} or has the form $\pi_{init};\pi_{per}^r;\pi_{per'}$, where $r \ge 0$ and $\pi_{per'} \le \pi_{per}$. Now define f^* such that, for each player i and history $\pi = \pi_{com};\pi''$,

$$f_i^*(\pi) = \begin{cases} f_i(\pi_{com};\pi'') & \text{if } \pi_{com} \le \pi_{init}, \\ f_i(\pi_{init};\pi_{per};\pi_{per'};\pi'') & \text{if } \pi_{com} = \pi_{init};\pi_{per}^r;\pi_{per'}. \end{cases}$$

Observe that $\rho(f^*) = \rho^* = \pi_{init};\pi_{per}^\omega$ and, hence, $\rho(f^*) \models \gamma_\rho$.

We prove that f^* is a Nash equilibrium of G (also see Figure 1 for an illustration of the argument). To this end, assume for contradiction that there is a strategy g_i for some player i such that $\rho(f^*) \not\models \gamma_i$ and $\rho(f^*_{-i}, g_i) \models \gamma_i$. Then, $\rho(f^*)$ and $\rho(f^*_{-i}, g_i)$ have a maximal common prefix π', for which either $\pi' \le \pi_{init};\pi_{per}$ or $\pi' = \pi_{init};\pi_{per}^k;\pi_{per'}$ for some $k \ge 1$.

If the former—that is, if i deviates from f^* before the first period π_{per} is completed—we have $\rho(f_{-i}, g_i) = \rho(f^*_{-i}, g_i)$ and can immediately conclude that f is not a Nash equilibrium, a contradiction.

If the latter $\rho(f^*_{-i}, g_i) = \pi_{init};\pi_{per}^k,\rho'$ for some $\rho' \in runs(\Phi)$. Thus, k is the number of times that the period π_{per} is completed before i deviates from f^* by playing g_i. We now define another strategy g_i^* for i on the basis of g_i in a similar way as f^* was constructed from f. For every history $\pi = \pi_{com};\pi''$, let

$$g_i^*(\pi) = \begin{cases} g_i(\pi) & \text{if } \pi_{com} \le \pi_{init}, \\ g_i(\pi_{init};\pi_{per}^{r+k};\pi_{per'};\pi'') & \text{if } \pi_{com} = \pi_{init};\pi_{per}^r;\pi_{per'}. \end{cases}$$

Thus, for every history π and $r = 0$, we have $g_i^*(\pi) = g_i(\pi) = f_i(\pi)$ if $\pi_{com} \le \pi_{init}$, and $g_i^*(\pi_{init};\pi'') = g(\pi_{init};\pi_{per}^k;\pi'')$ if

$\pi_{init} < \pi_{com} \leq \pi_{init};\pi_{per}$. Intuitively, g_i^* behaves in the first period exactly as g_i in the k-th period. Hence, $\rho(f_{-i}^*, g_i^*) = \pi_{init};\pi_{per'};\rho'$. Then, by virtue of Lemma 4.5 in [29] it follows that $\rho(f_{-i}^*, g_i^*) \models \gamma_i$. Moreover, $\rho(f_{-i}^*, g_i^*) = \rho(f_{-i}, g_i^*)$ and hence $\rho(f_{-i}, g_i^*) \models \gamma_i$. It follows that f is *not* a Nash equilibrium and a contradiction ensues, as desired. □

As an immediate consequence we find that every property that can be expressed by LTL in equilibrium can be weakly expressed by LTL itself. By virtue of Proposition 8, every non-empty set of equilibrium runs also contains an ultimately periodic run, which, by Proposition 3, is characterised by an LTL-formula.

COROLLARY 9. LTL $\geq^{\mathsf{w}}$ LTL^{NE}.

For the moment we leave whether LTL $\geq^{\mathsf{p}}$ LTL^{NE} and whether LTL $\geq$ LTL^{NE} as open questions. As a first step to resolve these issues, however, we show that LTL$^{NE} \not\geq L_{\omega\text{-reg}}$. Key to this issue are temporal properties that are *noncounting* (see, e.g., [23]). A property $X \subseteq runs(\Phi)$ is said to be **noncounting** if there exists an $n_0 > 0$ such that, for all $k \geq n_0$, histories π, π', and runs ρ,

$$\pi;\pi^k;\rho \in X \text{ if and only if } \pi;\pi^{k+1};\rho \in X.$$

As a prime example of a *counting* property over $runs(\Phi)$ consider the one defined by the ω-regular expression $(\emptyset;\emptyset)^*;\{p\}^\omega$. Although for every even $k \geq 0$ we have that the run $\emptyset^k;p^\omega$ belongs to this property but the run $\emptyset^{k+1};p^\omega$ does not. Indeed, Kučera and Strejček [19] have characterised the LTL-properties as those that are both ω-regular and non-counting.

