Expressiveness and Nash Equilibrium in Iterated Boolean Games

JULIAN GUTIERREZ, Monash University, Australia
PAUL HARRENSTEIN, University of Oxford, UK
GIUSEPPE PERELLI, Sapienza University of Rome, Italy
MICHAEL WOOLDRIDGE, University of Oxford, UK

We define and investigate a novel notion of expressiveness for temporal logics that is based on game theoretic equilibria of multi-agent systems. We use *iterated Boolean games* as our abstract model of multi-agent systems [Gutierrez et al. 2013, 2015a]. In such a game, 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, in the sense that the models of γ_i represent system behaviours that would satisfy i. Each player controls a subset of Boolean variables Φ_i, and at each round in the game, player i is at liberty to choose values for variables Φ_i in any way that she sees fit. Play continues for an infinite sequence of rounds, and so as players act they collectively trace out a model for LTL, which for every player will either satisfy or fail to satisfy their goal. Players are assumed to act strategically, taking into account the goals of other players, in an attempt to bring about computations satisfying their goal. In this setting, we apply the standard game-theoretic concept of (pure) Nash equilibria. The (possibly empty) set of Nash equilibria of an iterated Boolean game can be understood as inducing a set of computations, each computation representing one way the system could evolve if players chose strategies that together constitute a Nash equilibrium. Such a set of equilibrium computations expresses a temporal property—which may or may not be expressible within a particular LTL fragment. The new notion of expressiveness that we formally define and investigate is then as follows: *What temporal properties are characterised by the Nash equilibria of games in which agent goals are expressed in specific fragments of* LTL? We formally define and investigate this notion of expressiveness for a range of LTL fragments. For example, a very natural question is the following: *Suppose we have an iterated Boolean game in which every goal is represented using a particular fragment L of* LTL*: is it then always the case that the equilibria of the game can be characterised within L?* We show that this is not true in general.

This is an extensively revised and reorganised version of our paper that was presented at the *15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16)* [Gutierrez et al. 2016]. All authors acknowledge with gratitude the financial support of ERC Advanced Investigator Grant 291528 ("RACE") at the University of Oxford. Paul Harrenstein was also supported in part by ERC Starting Grant 639945 ("ACCORD") also at the University of Oxford. Michael Wooldridge and Paul Harrenstein furthermore acknowledge the financial support of the Alan Turing Institute in London. Giuseppe Perelli acknowledges the support of the project ERC Advanced Grant 834228 ("WhiteMech") and the EU ICT-48 2020 project TAILOR (No. 952215).
Authors' addresses: J. Gutierrez, Department of Data Science and AI. 26 Ancora Imparo Way, Clayton campus, Melbourne, Australia; email: julian.gutierrez@monash.edu; P. Harrenstein, Department of Computer Science, University of Oxford. Wolfson Building, Parks Road, Oxford OX1 3QD, UK; email: paul.harrenstein@cs.ox.ac.uk; G. Perelli, Department of Computer, Automation, and Business Engineering. Via Ariosto, 25, Rome, Italy; email: perelli@diag.uniroma1.it; M. Woldridge, Department of Computer Science, University of Oxford. Wolfson Building, Parks Road, Oxford OX1 3QD, UK; email: michael.wooldridge@cs.ox.ac.uk.

1529-3785/2021/06-ART8 $15.00
https://doi.org/10.1145/3439900

CCS Concepts: • **Theory of computation** → **Logic**; • **Computing methodologies** → *Artificial intelligence*; *Knowledge representation and reasoning*; **Multi-agent systems**;

Additional Key Words and Phrases: Logic, game theory, Nash equilibrium, concurrent games, multi-agent systems, expressiveness

ACM Reference format:
Julian Gutierrez, Paul Harrenstein, Giuseppe Perelli, and Michael Wooldridge. 2021. Expressiveness and Nash Equilibrium in Iterated Boolean Games. *ACM Trans. Comput. Logic* 22, 2, Article 8 (June 2021), 38 pages.
https://doi.org/10.1145/3439900

1 INTRODUCTION

Temporal logics are probably the most successful and widely used class of formalisms for the specification and verification of computer systems [Emerson 1990]. 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 [Clarke et al. 2000]. 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, and still generates some debate to the present day [Emerson and Halpern 1986; Vardi 2001].

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 [Gutierrez et al. 2013, 2015a]. In this model, each agent 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) [Pnueli 1977; Emerson 1990]. 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 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. For this setting we use the standard game-theoretic concept of Nash equilibrium [Osborne and Rubinstein 1994]: the Nash equilibria of an iterated Boolean game can be understood as characterising 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 article is as follows: 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 property is expressed by the equilibria of* G? In particular, it is very natural to ask whether the equilibria of a game with the players' goals given by formulas in L can be characterised within L itself. We formally define and investigate this novel notion of expressiveness, which we refer to as *expressiveness in equilibrium*. We do this for a range of known fragments of LTL, in particular, the maximal stutter-invariant fragment without a *next*-operator.

The problem of reasoning about 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* for example by adding new operators to existing temporal logics [Bulling et al. 2008; Gutierrez et al. 2014, 2017a]. Alternatively, one might use a very general formalism such as Strategy Logic to reason about equilibrium properties [Chatterjee et al. 2010]. Our approach—focussing on the temporal properties that Nash equilibrium can distinguish in logic-based games—is fundamentally different. In the present work, we are not primarily concerned with questions such as whether a particular LTL formula holds on some or all Nash equilibrium computations. Rather, we consider the temporal properties of the runs induced by the equilibria of a game in which the players' goals are specified using specific fragments of LTL.

As a motivating example, consider the following temporal variation of the well-known *Battle of the Sexes* game [Luce and Raiffa 1957], which we will refer to as *Boolean Ballet.*

Example 1.1 (Boolean Ballet). Suppose that two friends, whom we denote by i and j, go out every weekend. Both are dance aficionados and each weekend they have to decide separately whether to go to the ballet or to the discotheque, leading to a sequence of evenings going out. We say they go out together, if they decide to go to the same venue. Assume furthermore that i wishes always to go out together, be it to the ballet or to the disco, whereas j wants to go out to the ballet with i sometimes, but also wants to go to the disco alone at some other occasions.

In this game, the sequence in which i and j go to the ballet the first weekend and to the disco ever after is sustained by a Nash equilibrium, but the sequence where they go to the ballet the first *two* weekends and to the disco ever after is not. As a matter of fact, the Nash equilibria of this game give rise to precisely those sequences of evenings out where i and j always go out together, but to the ballet at most once.

We find that this situation can be modelled conveniently as an iterated Boolean game, for instance, by giving i and j each control over a propositional variable, p and q, respectively, and assuming that by setting their variable to true at a particular time, the respective friend goes to the ballet and to the disco otherwise. The two players' (binary) preferences can then be expressed using the temporal operators F ("eventually") and G ("always") only. Player i's goal could for instance be represented by the LTL-formula $\mathsf{G}(p \leftrightarrow q)$ and player j's by $\mathsf{F}(p \wedge q) \wedge \mathsf{F}(p \wedge \neg q)$. Also see Figure 1.

A salient feature of Boolean Ballet is that the players' goals are invariant under the repetition of evenings going out, whereas the set of sequences the Nash equilibria give rise to is not. This means, for instance, that i is just as satisfied when the friends go out together to the ballet and the disco on alternate weekends, as i would be if they were to go out, alternately, to the ballet two weekends in a row and then to the disco another two weekends in a row. By contrast, the two friends going out together to the ballet the first weekend and to the disco ever after is sustained by a Nash equilibrium, but going out to the ballet the first *two* weekends and to the disco every after is not! This phenomenon indicates that, for a given fragment of LTL, the traditional concept of expressiveness for temporal logics can be quite different from our notion of expressiveness in equilibrium. We will therefore explore in detail the issue of which temporal properties are characterised by the Nash equilibria of iterated Boolean games with the players' preferences formulated in various fragments of linear temporal logic.

Our work also has a natural bearing on settings where a designer uses temporal logic to specify the desired behaviour of a system. Consider, for instance, a situation in which a designer is to design a multi-agent system that is to behave accordingly to a given complex specification, and she has to distribute over several agents the complex task represented by the specification. Moreover, these agents may possess only limited computational capabilities, and the designer may want to allocate them tasks that are as simple or computationally light as possible. In our setting, this translates to the designer's specification possibly being formulated in an expressive but possibly computationally expensive fragment of LTL, and the agents having to be assigned objectives phrased in weaker fragments with better computational properties. The designer's task is then to find such "weak" goals for the agents so that the specification is satisfied in those Nash

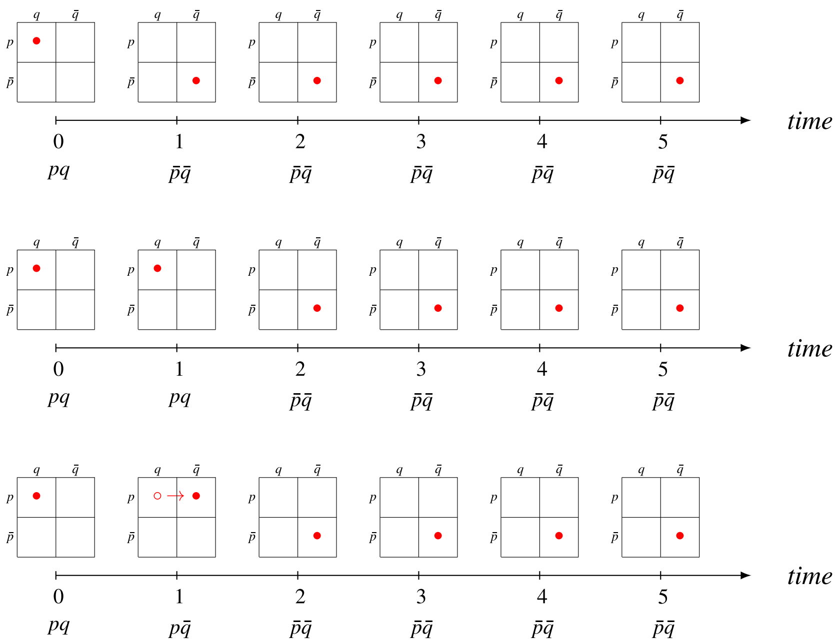

Fig. 1. Three plays of Boolean Ballet, where, at each time (weekend), the red dot indicates how the friends have decided which venue (p and q for ballet, and $\bar{p}$, respectively, $\bar{q}$ for discotheque) to go to. The first run is sustained by a Nash equilibrium where i's goal is satisfied and j cannot beneficially deviate. The second run is not sustained by a Nash equilibrium. To see this, observe that player j would like to deviate from any strategy profile underlying this run by playing along at time 0, only to sneak off to the discotheque in the second, leaving i alone stranded at the ballet. This is indicated by the third play.

equilibrium computations of the multi-agent system that satisfy the specification. The following simple example illustrates this idea.

Example 1.2 (The Rabbit Hunt). Consider the following coordination task for two agents, i and j who together have to catch a rabbit in the dark. One agent holds a torch and the other one has a gun. At each time, the first agent, i, can light the torch (p) or not ($\neg p$), and the second agent, j, can fire the gun q or refrain from doing so ($\neg q$). If i lights the torch without j firing the gun, the rabbit will be alarmed and dash off. Similarly, if j fires the gun without i lighting the torch, agent j will miss. The only way for the two of them to catch the rabbit is if they light the torch and fire the gun at the same time. This setting can be modelled as a situation where a system designer aims to fulfil the overall specification that the torch and the gun should not be put to use until some point in time, when the torch should be lighted and the gun be fired simultaneously. This specification can be formulated by the temporal logic formula $\gamma_0 = (\neg p \wedge \neg q)\, \mathsf{U} (p \wedge q)$, where U is the *until*-operator. Let us assume that the computational powers of the two agents are limited and that they can only perform tasks specified either as safety goals or as reachability goals, which only contain the *always*-operator G and the *eventuality*-operator F as temporal connectives, respectively.

Now, the system designer can find an implementation of the specification γ_0 by assigning to player i objective $\gamma_i = \mathsf{F}\, p$ and to player j objective $\gamma_j = \mathsf{G}(p \leftrightarrow q)$. Then, as γ_0, γ_i, and γ_j are consistent and neither i nor j would like to deviate from any run satisfying their goals, there clearly is a Nash equilibrium run in which γ_0 is satisfied. Moreover, there are no Nash equilibria in which γ_0 does not hold. Observe that player i can achieve his goal on his own by simply lighting the torch (that is, setting p to true) at some point. Accordingly, this will also happen in all equilibria. Moreover, due to the epistemic presuppositions inherent in the definition of Nash equilibrium, we may assume that player j knows player i's strategy and can at each time predict whether i is going to light the torch or not, that is, which truth-value of p assumes, and choose the value for q accordingly.[1] Thus, in every equilibrium run also, player j's goal γ_j will be satisfied. Now, let ρ be an equilibrium run and t the earliest time that i lights the torch, that is, the first time that p is set to true. Then, simultaneously, j fires the gun, that is, q is set to true at t as well. Moreover, at all previous times, i does not light the torch and neither is the gun fired as we know that player j will have her goal achieved. It follows that the specification γ_0 holds in all equilibria.

Clearly, there are also runs satisfying γ_0 that are not sustained by an equilibrium, for instance, every run in which both p and q hold at time $t = 0$ and $p \wedge \neg q$ at some later time. In that case, player j would like to deviate, so as to match the truth values of p and q at each time. This simply means that the specification γ_0 is weaker than its implementation in Nash equilibria, and arguably presents no problem to the system designer.

It may furthermore be worthwhile to observe that the conjunction of the players' goals $\gamma_i \wedge \gamma_j = \mathsf{F}\, p \wedge \mathsf{G}(p \leftrightarrow q)$ implies the specification γ_0, but that nevertheless by allocating this formula to both players, there are equilibrium runs that do not satisfy γ_0. For instance, if the players were myopically always to set p and q to false, respectively.

Our article is organised as follows: Following this introduction, we present the background technical concepts used throughout this article. In Section 3, we present the model of iterated Boolean games and provide a useful characterisation of the computations that are induced by their Nash equilibria. Our main contributions are to be found in Section 4, where we introduce the central concept of expressiveness in equilibrium, explore its ramifications for the full fragment of LTL, which we will also refer to as full LTL, propositional calculus, and the important maximal stutter-invariant fragment without the *next*-operator X. On basis of the game of Boolean Ballet (Example 1.1), we formally demonstrate that the temporal property expressed by a Boolean game with players' goals in a fragment of LTL need not necessarily be expressible in that fragment itself. This is also the case for the maximal stutter-invariant fragment, even though we also show that there are still LTL-properties that this fragment cannot express in equilibrium. This is in contrast to full LTL, for which we show that every non-empty temporal property expressible in full LTL is also expressed by an iterated Boolean game with LTL goals, and vice-versa. In Section 5, we study the contrasts between the weaker concepts of *projective expressiveness* and *projective expressiveness in equilibrium*. We prove that both full LTL and the maximal stutter-invariant fragment can projectively express in equilibrium every ω-regular temporal property. In Section 6, we argue how another weakening of the regular expressiveness notion, which we refer to as *weak expressiveness in equilibrium*, chimes in well with the incentive engineering perspective illustrated by the Rabbit Hunt (Example 1.2). We show how the very weak fragment $L_{\mathsf{X},\mathsf{F}^+}$, which only allows reachability goals to be formulated, can already weakly projectively express in equilibrium every ω-regular temporal property. We conclude by reviewing the related literature (Section 7) and suggesting a number of topics for future research (Section 8). This is an extensively revised and reorganised version of our paper that was presented at the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'16) [Gutierrez et al. 2016].

[1]The epistemic and rational preconditions for Nash equilibria to be played has been an area of intensive research (see, e.g., [Aumann and Brandenburger 1995; de Bruin 2010], and [Perea 2012]). These turn out to be very strong and involve common knowledge among the players what actions are going to be played. These results relate to Nash equilibrium in general and are not restricted to iterated Boolean games.

2 PRELIMINARIES

In this section, we provide the necessary preliminaries, introducing the central formal concepts, notions, and definitions we will employ throughout the article.

2.1 Words and Languages

Let Σ be an *alphabet* set of symbols or letters $a, b, c, \ldots$. A *word* w is finite sequence $w = a_0 \ldots a_n$ of letters and an ω-word is an infinite sequence $w = a_0 a_1 a_2, \ldots$ of letters. By ε, we denote the *empty word*. A *language* is subset of words over Σ and an ω-language is a subset of ω-words over Σ. For words w and w' over an alphabet Σ, we use the operations of *concatenation* (ww'), *finite iteration* (w^*), and *infinite iteration* (w^ω). For integers $k \geq 0$, we have a^k denote the k-fold iteration of symbol a.

Regular expressions over the alphabet Σ are defined as usual on basis of ε, $\emptyset$, and each $a \in \Sigma$ using the operation's *concatenation* $(X; Y)$, *union* $(X + Y)$, and *finite iteration* (X^*). Using *infinite iteration* (X^ω), we define an $\boldsymbol{\omega - regularexpression}$ over the alphabet Σ as a finite union of expressions of the form $X; Y^\omega$, where X and Y are a regular expressions and $Y \neq \varepsilon$. An ω-language is said to be $\boldsymbol{\omega - regular}$ if it is the language of some ω-regular expression.

For alphabets Σ_1 and Σ_2, a function $h : \Sigma_1 \to \Sigma_2^+$ defines a *homomorphism* $h : \Sigma_1^\omega \to \Sigma_2^\omega$ by $h(w) = h(a_0)h(a_1)h(a_2)h(a_3)\ldots$ for every ω-word $w = a_0 a_1 a_2 a_3 \ldots$. For an ω-language Λ over Σ_1, we write $h(\Lambda) = \{h(w) : w \in \Lambda\}$ for the corresponding ω-language over Σ_2. The class of ω-regular languages is known to be closed under homomorphisms, that is, if h is a homomorphism and Λ is an ω-regular language, so is $h(\Lambda)$ (see, e.g., [Perrin and Pin 2004], Proposition 3.3, who refer to homomorphisms as morphisms and to ω-regularity as ω-rationality).

An ω-language Λ over alphabet Σ is *stutter-invariant* if for all ω-words $w = a_0 a_1 a_2 a_3 \ldots$ and every sequence $k_0, k_1, k_2, \ldots$ of positive integers, we have $a_0 a_1 a_2 a_3 \ldots \in \Lambda$ if and only if $a_0^{k_0} a_1^{k_1} a_2^{k_2} a_3^{k_3} \ldots \in \Lambda$, where v^k denotes the k-fold iteration of v.

A finite word $w = a_0 \ldots a_k$ is a *prefix* of a finite word $w = b_0 \ldots b_m$ with $k \leq m$ or an ω-word $w' = b_0 b_1 b_2 b_3 \ldots$ if $a_t = b_t$ for all $t \leq k$. The set of prefixes of an ω-word w we denote by *prefix*(w). We say that a finite word w is the (unique) *maximal common prefix* of ω-words w' and w'' if w is a prefix of both w' and w'' and no prefix of greater length has this property. For instance, ab and abc are both common prefixes of the ω-words $abcccc\ldots$ and $abcddd\ldots$, but only abc is maximal.

For an excellent survey of infinite words, we refer the reader to [Perrin and Pin 2004].

2.2 Propositional Temporal Logic

We make extensive use of *Linear Temporal Logic* and the iterated Boolean games based on it. In this section, we present the core concepts of this framework along with a number of auxiliary notions.