THEOREM 10 ([19]). *A property $X \subseteq runs(\Phi)$ can be expressed by* LTL *if and only if X is ω-regular and noncounting.*

We find that for every LTL-game G the set $NE(G)$ of equilibrium runs is noncounting. As the proof of this result runs along similar lines as the one for Proposition 8, here we only give a sketch.

PROPOSITION 11. *For every* LTL*-game G, the set $NE(G)$ of equilibrium runs is noncounting.*

Sketch of proof: Consider an arbitrary LTL-game G. For every run $\rho \in NE(G)$ there is a subset $W \subseteq N$ such that ρ satisfies $\gamma_W = \bigwedge_{i\in W}\gamma_i \wedge \bigwedge_{i\in N\setminus W}\neg\gamma_i$. Moreover, by Theorem 10, there exists an $n_{\gamma_W} > 0$ by virtue of which $[\![\gamma_W]\!]$ is noncounting. At this point, consider $n_0 = \max_{W\subseteq N}\{n_{\gamma_W}\}$. The proof proceeds by showing that, for all $k \geq n_0$, histories π, π', and runs ρ,

$$\pi;\pi^k;\rho \in NE(G) \text{ if and only if } \pi;\pi^{k+1};\rho \in NE(G).$$

For the "only if"-direction, we may assume the existence of a Nash equilibrium f that sustains $\pi;\pi^k;\rho$ and on its basis construct a Nash equilibrium f^* that induces $\pi;\pi^{k+1};\rho \in NE(G)$. The definition of f^* is similar to the construction in Proposition 8. The "if"-direction is proven analogously. □

Thus, in particular, the ω-regular property $(\emptyset;\emptyset)^*;p^\omega$ can not be obtained as the set of equilibrium runs of any LTL-game. As an immediate consequence of Proposition 11 we thus obtain that

$$\text{LTL}^{NE} \not\geq L_{\omega\text{-reg}}.$$

In view of the characterisation result by Kučera and Strejček, Proposition 11 gives us one half of the proof that LTL $\geq$ LTL^{NE}. Indeed, it remains to be shown that the set of equilibrium runs of every LTL-game is an ω-regular set. We leave it here as an open question to be investigated in a future work.

3.2 The Next-Free Fragment

In this section we consider the fragment L_{U}, which does not contain the "next"-operator X, and show that we can characterise *every* LTL-property as the set of runs sustained by a Nash equilibrium in some L_{U}-game. For our construction we need at least two players and six additional propositional variables on which the LTL-property does not depend.

In [20] it is made a case for so-called stutter-invariant specifications.[3] Formally, we say that a temporal property $X \subseteq runs(\Phi)$ is **stutter-invariant** if, for all runs $\rho = v_0, v_1, v_2, \ldots$ and every sequence $k_0, k_1, k_2, \ldots$ of positive integers,

$$v_0, v_1, v_2, \ldots \in X \quad \text{if and only if} \quad v_0^{k_0}, v_1^{k_1}, v_2^{k_2}, \ldots \in X$$

where v^k denotes the k-fold iteration of v. Thus, for instance, the property defined by $(p;\bar{p})^\omega$, henceforth denoted by $toggle(p)$, does *not* define a stutter-free property, as it contains $p, \bar{p}, p, \bar{p}, \ldots$, but not $p, p, \bar{p}, \bar{p}, p, p, \bar{p}, \bar{p}, \ldots$. Peled and Wilke [25], furthermore, showed that L_{U} is the *largest* stutter-free fragment of LTL.