*Linear Temporal Logic (*LTL*).* We use the well-known framework of **Linear Temporal Logic (LTL)** (see, e.g., [Pnueli 1977], [Emerson 1990], [Demri et al. 2016]). The formulas of LTL are constructed in the usual fashion from a non-empty and finite set Φ of propositional variables $p, q, r, \ldots$ using the Boolean connectives *negation* $(\neg\varphi)$, *conjunction* $(\varphi \wedge \psi)$, 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* $(\varphi \,\mathsf{U}\, \psi)$. The connectives *truth* $(\top)$, *falsity* $(\bot)$, *implication* $(\varphi \to \psi)$, 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 $(\varphi \to \psi) \wedge (\psi \to \varphi)$, respectively. Where p is a propositional variable, we sometimes write $\bar{p}$ for $\neg p$. We also sometimes 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 if $p \in v$, and false otherwise. For a valuation $v \subseteq \Phi$, we have χ_v^Φ denote the *characteristic clause for* $\boldsymbol{v}$ 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. Accordingly, we, for instance, write $p\bar{q}r$ for valuation $\{p, r\}$ if $\Phi = \{p, q, r\}$.

The formulas of LTL are interpreted with respect to *runs* $\rho = v_0v_1v_2v_3\ldots$, 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 it is clear from the context. Thus, every LTL formula defines an ω-language over 2^Φ, which could also be characterised by other means, for instance, by automata or grammars.

The semantics of LTL then interprets LTL-formulas with respect to a run $\rho = v_0v_1v_2v_3\ldots$ and time index $t \in \mathbb{N}$ as follows:

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

We say that a run ρ *satisfies* a formula φ if $\rho, 0 \models \varphi$. The set of runs in $runs_\Phi$ that satisfy formula φ we denote by $runs_\Phi(\varphi)$. 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 $\{\rho|_\Psi \in runs_\Psi : \rho \in X\}$.

By a *history,* we understand a *finite* and possibly empty sequence $\pi = v_0, \ldots, v_k$ in $(2^\Phi)^*$. We let $length(\pi)$ denote the *length* of π.

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: $L_{\mathsf{G},\mathsf{F},\mathsf{X}}$ (sometimes also referred to as *“restricted* LTL*”*, for instance, by [Perrin and Pin 2004]), where only G and F and X are allowed, and with similar interpretations, the sublogics $L_{\mathsf{G},\mathsf{F}}$, $L_{\mathsf{X},\mathsf{G}^+}$, and $L_{\mathsf{X},\mathsf{F}^+}$, 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 $L_{\mathsf{G},\mathsf{F}}$ was briefly discussed in Example 1.1. In this article, we only tangentially touch upon the important sublogics L_{G^+}, which expresses safety objectives, L_{F^+}, which expresses reachability objectives, and L_X. We have $L_\emptyset$ denote propositional logic. Finally, by $L_{\omega\text{-}reg}$ we denote the set of ω-regular expressions over the alphabet 2^Φ and thus expresses exactly all ω-regular properties of runs. As such, $L_{\omega\text{-}reg}$ is an *extension*, rather than a fragment, of LTL [Wolper 1983]. Sometimes we refer explicitly to the set Φ of variables over which a fragment L is defined, and write $L(\Phi)$ for L.

For a detailed discussion and comparison of expressiveness of the various fragments of LTL, also see [Strejček 2004].

3 ITERATED BOOLEAN GAMES AND NASH EQUILIBRIUM

Boolean games were introduced by [Harrenstein et al. 2001] and further developed by, among others, [Bonzon et al. 2006]and [Endriss et al. 2011]. In this article, we adopt the framework of *iterated Boolean games* as proposed by [Gutierrez et al. 2015a], where players play a Boolean game over an infinite number of rounds and where each player’s goal is given by an LTL-formula.

3.1 Iterated Boolean Games

For L a fragment of LTL over Φ, an *L-iterated Boolean game* (over Φ) is a tuple

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

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 $\Phi_1, \ldots, \Phi_n$ to partition Φ, that is, $\Phi_1 \cup \cdots \cup \Phi_n = \Phi$ and $i \neq j$ implies $\Phi_i \cap \Phi_j = \emptyset$. Henceforth, we also refer to an L-iterated Boolean game as an ***L*-game**.

Example 3.1 (Boolean Ballet, cont'd). The Boolean Ballet game in the introduction can thus be formalised as the tuple $G_{bb} = (\{i, j\}, \{p, q\}, \Phi_i, \Phi_j, \gamma_i, \gamma_j)$, where $\Phi_i = \{p\}$ and $\Phi_j = \{p\}$, as well as $\gamma_i = \mathsf{G}(p \leftrightarrow q)$ and $\gamma_j = \mathsf{F}\, pq \wedge \mathsf{F}\, p\bar{q}$.

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 $f_i : (2^\Phi)^* \to 2^{\Phi_i}$ which associates with every history $\pi \in (2^\Phi)^*$ a choice $f_i(\pi) \in 2^{\Phi_i}$. A *strategy profile* is a tuple $f = (f_1, \ldots, f_n)$ that associates with each player i a strategy f_i and induces an infinite run $\rho(f) = v_0 v_1 v_2 v_3 \ldots$ defined as follows:

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

With a slight abuse of notation, we also write $f(\epsilon) = v_0$ and $f(v_0 \ldots v_t) = v_{t+1}$.

A strategy f_i, as defined above, can make player i's choice at time t dependent on the preceding history $v_0 \ldots v_{t-1}$. It seems, however, natural to assume a player can also simply choose to play a sequence of valuations in Φ_i no matter what choices the other players make. Such strategies have much less structure. Formally, we say that a player's strategy f_i is *naive* if $f_i(v_0 \ldots v_t) = f_i(v'_0 \ldots v'_t)$ for all $t \geq 0$, all histories $v_0 \ldots v_t$ and $v'_0 \ldots v'_t$ of equal length.

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 $\rho \models \gamma_i$ and $\rho' \not\models \gamma_i$. A player i *(weakly) prefers* run ρ to run ρ' if it is not the case that i strictly prefers run ρ' to ρ. 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.

3.2 Nash Equilibrium and Equilibrium Runs

It can easily be seen that with the players, strategies, and preferences defined as in the previous section, each iterated Boolean game defines a *strategic game* in the game-theoretic sense of the word [Shoham and Leyton-Brown 2008; Osborne and Rubinstein 1994; Maschler et al. 2013]. Accordingly, the usual game theoretic solution concepts are available for the analysis of iterated Boolean games. This in particular holds for *(pure) Nash equilibrium*, which in our present setting is a strategy profile $f^* = (f_1^*, \ldots, f_n^*)$ such that, for all players i and all of i's strategies g_i, we have that

$$\rho(f_{-i}^*, g_i) \models \gamma_i \text{ implies } \rho(f^*) \models \gamma_i,$$

where (f_{-i}^*, g_i) denotes the profile $(f_1^*, \ldots, f_{i-1}^*, g_i, f_{i+1}^*, \ldots, f_n^*)$.

We say a Nash equilibrium $f^* = (f_1^*, \ldots, f_n^*)$ is *naive* if, for every player i, strategy f_i^* is naive. We also have the following useful fact, which states that to establish whether a strategy profile is a Nash equilibrium, it suffices to investigate if it does not allow profitable deviations to *naive* strategies. It is important to note, however, this does not entail that every Nash equilibrium can be reduced to an equilibrium in naive strategies.

Lemma 3.2. *Let $G = (N, \Phi_1, \dots, \Phi_n, \gamma_1, \dots, \gamma_n)$ be an iterated Boolean game. Then, a strategy profile $f = (f_1, \dots, f_n)$ is a Nash equilibrium if and only if $\rho(f^*_{-i}, g_i) \models \gamma_i$ implies $\rho(f^*) \models \gamma_i$, for all players i and all* naive *strategies g_i.*

Proof. The "only if"-direction is trivial. For the "if"-direction, assume that $f = (f_1, \dots, f_n)$ is not a Nash equilibrium. Then, there is some player i and some strategy g_i such that $\rho(f) \not\models \gamma_i$ and $\rho(f_{-i}, g_i) \models \gamma_i$. Let $\rho(f_{-i}, g_i) = w_0 w_1 w_2, \dots$. Now define the *naive* strategy g'_i for player i such that $g'_i(\epsilon) = g_i(\epsilon)$ and $g'_i(v_0 \dots v_t) = g_i(w_0, \dots, w_t)$ for every history $v_0 \dots v_t$ of length $t + 1$. By a straightforward inductive argument over t, it can then easily be established that $\rho(f_{-i}, g_i) = \rho(f_{-i}, g'_i)$. It thus follows that $\rho(f_{-i}, g'_i) \models \gamma_i$ and $\rho(f) \not\models \gamma_i$ for some player i and some *naive* strategy g'_i, as desired. □

We say that a run $\rho \in runs_\Phi$ is *sustained* by a Nash equilibrium f^* whenever $\rho(f^*) = \rho$. We then refer to ρ as an *equilibrium run* . The set of equilibrium runs of a game G—rather than the set of equilibria itself—we denote by $NE(G)$.

The relationship between the strategy profiles that are Nash equilibria on the one hand and the equilibrium runs that are sustained by them on the other, is a complex one. Due to the deterministic character of strategies as we defined them, it is obvious that each Nash equilibrium sustains one run only, namely, the run it induces. It is also easy to construct examples showing that iterated Boolean game may have multiple equilibria, which moreover may sustain different runs. Neither should it come as a surprise that one run can be sustained by two different equilibria. However, whether a profile is a Nash equilibrium not only depends on its behaviour on prefixes of the run it induces, but also on its definition on histories that are not a prefix of the equilibrium run. Hence, it is quite possible for two profiles $f = (f_1, \dots, f_n)$ and $g = (g_1, \dots, g_n)$ to induce the same run, that is, $\rho(f) = \rho(g)$, even though f is a Nash equilibrium and g is not. Put slightly differently, the fact that a profile $f = (f_1, \dots, f_n)$ induces an equilibrium run ρ, does not imply that f has to be an equilibrium. To see this, consider the following example.

Example 3.3. Consider the following game with two players, i and j controlling variables p and q, respectively. Let further the players' goals be given by $\gamma_i = \mathsf{G}\, pq$ and $\gamma_j = \mathsf{G}\,\mathsf{F}\, p\bar{q}$. Now consider the strategy f_i for player i defined such that $f_i(\epsilon) = p$ and, for all histories $\pi = w_0 \dots w_k$ ($k \geq 0$),

$$f_i(w_0 \dots w_k) = \begin{cases} p & \text{if } w_t = pq \text{ for all } 0 \leq t \leq k \\ \bar{p} & \text{otherwise.} \end{cases}$$

Also consider the naive strategies g_i and g_j, for players i and j, respectively, such that $g_i(\pi) = p$ and $g_j(\pi) = q$ for all histories π. Then,

$$\rho(f_i, g_j) = \rho(g_i, g_j) = v_0 v_1 v_2 v_3 \dots,$$

where $v_t = pq$ for all $t \geq 0$. Thus, on this run, player i's goal is satisfied on ρ, but player j's goal is not. However, (f_i, g_j) is a Nash equilibrium, whereas (g_i, g_j) is not. To see the latter, observe that player j would like to deviate from (g_i, g_j), for instance, by playing the naive strategy g'_j defined such that $g'_j(\pi) = \bar{q}$ for all histories π. Note that $\rho(g_i, g'_j) = w_0 w_1 w_2 \dots$ such that $w_t = p\bar{q}$ for all $t \geq 0$, and accordingly $\rho(g_i, g'_j)$ satisfies j's goal.

To appreciate that (f_i, g_j) is a Nash equilibrium, assume for a contradiction that there is some strategy g''_j for player j such that $\rho(f_i, g''_j) = u_0 u_1 u_2 \dots$ satisfies j's goal γ_j. Then, there is a smallest $t \geq 0$ such that $v_t \neq u_t$. In particular, $u_t \neq pq$. By definition of f_i it then follows that $u_{t'} \not\models p\bar{q}$ for all $t' > t$. Hence, $\rho(f_i, g''_j) \not\models \mathsf{G}\,\mathsf{F}\, p\bar{q}$, that is, $\rho(f_i, g''_j)$ does not satisfies j's goal γ_j. Therefore, player j does not want to deviate from (f_i, g_j).

Perhaps more profoundly, we find that whether a run is sustained by an equilibrium cannot be reduced to this run satisfying particular players' goals and not satisfying those of other players. As the following example demonstrates, it is quite possible for two runs ρ and ρ' to satisfy the same players' goals—that is, $\rho \models \gamma_i$ if and only if $\rho' \models \gamma_i$, for all players i—even if ρ is sustained by a Nash equilibrium and ρ' is not.

Example 3.4. Again consider a two-player game where player i controls variable p and player j variable q. Let the players' goals be given by $\gamma_i = \mathsf{F}\, pq$ and $\gamma_j = \top$, respectively. Then, run $\rho = v_0v_1v_2\dots$ with $v_t = \bar{p}\bar{q}$ for all $t \geq 0$ is sustained by the Nash equilibrium where the two players play naive strategies that always set p and q to false. Thus, γ_2 is satisfied, but γ_1 is not. The latter is also true of run $\rho' = w_0w_1w_2\dots$ with $w_t = \bar{p}q$ for all $t \geq 0$. This run, however, is not sustained by any Nash equilibrium: for every strategy profile $f = (f_i, f_j)$ that induces run ρ', $f_i(\pi) = \bar{p}$ and $f_j(\pi) = q$ for every prefix π of ρ'. But then player i would want to deviate to a strategy that sets p to true at some point, that is, to a strategy g_i with $g_i(\pi) = p$ for some prefix π of ρ'.

We nevertheless find that we can obtain a convenient characterisation of the Nash equilibrium runs in an iterated Boolean game. To this end, we first introduce a couple of auxiliary concepts.

In Example 3.3, the crucial difference between player i's strategies f_i and g_i is that the former precludes player j achieving his goal no matter which strategy j chooses, whereas the latter does not. In an important sense, by playing f_i, player i threatens to *punish* player j, if j were to deviate and play a strategy that would result in another run than $\rho(f_i, g_j)$. In this way, player i can 'coerce' player j to play g_j, even if by doing so player j does not achieve his goal: by playing any other strategy player j would not achieve his aim either! Formally, for a given player j, we say a run ρ is *consistent with a punishment profile against* $\boldsymbol{j}$ if there is some strategy profile $f^j = (f_1^j, \dots, f_n^j)$ such that $\rho = \rho(f_1^j, \dots, f_n^j)$ and $\rho(f_{-j}^j, g_j) \not\models \gamma_j$ for all strategies g_j for player j. Observe that, for the punishment of player j to be effective, the other players may have to coordinate their strategies.

Lemma 3.5 below characterises an equilibrium run of an iterated Boolean game as a run ρ that is consistent with a punishment profile against each player who does not achieve their goal at ρ. Note that the punishment profile for each 'losing' player may be different. In the proof, we therefore face the challenge of combining the various punishment profiles against the losing players consistent with a run ρ into a single *equilibrium* profile $f^* = (f_1^*, \dots, f_n^*)$ sustaining ρ. For every prefix $v_0 \dots v_t$ of ρ, we can define $f_i(v_0 \dots v_t) = f_i^*(v_0 \dots v_t)$ for each player i. If some losing player j, however, deviates from f_j, the other players will want to jointly enact a punishment profile f_{-j}^j against j. Note that, as we are dealing with Nash equilibrium, we are only concerned with individual deviations by a single losing player.

A specific feature of iterated Boolean games is that a player's choice of strategy is immediately reflected in the run that ensues. If $\rho(f_{-j}, f_j) = v_0v_1v_2\dots$, then we know that $f_j(\epsilon) = v_0 \cap \Phi_j$ and $f_j(v_0 \dots v_t) = v_{t+1} \cap \Phi_j$ for all $t \geq 0$. Accordingly, deviations by a single player can be spotted straightforwardly and immediately: if $\rho(f) = v_0v_1v_2\dots$ and at some point a history $\pi = v_0 \dots v_{t-1}w_t$ such that both $w_t \cap \Phi_j \neq v_t \cap \Phi_j$ and $w_i \cap \Phi_i = v_i \cap \Phi_i$ for all players i distinct from j materialises, it is clear that the deviation from ρ can uniquely be attributed to player j.[2] Formally, we say that a history $w_0 \dots w_t$ is a $\boldsymbol{j - deviation}$ $\rho = v_0v_1v_2\dots$ by player j if there is an $s \leq t$ such that $w_0 \dots w_{s-1} = v_0 \dots v_{s-1}$ (on the understanding that $w_0 \dots w_{s-1} = v_0 \dots v_{s-1} = \epsilon$ if $s = 0$), $w_s \cap \Phi_j \neq v_s \cap \Phi_j$, and $w_s \cap \Phi_i = v_s \cap \Phi_i$ for all players i distinct from j, that is, if the first deviation from ρ can uniquely be attributed to j.

On this basis, we now obtain the following characterisation of Nash equilibrium runs in iterated Boolean games.

[2] Also compare the concept of *attributability* of a deviation in Gutierrez et al. [2015b].

Lemma 3.5. *Let $\rho = v_0v_1v_2\dots$ be a run in an iterated Boolean game G. Then, $\rho \in NE(G)$ if and only if, for each player j with $\rho \not\models \gamma_j$, run ρ is consistent with a punishment profile against j.*

Proof. For the "only if"-direction assume for contraposition that there is a player j with $\rho \not\models \gamma_j$ against whom there is no punishment profile consistent with ρ. That is, there is no strategy profile $f^j = (f_1^j, \dots, f_n^j)$ such that $\rho = \rho(f_1^j, \dots, f_n^j)$ and $\rho(f_{-j}^j, g_j) \not\models \gamma_i$ for all strategies g_j. Now consider an arbitrary strategy profile $f = (f_1, \dots, f_n)$ with $\rho = \rho(f_1, \dots, f_n)$. Then, there is some strategy g_j for j such that $(f_{-j}, g_j) \models \gamma_i$. Accordingly, f is no Nash equilibrium and, with f having been chosen arbitrarily, it follows that $\rho \notin NE(G)$.

For the opposite direction, assume $\rho = v_0v_1v_2\dots$ to be consistent with a punishment profile $f^j = (f_1^j, \dots, f_n^j)$ against each player j with $\rho \not\models \gamma_j$. Let, moreover, $f = (f_1, \dots, f_n)$ be a profile such that $\rho = \rho(f_1, \dots, f_n)$; observe that we may assume that such a profile exists. Now define $f^* = (f_1^*, \dots, f_n^*)$ as the strategy profile such that, for every player i and every history π,

$$f_i^*(\pi) = \begin{cases} f_i^j(\pi) & \text{if } \pi \text{ is a } j\text{-deviation from } \rho \text{ for some player } j \text{ with } \rho \not\models \gamma_j, \\ f_i(\pi) & \text{otherwise.} \end{cases}$$

Observe that f^* is well defined because, by definitionl, each history can be a j-deviation from ρ for at most one player j only. As for every player i, neither ϵ nor any prefix $v_0 \dots v_t$ of ρ is a j-deviation, we have $f_i(\epsilon) = f^*(\epsilon)$ and $f_i^*(v_0 \dots v_t) = f_i(v_0 \dots v_t)$ for every $t \geq 0$, and, hence, $\rho(f^*) = \rho(f) = \rho$. In a similar vein, observe that consistency with $\rho = v_0v_1v_2\dots$ implies that $f_i^j(\epsilon) = f_i(\epsilon)$ and $f_i^j(v_0 \dots v_t) = f_i(v_0 \dots v_t)$ for prefix $v_0 \dots v_t$ of ρ and all players i (including j).

We conclude the proof by showing that f^* is a Nash equilibrium. To this end, consider an arbitrary player j with $\rho(f^*) \not\models \gamma_j$. Then, also $\rho(f) \not\models \gamma_j$. Now also consider an arbitrary strategy g_j for j. We demonstrate that $\rho(f_{-j}^*, g_j) \not\models \gamma_j$, which implies that f^* is a Nash equilibrium. As $f^j = (f_1^j, \dots, f_n^j)$ is a punishment profile, $\rho(f_{-j}^j, g_j) \not\models \gamma_j$. Hence, it suffices to show that $\rho(f_{-j}^*, g_j) = \rho(f_{-j}^j, g_j)$.

To this end, let $\rho(f_{-j}^*, g_j) = w_0w_1w_2\dots$ and $\rho(f_{-j}^j, g_j) = u_0u_1u_2\dots$. We prove by induction that $w_t = u_t$ for every $t \geq 0$. For the basis $t = 0$, observe that ϵ is not a j-deviation from ρ. Hence,

$$\begin{aligned} w_0 &= f_1^*(\epsilon) \cup \dots \cup g_j(\epsilon) \cup \dots \cup f_n^*(\epsilon) \\ &= f_1(\epsilon) \cup \dots \cup g_j(\epsilon) \cup \dots \cup f_n(\epsilon) \\ &= f_1^j(\epsilon) \cup \dots \cup g_j(\epsilon) \cup \dots \cup f_n^j(\epsilon) \\ &= u_0. \end{aligned}$$

For the induction step, we may assume that $w_0 \dots w_t = u_0 \dots u_t$ to prove that $w_{t+1} = u_{t+1}$. First assume that $w_0 \dots w_t = v_0 \dots v_t$. Then obviously also $u_0 \dots u_t = v_0 \dots v_t$. Moreover, $w_0 \dots w_t$ is a i-deviation from ρ for no player i. Hence,

$$\begin{aligned} w_{t+1} &= f_1^*(w_0 \dots w_t) \cup \dots \cup g_j(w_0 \dots w_t) \cup \dots \cup f_n^*(w_0 \dots w_t) \\ &= f_1^*(v_0 \dots v_t) \cup \dots \cup g_j(v_0 \dots v_t) \cup \dots \cup f_n^*(v_0 \dots v_t) \\ &= f_1(v_0 \dots v_t) \cup \dots \cup g_j(v_0 \dots v_t) \cup \dots \cup f_n(v_0 \dots v_t) \\ &= f_1^j(v_0 \dots v_t) \cup \dots \cup g_j(v_0 \dots v_t) \cup \dots \cup f_n^j(v_0 \dots v_t) \\ &=_{i.h} f_1^j(u_0 \dots u_t) \cup \dots \cup g_j(u_0 \dots u_t) \cup \dots \cup f_n^j(u_0 \dots u_t) \\ &= u_{t+1}. \end{aligned}$$

Finally, assume that $w_0 \dots w_t \neq v_0 \dots v_t$. Consider the smallest $0 \leq s \leq t$ such that $w_s \neq u_s$. Then, $w_0 \dots w_{s-1} = v_0 \dots v_{s-1}$, on the understanding that $w_0 \dots w_{s-1} = \epsilon$ if $s = 0$. Accordingly,

$w_0 \dots w_{s-1}$ is a k-deviation from ρ for no player k. Hence, for all i distinct from j,

$$f_i^*(w_0 \dots w_{s-1}) = f_i(w_0 \dots w_{s-1}) = f_i(v_0 \dots v_{s-1}) = v_s \cap \Phi_i$$

It follows that $w_0 \dots w_t$ is a j-deviation from ρ. Consequently,

$$\begin{aligned} w_{t+1} &= f_1^*(w_0 \dots w_t) \cup \dots \cup g_j(w_0 \dots w_t) \cup \dots \cup f_n^*(w_0 \dots w_t) \\ &= f_1^j(w_0 \dots w_t) \cup \dots \cup g_j(w_0 \dots w_t) \cup \dots \cup f_n^j(w_0 \dots w_t) \\ &=_{i.h.} f_1^j(u_0 \dots u_t) \cup \dots \cup g_j(u_0 \dots u_t) \cup \dots \cup f_n^j(u_0 \dots u_t) \\ &= u_{t+1}. \end{aligned}$$

This concludes the proof. □

4 EXPRESSIVENESS, ITERATED BOOLEAN GAMES, AND NASH EQUILIBRIUM

In this section, we introduce the central concept of this article: *expressiveness in equilibrium.* This expressiveness notion is based on the equilibria of iterated Boolean games. We find that some LTL-fragments can express properties in equilibrium that they cannot normally express, showing that the two expressiveness notions are to be distinguished. For propositional logic and LTL, however, we show that the two notions are equivalent. Finally, we focus on the next-free fragment $L_{\mathbf{U}}$, which is known to be the largest stutter-invariant fragment of LTL. We demonstrate that the expressive strength in equilibrium of this important fragment lies strictly between that of full LTL and $L_{\mathbf{U}}$ itself.

4.1 Expressiveness in Equilibrium

Given a suitable model-theoretic semantics, the expressive power of a logic can be measured in terms of the sets of models it characterises. In the case of temporal propositional logics, the models are infinite sequences of valuations, and by a *linear time property* we understand any subset $X \subseteq runs_\Phi$ of runs over a set of propositional variables Φ. Thus, we say that an LTL-fragment L *expresses* property $X \subseteq runs_\Phi$ if there is some formula φ in the fragment L such that X coincides with the set of runs that satisfy φ, that is, if $X = runs_\Phi(\varphi)$. We also say in this case that L *expresses property X in formulas* so as to distinguish the concept from expressiveness in equilibrium to be defined below.

Alternatively, each temporal logic formula over propositional variables Φ can be seen to define an *ω-language* over 2^Φ. Accordingly, every fragment of LTL defines a model of computation, accepting exactly those ω-languages consisting of those ω-words $\rho = v_0 v_1 v_2 \dots$ over 2^Φ for which there is a formula φ in L such that ρ satisfies φ.

We use the notation $L \geq L'$ to denote that every property expressible in fragment L can also be expressed in fragment L'. We write $L \equiv L'$ if both $L \geq L'$ and $L' \geq L$, and $L > L'$ if $L \geq L'$ but not $L' \geq L$. We extend these notations in the natural way to models of computation in general, and it is easy to see that the relation $\geq$ is transitive.

Like the formulas of a temporal logic, the Nash equilibria of an iterated Boolean game similarly define a linear time property, namely, the set of runs these Nash equilibria sustain. Formally, we say that LTL-fragment L *expresses in equilibrium* property $X \subseteq runs_\Phi$, if there is some L-game G such that $X = NE(G)$. Phrased slightly differently, the equilibrium runs of each iterated L-game define an ω-language over 2^Φ. Accordingly, for each LTL-fragment L, the class of iterated L-games also defines its own model of computation, which we denote by L^{NE}. On this basis of expressible properties, respectively, accepted languages, we can now compare the expressiveness of fragments of linear temporal languages with the expressiveness of the classes of iterated Boolean games based on LTL-fragments.

Intuitively, one might expect that, for LTL-fragments L in general, expressiveness in equilibrium is a stronger notion than expressiveness in formulas, that is, if L can express property X in formulas, it can also express property X in equilibrium. We find that this intuition is generally vindicated, apart from the one notable exception where the set of propositional variables is a singleton $\{p\}$. Then, the empty property ϕ, consisting of no runs whatsoever, is obviously expressed by any unsatisfiable formula of L over p, for instance, $p \wedge \neg p$. In this borderline case, however, the empty property cannot be expressed by L in equilibrium.

To see this, consider an arbitrary L-game over $\Phi = \{p\}$. Then, control over the single variable p can be assigned to a single player i only. Moreover, if the set of equilibrium runs of this game is to be empty, it will also have to be this single player who wants to deviate from *every* strategy profile. Now, if there is a run satisfying i's goal, then i does not want to deviate from any strategy profile that underlies it. Consequently, this run is sustained by an equilibrium. If, on the other hand, i's goal is unsatisfiable, then there is no way she can deviate from *any* strategy profile to have her goal satisfied. So, in that case, all runs are sustained by a Nash equilibrium. An empty set of equilibrium runs can therefore not be achieved in this case.

In case Φ contains at least two variables, we can construct a two-player L-game with $runs_\Phi(\varphi)$ as the set of equilibrium runs, for every formula φ in an LTL-fragment L. The underlying idea is to assign control over one variable p to the one player and control over another variable q to the other. Then, define the players' preferences such that they both aim to satisfy the formula φ, but failing that, the first player wants to have the truth-values of p and q to be matched at time $t = 0$, whereas the second player wants them to be unmatched. Thus, if a run satisfies φ, both players are satisfied and neither wants to deviate. If, on the other hand, a run does not satisfy φ, the two players will be caught up in a *matching pennies* game [Osborne and Rubinstein 1994] on p and q, for which there is no (pure) Nash equilibrium. Formally, for any formula φ in a fragment L, define the two-player game G_φ^{mp}. Let the two players, i and j, such that i controls p and j all other variables, including q, that is, $\Phi_i = \Phi \setminus \{q\}$ and $\Phi_j = \{q\}$. Let furthermore the players' goals be given by:

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

By showing that $runs_\Phi(\varphi)$ is exactly the set of equilibrium runs of G_φ^{mp}, we obtain the following proposition:

Proposition 4.1. *Let L be an* LTL*-fragment on Φ with $|\Phi| \geq 2$. Then, every temporal property that can be expressed by L in formulas can also be expressed by L in equilibrium, that is, $L^{NE} \geq L$.*

Proof. Consider an arbitrary property X that is expressed by a formula φ in fragment L, and construct the two-player game G_φ^{mp}. It suffices to show that $NE(G_\varphi^{mp}) = runs_\Phi(\varphi)$. To this end, consider an arbitrary run ρ along with a strategy profile $f = (f_i, f_j)$ with $\rho = \rho(f)$. If ρ satisfies φ, both players' goals are satisfied as well. It follows that f is an equilibrium and $\rho \in NE(G_\varphi^{mp})$. If, on the other hand, ρ does not satisfy φ, consider $f_i(\epsilon)$ and $f_j(\epsilon)$. If both $p \in f_i(\epsilon)$ and $q \in f_j(\epsilon)$, player j can get his unsatisfied goal satisfied by deviating to any strategy g_j with $q \notin g_j(\epsilon)$, as in that case $\rho(f_i, g_j) \models p \leftrightarrow \bar{q}$. Player j similarly wants to deviate from f_j if $p \notin f_i(\epsilon)$ and $q \notin f_j(\epsilon)$. If it is not the case that $p \in f_i(\epsilon)$ if and only if $q \in f_j(\epsilon)$, it is player i who wants to deviate. In each of these cases, it follows that $f = (f_i, f_j)$ is not an equilibrium of G_φ^{mp}, as desired. □

Since we have established that, in virtually all cases, properties that can be expressed by a fragment L in formulas can also expressed by L in equilibrium, one may wonder if the inverse statement holds as well. The game Boolean Ballet, as presented informally in Example 1.1, shows that this is not generally the case: There exist LTL-fragments that cannot express certain properties that they can express in equilibrium, that is, not all properties of the players' goals are inherited by the equilibrium runs of iterated Boolean games. This is in particular the case for the fragment $L_{\mathsf{F},\mathsf{G}}$, which, as a sublogic LTL-fragment L_{U}, is known to be stutter-invariant. In the proof of this result, we buttress our informal reasoning in the introduction by a formally precise argument.

Proposition 4.2. *Let Φ be a set of propositional variables with $|\Phi| \geq 2$. Then, it is not generally the case for every LTL-fragment on Φ that every temporal property that can be expressed by L in equilibrium, can also be expressed by L in formulas. In particular, this does not hold for the* LTL-*fragment $L_{\mathsf{F},\mathsf{G}}$, and, hence, $L^{NE}_{\mathsf{F},\mathsf{G}} > L_{\mathsf{F},\mathsf{G}}$.*

Proof. As we may assume $|\Phi| \geq 2$, by Proposition 4.1, we immediately have that every property that $L_{\mathsf{F},\mathsf{G}}$ can express, $L_{\mathsf{F},\mathsf{G}}$ can also express in equilibrium. To show that there is a property that $L^{NE}_{\mathsf{F},\mathsf{G}}$ can express, but $L_{\mathsf{F},\mathsf{G}}$ cannot, consider the two-person iterated Boolean game G_{bb} referred to as "Boolean Ballet" in Example 1.1 in the introduction. (For expositional convenience, we prove the statement for the case where $\Phi = \{p, q\}$. The argument, however, can be easily extended to the more general case where $\{p, q\} \subseteq \Phi$.) Recall from Example 3.1 that

$$\Phi_i = \{p\} \qquad\qquad \Phi_j = \{q\}$$
$$\gamma_i = \mathsf{G}(p \leftrightarrow q) \qquad\qquad \gamma_j = \mathsf{F}\, pq \wedge \mathsf{F}\, p\bar{q}.$$

We show that for every run $\rho = v_0 v_1 v_2 \ldots$, we have that $\rho \in NE(G_{bb})$ if and only if both $v_t = pq$ or $v_t = \bar{p}\bar{q}$ for all $t \geq 0$, and $v_s = pq$ for at most one $s \geq 0$. As the Nash equilibria thus define a temporal property that is not stutter-invariant—observe that $pq\ \bar{p}\bar{q}\ \bar{p}\bar{q}\ \bar{p}\bar{q}\ \ldots$ is an equilibrium, but $pq\ pq\ \bar{p}\bar{q}\ \bar{p}\bar{q}\ \bar{p}\bar{q}\ \ldots$ is not—this suffices for a proof.

To this end, first consider an arbitrary run $\rho = v_0 v_1 v_2 \ldots$ such that $v_t = pq$ or $v_t = \bar{p}\bar{q}$ for all $t \geq 0$ and $v_s = pq$ for at most one $s \geq 0$. Also consider the naive strategies f_i and f_j defined such that, for every $t \geq 0$ and history $w_0 \ldots w_{t-1}$,

$$f_i(w_0 \ldots w_t) = v_{t+1} \cap \Phi_i \qquad \text{and} \qquad f_j(w_0 \ldots w_t) = v_{t+1} \cap \Phi_j,$$

(on the understanding that $w_0 \ldots w_{t-1} = \epsilon$ for $t = 0$). It is then easily appreciated that $\rho(f_i, f_j) = v_0 v_i v_j \ldots$. To see that $f = (f_i, f_j)$ is a Nash equilibrium, first observe that $\rho(f) \models \gamma_i$. Thus, player i has her goal satisfied and she does not want to deviate. For player j, however, $\rho(f) \not\models \gamma_j$. Now, assume for contradiction that there is a strategy g_j such that $\rho(f_i, g_j) \models \gamma_j$, that is, that player j would like to deviate from f and play g_j. Let $\rho(f_i, g_j) = v'_0 v'_1 v'_2, \ldots$. Then, a $t^* \geq 0$ exists such that $v'_{t^*} = p\bar{q}$. Moreover, as we defined f_i as a naive strategy, $v_{t^*} = pq$ and $v_t = \bar{p}\bar{q}$ for all $t \neq t^*$. Again because f_i is naive, $v'_t = \bar{p}q$ or $v'_t = \bar{p}\bar{q}$ for all $t \neq t^*$. It follows that $\rho(f_i, g_j) \not\models \mathsf{F}\, p\bar{q}$ and, hence, $\rho(f_i, g_j) \not\models \gamma_j$, a contradiction.

For the opposite direction, assume that $f = (f_i, f_j)$ is a strategy profile such that $\rho(f_i, f_j) = v_0 v_1 v_2 \ldots$, and, for contraposition, that either (i) it is not the case that $v_t = pq$ or $v_t = \bar{p}\bar{q}$ for all $t \geq 0$ or (ii) there are $t^{**} > t^* \geq 0$ with $v_{t^*} = v_{t^{**}} = pq$. We show that f is not a Nash equilibrium.

If (i) , observe that $\rho(f_i, f_j) \not\models \gamma_i$. Now, define strategy g_i such that, for all $t \geq 0$ and histories $w_0 \ldots w_{t-1}$,

$$g_i(w_0 \ldots w_{t-1}) = \begin{cases} p & \text{if } f_j(w_0 \ldots w_{t-1}) = q, \\ \bar{p} & \text{if } f_j(w_0 \ldots w_{t-1}) = \bar{q} \end{cases}$$

(again on the understanding that $w_0 \ldots w_{t-1} = \epsilon$ for $t = 0$). Intuitively, g_i inspects strategy f_j to predict at each time t to what truth-value player j sets q, and matches p's truth-value accordingly. Let $\rho(g_i, f_j) = w'_0 w'_1 w'_2 \ldots$. By means of a straightforward inductive argument, it can then be shown that $w'_t = pq$ or $w'_t = \bar{p}\bar{q}$ for all $t \geq 0$, and hence $\rho(g_i, f_j) \models \gamma_i$. Accordingly, f is not a Nash equilibrium.

If (ii), observe that we may assume that $v_t = pq$ or $v_t = \bar{p}\bar{q}$ for all $t \geq 0$. Then, obviously, $\rho(f_i, f_j) \not\models \gamma_j$. Now define g_j such that for all $t \geq 0$ and histories $w_0, \ldots, w_{t-1}$ (on the understanding that $w_0 \ldots w_{t-1} = \epsilon$ for $t = 0$):

$$g_j(w_0 \ldots w_{t-1}) = \begin{cases} f_j(w_0 \ldots w_{t-1}) & \text{if } t \leq t^{**}, \\ \bar{q} & \text{otherwise.} \end{cases}$$

Thus, by using strategy g_j, player j makes the same choices as with f_j up to time $t^{**} - 1$, where it deviates by choosing $\bar{q}$ instead of q. Let $\rho(f_i, g_j) = w_0'' w_1'' w_2'', \ldots$. By means of another straightforward inductive argument it can then be show that $w_t'' = v_t$ for all $0 \leq t \leq t^{**} - 1$. Hence, $w_0'' \ldots w_{t^{**}-1}'' = v_0 \ldots v_{t^{**}-1}$. In particular, it holds that $w_{t^*}'' = pq$. For t^{**}, moreover, we have both

$$f_i(w_0'' \ldots w_{t^{**}-1}'') = f_i(v_0 \ldots v_{t^{**}-1}) = p \qquad \text{and} \qquad g_j(w_0'' \ldots w_{t^{**}-1}'') = \bar{q},$$

that is, $w_{t^{**}}'' = p\bar{q}$. Accordingly, $\rho(f_j, g_j) \models \gamma_j$. Hence, as player j can thus deviate from f by playing g_j and satisfy his goal, we may conclude that f is not a Nash equilibrium, as desired. □

We conclude this section with the easy but useful observation concerning the relative expressive power of different fragments: if a fragment L_1 is at least as expressive as L_2 in formulas, then L_1 is also at least as expressive as L_2 in equilibrium.

Proposition 4.3. *For fragments L_1 and L_2, we have $L_1 \geq L_2$ implies $L_1^{NE} \geq L_2^{NE}$.*

Proof. Assume $L_1 \geq L_2$ and consider an arbitrary temporal property X such that $X = NE(G_2)$ for some L_2-game $G_2 = (N, \Phi, \Phi_1, \ldots, \Phi_n, \gamma_1, \ldots, \gamma_n)$ with $\gamma_i \in L_1$ for all players i. Then, for each player i, there is an formula $\gamma_i' \in L_1$ such that $runs_\Phi(\gamma_i) = runs_\Phi(\gamma_i')$. Let $G_1 = (N, \Phi, \Phi_1, \ldots, \Phi_n, \gamma_1', \ldots, \gamma_n')$. Note that the strategies for each player in G_1 and G_2 are identical. Then, clearly, $NE(G_1) = NE(G_2) = X$. Consider the following equivalences for every $\rho \in runs_\Phi$.