As $toggle(p)$ is expressed by the LTL-formula $p \wedge \mathsf{G}(p \leftrightarrow \mathsf{X}\bar{p})$, we immediately obtain $L_{\mathsf{U}} \not\geq$ LTL as well as $L_{\mathsf{U}} \not\geq^{\mathsf{w}}$ LTL. Some reflection reveals moreover that for every $\Psi \subseteq \Phi$, if X is a stutter-invariant property on $runs(\Phi)$, then so is $X|_\Psi$, that is, property X projected to Ψ. Hence, L_{U} does not even projectively express LTL:

$$L_{\mathsf{U}} \not\geq^{\mathsf{p}} \text{LTL}.$$

We find, however, that property $toggle(p)$ can be projectively expressed in equilibrium. To see this, consider the two-player game G_{toggle} with two players, i and j, who control $\Phi_i = \{p, s\}$ and $\Phi_j = \{q, r\}$, respectively, and whose goals are given by:

$$\gamma_i = p \wedge (r \leftrightarrow \bar{s}),$$
$$\gamma_j = \mathsf{F}((pq \wedge (pq \,\mathsf{U}\, p\bar{q})) \vee (\bar{p}q \wedge (\bar{p}q \,\mathsf{U}\, \bar{p}\bar{q}))) \wedge (r \leftrightarrow s).$$

Intuitively, player j's goal can only be satisfied if p is true at two subsequent time points in the future. But even if this were so, j still has to win a matching pennies game on r and s against player i. As player i would then deviate by choosing appropriate values for p and s, player j will not achieve its goal in any equilibrium. This being established, the first conjunct of γ_j will not be satisfied in any equilibrium either. Otherwise, j would deviate and win the matching pennies game on r and s and thus achieve its goal. Therefore, in no equilibrium will p subsequently assume the same truth-value. On the other hand, if p is initially true and then toggles truth-values indefinitely, there is no way player j can deviate and get its goal achieved. Accordingly, he can just as well let player i win the matching pennies game on r and s. Player i will then achieve its goal and have no incentive to deviate either.

LEMMA 12. *Let $\rho \in NE(G_{toggle})$. Then, $\rho, t \models p$ if and only if t is even and, hence, $\{\rho|_{\{p\}} : \rho \in NE(G_{toggle})\} = toggle(p)$.*

Lemma 12 not only shows that L_{U}^{NE} can projectively express $(p;\bar{p})^\omega$ in equilibrium—and thus that $L_{\mathsf{U}}^{NE} > L_{\mathsf{U}}$—it can also be leveraged to prove that the next-free fragment can actually projectively express *every* LTL-property. To this end, let $p \notin \Phi$. We then

[3] On this issue Peled and Wilke [25] write "A specification that is not invariant under stuttering will not allow refinement and will thus be useless for hierarchical (modular) reasoning." Moreover, stutter-invariance enables state space reductions in model checking of concurrent systems (also see [25]).

define a *translation* $\tau\colon \mathrm{LTL}(\Phi) \to L_{\mathsf{U}}(\Phi \cup \{p\})$ such that,

$$\begin{aligned} q^\tau &= q \\ (\neg\varphi)^\tau &= \neg(\varphi^\tau) \\ (\varphi \wedge \psi)^\tau &= \varphi^\tau \wedge \psi^\tau \\ (\varphi \vee \psi)^\tau &= \varphi^\tau \vee \psi^\tau \\ (\varphi \mathbin{\mathsf{U}} \psi)^\tau &= \varphi^\tau \mathbin{\mathsf{U}} \psi^\tau \\ (\mathsf{X}\,\varphi)^\tau &= (p \to p \mathbin{\mathsf{U}} (\bar{p} \wedge \varphi^\tau)) \wedge (\bar{p} \to \bar{p} \mathbin{\mathsf{U}} (p \wedge \varphi^\tau)), \end{aligned}$$

where $q \in \Phi$. Thus, on every run $\rho \in \Phi \cup \{p\}$ with $\rho|_{\{p\}} = (p;\bar{p})^\omega$, each formula φ and its translation φ^τ will have the same truth-value.

LEMMA 13. *Let $\varphi \in \mathrm{LTL}(\Phi)$, p a propositional variable not in Φ, and $\rho = v_0, v_1, v_2, \dots$ a run in $runs(\Phi \cup \{p\})$ such that $\rho, t \models p$ if and only if t is even. Then, $\rho \models \varphi$ if and only if $\rho \models \varphi^\tau$.*

Proof: The proof proceeds by structural induction on φ. The basis is immediate and the induction hypothesis covers all inductive cases apart from $\varphi = \mathsf{X}\,\psi$. Consider an arbitrary $t \geq 0$ and assume $\rho, t \models \mathsf{X}\,\psi$. Then, $\rho, t+1 \models \psi$ and by the induction hypothesis also $\rho, t+1 \models \psi^\tau$. Now either $\rho, t \models p$ or $\rho, t \models \bar{p}$. First assume the former. Then immediately $\rho, t \models \bar{p} \to \bar{p} \mathbin{\mathsf{U}} (p \wedge \psi^\tau)$. Moreover, $\rho, t+1 \models \bar{p}$ by definition of ρ and, therefore, $\rho, t \models p \mathbin{\mathsf{U}} (\bar{p} \wedge \psi^\tau)$ and also $\rho, t \models p \to (p \mathbin{\mathsf{U}} (\bar{p} \wedge \psi^\tau))$. We may conclude that $\rho, t \models (\mathsf{X}\,\psi)^\tau$, as desired. The argument if $\rho, t \models \bar{p}$ is analogous.

For the opposite direction, assume $\rho, t \not\models \mathsf{X}\,\psi$. Then, $\rho, t+1 \not\models \psi$ and by the induction hypothesis $\rho, t+1 \not\models \psi^\tau$. Now, either $\rho, t \models p$ or $\rho, t \models \bar{p}$. If the former, both $\rho, t \not\models \bar{p} \wedge \psi^\tau$ and $\rho, t+1 \not\models \bar{p} \wedge \psi^\tau$. It follows that $\rho, t \not\models p \mathbin{\mathsf{U}} (\bar{p} \wedge \psi^\tau)$, $\rho, t \not\models p \to (p \mathbin{\mathsf{U}} (\bar{p} \wedge \psi^\tau))$, and eventually $\rho, t \not\models (\mathsf{X}\,\psi)^\tau$. As the argument if $\rho, t \models \bar{p}$ is analogous, we may conclude the proof. □

To obtain the main result of this section, we construct for each LTL-formula φ a four-player L_{U}-game with six additional variables. Intuitively, two players play the game $G^{mp}_{\varphi^\tau}$ ensuring that φ^τ holds at precisely the equilibrium runs. Two other players play G_{toggle} ensuring that an additional variable p alternately assumes the truth values true and false, and φ^τ evaluates as intended.

THEOREM 14. $L^{NE}_{\mathsf{U}} \geq^{\mathsf{p}} \mathrm{LTL}$.

Proof: Let $\varphi \in \mathrm{LTL}(\Phi)$ and $\Psi = \{p, q, r, s, x, y\}$ a set of auxiliary variables disjoint from Φ. Now construct L_{U}-game G on $\Phi \cup \Psi$ with four players, 1, 2, 3, and 4, such that $\Phi_1 = \{p, s\}$, $\Phi_2 = \{q, r\}$, $\Phi_3 = \Phi \cup \{x\}$, and $\Phi_4 = \{y\}$. Let the players' goals, moreover, be given by:

$$\begin{aligned} \gamma_1 &= p \wedge (r \leftrightarrow \bar{s}), \\ \gamma_2 &= \mathsf{F}((pq \wedge (pq \mathbin{\mathsf{U}} p\bar{q})) \vee (\bar{p}q \wedge (\bar{p}q \mathbin{\mathsf{U}} \bar{p}\bar{q}))) \wedge (r \leftrightarrow s), \\ \gamma_3 &= \varphi^\tau \vee (x \leftrightarrow y), \\ \gamma_4 &= \varphi^\tau \vee (x \leftrightarrow \bar{y}). \end{aligned}$$

Thus, players 1 and 2 play G_{toggle} and Lemma 12 ensures that in every equilibrium run $\rho \in \Phi \cup \Psi$, $\rho, t \models p$ if and only if t is even. Players 3 and 4—quite independently from 1 and 2—play $G^{mp}_{\varphi^\tau}$. Thus, $\rho \models \varphi^\tau$ if and only if ρ is an equilibrium run of G. Lemma 13 then yields the result. □

The following result is then an immediate consequence of Theorem 14, Proposition 4, and transitivity of $\geq^{\mathsf{p}}$.

COROLLARY 15. $L^{NE}_{\mathsf{U}} \geq^{\mathsf{p}} L_{\omega\text{-}reg}$.

It is worth noting that the size of φ^τ is exponential in the number of nestings of the X-operator, that is, even if L_{U} can projectively express every LTL-property in equilibrium, this may come at the cost of having exponentially longer goals for the players. Whether this exponential blowup is inevitable, we leave as an open question.