$\rho \in NE(G_1)$
iff $\rho = \rho(f^*)$ and f^* a Nash equilibrium of G_1
iff $\rho = \rho(f^*)$ and $\rho(f_{-i}^*, f_i) \models \gamma_i$ implies $\rho(f^*) \models \gamma_i$, for all players~$i$ and strategies f_i
iff $\rho = \rho(f^*)$ and $\rho(f_{-i}^*, f_i) \models \gamma_i'$ implies $\rho(f^*) \models \gamma_i'$, for all players~$i$ and strategies f_i
iff $\rho = \rho(f^*)$ and f^* a Nash equilibrium of G_2
iff $\rho \in NE(G_2)$.

This concludes the proof. □

4.2 Propositional Logic: The Empty Fragment L_ϕ

In the previous section, we saw that an LTL-fragment can be more expressive in equilibrium than it is in formulas. This, however, is not necessarily the case. A prime example is *propositional logic*, that is, the empty LTL-fragment L_ϕ with no temporal operators. A fundamental observation in the literature on Boolean games is that for any given Boolean game there is a formula characterising the set of Nash equilibrium outcomes. More formally, given an L_ϕ-game $(N, \Phi, \Phi_1, \ldots, \Phi_n, \gamma_1, \ldots, \gamma_n)$, the set of equilibrium runs is characterised by

$$\eta(G) = \textstyle\bigwedge_{i \in N} (\bigvee_{\theta:\Phi_i \to \{\top,\bot\}} \theta(\gamma_i) \to \gamma_i),$$

where $\theta(\gamma_i)$ results from γ_i by replacing every occurrence of a propositional variable p in Φ_i by the Boolean $\top$ or $\bot$ as specified by $\theta(p)$ [Bonzon et al. 2006]. As an immediate consequence, we find that L_ϕ can express every property that it can express in equilibria, and—provided that Φ contains at least two propositional variables—*vice-versa*. In other words, the players can distinguish the same temporal properties with their goals, as L_ϕ-games do by means of their equilibrium runs.

Proposition 4.4. *Let L be an* LTL*-fragment on Φ with $|\Phi| \geq 2$. Then, L_ϕ can express property X in formulas if and only if L_ϕ can express X in equilibrium, that is, $L_\phi \equiv L_\phi^{NE}$.*

Proof. The "only if"-direction is immediate by Proposition 4.1. For the "if"-direction, let X be an arbitrary temporal property that can be expressed by L_ϕ in equilibrium. Then, there is an L_ϕ-game G with $X = NE(G)$. Dealing with L_ϕ, it is then well known that $X = runs_\Phi(\eta(G))$. Conclude by observing that $\eta(G)$ is a formula of L_ϕ. □

It may be worth observing that proof of Proposition 4.4 shows something in addition to the equi-expressiveness of propositional logic in formulas and in equilibrium: the formula scheme $\eta(G)$ provides a *uniform* way to find a propositional formula characterising the equilibrium runs of a Boolean game.

4.3 LTL: The Full Fragment

We now turn to the question how expressive LTL is with respect to LTL^{NE}, that is, to what extent the temporal properties defined by the Nash equilibria of LTL-games can be expressed in formulas by LTL itself. Again under the assumption that there be at least two propositional variables, we show that LTL is just as expressive in formulas as it is in equilibrium, that is, LTL $\equiv$ LTL^{NE} (see Theorem 4.8, below). Put slightly differently, for LTL-games, the structural properties of the players' goals are fully reflected as properties of the game's equilibrium runs.

Key to this issue are those temporal properties (or, more generally, ω-languages) that are *non-counting* [Mateescu and Salomaa 1997; Strejček 2004]. Formally, and adapted to our setting, a temporal property X over 2^Φ is *non-counting* if there is an *index* $n_0 > 0$ such that for all $k \geq n_0$ and histories π, π_0, π_1, and π_2, the following equivalences hold:

(i) $\pi_0\ \pi^k\ \pi_1\ \pi_2^\omega \in X$ if and only if $\pi_0\ \pi^{k+1}\ \pi_1\ \pi_2^\omega \in X$, and
(ii) $\pi_0\ (\pi_1\ \pi^k\ \pi_2)^\omega \in X$ if and only if $\pi_0\ (\pi_1\ \pi^{k+1}\ \pi_2)^\omega \in X$.

Intuitively, the property of being non-counting can be seen as a weak form of stutter-invariance, where a string can only be "stuttered" after it has been stuttered a sufficiently large number of times already. As a prime example of a *counting* property over $runs_\Phi$ consider the one consisting of all runs $v_0v_1v_2\ldots$ for which there is an *odd* $t \geq 0$ such that $v_{t'} = \{p\}$ for all $t' \leq t$ and $v_{t''} = \{q\}$ for all $t'' > t$, that is the property defined by the ω-regular expression $\boldsymbol{(\{p\};\{p\})^*;\{q\}^\omega}$. Whereas for every even $k \geq 0$, we have that the run $\{p\}^k\ \{q\}^\omega$ belongs to this property, the run $\{p\}^{k+1}\ \{q\}^\omega$ does not. Kučera and Strejček [2005] have characterised the LTL-properties as those that are both ω-regular and non-counting.

Theorem 4.5 ([Kučera and Strejček 2005]). *A property $X \subseteq runs_\Phi$ can be expressed by* LTL *if and only if X is ω-regular and non-counting.*

In the remainder of this section, we therefore demonstrate that every LTL-game G the set $NE(G)$ of equilibrium runs is both non-counting and ω-regular, proving the former first.

Proposition 4.6. *For every* LTL*-game G, the set $NE(G)$ of equilibrium runs is non-counting.*

Proof. Consider an arbitrary LTL-game $G = (N, \Phi, \Phi_1, \ldots, \Phi_n, \gamma_1, \ldots, \gamma_n)$. As the goal γ_i of each player i as well as its negation $\neg\gamma_i$ define LTL-properties, by Theorem 4.5, there are indices $n_{\gamma_i}, n_{\neg\gamma_i} \geq 0$ by virtue of which, respectively, $runs_\Phi(\gamma_i)$ and $runs_\Phi(\neg\gamma_i)$ are non-counting properties. Now let

$$n_0 = 1 + \max\{n_{\gamma_i}, n_{\neg\gamma_i} : i \in N\}.$$

Thus, for every player i, we have $n_{\gamma_i}, n_{\neg\gamma_i} < n_0$, that is, the index n_0 is chosen sufficiently large so as to apply to all the player's goals and their negations. To show that the set $NE(G)$ is non-counting with respect to index n_0, we consider an arbitrary $k \geq n_0$, and arbitrary histories π, π_0, π_1, and π_2, and show that both:

$$\begin{array}{llll} \text{(i)} & \pi_0\ \pi^k\ \pi_1\ \pi_2^\omega \in NE(G) & \text{if and only if} & \pi_0\ \pi^{k+1}\ \pi_1\ \pi_2^\omega \in NE(G), \\ \text{(ii)} & \pi_0\ (\pi_1\ \pi^k\ \pi_2)^\omega \in NE(G) & \text{if and only if} & \pi_0\ (\pi_1\ \pi^{k+1}\ \pi_2)^\omega \in NE(G). \end{array}$$

Thus, we have to provide proofs for four statements—the "if"-direction and "only if"-direction of both (i) and (ii) .

Each of these proofs has a similar structure. Both in (i) and (ii), the runs in question differ only in that one of them involves a $k + 1$st iteration of π whereas the other does not. Let the latter be denoted by ρ and the former by ρ'. Now, given an equilibrium f^* sustaining ρ, we can define an equilibrium g^* sustaining ρ', where, intuitively, g^* is very much like f^*, be it that g^* behaves at the $k + 1$st iteration of π in the same way as it (and f^*!) does at the kth iteration of π. Similarly, given an equilibrium f^* for ρ', we define an equilibrium g^* for ρ, where g^* is very much like f^*, be it that g^* "ignores" how f^* behaves at the the $k + 1$st iteration of π. Because their goals are assumed to be non-counting, the players will be indifferent between whether f^* or g^* is being played. Moreover, g^* is defined such that any deviation from it by some player i has a corresponding deviation from f^* by the same player with exactly the same effects for the satisfaction of this player's goal. Hence, a player deviating from g^* will be punished in the same way as the corresponding deviation from f^*. Hence, as f^* is an equilibrium, so is g^*. In the remainder of the proof, we provide formally precise definitions of g^* as it is defined on basis of f^* and prove that g^* is an equilibrium sustaining the appropriate run.

For the "only if"-direction of (i), let $\rho = \pi_0\ \pi^k\ \pi_1\ \pi_2^\omega$ and $\rho' = \pi_0\ \pi^{k+1}\ \pi_1\ \pi_2^\omega$. As γ_i defines a non-counting property for every player i, by choice of n_0, we have $\rho \models \gamma_i$ if and only if $\rho' \models \gamma_i$ for every player i. Assume that there is a Nash equilibrium $f^* = (f_1^*, \ldots, f_n^*)$ that sustains ρ, that is, $\rho(f^*) = \pi_0\ \pi^k\ \pi_1\ \pi_2^\omega$. On this basis, we construct another Nash equilibrium $g^* = (g_1^*, \ldots, g_n^*)$ of G that sustains ρ', that is, $\rho(g^*) = \pi_0\ \pi^{k+1}\ \pi_1\ \pi_2^\omega$, and $\rho(g^*) \in NE(G)$.

To this end, first observe that every history η is of the form $\pi_{com}\, \pi_{dev}$, where π_{com} is the *unique* maximal common prefix of ρ' and η, and π_{dev} is a tail piece of η that "deviates" from ρ'. Note that both π_{com} and π_{dev} may be the empty sequence ϵ. For every history $\eta = \pi_{com}\, \pi_{dev}$, there are then two possibilities: $\pi_{com} < \pi_0\ \pi^k$ or $\pi_{com} = \pi_0\ \pi^k\ \pi'$ for some $\pi' \leq \pi\ \pi_1\ \pi_2^\omega$. That is, η differs (if at all) from ρ' either before the kth iteration of π in ρ' is completed, or thereafter. Now define $g^* = (g_1^*, \ldots, g_n^*)$ such that, for each player i and history $\eta = \pi_{com}\, \pi_{dev}$,

$$g_i^*(\eta) = \begin{cases} f_i^*(\pi_{com}\, \pi_{dev}) & \text{if } \pi_{com} < \pi_0\ \pi^k, \\ f_i^*(\pi_0\ \pi^{k-1}\ \pi'\ \pi_{dev}) & \text{if } \pi_{com} = \pi_0\ \pi^k\ \pi' \text{ for some } \pi' \leq \pi\ \pi_1\ \pi_2^\omega. \end{cases}$$

Observe that g_i^* is well defined for each player i because we may assume that $k \geq 1$. Intuitively, the strategy profiles f^* and g^* are the same, apart from that, by playing g^*, the players repeat the kth iteration of π. Observe that this also includes the behaviour of g_i^* on histories that are not prefixes of the run $\rho' = \pi_0\ \pi^{k+1}\ \pi_1\ \pi_2^\omega$. By means of an inductive argument over the length of $\rho(g^*)$, it can then be shown that $\rho(g^*) = \pi_0\ \pi^{k+1}\ \pi_1\ \pi_2^\omega$.

It remains to be shown that $g^* = (g_1^*, \ldots, g_n^*)$ is also a Nash equilibrium. To this end, assume, for contradiction, that there is a strategy g_j' for some player j such that $\rho(g^*) \not\models \gamma_j$ and $\rho(g_{-j}^*, g_j') \models \gamma_i$. Recall that γ_j defines a non-counting property for index n_0. As $\rho(f^*) = \pi_0\pi^k\pi_1\pi_2^\omega$ and $\rho(g^*) = \pi_0\pi^{k+1}\pi_1\pi_2^\omega$, we thus obtain that $\rho(f^*) \not\models \gamma_j$.

To see that player j also wants to deviate from f^*, let $\rho(g_{-j}^*, g_j') = \pi'_{com}\rho'_{dev}$, where π'_{com} is the maximal common prefix of $\rho(g^*)$ and $\rho(g_{-j}^*, g_j')$, and ρ'_{com} a "deviant" continuation of the run.

Observe that π'_{com} is finite because $\rho(g^*) \neq \rho(g^*_{-j}, g'_j)$. We distinguish two cases: either $\pi'_{com} < \pi_0\pi^k$ or $\pi'_{com} = \pi_0\pi^k\pi'$ for some $\pi' \leq \pi\pi_1\pi_2^\omega$, that is, by playing g'_j player j deviates from g^* either before or after the kth iteration of π is completed.

If the former, we find that $\rho(f^*_{-j}, g'_j) = \rho(g^*_{-j}, g'_j)$, and hence immediately $\rho(f^*_{-j}, g'_j) \models \gamma_j$.

If the latter, $\rho(g^*_{-j}, g'_j) = \pi_0\pi^k\pi'\rho_{dev}$ for some $\pi' \leq \pi\pi_1\pi_2^\omega$ and some "deviant" run ρ_{dev}. We define a strategy f'_j for player j such that $\rho(f^*_{-j}, f'_j) = \pi_0\pi^{k-1}\pi'\rho_{dev}$. To this end, first observe that every history is of the form $\eta = \pi''_{com}\pi''_{dev}$, where π''_{com} is the maximal common prefix of η and $\pi_0\pi^{k-1}\pi'\rho_{dev}$, and π''_{dev} some "deviant" history. Now define strategy f'_j for player j such that, for every $\eta = \pi''_{com}\pi''_{dev}$,

$$f'_j(\pi''_{com}\pi''_{dev}) = \begin{cases} g'_j(\pi''_{com}\pi''_{dev}) & \text{if } \pi''_{com} < \pi_0\pi^{k-1}, \\ g'_j(\pi_0\pi^k\pi'\pi''_{dev}) & \text{if } \pi''_{com} = \pi_0\pi^{k-1}\pi' \text{ with } \pi' \leq \pi\pi_1\pi_2^\omega. \end{cases}$$

It can then be shown by an inductive argument over the length of $\rho(f^*_{-j}, f'_j)$ that $\rho(f^*_{-j}, f'_j) = \pi_0\pi^{k-1}\pi'\rho_{dev}$. As, by definition of n_0, we have $k - 1 \geq n_{\gamma_j}$, and, moreover, $\pi_0\pi^k\pi'\rho_{dev} \models \gamma_j$, it would follow that also $\rho(f^*_{-j}, f'_j) \models \gamma_j$.

In either case, it would follow that f^* is not an equilibrium, a contradiction, as desired.

The "if"-direction of (i) has basically the same structure as the "only if"-direction. Let $f^* = (f^*_1, \ldots, f^*_n)$ be a Nash equilibrium of G with $\rho(f^*) = \pi_0\pi^{k+1}\pi_1\pi_2^\omega \in NE(G)$. We define strategy profile $g^* = (g^*_1, \ldots, g^*_n)$ such that $\rho(g^*) = \pi_0\pi^k\pi_1\pi_2^\omega$ and show that g^* is a Nash equilibrium of G as well.

Observe that every history η is of the form $\pi_{com}\pi_{dev}$, where π_{com} is the largest common prefix of $\pi_0\pi^k\pi_1\pi_2^\omega$. Then, define, for every player i and every history $\eta = \pi$,

$$g^*_i(\pi_{com}\pi_{dev}) = \begin{cases} f^*(\pi_{com}\pi_{dev}) & \text{if } \pi_{com} < \pi_0\pi^k, \\ f^*(\pi_0\pi^k\pi\pi'_{dev}) & \text{if } \pi_{com} = \pi_0\pi^k\pi' \text{ for some } \pi' \leq \pi_1\pi_2^\omega. \end{cases}$$

By means of an inductive argument over the length of $\rho(g^*)$, it can then be proven that $\rho(g^*) = \pi_0\pi^k\pi_1\pi_2^*$. To demonstrate that g^* is also a Nash equilibrium, assume that g^* is not. Accordingly, there is some player j with $\rho(g^*) \not\models \gamma_j$ and some strategy g'_j for j such that $\rho(g^*_{-j}, g'_j) \models \gamma_j$. As $\neg\gamma_i$ is a non-counting property with an index not larger than k, it follows that $\rho(f^*) \not\models \gamma_j$.

If, by playing g'_j, player j unilaterally deviates from g^* before the kth iteration of π is completed—that is, if $\pi'_{com} < \pi_0\pi^k$, where π'_{com} is the maximal common prefix of $\rho(g^*)$ and $\rho(g^*_{-j}, g'_j)$—it can be shown by an inductive argument over the length of $\rho(f^*_{-j}, g'_j)$ that $\rho(f^*_{-j}, g'_j) = \rho(g^*_{-j}, g'_j)$. Hence, $\rho(f^*_{-j}, g'_j) \models \gamma_j$. It would thus follow that f^* were not a Nash equilibrium, a contradiction.

Now assume that, by playing g'_j, player j unilaterally deviates from g^* after the kth iteration of π is completed—that is, that $\pi'_{com} \geq \pi_0\pi^k$, where π'_{com} is the maximal common prefix of $\rho(g*)$ and $\rho(g^*_{-j}, g'_j)$. Then $\rho(g^*_{-j}, g'_j) = \pi_0\pi^k\pi''\rho_{dev}$, where $\pi'' \leq \pi_1\pi_2^\omega$. We now define strategy f'_j such that for every history η of the form $\pi''_{com}\pi''_{dev}$ where π''_{com} is the maximal common prefix of η and $\pi_0\pi^{k+1}\pi_1\pi_2^\omega$,

$$f'_j(\pi''_{com}\pi''_{dev}) = \begin{cases} g'_j(\pi''_{com}\pi''_{dev}) & \text{if } \pi''_{com} < \pi_0\pi, \\ g'_j(\pi_0, \pi', \pi''_{dev}) & \text{if } \pi''_{com} = \pi_0\pi\pi' \text{ where } \pi' \leq \pi^k\pi_1\pi_2^\omega. \end{cases}$$

Defined thus, we find by an inductive argument over the length of $\rho(f^*_{-j}, f'_j)$ that $\rho(f^*_{-j}, f'_j) = \pi_0\pi^{k+1}\pi''\rho_{dev}$. As $k > n_{\gamma_j}$, it follows that $\rho(f^*_{-j}, f'_j) \models \gamma_j$, and again we may conclude that f^* is not a Nash equilibrium, a contradiction.

For the "only if"-direction of (ii), assume that $f^* = (f_1^*, \ldots, f_n^*)$ is an equilibrium with $\rho(f^*) = \pi_0(\pi_1\pi^k\pi_2)^\omega$. We define profile $g^* = (g_1^*, \ldots, g_n^*)$ such that $\rho(g^*) = \pi_0(\pi_1\pi^{k+1}\pi_2)^\omega$, and show that g^* is an equilibrium as well. In order to define g^*, first observe that every history is of the form $\eta = \pi_{com}\pi_{dev}$, where π_{com} is the maximal common prefix of η and $\pi_0(\pi_1\pi^{k+1}\pi_2)^\omega$, and π_{dev} a "deviant" continuation. Moreover, either $\pi_{com} < \pi_0$ or $\pi_{com} = \pi_0(\pi_1\pi^{k+1}\pi_2)^\ell\pi'$ for $\ell \geq 0$ and some $\pi' < \pi_1\pi^{k+1}\pi_2$. Given $\eta = \pi_{com}\pi_{dev}$, define $\check{\pi}_{com}$ such that $\check{\pi}_{com} = \pi_{com}$ if the former and

$$\check{\pi}_{com} = \begin{cases} \pi_0(\pi_1\pi^k\pi_2)^\ell\pi' & \text{if } \pi' < \pi_1\pi^k \\ \pi_0(\pi_1\pi^k\pi_2)^\ell\pi_1\pi^{k-1}\pi'' & \text{if } \pi' = \pi_1\pi^k\pi'' \text{ for some } \pi'' \leq \pi\pi_2, \end{cases}$$

if the latter. Subsequently, define for every player i and history $\eta = \pi_{com}\pi_{dev}$,

$$g_i^*(\pi_{com}\pi_{dev}) = f_i^*(\check{\pi}_{com}\pi_{dev}).$$

Defined thus, $\rho(g^*) = \pi_0(\pi_1\pi^{k+1}\pi_2)^\omega$.

Now assume for a contradiction that g^* is not a Nash equilibrium. Then, there is a player j with $\rho(g^*) \not\models \gamma_j$ and some strategy g_j' for j such that $\rho(g_{-j}^*, g_j') \models \gamma_j$. As $n_{\neg\gamma_j} < k$, then also $\rho(f^*) \not\models \gamma_j$. Let $\rho(g_{-j}^*, g_j') = \pi_{com}'\rho_{dev}'$, where π_{com}' is the maximum common prefix of $\rho(g^*)$ and $\rho(g_{-j}^*, g_j')$, and ρ_{dev}' a "deviant" continuation.

If $\pi_{com}' < \pi_0\pi_1\pi^k$, then $\rho(g_{-j}^*, g_j') = \rho(f_{-j}^*, g_j')$. Hence, $\rho(f_{-j}, g_j') \not\models \gamma_j$, which would signify that f^* were not a Nash equilibrium, a contradiction. If, on the other hand, $\pi_{com}' = \pi_0(\pi_1\pi^{k+1}\pi_2)^m\pi'$ for some $m \geq 0$ and $\pi \leq \pi_1\pi^{k+1}\pi_2$. Then, $\rho(g_{-j}^*, g_j') = \pi_0(\pi_1\pi^{k+1}\pi_1)^m\pi'\rho_{dev}'$. We define a strategy f_j' for j such that $\rho(f_{-j}^*, f_j') = \pi_0(\pi_1\pi^k\pi_1)^m\pi'\rho_{dev}'$. Now every history is of the form $\eta = \pi_{com}''\pi_{dev}''$, where π_{com}'' is the maximal common prefix of η and $\pi_0(\pi_1\pi^{k+1}\pi_1)^m\pi'\rho_{dev}'$. Given $\eta = \pi_{com}''\pi_{dev}''$, define

$$\hat{\pi}_{com}'' = \begin{cases} \pi_{com}'' & \text{if } \pi_{com}'' < \pi_0, \\ \pi_0(\pi_1\pi^{k+1}\pi_2)^\ell\pi'' & \text{if} \pi_{com}'' = \pi_0(\pi_1\pi^k\pi_2)^\ell\pi'', 0 \leq \ell \leq m, \\ & \quad \text{and } \pi'' \leq \pi_1\pi^k\pi_2. \end{cases}$$

Subsequently, define f_j' such that for every history $\eta = \pi_{com}''\pi_{dev}''$,

$$f_j'(\pi_{com}''\pi_{dev}'') = g_j'(\hat{\pi}_{com}''\pi_{dev}'').$$

Defined thus, $\rho(f_{-j}^*, f_j') = \pi_0(\pi_1\pi^k\pi_1)^m\pi'\rho_{dev}'$. As $n_{\gamma_j} < k$, it now follows that $\rho(f_{-j}^*, f_j') \models \gamma_j$. This would again signify that f^* is not a Nash equilibrium, a contradiction.

Finally, for the "if"-direction of (ii) , let $f^* = (f_1^*, \ldots, f_n^*)$ be a Nash equilibrium such that $\rho(f^*) = \pi_0(\pi_1\pi^{k+1}\pi_2)^\omega$. We define another strategy profile $g^* = (g_1^*, \ldots, g_n^*)$ with $\rho(g^*) = \pi_0(\pi_1\pi^k\pi_2)^\omega$ and show that it is a Nash equilibrium as well. Note that every history has the form $\eta = \pi_{com}\pi_{dev}$, where π_{com} is the maximal common prefix of η and $\pi_0(\pi_1\pi^k\pi_2)^\omega$, whereas π_{dev} is some "deviant" continuation. Given $\eta = \pi_{com}\pi_{dev}$, moreover, either $\pi_{com} < \pi_0$ or $\pi_{com} = \pi_0(\pi_1\pi^k\pi_2)^\ell\pi'$ for some $0 \leq \ell \leq m$ and $\pi' \leq \pi_1\pi^k\pi_2$. If the former, define $\hat{\pi}_{com} = \pi_{com}$. If the latter, that is, if $\pi_{com} = \pi_0(\pi_1\pi^k\pi_2)^\ell\pi'$, let

$$\hat{\pi}_{com} = \begin{cases} \pi_0(\pi_1\pi^{k+1}\pi_2)^\ell\pi' & \text{if } \pi' < \pi_1\pi^k, \\ \pi_0(\pi_1\pi^{k+1}\pi_2)^\ell\pi_1\pi^{k+1}\pi' & \text{if } \pi' \leq \pi_2. \end{cases}$$

Subsequently, define for every player and every history $\eta = \pi_{com}\pi_{dev}$,

$$g_i^*(\pi_{com}\pi_{dev}) = f_i^*(\hat{\pi}_{com}\pi_{dev}).$$

Then, $\rho(g^*) = \pi_0(\pi_1\pi^k\pi_2)^\omega$.

Now assume for a contradiction that g^* is not a Nash equilibrium. Then, there is some player j with $\rho(g^*) \not\models \gamma_j$ and some strategy g_j' for j such that $\rho(g_{-j}^*, g_j') \models \gamma$. As $n_{\neg\gamma_j} < k$, then also

$\rho(f^*) \not\models \gamma_j$. Let $\rho(g^*_{-j}, g'_j) = \pi'_{com}\rho'_{dev}$, where π'_{com} is the maximal common prefix of $\rho(g^*)$ and $\rho(g^*_{-j}, g'_j)$, and ρ'_{dev} some "deviant" continuation.

If $\pi'_{com} < \pi_0$, then $\rho(f^*_{-j}, g'_j) = \rho(g^*_{-j}, g'_j)$. Hence, $\rho(f^*_{-j}, g'_j) \models \gamma_j$, and it would follow that f^* is not a Nash equilibrium. If, on the other hand, $\pi'_{com} = \pi_0(\pi_1\pi^k\pi_2)^\ell\pi'$ for some $\ell \geq 0$ and some $\pi' \leq \pi_1\pi^k\pi_2$, then we define a strategy f'_j for j such that $\rho(f^*_{-j}, f'_j) = \pi_0(\pi_1\pi^{k+1}\pi_2)^\ell\pi'\rho'_{dev}$. To this end, note that every history is of the form $\eta = \pi''_{com}\pi''_{dev}$, where π''_{com} is the maximal common prefix of η and $\pi_0(\pi_1\pi^{k+1}\pi_2)^m\pi'\rho'_{dev}$, and π''_{dev} some "deviant" continuation. Given $\eta = \pi''_{com}\pi''_{dev}$, define $\check{\pi}''_{com} = \pi''_{com}$ if $\pi''_{com} < \pi_0\pi^k$. Otherwise, that is if $\pi''_{com} = \pi_0(\pi_1\pi^{k+1}\pi_1)^m\pi''$ for some $0 \leq \ell \leq m$ and $\pi'' \leq \pi_1\pi^{k+1}\pi_2$, let

$$\check{\pi}''_{com} = \begin{cases} \pi_0(\pi_1\pi^k\pi_2)^\ell\pi'' & \text{if } \pi'' < \pi_1\pi^k, \\ \pi_0(\pi_1\pi^k\pi_2)^\ell\pi_1\pi^{k-1}\pi''' & \text{if } \pi'' = \pi_0\pi^k\pi''' \text{ and } \pi''' \leq \pi^k\pi_2. \end{cases}$$

Now define for all histories $\eta = \pi''_{com}\pi''_d ev$,

$$f'_j(\pi''_{com}\pi''_{dev}) = g'_j(\check{\pi}''_{com}\pi''_{dev}).$$

Then, $\rho(f^*_{-j}, f'_j) = \pi_0(\pi_1\pi^{k+1}\pi_2)^\ell\rho'_{dev}$. As $n_{\gamma_j} < k$, it follows that $\rho(f^*_{-j}, f'_j) \models \gamma_j$, which implies that f^* is not a Nash equilibrium, a contradiction. □

As an immediate consequence of Proposition 4.6, we thus find that LTL cannot express in equilibrium every ω-regular property, that is, $\text{LTL}^{NE} \not\geq L_{\omega\text{-}reg}$. For instance, the property expressed by the ω-regular expression $\boldsymbol{(\phi;\phi)^*;\{p\}^\omega}$ cannot be obtained as the set of equilibrium runs of any LTL-game.

In view of the characterisation result by Kučera and Strejček, Proposition 4.6 gives us one half of the proof that $\text{LTL} \geq \text{LTL}^{NE}$. To prove the second half—namely, that the set of equilibrium runs of every LTL-game is an ω-regular set—we adapt a result by Gutierrez et al. [2017b], which provides us with a construction of a parity automata that recognises all words corresponding to the equilibrium runs of a given LTL-game. We saw above how Lemma 3.5 characterised the equilibrium runs ρ of an iterated Boolean game G as those for which there is a bi-partition $\{N_0, N_1\}$ of the players such that N_1 is the group of agents that have their goal achieved, whereas N_0 consists of those players who do not have their goal achieved but for whom run ρ is consistent with a punishment strategy against each of them. We can now leverage this result apply the automata-theoretic approach to linear temporal logic as proposed by Vardi [1996] so as to obtain the following proposition:

Proposition 4.7 (after [Gutierrez et al. 2017b]). *For every* LTL-*game G, the set $NE(G)$ of equilibrium runs is ω-regular.*

Proof Sketch. The proof invokes the concepts of deterministic and non-deterministic Rabin word- and tree automata. For the formal definitions of these, the reader is referred to [Kupferman 2018; Löding 2012; Perrin and Pin 2004; Strejček 2004].

Let $\{N_0, N_1\}$ be any bipartition of the set of players. For every player i in N_1, we can construct a non-deterministic Rabin word automaton A_{γ_i} that recognises all runs satisfying γ_i. In a slightly more complicated way we can also construct for each player $j \in N_0$ a non-deterministic Rabin word automaton B_{γ_j} that recognises all runs that are consistent with a punishment strategy against j. First, consider the non-deterministic Rabin word automaton $\overline{A}_{\gamma_j}$ that recognises all runs *not* satisfying γ_j. By virtue of the classic synthesis algorithm for LTL provided in [Pnueli and Rosner 1989], there exists a non-deterministic Rabin automaton on *trees* $\overline{A}'_{\gamma_j}$ that recognises a tree iff the strategy represents a winning strategy for the coalition $N \setminus \{j\}$ against j. Now observe that the branches of the trees recognised by $\overline{A}'_{\gamma_j}$ correspond exactly to those runs of G that are consistent with a punishment strategy against j. Moreover, such set of all such brunches is an ω-regular language [Niwinski and Walukiewicz 1998] and so can be recognised by a non-deterministic Rabin word-automaton B_{γ_j}. As the languages recognised by non-deterministic Rabin word-automaton are closed under union and intersection, there exists a non-deterministic Rabin word automaton A_{NE} that recognises exactly those runs ρ of G for which there is some bipartition $\{N_0, N_1\}$ of the players such that ρ satisfies the goal of each player in N_1 and is consistent with a punishment strategy against each player in N_0. By Lemma 3.5, it follows that A_{NE} recognises the Nash equilibrium runs of G, and we may conclude that $NE(G)$ is ω-regular. □

Together with Theorem 4.5, Propositions 4.6 and 4.7 immediately yield that LTL can express every temporal property in formulas that LTL can express in equilibrium. That also the opposite direction holds is an immediate consequence of Proposition 4.1, provided that there are least two propositional variables. Hence, we obtain the main result of this subsection.

Theorem 4.8. *Let* $|\Phi| \geq 2$. *Then,* $\mathrm{LTL}(\Phi)$ *is just as expressive in formulas as it is in equilibrium, that is,* $\mathrm{LTL}(\Phi) \equiv \mathrm{LTL}^{NE}(\Phi)$.

Proof. Having assumed $|\Phi| \geq 2$, we obtain $\mathrm{LTL}^{NE}(\Phi) \geq \mathrm{LTL}(\Phi)$ as an immediate consequence of Proposition 4.1. To see that also $\mathrm{LTL} \geq \mathrm{LTL}^{NE}$, consider an arbitrary temporal property X that is expressed in equilibrium by LTL, that is, X is expressed by LTL^{NE}. Then, there is an iterated Boolean game G such that X is the set of Nash equilibrium runs of G. Propositions 4.6 and 4.7 yield that X is non-counting and ω-regular, respectively. By virtue of Theorem 4.5, we may then conclude that LTL can express X in formulas as well. □

4.4 Stutter-Invariant Specifications: The Next-Free Fragment L_{U}

We conclude this section by investigating the expressive power in equilibrium of the next-free or maximally stutter-invariant fragment L_{U}. In the verification and model checking literature, this fragment plays a prominent role because it guarantees stutter-invariant specifications of programs and concurrent systems, without losing expressiveness otherwise. We find that nevertheless there are stutter-sensitive properties that L_{U} can express in equilibrium, even though it cannot express all properties that are expressed by LTL in formulas. That is, $\mathrm{LTL} > L_{\mathsf{U}}^{NE} > L_{\mathsf{U}}$ (provided that there are at least two propositional variables). Accordingly, for L_{U}-games, the stutter-invariance of the players' goals is not fully reflected in the equilibrium runs.

Recall that, formally, a temporal property $X \subseteq runs_\Phi$ is said to be *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,

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

where v^k denotes the k-fold iteration of v. Thus, for instance, consider the property, henceforth denoted by *toggle*(p), consisting of all runs $\rho = v_0 v_1 v_2 \ldots$ such that v_t satisfies p if and only if t is even. This property does *not* define a stutter-free property, as it contains for instance the run $\{p\}\,\phi\,\{p\}\,\phi\,\ldots$, but not the run $\{p\}\,\{p\}\,\phi\,\phi\,\{p\}\,\{p\}\,\phi\,\phi\,\ldots$. Note that *toggle*$(p)$ is expressed by the LTL-formula $p \wedge \mathsf{G}(p \leftrightarrow \mathsf{X}\,\neg p)$; it cannot be expressed by a formula of L_{U} though. In a seminal paper, [Peled and Wilke 1997] showed that the so-called next-free fragment L_{U} is the *largest* stutter-free fragment of LTL, that is, L_{U} can express every stutter-free property that full LTL can express in formulas (also see [Etessami 2000]).

Leslie Lamport has made a case in favour of stutter-invariant specifications and, in particular, emphatically argued against the inclusion of the next-operator X in the syntax of temporal logic [Lamport 1983]. The main reason for his opposition is the observation that temporal logics are meant to facilitate reasoning about abstract specifications of programs rather than about their concrete implementations, and that the next-operator enables reasoning about irrelevant aspects of the implementation of a program specification. Lamport gives the example of requiring that a program takes exactly 17 steps to implement a queue, which relates to the specification, but can only be expressed by stutter-sensitive fragments. He furthermore alleged that stutter-invariant specifications still enable reasoning about "*the next state in which a significant change occurs* [...] *at the level of detail of the specification*" [Lamport 1983, page 661]. Accordingly, including the next-operator would run counter to modular or hierarchical approaches towards program specification: "*you will see that increasing the expressiveness of our temporal logic with a* next*operator would destroy the entire logical foundation for its use in hierarchical methods*" (ibid., page 661).

Elaborating on this topic, [Peled and Wilke 1997] furthermore point to the advantages of stutter-invariant specifications of concurrent programs, where the different ways in which processes can be interleaved is seen as a essentially being part of the implementation. This invariance under arbitrary interleavings is also the technical basis for state-space reduction techniques for model checking concurrent systems.

We find that the stutter-invariance of properties expressed by L_{U} in formulas does not extend to the temporal properties that can be expressed by L_{U} in equilibrium. To see this, let $G_{toggle(p)}$ be the iterated Boolean game on a set Φ of propositional variables containing p and q, and two players, i and j, such that:

$$
\begin{aligned}
\Phi_i &= \{p\}\\
\Phi_j &= \Phi \setminus \{p\}\\
\gamma_i &= p \wedge \neg\gamma_j\\
\gamma_j &= \mathsf{F}((pq \wedge pq \,\mathsf{U}\, p\bar{q}) \vee (\bar{p}q \wedge \bar{p}q \,\mathsf{U}\, \bar{p}\bar{q}) \vee (p\bar{q} \wedge p\bar{q} \,\mathsf{U}\, pq) \vee (\bar{p}\bar{q} \wedge \bar{p}\bar{q} \,\mathsf{U}\, \bar{p}q)).
\end{aligned}
$$

Note that, for γ_j to be satisfied, variable p has to assume the same truth-value at at least two consecutive times. The only way player i can prevent this from happening, and thus hope to satisfy her goal, is by single-handedly toggling the truth-value of p starting by setting p to true. On this basis, we can show that the equilibria runs of $G_{toggle(p)}$ are given exactly by $toggle(p)$. Hence, L_{U} can express stutter-sensitive properties in equilibrium.

This is interesting, because, intuitively, the Nash equilibria of a multi-agent system of which the players are assumed to interact strategically in pursuit of their L_{U}-goals does not seem to pertain to the specifics of how the constituent agents are implemented. As such, the result below may cast a new light on the desirability of stutter-invariant specifications of multi-agent systems. Later, we will see, however, that the expressiveness of L_{U} in equilibrium does not quite equal the expressive power of L_{U} in formulas with the next-operator X added, that is, the expressive power of full LTL in formulas.

Proposition 4.9. *Let $p, q \in \Phi$. Then, $NE(G_{toggle}(p)) = toggle(p)$. Hence, $L_{\mathsf{U}}(\Phi)$ expresses $toggle(p)$ in equilibrium.*

Proof. It suffices to show that $NE(G_{toggle(p)}) = toggle(p)$. To this end, consider an arbitrary run $\rho \in runs_\Phi$ that toggles p, that is, such that $\rho, t \models p$ if and only if t is even. We first observe that $\rho \not\models \gamma_j$. To see this, consider an arbitrary $t \geq 0$ and assume $\rho, t \models pq$. Then, $\rho, t+1 \not\models p$. Now consider an arbitrary $t' \geq t$ with $\rho, t' \models p\bar{q}$. Hence, note that $t' \geq t+2$, and, as $\rho, t+1 \not\models pq$, we obtain $\rho, t \not\models pq \wedge pq \,\mathsf{U}\, p\bar{q}$. Assuming instead that $\rho, t \models \bar{p}q$, $\rho, t \models p\bar{q}$, or $\rho, t \models \bar{p}\bar{q}$, an analogous argument yields $\rho, t \not\models \bar{p}q \wedge \bar{p}q \,\mathsf{U}\, \bar{p}\bar{q}$, $\rho, t \not\models p\bar{q} \wedge p\bar{q} \,\mathsf{U}\, pq$, and $\rho, t \not\models \bar{p}\bar{q} \wedge \bar{p}\bar{q} \,\mathsf{U}\, \bar{p}q$, respectively. We may conclude that $\rho \not\models \gamma_j$.