3.3 The Positive Future Fragment

We now consider the fragment $L_{\mathsf{X},\mathsf{F}^+}$, where the F-operator cannot occur within the scope of a negation. We find that also this, rather weak, fragment can express more in equilibrium than it can by itself. First we have the following lemma, which intuitively says that every $L_{\mathsf{X}\,\mathsf{F}^+}$-formula, when satisfied, will be satisfied after a finite number of rounds. Formally, a temporal property $X \subseteq runs(\Phi)$ is **tail-invariant** if $\rho \in X$ implies the existence of a prefix $\pi \in prefix(\rho)$ such that $\pi; \rho' \in X$ for all $\rho' \in runs(\Phi)$.

LEMMA 16. *Every temporal property $X \subseteq runs(\Phi)$ that can be expressed in $L_{\mathsf{X},\mathsf{F}^+}$ is tail-invariant.*

Sketch of proof: Let $\varphi \in L_{\mathsf{X},\mathsf{F}}$. As the F-operator occurs within the scope of a negation symbol $\neg$, exploiting the equivalence of $\neg\,\mathsf{X}\,\varphi$ and $\mathsf{X}\,\neg\varphi$, we can transform φ to an equivalent formula in which all negation symbols occur in front of propositional variables.

Let $\varphi \in L_{\mathsf{X},\mathsf{F}^+}$ and assume that φ is in this normal form. Consider an arbitrary run $\rho = v_0, v_1, \dots$ in $runs(\Phi)$ such that $\rho \models \varphi$. We have to show that there is a prefix $\pi \in prefix(\rho)$ such that $\pi; \rho' \models \varphi$ for all $\rho' \in runs(\Phi)$. Define inductively for every formula $\psi \in L_{\mathsf{X},\mathsf{F}^+}$ and every $t \geq 0$, $\kappa_{\rho,t}(\psi)$ the following:

$$\begin{aligned} \kappa_{\rho,t}(p) &= \kappa_{\rho,t}(\bar{p}) = 0 \\ \kappa_{\rho,t}(\chi_1 \wedge \chi_2) &= \kappa_{\rho,t}(\chi_1 \vee \chi_2) = \max(\kappa_{\rho,t}(\chi_1), \kappa_{\rho,t}(\chi_2)) \\ \kappa_{\rho,t}(\mathsf{X}\,\chi) &= \kappa_{\rho,t}(\chi) + 1 \\ \kappa_{\rho,t}(\mathsf{F}\,\chi) &= \begin{cases} t' - t + \kappa_{\rho,t'}(\chi) & \text{if } \rho, t \models \mathsf{F}\,\chi, \\ 0 & \text{otherwise,} \end{cases} \end{aligned}$$

where $t' = \min\{t'' \geq t \colon \rho, t'' \models \chi\}$. By a structural induction on ψ it can then be shown that $\rho, t \models \psi$ implies the existence of a prefix $\pi \in prefix(\rho)$ with $length(\pi) \leq t + \kappa_{\rho,t}(\psi)$ such that $\pi_\psi; \rho', t \models \psi$ for all $\rho' \in runs(\Phi)$. This holds in particular for φ, which yields the result. □

The property defined by the LTL-formula $\mathsf{G}\,p$ may serve as the quintessential property that is *not* tail-invariant, and is neither expressible nor weakly expressible in $L_{\mathsf{X},\mathsf{F}^+}$. Observing that tail-invariance of $X \subseteq runs(\Phi)$ implies tail-invariance of $X|_\Psi$ for every $\Psi \subseteq \Phi$, we may even conclude that $\mathsf{G}\,p$ cannot even be projectively expressed by $L_{\mathsf{X},\mathsf{F}^+}$. Yet, as $\mathsf{G}\,p$ is LTL-equivalent to $\neg\,\mathsf{F}\,\bar{p}$, in virtue of Lemma 6, we find that $L_{\mathsf{X},\mathsf{F}^+}$ can projectively express $\mathsf{G}\,p$ in equilibrium. Hence,

$$L^{NE}_{\mathsf{X},\mathsf{F}^+} >^{\mathsf{p}} L_{\mathsf{X},\mathsf{F}^+}.$$

Leveraging the same ideas along with the fact that every ultimately periodic run can be characterised in $L_{\mathsf{X},\mathsf{F},\mathsf{G}}$ with only one occurrence of the G-operator, we also obtain the following expressiveness result for $L_{\mathsf{X},\mathsf{F}^+}$ with respect to LTL.