Now, consider an arbitrary run $\rho = v_0v_1v_2v_3\dots$ in $runs_\Phi$ that toggles p. We prove that ρ is sustained by an equilibrium, that is, $\rho \in NE(G_{toggle(p)})$. To this end, define strategy $f_i^{toggle(p)}$ for player i such that $f_i^{toggle(p)}(\epsilon) = \{p\}$ and for paths $v_0 \dots v_t$ with $t \geq 0$,

$$f_i^{toggle(p)}(v_0 \dots v_t) = \begin{cases} \{p\} & \text{if } t \text{ is odd,} \\ \phi & \text{otherwise.} \end{cases}$$

For player j define strategy f_j such that $f_j(\epsilon) = v_0 \cap \Phi_j$ and $f_j(v_0, \dots, v_k) = v_{k+1} \cap \Phi_j$, if $k \geq 0$. Defined thus, it can easily be seen that $\rho(f_i^{toggle(p)}, f_j) = v_0v_1v_2\dots$. Also, observe that by definition of f_i, we find that $\rho(f_i^{toggle(p)}, g_j)$ toggles p for *every* strategy g_j for player j. Hence, $\rho(f_i^{toggle(p)}, g_j) \not\models \gamma_j$ for every strategy g_j for player j as well. It follows that player j does not want to deviate from $(f_i^{toggle(p)}, f_j)$. Moreover, $\rho(f_i^{toggle(p)}, f_j) \models p$ as well as $\rho(f_i^{toggle(p)}, f_j) \not\models \gamma_j$. Hence, $\rho(f_i^{toggle(p)}, f_j) \models \gamma_i$, and player i does not want to deviate from $(f_i^{toggle(p)}, f_j)$ either. It follows that $(f_i^{toggle(p)}, f_j)$ is a Nash equilibrium sustaining ρ.

To conclude the proof, consider an arbitrary run $\rho' = v_0'v_1'v_2'\dots$ that does not toggle p and equally arbitrary strategy profile $g = (g_i, g_j)$ with $\rho(g_i, g_j) = \rho'$. Then, there is a $t \geq 0$ such that either (i) both $\rho, t \models p$ and $\rho', t+1 \models p$ or (ii) both $\rho', t \models \bar{p}$ and $\rho', t+1 \models \bar{p}$.

If (i) , and $\rho' \not\models \gamma_i$, then $\rho(f_i^{toggle(p)}, g_j) \models \gamma_i$, that is, player i wants to deviate from $g = (g_i, g_j)$ and get her goal achieved. On the other hand, if $\rho' \models \gamma_i$, then $\rho' \not\models \gamma_j$. In that case, we may assume, without loss of generality, that $\rho', t \models pq$, and define strategy g_j' for player j such that for all $g_j'(\epsilon) = g_j(\epsilon)$ and, for paths $w_0 \dots w_k$ with $k \geq 0$,

$$g_j'(w_0 \dots w_k) = \begin{cases} g_j(w_0 \dots w_k) & \text{if } k < t, \\ \phi & \text{otherwise.} \end{cases}$$

Let $\rho(g_i, g_j') = v_0''v_1''v_2''\dots$. Now observe that, $v_0'' \dots v_t'' = v_0' \dots v_t'$. Hence, $\rho(g_i, g_j'), t \models pq$. Observe moreover that both $p \in g_i(v_0'' \dots v_t'')$ and $q \notin g_j'(v_0'' \dots v_t'')$. Therefore, also $\rho(g_i, g_j'), t+1 \models p\bar{q}$. Accordingly, $\rho(g_i, g_j'), t \models pq \wedge pq \,\mathsf{U}\, p\bar{q}$, and hence $\rho(g_i, g_j') \models \gamma_j$. It follows that $g = (g_i, g_j)$ is not a Nash equilibrium. If (ii) , the argumentation is analogous, and, with $g = (g_i, g_j)$ having been chosen arbitrarily, we may conclude that $\rho' \notin NE(G^{toggle(p)})$, as desired. □

As an almost immediate consequence of Proposition 4.9, we obtain that the stutter-invariant fragment L_{U} is strictly more expressive in equilibrium than it is in formulas.

Theorem 4.10. *Let $|\Phi| \geq 2$. Then, L_{U} is strictly more expressive in equilibrium than it is in formulas, that is, $L_{\mathsf{U}}^{NE}(\Phi) > L_{\mathsf{U}}(\Phi)$.*

Proof. As we may assume that $|\Phi| \geq 2$, by Proposition 4.1, we immediately obtain $L_{\mathsf{U}}^{NE} \geq L_{\mathsf{U}}$. For p a propositional variable in Φ, moreover, the stutter-sensitive property *toggle*(p) cannot be expressed by L_{U} as this fragment is the largest stutter-invariant fragment of LTL. By Proposition 4.9, however, *toggle*(p) can be expressed in equilibrium by L_{U}. Hence, $L_{\mathsf{U}}^{NE} > L_{\mathsf{U}}$. □

Even though L_{U} can express in equilibrium some stutter-sensitive LTL-properties, its expressive power in equilibrium does not match that of full LTL in formulas. We thus find, quite strikingly, that L_{U} cannot even express in equilibrium the property characterised by the LTL-formula $\mathsf{X}\, p$. The main idea behind the proof of this statement is a *reductio ad absurdum* argument that runs as follows: Assume for contradiction that $\mathsf{X}\, p$ were expressible and consider some run $\rho = v_0v_0v_1v_2\dots$ that stutters a valuation v_0 at time $t = 1$. Assume furthermore $p \in v_0$ and $p \notin v_1$. As ρ satisfies $\mathsf{X}\, p$, it is sustained by some Nash equilibrium $f^* = (f^*, \dots, f_n^*)$. Now consider $\rho' = v_0v_1v_2v_3\dots$, which is exactly like ρ, but does not stutter the first valuation. This run *as well as all deviations from it* satisfy exactly the same players' goals, because the latter are phrased in the stutter-invariant fragment L_{U}. Then, $\rho' = \rho(f)$, where $f = (f_1, \dots, f_n)$ is the strategy profile that is exactly like f^*, but behaves as if the first occurrence of v_0 has already been produced. As $\mathsf{X}\,p$ is not satisfied at ρ', it follows that f is not an equilibrium. Accordingly, there is some player j who has an incentive to deviate from f and play some strategy g_j instead. It can the be show that g_j can be adapted in such a way as to secure a profitable deviation by j from f^* as well, which yields the desired contradiction.

Proposition 4.11. *Let p be a propositional variable contained in Φ. Then, the fragment $L_{\mathsf{U}}(\Phi)$ cannot express in equilibrium the temporal property characterised by the* LTL*-formula $\mathsf{X}\,p$.*

Proof. Assume for a contradiction that G is an L_{U}-game such that $NE(G) = runs_\Phi(\mathsf{X}\,p)$. Consider a run $\rho = v_0v_0v_1v_2\dots$ where $p \in v_0$ and $p \notin v_1$. Thus, $\rho \in runs_\Phi(\mathsf{X}\,p)$, and, as a consequence, $\rho \in NE(G)$. Accordingly, there is an equilibrium $f^* = (f_1^*, \dots, f_n^*)$ with $\rho = \rho(f^*)$. Observe that $f_i^*(\epsilon) = f_i^*(v_0)$ for all players i. Now define the strategy profile $f = (f_1, \dots, f_n)$ such that, for every player i, we have $f_i(\epsilon) = f_i^*(v_0)$ and, for every $x_0, \dots, x_k$ with $k \geq 0$,

$$f_i(x_0, \dots, x_k) = \begin{cases} f_i^*(v_0, x_0, \dots, x_k) & \text{if } x_0 = v_0, \\ f^*(x_0, \dots, x_k) & \text{otherwise.} \end{cases}$$

A straightforward induction on the length of $\rho(f)$ establishes that $\rho(f) = v_0v_1v_2v_3\dots$. Having assumed that $p \notin v_1$, clearly $\rho(f) \notin runs_\Phi(\mathsf{X}\,p)$. Hence, $f = (f_1, \dots, f_n)$ is not an equilibrium of G and there is some player j such that $\rho(f) \not\models \gamma_j$ and some strategy g_j for j such that $\rho(f_{-j}, g_j) \models \gamma_j$. As γ_j expresses a stutter-invariant property, observe that now also $\rho(f^*) \not\models \gamma_j$. Let $\rho(f_{-j}, g_j) = u_0u_1u_2\dots$.

At this point, define strategy g_j' such that $g_j'(\epsilon) = g_j(\epsilon)$ and, for every path $x_0 \dots x_k$,

$$g_j'(x_0 \dots x_k) = \begin{cases} g_j(x_1 \dots x_k) & \text{if } x_0 = v_0, \\ g_j(x_0 \dots x_k) & \text{otherwise,} \end{cases}$$

on the understanding that $x_1 \dots x_0 = \epsilon$. Let $\rho(f_{-j}^*, g_j') = w_0w_1w_2w_3\dots$.

At this point, we distinguish two cases. In either one, we show that $\rho(f_{-j}^*, g_j') \models \gamma_j$. Having already seen that $\rho(f^*) \not\models \gamma_j$, this then contradicts that f^* is a Nash equilibrium.

First assume that $g_j(\epsilon) = f_j(\epsilon)$, that is, player j does not deviate from f_j immediately. Hence, $u_0 = v_0$. We show that $\rho(f_{-j}^*, g_j') = u_0u_0u_1u_2\dots$, that is, $w_0 = u_0$ and $w_t = u_{t-1}$ for all $t \geq 1$. Since, $\rho(f_{-j}, g_j) \models \gamma_j$ and γ_j is a stutter-invariant property, it follows that also $\rho(f_{-j}^*, g_j') \models \gamma_j$, a contradiction.

Observe that both $g_j'(\epsilon) = g_j(\epsilon) = f_j(\epsilon)$ and $f_i^*(\epsilon) = f_i^*(v_0) = f_i(\epsilon)$ for all $i \neq j$. Hence, $w_0 = v_0$, and therefore $w_0 = u_0$.

We now show by induction that $w_t = u_{t-1}$ for all $t \geq 1$. If $t = 1$, observe that $g_j'(w_0) = g_j(\epsilon)$, as $w_0 = v_0$. Moreover, $f_i^*(w_0) = f_i^*(v_0) = f_i(\epsilon)$ for all $i \neq j$. It follows that $w_1 = u_0$, as desired. For the induction step, we may assume that $w_0 \dots w_t = u_0u_0 \dots u_{t-1}$. Then, $g_j'(w_0 \dots w_t) = g_j(w_1 \dots w_t)$ as $w_0 = v_0$, and hence, by the induction hypothesis, $g_j'(w_0 \dots w_t) = g_j(u_0 \dots u_{t-1})$. Moreover, for all players $i \neq j$, we have $f_i^*(w_0 \dots w_t) = f_i(w_1 \dots w_t)$ because $w_0 = v_0$, and hence, by the induction hypothesis, $f_i^*(w_0 \dots w_t) = f_i(u_0 \dots u_{t-1})$. It follows that $w_{t+1} = u_t$, as desired.

Finally, assume that $g_j(\epsilon) \neq f_j(\epsilon)$, and hence $u_0 \neq v_0$. We prove by induction that $w_t = u_t$ for all $t \geq 0$, that is, $\rho(f_{-j}^*, g_j') = \rho(f_{-j}, g_j)$. It then immediately follows that $\rho(f_{-j}^*, g_j') \models \gamma_j$. For the basis with $t = 0$, observe that both $g_j'(\epsilon) = g_j(\epsilon)$ and $f_i^*(\epsilon) = f_i^*(v_0) = f_i(\epsilon)$. It follows that $w_0 = u_0$. For the induction step, we may assume that $w_0 \dots w_t = u_0 \dots u_t$. Recall that $u_0 \neq v_0$. Hence, $g_j'(w_0 \dots w_t) = g'(u_0 \dots u_t) = g_j(u_0 \dots u_t)$, as well as $f_i^*(w_0 \dots w_t) = f_i^*(u_0 \dots u_t) = f_i(u_0 \dots u_t)$ for all $i \neq j$. We may conclude that $w_{t+1} = u_{t+1}$, as desired. □

We are now in a position to conclude this section by showing that LTL is still strictly more expressive in formulas than L_{U} in equilibrium.

Theorem 4.12. *Let* $|\Phi| \geq 2$. *Then,* LTL *is strictly more expressive in formulas than* L_{U} *in equilibrium, that is,* $\text{LTL}(\Phi) > L_{\text{U}}^{NE}(\Phi)$.

Proof. As $\text{LTL} \geq L_{\text{U}}$, by Proposition 4.3 also $\text{LTL}^{NE} \geq L_{\text{U}}^{NE}$. Having assumed that $|\Phi| \geq 2$, Theorem 4.8 then entails that $\text{LTL} \geq L_{\text{U}}^{NE}$. Proposition 4.11, moreover, shows that there are properties expressible by LTL in formulas that cannot be expressed by L_{U} in equilibrium. Hence, $\text{LTL} > L_{\text{U}}^{NE}$. □

5 PROJECTIVE EXPRESSIVENESS

In this section, we consider *projective expressiveness in formulas* along with the attendant concept of *projective expressiveness in equilibrium*. Projective expressiveness is weaker than regular expressiveness in that a property being expressible (either in formulas or in equilibrium) implies that property to be projectively expressible, but not necessarily the other way round. Projective expressiveness captures the idea that one may be able to express more properties of runs over Φ, either in formulas or in equilibrium, by having at one's disposal an auxiliary set of variables not in Φ.

On an intuitive level, these auxiliary propositional variables could seen as playing a similar role of additional tools that facilitate a player or group of players performing a complicated task, but are not part of the task itself. Examples could be, for instance, a flight attendant who is only able to count all the passengers of a large aircraft accurately by means of a mechanical hand counter, a beginning singer who only able to sing rhythmically by simultaneously clapping her hands, or an orchestra that is only able to harmonise by virtue of a conductor directing them.

The roots of *projective expressiveness* go back to the work of Beth [1953] and Craig [1957] in the 1950s on definability in first-order logic. The concept has numerous applications in model theory [Chang and Keisler 1990; Hodges 1993] and has recently also been studied in the context of modal and temporal logics [Gheerbrant and ten Cate 2009; Halpern et al.2009].

Apart from this technical perspective, there is also a natural connection with mechanism design scenarios. Suppose a system's specification is represented in a language $L(\Phi)$ and requires the system's runs to satisfy a certain temporal property. Also assume that one finds oneself unable to design the system in such a way that it only manipulates the variables in Φ. In such a case, one may consider redesigning the system so as to have it satisfy its specification, by having it perform some useful auxiliary tasks on the side, whose specification may involve variables not in Φ. The concept of projective expressiveness, either in formulas or in equilibrium, allows us to address this issue in a formal setting. This also touches on the interesting question, which we have to leave for future investigation, concerning the minimal number additional variables needed to let the system operate according to its original specification, that is, concerning the potential succinctness of the specification of an extended system.

We find that projective expressiveness in equilibrium is a very powerful concept, and as the main result of this section, we show that the next-free fragment L_{U} can express in equilibrium every ω-regular property.

5.1 Projective Expressiveness in Formulas

Projective expressiveness abstracts away from the propositional variables available in the language. Rather than requiring that a temporal property X *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 (in formulas)* 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 $X = runs_{\Phi \cup \Psi}(\varphi)|_{\Phi}$. Our concept of projective expressiveness should be distinguished from that of *projective properties* as defined in Peled [1997]. There, a property over $(\Sigma_1 \times \cdots \times \Sigma_n)^\omega$ is said to be *projective* whenever two runs are in the property if and only if their *stutter-free* projections on each Σ_i^ω are identical. That means, that each of the n components behave in a stutterinvariant fashion. Peled writes: "Thus, for the temporal logics ETL and LTL properties using the correspondence between properties expressed in these logics and ω-regular and star-free ω-regular, respectively. Thus, for the temporal logics LTL and ETL, the projective properties are always expressible as a Boolean combination of local (stuttering-closed) properties."

Projective expressiveness in formulas should be carefully distinguished from standard expressiveness in formulas. Thus, recall that the properties that LTL can express in formulas 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 [Wolper 1983]. Observe that for *even*(p) to hold, p need *not* necessarily be set false at odd states, and in this way is to be distinguished from *toggle*(p). It is possible, however, to express *even*(p) projectively, for instance, by $q \wedge \mathsf{G}(q \leftrightarrow \mathsf{X}\neg q) \wedge \mathsf{G}(q \rightarrow p)$, where $q \notin \Phi$. It is known from the literature that every *ω-regular* property can be projectively expressed by LTL, if one can use an *unbounded* number of additional propositional variables [Gheerbrant and ten Cate 2009; Thomas 1997]. We reformulate this result for our setting, and show that for every non-deterministic Büchi automaton A on alphabet Φ and with states Q there is a formula $\varphi_A \in \text{LTL}(\Phi \cup Q)$ such that the language Λ_A accepted by A coincides with $runs_{\Phi\cup Q}(\varphi_A)|_\Phi$.

Proposition 5.1. LTL *can projectively express all ω-regular properties.*

sketch Let $A = (Q, 2^\Phi, \delta, Q_0, F)$ be a nondeterministic Büchi automaton (see, for instance, [Baier and Katoen 2008]). We construct a formula φ_A in LTL($\Phi \cup Q$), where Φ and Q are disjoint, as follows:

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

Then, set $\varphi^A = \varphi^A_{init} \wedge \varphi^A_{trans} \wedge \varphi^A_{accept} \wedge \varphi^A_{invar}$. It can then be shown that the ω-language that is accepted by A is given by $\{\rho|_\Phi : \rho \in runs_{\Phi\cup Q}(\varphi_A)\} = runs_{\Phi\cup Q}(\varphi_A)|_\Phi$. Recalling that the class of ω-regular languages over 2^Φ is exactly the class of languages accepted by some nondeterministic Büchi automaton, it follows that LTL can express projectively in formulas every ω-regular property. □

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 in the syntax suitable *grammar-operators* as proposed in Wolper [1983].

Moreover, all properties that LTL can express projectively in formulas are ω-regular. This is an immediate consequence of the ω-regular languages being closed under homomorphisms (see, preliminaries or [Perrin and Pin 2004], Proposition 3.3) and projection being a special kind of homomorphism. For easy reference, we state the following standard result:

Lemma 5.2. *Let $X \subseteq runs_\Phi$ be an ω-regular property, and $\Psi \subseteq \Phi$. Then, $X|_\Psi$ is also ω-regular.*

Proof. Consider the homomorphism from $runs_\Phi$ to $runs_\Psi$ defined by the function $h : 2^\Phi \rightarrow 2^\Psi$ such that $h(v) = v \cap \Psi$ for every $v \in 2^\Phi$. As ω-regular languages are closed under homomorphisms, $h(X)$ is ω-regular as well. Conclude the proof by observing that $h(X) = X|_\Psi$. □

The following proposition now follows almost immediately.

Proposition 5.3. *All temporal properties that are projectively expressed by* LTL *in formulas are ω-regular.*

Proof. Consider an arbitrary temporal property $X \subseteq runs_\Phi$ that LTL expresses projectively. Then, there is some auxiliary set of variables Ψ and a formula $\varphi \in \text{LTL}(\Phi \cup \Psi)$ such that $X = runs_{\Phi\cup\Psi}(\varphi)|_\Phi$. Observe that $runs_{\Phi\cup\Psi}(\varphi)$ is an ω-regular property, and, hence, by Lemma 5.2, so is X. □

5.2 Projective Expressiveness in Equilibrium

Projective expressiveness in formulas can now be extended to *projective expressiveness in equilibrium* analogously to how expressiveness in formulas was extended to expressiveness in equilibrium. Thus, for a given set Φ of propositional variables, we say that LTL-fragment $L(\Phi)$ *projectively expresses in equilibrium* temporal property $X \subseteq runs_\Phi$, if there is some set of Ψ of auxiliary propositional variables and an $L(\Phi \cup \Psi)$-game G such that $X = NE(G)|_\Phi$. Recall that an $L(\Phi \cup \Psi)$-game is an iterated Boolean game whose players' goals are given by formulas of $L(\Phi \cup \Psi)$, and that $NE(G)|_\Phi = \{\rho|_\Phi : \rho \in NE(G)\}$. Note that, as one can always set $\Psi = \phi$, a property being expressible in equilibrium implies that that property can also be expressed *projectively* in equilibrium.