PROPOSITION 17. $L_{\mathsf{X},\mathsf{F}^+} \geq^{\mathsf{wp}} \mathrm{LTL}$.

Proof: Let $\varphi \in \mathrm{LTL}$. If φ is unsatisfiable, Lemma 6 yields the result immediately as $p \vee \bar{p}$ is a formula in $L_{\mathsf{X},\mathsf{F}^+}$. On the other hand, if φ is satisfiable then by Theorem 2 and Proposition 3 there is an ultimately periodic run $\rho = v_0, v_1, \dots$ with starting index s and period p such that is characterised by the LTL-formula χ_ρ given by

$$\textstyle\bigwedge_{0 \leq t \leq s+p} \mathsf{X}^t \chi_{v_t} \wedge \mathsf{X}^s\,\mathsf{G}(\bigwedge_{v \in 2^\Phi} (\chi_v \to \mathsf{X}^p \chi_v)).$$

By suitably applying the laws of propositional logic, the duality of F and G, as well as the equivalence of $\neg\,\mathsf{X}\,\varphi$ and $\mathsf{X}\,\neg\varphi$, we find that the *negation* of χ_ρ is equivalent to

$$\textstyle\bigvee_{0 \leq t \leq s+p} \mathsf{X}^i \neg\chi_{v_i} \vee \mathsf{X}^s\,\mathsf{F}(\bigvee_{v \in 2^\Phi} (\chi_v \wedge \mathsf{X}^p \neg\chi_v)),$$

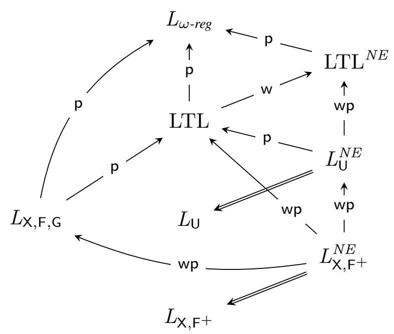

Figure 2: Relation between expressiveness and expressiveness in equilibrium of LTL**-fragments assuming** $|\Phi| \geq 2$**. Double arrows relate to strict expressiveness (**$L_1 > L_2$**).**

which is included in the fragment $L_{\mathsf{X,F+}}$. By Lemma 6 we know that $L_{\mathsf{X,F+}}$ can projectively express χ_ρ in equilibrium. Hence, $L_{\mathsf{X,F+}}$ can weakly projectively express φ in equilibrium, as desired. $\square$

4. RELATED WORK

The expressive power of LTL and many of its syntactic fragments has been a research topic for decades, with work showing connections with other temporal logic languages as well as results classifying the power of LTL and its fragments [27]. The most basic and well known classifications are with respect to sub-languages, that is, LTL fragments where only some operators are allowed. However, more refined studies have also been conducted, for instance, LTL fragments with respect to the allowed number of propositional variables or the number of nested temporal operators [9].

Most of these studies have focused not only in the expressive power of the resulting sublogics but also in the implications of imposing such restrictions in the complexity of the model checking and satisfiability problems of such sublogics. These studies have also made it possible to understand connections between LTL fragments and standard automata models over infinite words—which in turn also easily show how to define different automata-theoretic decision procedures for each LTL sublanguage at hand [32].

Despite the very many studies about the expressive power of LTL and related sublanguages, to the best of our knowledge, there are no results on the expressive power of LTL or its fragments with respect to the classes of runs that can be sustained by some Nash equilibrium. In this paper, we study precisely that issue and provide the first known results in the literature. The results are rather promising: they show that even though some LTL sublanguage, say L_1, may be strictly more expressive than other LTL sublanguage, say L_2, when interpreted over the full class of ω-regular runs, such two sublogics L_1 and L_2 become equi-expressive when interpreted over a class of runs that can be sustained by some Nash equilibrium in a given class of games, as many of our results show.

As this kind of result can usually only be obtained by adding extra propositional variables to the "weaker" language, we also studied the expressive power, and game-theoretic implications, of allowing languages interpreted over different sets of propositional variables (projective expressiveness). Again, the results were promising in the sense that they show that generally weaker LTL sublogics can be made as expressive as generally stronger LTL sublogics by the addition of fresh propositional variables to the weaker language, a notion that goes back to the 1950s [3, 8] and has proven useful in a number of settings (see, for instance, [18, 6, 13, 16]).

5. CONCLUSION

In this paper, we have explored the temporal properties that are characterised by the equilibrium runs of iterated Boolean games, where the players' dichotomous preferences are represented by formulae in fragments of Linear Time Logic (LTL). The Nash equilibria of an iterated Boolean game are fully determined by the goals of the players and the way control of the propositional variables is distributed over the players. In particular, they are not dependent on an additional underlying game structure—like, for instance, *concurrent game structures* (see, for instance, [1]). This enabled us to focus on the *logical* aspects of Nash equilibrium and accordingly we formulated our research issue in terms of expressiveness.

We investigated the concept of expressiveness in equilibrium for a number of fragments of LTL. We found that for a given fragment every (non-trivial) property that can be (projectively) expressed can also be (projectively) expressed in equilibrium, but not generally the other way round. For an overview of our results see Figure 2. In the future, we plan to explore more fully the links that remain missing. In particular, whether $\mathrm{LTL} \geq \mathrm{LTL}^{NE}$ is still open. Other questions for future research concern the minimal number of players and additional propositional variables that may be needed to express or projectively express temporal properties in a particular fragment.

Apart from specific fragments, the concept of *expressiveness in equilibrium* gives rise to a number of more abstract and conceptual questions. First, most of our game constructions involve a "matching pennies" game like G^{mp}_φ. These games establish a crucial link between runs that satisfy a given formula and equilibrium runs in iterated Boolean games. This feature, however, is due to one player trying to achieve $p \leftrightarrow q$ and another $p \leftrightarrow \bar{q}$, and as such is largely of a non-temporal nature. An interesting question is if this is peculiar to the results in this paper or that it points at a more fundamental connection with the concept of expressiveness in equilibrium.

Second, our investigations in a considerable measure relied on projective expressiveness, as our game constructions usually involve additional propositional variables. Expressiveness in equilibrium does not afford such freedom and one would suspect a narrower gap between expressiveness and expressiveness in equilibrium that between the respective projective notions for most fragments. An important question is how narrow this gap is.

Third, as in most work on model checking and concurrent game structures, the strategies of the players are functions from histories to choices whereas, in stark contrast, the preferences of the players are represented by logical formulae. As such the former are much finer grained than the latter. One may wonder to what extent the gap between the standard notion of expressiveness and expressiveness in equilibrium can be attributed to this "mismatch." For instance, would the phenomenon also occur in iterated Boolean games, if the players' strategies are similarly specified by temporal formulae?

Fourth, all our notions of expressiveness naturally extend to game-theoretic solution concepts other than Nash Equilibrium and we aim to investigate some of these concepts in future work.

Finally, the original motivation for our work was the problem of allocating tasks to agents so that a given property is satisfied in all resulting equilibria. We hope that our research on expressiveness in equilibrium shed some light on this issue and to what extent the design of multi-agent systems in practice can be simplified by providing "lean" temporal specifications of individual agents who can be assumed to play strategies that together form an equilibrium.

Acknowledgments

The authors are financially supported by the ERC under Advanced Investigator Grant 291528 ("RACE") at Oxford.

REFERENCES

[1] R. Alur, T. A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49(5):672–713, 2002.

[2] C. Baier and J.-P. Katoen. *Principles of Model Checking*. The MIT Press, 2008.

[3] E. W. Beth. On Padoa's method in the theory of definition. *Indagationes Mathematicae*, 15:330–339, 1953.

[4] E. Bonzon, M. Lagasquie, J. Lang, and B. Zanuttini. Boolean games revisited. In *Proceedings of the Seventeenth European Conference on Artificial Intelligence (ECAI-2006)*, 2006.

[5] N. Bulling, W. Jamroga, and J. Dix. Reasoning about temporal properties of rational play. *Annals of Mathematics and Artificial Intelligence*, 53(1):51–114, 2008.

[6] C. C. Chang and H. J. Keisler. *Model Theory*, volume 73 of *Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Company, third edition, 1990.

[7] E. M. Clarke, O. Grumberg, and D. A. Peled. *Model Checking*. The MIT Press: Cambridge, MA, 2000.

[8] W. Craig. Three uses of the Herbrand-Gentzen theorem in relating model theory and proof theory. *Journal of Symbolic Logic*, 22(3):269–285, 1957.