In the previous section, Proposition 4.1 established that expressiveness in formulas implies expressiveness in equilibrium, provided that there are at least two propositional variables. For projective expressiveness this still holds, but without the condition of there being at least two propositional variables. The reason is that the two propositional variables needed in the proof of Proposition 4.1 can be absorbed by the set of auxiliary variables.

Proposition 5.4. *Let L be an* LTL*-fragment on Φ and $X \subseteq runs_\Phi$ a temporal property over 2^Φ. Then, if L can projectively express X in formulas, then L can also projectively express X in equilibrium.*

Proof. Assume that $L(\Phi)$ projectively expresses $X \subseteq runs_\Phi$ in formulas. Then, there is an auxiliary set Ψ of propositional variables and a formula $\varphi \in L(\Phi \cup \Psi)$ such that $X = runs_{\Phi\cup\Psi}(\varphi)|_\Phi$. Let p and q be two variables disjoint from $\Phi \cup \Psi$, and let X' be defined as the set of all runs $\rho \in runs_{\Phi\cup\Psi\cup\{p,q\}}$ such that $\rho \models \varphi$. Hence, X' is expressed by $L(\Phi \cup \Psi \cup \{p, q\})$ in formulas. By Proposition 4.1, we find that X' can be expressed in equilibrium by $L(\Phi \cup \Psi \cup \{p, q\})$. Observing that $X = X|_\Phi$ finally yields that X can be projectively expressed by $L(\Phi)$ in equilibrium. □

Using a construction similar to the matching-pennies game G^{mp}_φ, we can furthermore show that, if a fragment L can express a property X in formulas, it can also projectively express the complement of X in equilibrium. This result is especially relevant for fragments that are not closed under negation.

Proposition 5.5. *Let $X \subseteq runs_\Phi$ be a temporal property that can be expressed in formulas by fragment L. Then, L can projectively express $runs_\Phi \setminus X$ in equilibrium.*

Proof. Let $p, q \notin \Phi$ and let X be expressed by the L-formula φ. Consider the two-player L-game G with $\Phi_i = \{p\}$, $\Phi_j = \Phi \cup \{q\}$ and the players' goals being given by $\gamma_i = \varphi \wedge (p \leftrightarrow q)$ and $\gamma_j = \bar{p} \leftrightarrow q$. 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 $\rho(f) \not\models \gamma_i$, player i would deviate by choosing the opposite value for p in the first round. If, on the other hand, $\rho(f) \models \gamma_i$, player j would deviate by choosing the opposite value for q in the first round. Hence, $NE(G)|_\Phi \subseteq runs_\Phi(\neg\varphi)$.

To see that the inclusion also holds in the opposite direction, assume that $\rho \not\models \varphi$ and that profile f induces ρ. Let ρ' be such that $\rho|_\Phi = \rho'|_\Phi$ and $\rho' \models \bar{p} \leftrightarrow q$. Assume further that f induces ρ'.

Thus, player j has its goal achieved at ρ' and will not deviate from f. As, moreover, player i controls no variables occurring in φ, no deviation from f will satisfy her goal. It follows that f is a Nash equilibrium and $\rho' \in NE(G)$. Since, $\rho|_\Phi = \rho'|_\Phi$, we may conclude that $\rho|_\Phi \in NE(G)|_\Phi$, as desired. □

Previously, we have shown how LTL is equally expressive in formulas as it is in equilibrium. This required a rather elaborate proof, especially for showing that LTL-expressible properties are non-counting. We find that the analogous result for projective expressiveness of LTL can now be show rather straightforwardly by reference to Lemma 5.2 and Propositions 4.7, 5.1, and 5.4.

Theorem 5.6. *Let $X \subseteq runs_\Phi$ be a temporal property. Then,* LTL *can projectively express X in formulas if and only if* LTL *can projectively express X in equilibrium. Hence,* LTL *can projectively express in equilibrium all and only all ω-regular temporal properties.*

Proof. By virtue of Propositions 5.1 and 4.7, it suffices to prove the first statement. The "only if"-direction follows immediately from Proposition 5.4. For the "if"-direction, assume that LTL can projectively express X in equilibrium. Then, there is a set Ψ of auxiliary propositional variables and an LTL($\Phi \cup \Psi$)-game G such that $NE(G)|_\Phi = X$. By Proposition 4.7, we know that the temporal property $NE(G)$ is ω-regular. Lemma 5.2 then yields that $NE(G)|_\Phi$ is ω-regular property as well. By virtue of Proposition 5.1, we may conclude that LTL can projectively express $X = NE(G)|_\Phi$ in formulas. □

5.3 The Next-Free Fragment L_U Revisited

In Section 4.4, we saw that L_U can express in equilibrium some stutter-sensitive properties, in particular *toggle*(p), something it cannot do in formulas. On the other hand, we also found that the expressive power of L_U in equilibrium does not quite match the expressive power of full LTL in formulas (or, in reference to Theorem 4.8, that of LTL in equilibrium for that matter). In this section, however, we show that *projectively* L_U can express in equilibrium every temporal property that LTL can express projectively in equilibrium, that is, all and only all ω-regular properties. This equivalence, however, does not hold for L_U's projective expressiveness in formulas.

To appreciate the latter statement, it can easily be seen that the temporal properties that L_U can projectively express in formulas, are closed under "positive" stuttering in the following sense. For every property X that L_U projectively expresses in formulas, it holds that, if run $v_0v_1v_2v_3\dots$ is in X, then $v_0^{k_0}v_1^{k_1}v_2^{k_2}v_3^{k_3}\dots$ is also in X for all $k_0, k_1, k_2, \dots \geq 1$. (Observe that this captures one-half of the definition of stutter-invariance.) Let $X \subseteq runs_\Phi$ be a property that L_U can projectively express in formulas. Then, there is some auxiliary set of variables Ψ and some property $Y \subseteq runs_{\Phi\cup\Psi}$ with $Y|_\Phi = X$ that L_U can express in formulas. Now, let $v_0v_1v_2v_3\dots$ be in X. Then, there is some run $w_0w_1w_2w_3\dots$ in Y such that $w_t \cap \Phi = v_t$ for all $t \geq 0$. As $L_U(\Phi \cup \Psi)$ can express Y in formulas, it follows that $w_0^{k^0}w_1^{k_1}w_2^{k_2}w_3^{k_3}\dots$ must be in Y as well. Then, moreover, it follows that $v_0^{k^0}v_1^{k_1}v_2^{k_2}v_3^{k_3}\dots$ in $Y|_\Phi = X$, as desired. Accordingly, L_U cannot projectively express in formulas some stutter-sensitive properties like *toggle*(p).

The situation is quite different for the properties that L_U can projectively express in equilibrium, which, as the main result of this section, we find are exactly the ω-regular ones. To obtain this result, we leverage the fact established by Proposition 4.9 that L_U can express in equilibrium the property *toggle*(p).

Let p be a variable not contained in Φ. We first define the *translation* $\tau : \mathrm{LTL}(\Phi) \to L_{\mathsf{U}}(\Phi \cup \{p\})$ inductively such that $q^\tau = q$ for all $q \in \Phi$, and

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

Then, 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. Intuitively, the perpetual toggling of p's truth value allows us to express $\mathsf{X}\,\varphi$ using the temporal operator U only.

Lemma 5.7. *Let φ be an* LTL(Φ)*-formula, p a propositional variable not in Φ, and $\rho = v_0 v_1 v_2 \ldots$ 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 $\rho, t \models \bar{p} \to \bar{p}\,\mathsf{U}(p \wedge \psi^\tau)$. Moreover, $\rho, t+1 \models \bar{p}$ by definition of ρ and, therefore, $\rho, t \models p\,\mathsf{U}(\bar{p} \wedge \psi^\tau)$ and also $\rho, t \models p \to (p\,\mathsf{U}(\bar{p} \wedge \psi^\tau))$. 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 $\rho, t \not\models p\,\mathsf{U}(\bar{p} \wedge \psi^\tau)$, $\rho, t \not\models p \to (p\,\mathsf{U}(\bar{p} \wedge \psi^\tau))$, and eventually $\rho, t \not\models (\mathsf{X}\,\psi)^\tau$. As the argument showing that $\rho, t \models \bar{p}$ 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 four additional variables. Intuitively, two of the players play the "matching pennies"-like game $G^{mp}_{\varphi^\tau}$, as it was employed in the proof of Proposition 4.1. This guarantees that φ^τ holds at all and only the equilibrium runs of the game. The other two players play the game $G_{toggle(p)}$, which played an essential role in the proof of Proposition 4.9. This then ensures that an additional variable p alternately assumes the truth values true and false, and thus that φ^τ evaluates as intended, namely, as equivalent to φ.

Theorem 5.8. *The fragment L_{U} can projectively express in equilibrium every temporal property that* LTL *can express in formulas.*

Proof. Let X be a property expressible by LTL in formulas. Then, there is some LTL(Φ)-formula φ with $X = \{\rho \in runs_\Phi : \rho \models \varphi\}$. Let furthermore $\Psi = \{p, q, r, s\}$ a set of auxiliary variables disjoint from Φ, and construct L_{U}-game G on $\Phi \cup \Psi$ with four players, 1, 2, 3, and 4, such that

$$\Phi_1 = \{p\}, \qquad \Phi_2 = \{q\}, \qquad \Phi_3 = \Phi \cup \{r\}, \qquad \Phi_4 = \{s\}.$$

Let the players' goals, moreover, be given by:

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

Thus, players 1 and 2 play $G_{toggle(p)}$. Proposition 4.9 ensures that in every equilibrium run ρ in $runs_{\Phi\cup\Psi}$, we have $\rho, t \models p$ if and only if t is even. Accordingly, by virtue of Lemma 5.7, on all equilibrium runs ρ in $NE(G)$, we have $\rho \models \varphi^\tau$ if and only if $\rho \models \varphi$.

Players 3 and 4—quite independently from 1 and 2—play $G^{mp}_{\varphi^\tau}$. In virtue of Proposition 4.1, we obtain, for every run $\rho \in runs_{\Phi\cup\Psi}$, that $\rho \models \varphi^\tau$ if and only if ρ is an equilibrium run in $NE(G)$. Hence, $\rho \models \varphi$ if and only if $\rho \in NE(G)$, for all runs $\rho \in runs_{\Phi\cup\Psi}$. Finally, as φ only depends on Φ, we also have that $\rho|_\Phi \models \varphi^\tau$ for every equilibrium run $\rho \in NE(G)$, and $\rho|_\Phi \not\models \varphi^\tau$ for every non-equilibrium run $\rho \notin NE(G)$. It follows that $NE(G)|_\Phi = X$, as desired. □

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 for future research.

Corollary 5.9. *Let $X \subseteq runs_\Phi$ be a temporal property. Then, L_U can projectively express X in equilibrium if and only if X is ω-regular.*

Proof. For the "only if"-direction, let X be a temporal property in $runs_\Phi$ that $L_U(\Phi)$ can projectively express in equilibrium. Then, LTL(Φ) can also projectively express X in equilibrium, and, by Theorem 5.6, it follows that X is ω-regular.

For the "only if"-direction, consider an arbitrary ω-regular property $X \subseteq runs_\Phi$. By Proposition 5.1, we know that LTL(Φ) can projectively express X in formulas. Hence, there is a set Ψ of auxiliary variables and a property $Y \subseteq runs_{\Phi\cup\Psi}$ that LTL($\Phi \cup \Psi$) can express in formulas with $X = Y|_\Phi$. Theorem 5.8 then gives us that $L_U(\Phi \cup \Psi)$ can projectively express Y in equilibrium. Accordingly, there is an auxiliary set Θ of variables and a temporal property $Z \subseteq runs_{\Phi\cup\Psi\cup\Theta}$ that is expressed by $L_U(\Phi \cup \Psi \cup \Theta)$ in equilibrium with $Z|_{\Phi\cup\Psi} = Y$. We may conclude the proof by observing that $X = Z|_\Phi$. Hence, $L_U(\Phi)$ can projectively express in equilibrium property X, as desired. □

6 LOGICAL INCENTIVE ENGINEERING: WEAK EXPRESSIVENESS

In this section, we consider another weaker notion of expressiveness pertaining to non-empty temporal properties, which we will refer to as *weak expressiveness in formulas.* We will then extend this expressiveness notation to *weak expressiveness in equilibria* and argue for its relevance and naturalness in incentive design settings.

Formally, a fragment $L(\Phi)$ is said to *weakly express (in formulas)* the non-empty property X if it can express in formulas a stronger non-empty property, that is, if there is a *satisfiable* formula φ of $L(\Phi)$ with $runs_\Phi(\varphi) \subseteq X$. This concept of weak expressiveness in formulas is much weaker than the standard notion of expressiveness in formulas, in the sense that, if a fragment can express a property, it can also weakly express this property. To appreciate how much weaker a notion weak expressiveness is, consider the seminal *Ultimately Periodic Model Theorem* by Sistla and Clarke in which every satisfiable formula is satisfied on an ultimately periodic run [Sistla and Clarke 1985]. Here, a run $\rho = v_0v_1v_2v_3\ldots$ is said to be *ultimately periodic* if there are integers $s, p \geq 0$, called the *starting index* and *period*, respectively, such that for every $t \geq s$, we have that $v_t = v_{t+p}$.

Theorem 6.1 ([Sistla and Clarke 1985]). *Every satisfiable* LTL-*formula is satisfied by an ultimately periodic run.*

Consequently, for an LTL-fragment L to weakly express *every* non-empty LTL-property in formulas, it already suffices to be able to express $\{\rho\}$ for every ultimately periodic run ρ. Now observe that any given ultimately periodic run $\rho = v_0v_1v_2v_3\ldots$ over 2^Φ with starting index s and period p

is characterised by the formula

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

where $\mathsf{X}^k \varphi = \overbrace{\mathsf{X} \cdots \mathsf{X}}^{k \text{ times}} \varphi$. Accordingly, we find that every LTL-fragment subsuming $L_{\mathsf{X},\mathsf{F},\mathsf{G}}$, can weakly express in formulas every property that full LTL can express in formulas.

We extend the concept of weak expressiveness in formulas to weak expressiveness in equilibrium. Accordingly, $L(\Phi)$ is said to *weakly express X in equilibrium* if there is an $L(\Phi)$-game G with $\phi \subsetneq X \subseteq NE(G)$. To complete the picture, we also introduce projective versions of weak expressiveness. Thus, a fragment $L(\Phi)$ is said to be able to *projectively weakly express in formulas* 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 $\rho \in runs_\Phi(\varphi)$ implies $\rho|_\Phi \in X$. Furthermore, L *projectively weakly expresses X in equilibrium* if there is some finite set Ψ of auxiliary variables and some L-game G with the players' goals defined over the variables $\Phi \cup \Psi$ such that $\phi \subsetneq X \subseteq NE(G)|_\Phi$.

Weak expressiveness in equilibrium is quite a natural notion in settings where a system designer has to design a multi-agent system that has to behave according to a given LTL-specification φ. If she manages to do so in such a way that the system behaves (in equilibrium) according to a *stronger* but consistent specification, she should still be satisfied with her work: the additional features for the runs generated can be seen as belonging to the specifics of the implementation.

Furthermore, in designing her system, the designer may aim to distribute a possibly complex task over several agents each with limited computational capabilities by allocating to them tasks that are as "simple" as possible. The designer's objective as well as the specifications of the allocated tasks could be given by LTL-formulas in another fragment of LTL. Recall Example 1.2, the Rabbit Hunt, from the introduction, where we saw how an LTL-specification essentially involving the *until*-operator U could be implemented as the equilibria runs of an iterated Boolean game with players that have safety and reachability objectives that could be expressed in the weaker fragment $L_{\mathsf{F},\mathsf{G}}$.

From this perspective, the issue is closely related to the A-Nash problem for iterated Boolean games, which askes if all Nash equilibria of a given iterated Boolean game satisfy a given LTL-specification (see, e.g., [Gutierrez et al. 2013]). In this way, an algorithm to solve the question whether a certain (weak) fragment L of LTL can weakly express in equilibrium a given temporal property expressed by an LTL-formula φ corresponds to a task allocation design method, where players are assigned goals specified in fragment L in such a way that the A-Nash problem is solved for φ and the resulting game.

To illustrate how powerful the concept of (projective) weak expressiveness can be in incentive design settings, we will focus in the remainder of this section on the very weak fragment $L_{\mathsf{X},\mathsf{F}^+}$, where the F-operator cannot occur within the scope of a negation. We find that also this fragment is more expressive in equilibrium than it is in formulas. Before we formally show this, we first have the following lemma, which intuitively says that every $L_{\mathsf{X},\mathsf{F}^+}$-formula will be satisfied after a finite number of rounds on every satisfying run. Here, a temporal property $X \subseteq runs_\Phi$ is said to be *tail-invariant* if $\rho \in X$ implies the existence of a prefix $\pi \in prefix(\rho)$ such that $\pi; \rho' \in X$ for all $\rho' \in runs_\Phi$.

Proposition 6.2. *Every temporal property $X \subseteq runs_\Phi$ that can be expressed in $L_{\mathsf{X},\mathsf{F}^+}$ is tail-invariant.*

Proof Sketch. Let $\varphi \in L_{\mathsf{X},\mathsf{F}^+}$. As the F-operator does not 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. Hence, we may assume that φ is in this normal form. Consider an arbitrary run $\rho = v_0 v_1 v_2 v_3 \dots$ 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 $t \geq 0$, the integer $\kappa_{\rho,t}(\psi)$ as follows:

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

where $t' = \min\{t'' \geq t : \rho, t'' \models \chi\}$. 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 and $t = 0$, 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 $\neg\,\mathsf{F}\,\bar{p}$, by virtue of Proposition 5.5, we find that $L_{\mathsf{X},\mathsf{F}^+}$ can projectively express $\mathsf{G}\,p$ in equilibrium. That is, $L_{\mathsf{X},\mathsf{F}^+}$ can projectively express in equilibrium strictly more temporal properties than it can express projectively in formulas.

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.

Theorem 6.3. *The fragment $L_{\mathsf{X},\mathsf{F}^+}$ can weakly projectively express in equilibrium every property that* LTL *can express in formulas.*

Proof. Let $\varphi \in$ LTL. If φ is unsatisfiable, we are done immediately, because as $p \wedge \neg p$ is a formula in $L_{\mathsf{X},\mathsf{F}^+}$. On the other hand, if φ is satisfiable, then by Theorem 6.1, there is an ultimately periodic run $\rho = v_0 v_1 v_2 v_3 \dots$ with starting index s and period p such that is characterised by the LTL-formula χ_ρ given by

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

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* $\neg\chi_\rho$ of χ_ρ is equivalent to

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

which is included in the fragment $L_{\mathsf{X},\mathsf{F}^+}$. By Proposition 5.5, we know that $L_{\mathsf{X},\mathsf{F}^+}$ can therefore projectively express χ_ρ in equilibrium. Hence, $L_{\mathsf{X},\mathsf{F}^+}$ can weakly projectively express φ in equilibrium, as desired. □

As a corollary of Theorem 6.3, we find that $L_{\mathsf{X},\mathsf{F}^+}$ can even weakly express in equilibrium every ω-regular expression. The proof is analogous to that of the "if"-direction of Corollary 5.9.

Corollary 6.4. *The fragment $L_{\mathsf{X},\mathsf{F}^+}$ can weakly projectively express in equilibrium every ω-regular property.*

Proof. Consider an arbitrary ω-regular property $X \subseteq runs_\Phi$. By Proposition 5.1, we know that LTL can projectively express X in formulas. Hence, there is a set Ψ of auxiliary variables and a property $Y \subseteq runs_{\Phi\cup\Psi}$ that LTL can express in formulas with $X = Y|_{\Phi}$. Theorem 6.3 then gives us that $L_{\mathsf{X},\mathsf{F}^+}$ can weakly projectively express Y in equilibrium. Accordingly, there is an auxiliary set Θ of variables and a temporal property $Z \subseteq runs_{\Phi\cup\Psi\cup\Theta}$ that is expressed by $L_{\mathsf{X},\mathsf{F}^+}$ in equilibrium and that is such that $Z|_{\Phi\cup\Psi} = Y$. We may conclude the proof by observing that $X = Z|_{\Phi}$. Hence, $L_{\mathsf{X},\mathsf{F}^+}$ can weakly projectively express in equilibrium property X, as desired. □

7 RELATED WORK AND SUMMARY

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 [Rabinovich 2002; Strejček 2004]. 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 [Demri and Schnoebelen 2002].

Most of these studies have focused not only on the expressive power of the resulting sublogics but also on the implications of imposing such restrictions in the complexity of the model checking and satisfiability problems of such sublogics. These studies have also enabled a profound understanding of the connections between the various fragments of LTL 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 [Vardi and Wolper 1994].

Despite the abundance of studies on the expressive power of LTL and its fragments, to the best of our knowledge, there are no results on the expressive power of these logics with respect to the classes of runs that can be sustained by some Nash equilibrium. In this article, we have addressed precisely this issue, introduced the concept of expressiveness in equilibrium, and provided the first known results in the literature on this topic. The results are rather promising and perhaps even surprising: they show that even though one LTL-fragment may be more expressive in formulas than another LTL-fragment, they may be equally expressive in equilibrium.

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 fragments can be made as expressive as generally stronger LTL fragments by the addition of fresh propositional variables to the weaker language.

8 FINAL REMARKS AND FUTURE RESEARCH

In this article, we proposed the concept of expressiveness in equilibrium for linear temporal logics. Thus, 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 formulas in various fragments of Linear Time Logic (LTL). We also introduced weak and projective variations of this concept. See Figure 2 for a schematic overview of our main findings in Sections 4 and 5.

The Nash equilibria of an iterated Boolean game are fully determined by which goals the players have and which propositional variables they control. In particular, they are not dependent on an additional underlying game structure—like, for instance, *concurrent game structures* [Alur et al. 2002]. This enabled us to focus on the *logical* aspects of Nash equilibrium and accordingly we formulated our research issue in terms of expressiveness.

Apart from specific fragments, the concept of *expressiveness in equilibrium* gives rise to a number of questions, both theoretical and conceptual, for future research.

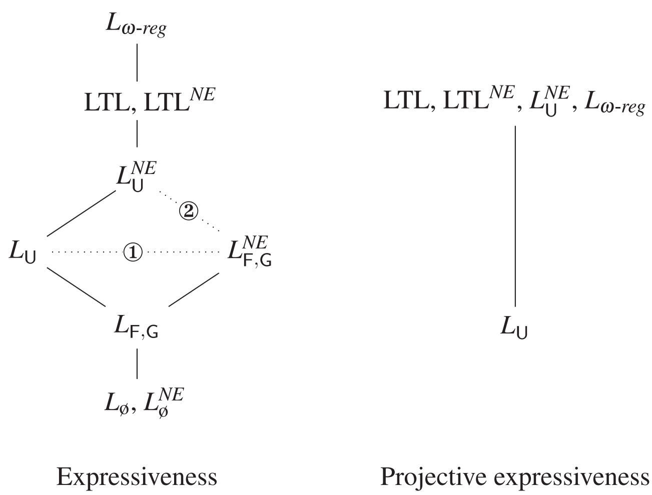

Fig. 2. Overview of our main results. A continuous line between two nodes indicates that the higher fragment is strictly more expressive than the lower fragment. Dotted line indicated by ①: as $L_{F,G}$ can express in equilibrium some stutter-sensitive properties, it is not less expressive in equilibrium than L_U in formulas. We conjecture that there are also properties that L_U can express in formulas that $L_{F,G}$ cannot express in equilibrium, which would imply that $L_{F,G}^{NE}$ and L_U are incomparable. Dotted line indicated by ②: as L_U is at least as expressive in formulas as $L_{F,G}$ is in formulas, by virtue of Proposition 4.3, we know that L_U is also at least as expressive in equilibria as $L_{F,G}$ is in equilibrium. We conjecture that L_U is strictly more expressive in equilibria than $L_{F,G}$ is in equilibrium.

First, in this article, we focussed mainly on the concepts of expressiveness in equilibrium and projective expressiveness in equilibrium. To highlight their formal features and how these notions relate to concepts of expressiveness in formulas, we concentrated on the full fragment of LTL and the important maximal stutter-invariant fragment L_U. To obtain a complete picture of (projective) expressiveness in equilibrium, we aim to explore a wider range of fragments. To this end, the hierarchy of temporal properties as proposed by Manna and Pnueli [1990], Chang et al. [1992], and Černa and Pelánek [2003] may provide the necessary structure for this research. In particular, this hierarchy would allow us to address in a principled fashion fragments that can express *safety*, *guarantee* or *reachability obligation*, *response*, *persistence*, and *reactivity* properties [Strejček 2004]. This would also cover the important fragments L_{F^+} and L_{G^+}, which importantly capture guarantee goals and safety goals, respectively. Leveraging Proposition 4.1 and using a similar construction as in the proof for Proposition 5.5, it is not hard to show that both $L_{F^+}^{NE} > L_{F^+}$ and $L_{G^+}^{NE} > L_{G^+}$, provided that $|\Phi| \geq 3$. A more complete picture, however, is desirable.

Second, another theoretical issue concerns the closure of the ω-languages expressible in equilibrium under such operations as complement, intersection, and union. For some fragments, like full LTL and propositional calculus, the issue is trivial because these fragments are equally expressive in formulas as they are in equilibrium. For other fragments, like, for instance, the maximal stutter-invariant fragment L_U, the question is still very much open.

Third, in this article, we considered expressiveness in terms of ω-runs. There is, however, also a considerable line of research that concerns the expressiveness of temporal logics when semantically interpreted on finite runs [Giacomo and Vardi 2013]. The model of iterated Boolean games can likewise be adapted in such a way that their outcomes are finite runs rather than infinite runs, as, for instance, in Gutierrez et al. [2017b]. In that way, expressiveness in formulas and expressiveness in equilibrium with respect to finite runs can again be compared. This would potentially allow us to leverage the work on *forbidden fragments* as originally introduced by Cohen et al. [1993], which pertains to finite runs only and which furnishes us with some interesting characterisations of LTL-fragments [Strejček 2004].

Fourth, many of our game constructions involve a "matching pennies" game like G_φ^{mp}. 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 article or that it points at a more fundamental connection with the concept of expressiveness in equilibrium. A related question for future research concerns the dependence on the number of players needed so as to be able to express a given property in equilibrium. Thus, one could investigate (bounds on) the minimal number of players and additional propositional variables that may be needed to express or projectively express temporal properties in a particular fragment, or one could examine expressivity in equilibrium if one can only dispose of a constant number of players. In particular, it would be interesting to know if every property that can be expressed by a fragment L in equilibrium coincides with the equilibrium runs of a two-player L-game.

Fifth, 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 formulas. As such, the former are much finer grained than the latter. One may wonder to what extent the gap between expressiveness in formulas 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 or abilities are similarly specified by temporal formulas? The work by Jamroga et al. [2019] on natural strategic ability suggests one way in which this may be achieved. More generally, the ramifications of restricting attention to specific classes of strategies, e.g., finite-state strategies, for expressiveness in equilibrium for a particular fragment is an natural line of future research.

Sixth, all our notions of expressiveness naturally extend to game-theoretic solution concepts other than Nash Equilibrium such as, for instance, subgame perfect equilibrium and dominant strategy equilibrium. Investigating some of these concepts suggests itself as a natural direction of future research.

Seventh, the topic of expressiveness in equilibrium is not restricted to iterated Boolean games. It could also be investigated in the context of other and richter types of transition structure like *reactive modules* [Alur and Henzinger 1999], *concurrent game structures* [Alur et al. 2002], *simple reactive modules* [Hoek et al. 2006]. This line of research also suggests an alternative concept of expressive power in equilibrium. Analogous to the concept of *distinguishability* of temporal languages, which measures the expressive power of a temporal logic by the models (interactive transition structures) its formulas can distinguish (cf., [Wang and Dechesne 2009]), one could define *distinguishability in equilibrium* as follows: For a given interactive transition structure M of a particular kind and formula φ of temporal logic L, let $M \models^{NE} \varphi$ denote that φ holds on all equilibrium runs of M. Then, a formula $\varphi \in L$ could be defined as an *equilibria-separating formula* of a temporal language L for structures M_1 and M_2—in formulas $M_1 \|_L^{NE} M_2$—if $M_1 \models^{NE} \varphi$ and $M_2 \not\models^{NE} \varphi$, or *vice-versa*, that is, if L can distinguish M_1 and M_2 by their equilibrium runs. A temporal language L_1 is said to be at least as *distinguishable in equilibria* as temporal language L_2 if $M_1 \|_{L_1}^{NE} M_2$ implies $M_1 \|_{L_2}^{NE} M_2$ for all interactive structures M_1 and M_2 of the relevant kind.[3]

[3]We are thankful to an anonymous reviewer for this suggestion.

Finally, in Section 6, we briefly dwelt on the topic of allocating tasks to agents so that a given property is satisfied in all resulting equilibria. We hope that our study of expressiveness in equilibrium has shed some light on this issue and will have practical ramifications for the design of multi-agent systems by providing "lean" temporal specifications of individual agents who can be assumed to play strategies that together form an equilibrium.

ACKNOWLEDGMENTS

The authors would also like to thank Johan van Benthem, Moshe Vardi, Sasha Rubin, and an anonymous reviewer for helpful discussions and insightful suggestions.

REFERENCES

R. Alur and T. A. Henzinger. 1999. Reactive modules. *Formal Methods in System Design* 15, 11 (July 1999), 7–48.

R. Alur, T. A. Henzinger, and O. Kupferman. 2002. Alternating-time temporal logic. *J. ACM* 49, 5 (Sept. 2002), 672–713.

R. J. Aumann and A. Brandenburger. 1995. Epistemic conditions for Nash equilibrium. *Econometrica* 63, 5 (1995), 1161–1180.

C. Baier and J.-P. Katoen. 2008. *Principles of Model Checking*. MIT Press.

E. W. Beth. 1953. On Padoa's method in the theory of definition. *Indagationes Mathematicae* 15 (1953), 330–339.

E. Bonzon, M.-C. Lagasquie, J. Lang, and B. Zanuttini. 2006. Boolean games revisited. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI'06)*, G. Brewka, S. Coradeschi, A. Perini, and P. Traverso (Eds.). 265–269.

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

I. Černa and R. Pelánek. 2003. Relating hierarchy of temporal properties to model checking. In *Proceedings of the 30th Symposium on the Mathematical Foundations of Computer Science (MFCS'03) (Lecture Notes in Computer Science (LNCS))*, vol. 2747,B. Rovan and Peter Vojtáš (Eds.). Springer-Verlag, 318–327.

C. C. Chang and H. J. Keisler. 1990. *Model Theory* (3rd ed.). *Studies in Logic and the Foundations of Mathematics*, vol. 73. North-Holland Publishing Company.

E. Chang, Z. Manna, and A. Pnueli. 1992. Characterization of temporal property classes. In *Proceedings of the 19th International Colloquium on Automata, Languages and Programming (ICALP'92) (Lecture Notes in Computer Science (LNCS))*, W. Kuich (Ed.), vol. 623. Springer-Verlag, 474–486.

K. Chatterjee, T. A. Henzinger, and N. Piterman. 2010. Strategy logic. *Information and Computation* 208, 6 (June 2010), 677–693.

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

J. Cohen, D. Perrin, and J-É. Pin. 1993. On the expressive power of termporal logic. *J. Comput. System Sci.* 46 (1993), 271–294.

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

B. de Bruin. 2010. *Explaining Games: The Epistemic Programme in Game Theory*. Synthese Library, vol. 346. Springer.

S. Demri, V. Goranko, and M. Lange. 2016. *Temporal Logics in Computer Science: Finite State Systems*. Cambridge Tracts in Theoretical Computer Science, vol. 58. Cambridge University Press.

S. Demri and P. Schnoebelen. 2002. The complexity of propositional linear temporal logics in simple cases. *Inf. Comput.* 174, 1 (2002), 84–103.

E. A. Emerson. 1990. Temporal and modal logic. In *Handbook of Theoretical Computer Science Volume B: Formal Models and Semantics*. Elsevier Science Publishers, 996–1072.

E. A. Emerson and J. Y. Halpern. 1986. 'Sometimes' and 'Not Never' revisited: On branching time versus linear time temporal logic. *J. ACM* 33, 1 (1986), 151–178.

U. Endriss, S. Kraus, J. Lang, and M. Wooldridge. 2011. Designing incentives for Boolean games. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'11)*, K. Turner, P. Yolum, L. Sonenberg, and P. Stone (Eds.), Vol. 1. 79–86.

K. Etessami. 2000. A note on a question of Peled and Wilke on stutter-invariant LTL. *Inform. Process. Lett.* 75, 6 (2000), 261–263.

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

G. Giacomo and M. Vardi. 2013. Linear temporal logic and linear dynamic logic on finite traces. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, S. Abramsky, G. Gottlob, and P. G. Kolaitis (Eds.). 854–860.

J. Gutierrez, P. Harrenstein, G. Perelli, and M. Wooldridge. 2016. Expressiveness and Nash equilibrium in iterated Boolean games. In *Proceedings of the 15th International Conference on Autonomous Agents & Multiagent Systems (AAMAS'16)*, J. Thangarajah, K. Tuyls, C. Jonker, and S. Marsella (Eds.). 707–715.

J. Gutierrez, P. Harrenstein, and M. Wooldridge. 2013. Iterated Boolean games. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, F. Rossi (Ed.). 932–938.

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

J. Gutierrez, P. Harrenstein, and M. Wooldridge. 2015b. Expressiveness and complexity results for strategic reasoning. In *Proceedings of the CONCUR'15 (LIPIcs)*, vol. 42. Schloss Dagstuhl, 268–282.

Julian Gutierrez, Paul Harrenstein, and Michael Wooldridge. 2015a. Iterated Boolean games. *Information and Computation* 242 (2015), 53–79.

J. Gutierrez, P. Harrenstein, and M. Wooldridge. 2017a. Reasoning about equilibria in game-like concurrent systems. *Annals of Pure and Applied Logic* 168, 2 (2017), 373–403.

J. Gutierrez, G. Perelli, and M. Wooldridge. 2017b. Iterated games with LDL goals over finite traces. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems (AAMAS'17)*. 696–704.

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

P. Harrenstein, W. van der Hoek, J.-J.Ch. Meyer, and C. Witteveen. 2001. Boolean games. In *Proceedings of the 8th Conference on Theoretical Aspects of Rationality and Knowledge (TARK VIII)*, J. van Benthem (Ed.). Siena, Italy, 287–298.

W. Hodges. 1993. *Model Theory*. Encyclopedia of Mathematics and its Applications, vol. 42. Cambridge University Press.

W. van der Hoek, A. Lomuscio, and M. Wooldridge. 2006. On the complexity of practical ATL model checking. In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'06)*. 201–208.

W. Jamroga, V. Malvone, and A. Murano. 2019. Natural strategic ability. *Artificial Intelligence* 277, 103170 (2019), 1–33. DOI: https://doi.org/10.1016/j.artint.2019.103170

O. Kupferman. 2018. Automata theory and model checking. In *Handbook of Model Checking*, Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and Roderick Bloem (Eds.). Springer-Verlag, 107–151.

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

L. Lamport. 1983. What good is temporal logic? In *Information Processing '83: Proceedings of the 9th IFIP Congress on Information Processing*, R. E. A. Mason (Ed.). North-Holland, 657–667.

Christof Löding. 2012. Basics on tree automata. In *Modern Applications of Automata Theory*, D. D'Souza and P Shankar (Eds.). IISc Research Monographs Series, Vol. 2. World Scientific, 79–110.

R. D. Luce and H. Raiffa. 1957. *Games and Decisions*. John Wiley & Sons.

Z. Manna and A. Pnueli. 1990. A hierarchy of temporal properties. In *Proceedings of the ACM Symposium on Principels of Distributed Computing (PODC'90)*. ACM Press, 377–410.

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

A. Mateescu and A. Salomaa. 1997. Formal languages: An introduction and a synopsis. In *Handbook of Formal Languages*, G. Rozenberg and A. Salomaa (Eds.). Vol. I: Word, Language, Grammar. Springer, Chapter 1.

D. Niwinski and I. Walukiewicz. 1998. Relating hierarchies of word and tree automata. In *Annual Symposium on Theoretical Aspects of Computer Science 1998 (STACS'98) (Lecture Notes in Computer Science (LNCS))*, vol. 1373. Springer, 320–331.

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

D. Peled. 1997. On projective and separable properties. *Theoretical Computer Science* 186 (1997), 135–157.

D. A. Peled and T. Wilke. 1997. Stutter-invariant temporal properties are expressible without next-time operator. *Inform. Process. Lett.* 63, 5 (1997), 243–246.

A. Perea. 2012. *Epistemic Game Theory: Reasoning and Choice*. Cambridge University Press, U.K.

D. Perrin and J-É. Pin. 2004. *Infinite Words: Automata, Semigroups, Logic and Games*. Pure and Applied Mathematics, Vol. 141. Elsevier Academic Press.

A. Pnueli. 1977. The temporal logic of programs. In *Proceedings of the 18th IEEE Symposium on the Foundations of Computer Science*. 46–57.

A. Pnueli and R. Rosner. 1989. On the synthesis of a reactive module. In *Proceedings of the 16th ACM Symposium on the Principles of Programming Languages (POPL)*. 179–190.

Alexander Moshe Rabinovich. 2002. Expressive power of temporal logics. In *Proceedings of the 13th International Conference on Concurrency Theory (CONCUR 2002) (Brno, Czech Republic, August 20-23, 2002) (Lecture Notes in Computer Science)*, vol. 2421, Lubos Brim, Petr Jancar, Mojmír Kretínský, and Antonín Kucera (Eds.). Springer, 57–75.

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

A. P. Sistla and E. M. Clarke. 1985. The complexity of propositional linear temporal logics. *J. ACM* 32, 3 (1985), 733–749.

J. Strejček. 2004. *Linear Temporal Logic: Expressiveness and Model Checking*. Ph.D. dissertation. Masaryk University of Brno.

W. Thomas. 1997. Languages, automata, and logic. In *Handbook of Formal Languages*, G. Rozenberg and A. Salomaa (Eds.). Vol. III: Beyond words. Springer, Chapter 5.

M. Y. Vardi. 1996. An automata-theoretic approach to linear temporal logic. In *Logics for Concurrency: Structure vs Automata*, A. Moller and G. Birtwistle (Eds.). *Lecture Notes in Computer Science (LNCS)*, vol. 1043. Springer, 238–266.

M. Y. Vardi. 2001. Branching vs. linear time: Final showdown. In *Proceedings of the 2001 Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS'01)*, T. Margaria and W. Yi (Eds.), *Lecture Notes in Computer Science*, vol. 2031. 1–22.

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

Y. Wang and F. Dechesne. 2009. On expressive power and class invariance. https://arxiv.org/abs/0905.4332 (2009).

P. Wolper. 1983. Temporal logic can be more expressive. *Information and Control* 56, 1–2 (1983), 72–99.

Received January 2020; revised September 2020; accepted November 2020