[9] S. Demri and P. Schnoebelen. The complexity of propositional linear temporal logics in simple cases. *Information and Computation*, 174(1):84–103, 2002.

[10] E. A. Emerson. Temporal and modal logic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science Volume B: Formal Models and Semantics*, pages 996–1072. Elsevier, 1990.

[11] E. A. Emerson and J. Y. Halpern. 'Sometimes' and 'not never' revisited: on branching time versus linear time temporal logic. *Journal of the ACM*, 33(1):151–178, 1986.

[12] U. Endriss, S. Kraus, J. Lang, and M. Wooldridge. Designing incentives for Boolean games. In R. Turner, Yolum, E. A. Sonenberg, and Stone, editors, *10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011)*, volume 1, pages 79–86, 2011.

[13] A. P. Gheerbrant and B. ten Cate. Craig interpolation for linear temporal languages. In E. Grädel and R. Kahle, editors, *Proceedings of 23rd international Workshop, CSL 2009, 18th Annual Conference of the EACSL, Coimbra, Portugal, September 7-11, 2009*, volume 5771 of *Lecture Notes in Computer Science*, pages 287–301. Springer, 2009.

[14] J. Gutierrez, P. Harrenstein, and M. Wooldridge. Reasoning about equilibria in game-like concurrent systems. In *14th Conference on Principles of Knowledge Representation and Reasoning (KR-2014)*, pages 408–417, 2014.

[15] J. Gutierrez, P. Harrenstein, and M. Wooldridge. Iterated Boolean games. *Information and Computation*, 242:53–79, 2015.

[16] J. Y. Halpern, D. Samet, and E. Segev. On definability in multimodal logic. *The Review of Symbolic Logic*, 2(3):451–468, 2009.

[17] P. Harrenstein, W. van der Hoek, J.-J. Meyer, and C. Witteveen. Boolean games. In J. van Benthem, editor, *Proceeding of the Eighth Conference on Theoretical Aspects of Rationality and Knowledge (TARK VIII)*, pages 287–298, 2001.

[18] W. Hodges. *Model Theory*. Cambridge University Press, 1993.

[19] A. Kučera and J. Strejček. The stuttering principle revisited. *Acta Informatica*, 41(7–8):415–434, 2005.

[20] L. Lamport. Specifying concurrent program modules. *ACM Transactions on Programming Languages and Systems*, 5(2):190–222, 1983.

[21] R. D. Luce and H. Raiffa. *Games and Decisions*. John Wiley & Sons, 1957.

[22] M. Maschler, E. Solan, and S. Zamir. *Game Theory*. Cambridge U.P., 2013.

[23] A. Mateescu and A. Salomaa. Formal languages: An introduction and a synopsis. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume I: Word, Language, Grammar, chapter 1. Springer, 1997.

[24] M. J. Osborne and A. Rubinstein. *A Course in Game Theory*. The MIT Press: Cambridge, MA, 1994.

[25] D. A. Peled and T. Wilke. Stutter-invariant temporal properties are expressible without next-time operator. *Information Processing Letters*, 63(5):243–246, 1997.

[26] A. Pnueli. The temporal logic of programs. In *Proceedings of the Eighteenth IEEE Symposium on the Foundations of Computer Science*, pages 46–57, 1977.

[27] A. M. Rabinovich. Expressive power of temporal logics. In L. Brim, P. Jancar, M. Kretínský, and A. Kucera, editors, *CONCUR 2002—Concurrency Theory*, volume 2421 of *Lecture Notes in Computer Science*, pages 57–75. Springer, 2002.

[28] Y. Shoham and K. Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press: Cambridge, England, 2008.

[29] A. P. Sistla and E. M. Clarke. The complexity of propositional linear temporal logics. *Journal of the ACM*, 32(3):733–749, 1985.

[30] W. Thomas. Languages, automata, and logic. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume III: Beyond works, chapter 5. Springer, 1997.

[31] M. Y. Vardi. Branching vs. linear time: Final showdown. In T. Margaria and W. Yi, editors, *Proceedings of the 2001 Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2001* volume 2031 of *Lecture Notes in Computer Science*, pages 1–22. Springer, 2001.

[32] M. Y. Vardi and P. Wolper. Reasoning about infinite computations. *Information and Computation*, 115(1):1–37, 1994.

[33] P. Wolper. Temporal logic can be more expressive. *Information and Control*, 56, 1983.