A Compositional Treatment of Iterated Open Games

Neil Ghani, Clemens Kupke, Alasdair Lambert, Fredrik Nordvall Forsberg

Department of Computer and Information Sciences, University of Strathclyde

Abstract

Compositional Game Theory is a new, recently introduced model of economic games based upon the computer science idea of compositionality. In it, complex and irregular games can be built up from smaller and simpler games, and the equilibria of these complex games can be defined recursively from the equilibria of their simpler subgames. This paper extends the model by providing a final coalgebra semantics for infinite games. In the course of this, we introduce a new operator on games to model the economic concept of subgame perfection.

Keywords: Compositional game theory; Final coalgebra semantics; Infinite iterated games; subgame perfection.

As one of Don's PhD students, I have always admired Don's commitment to (i) fundamental research which would stand the test of time; (ii) taking seriously the real-world problems that inspire research; and (iii) the power of categorical tools for devising compositional techniques for treating complex phenomena. It is therefore a great pleasure to dedicate this work to him, as it embodies exactly these qualities: using category theory to devise a compositional treatment of infinite games that occur in economic modelling. Additionally, Don showed me great kindness and support throughout my career, e.g. by funding part of my post-doc even though my research was not directly related to his. This selfless support for future generations is something I am extremely grateful for, and have tried to replicate in my own dealings with younger academics.

Professor Neil Ghani

1. Introduction

Compositionality, where one sees complex systems as being built from smaller subsystems, is widely regarded within computer science as a key enabling technique for scalability. Since the subsystems are smaller, they are easier to reason about, and compositionality also promotes modularity and reuse; a particular system can be a subsystem of many different supersystems. Can compositionality be applied also to economic games? In general, not all reasoning is compositional, especially if significant emergent behaviour is present in a large system but not in its subsystems. This is unfortunately the case for economic games. For example, if σ is an optimal strategy for a game $\mathcal{G}$, then is σ part of an optimal strategy for $\mathcal{G} * \mathcal{H}$, where $\mathcal{G} * \mathcal{H}$ is a super-game built from $\mathcal{G}$ and $\mathcal{H}$? Clearly not, e.g. the *Iterated Prisoner's Dilemma* has equilibria — such as cooperative equilibria — that do not arise from repeatedly playing the Nash equilibrium from the *Prisoner's Dilemma* (Axelrod and Dion, 1988).

However, Ghani et al. (2018) produced a compositional model of game theory which included a limited set of operators for building new games from old. One shortcoming was that this did not treat the infinite iteration of games, or more generally contain an operator to compositionally build infinite iterations of games. This paper addresses that problem. Within programming language theory, these sort of issues are tackled by final coalgebra semantics (Rutten and Turi, 1994) and we follow this practice, with the added benefit of bringing related bisimulation techniques to the game theory community. We highlight two relationships between our work and traditional approaches:

- Each round of an infinite game produces utility. Traditionally, this infinite sequence of staged utilities is combined into a single utility in one of a number of ad hoc manners. We take the bolder approach of not requiring the choice of a single mechanism for combining utilities.

- The coalgebraic approach we advocate dovetails well with the economic concept of *subgame perfection*, where a strategy must be an optimal response in all subgames of the supergame (Shubik, 1984).

The general approach of Compositional Game Theory deals with a new concept of coutility. However this paper makes the simplifying assumption that the coutility function is the identity. Despite this, our approach covers many games, as those occurring in the traditional literature do not possess coutility.

Related Work. An introduction to the economic treatment of iterated games can be found in Mailath and Samuelson (2006). The fundamental concept of game theory is that of Nash equilibrium (Nash, 1951), which has been adapted for the study of repeated and dynamic games to the concept of subgame perfect equilibrium first introduced by Selten (1965). Significantly influential work on using logical methods and coalgebraic reasoning in economics include Lescanne (2012) and Abramsky and Winschel (2017). Open games are also closely related to the 'partially defined games' of Oliva and Powell (2015).

Structure of the paper. Section 2 consists of preliminaries and a summary of previous work on open games; Section 3 introduces an operator for dealing with subgame perfection; Section 4 introduces morphisms between games, and Section 5 consists of our final coalgebra semantics for infinite open games. In Section 6, we show how bisimulation and coinduction can be used to reason about infinitely repeated games. Finally Section 7 contains concluding remarks and discussions of further work.

2. Preliminaries

The key concept of Ghani et al. (2018) is the following:

Definition 1 (Open Game)**.** Let X, Y, R and S be sets. An *open game* $\mathcal{G} = (\Sigma_{\mathcal{G}}, P_{\mathcal{G}}, C_{\mathcal{G}}, E_{\mathcal{G}}) : (X, S) \to (Y, R)$ consists of:

- a set $\Sigma_{\mathcal{G}}$, called the set of *strategy profiles* of $\mathcal{G}$,
- a function $P_{\mathcal{G}} : \Sigma_{\mathcal{G}} \times X \to Y$, called the *play function* of $\mathcal{G}$,
- a function $C_{\mathcal{G}} : \Sigma_{\mathcal{G}} \times X \times R \to S$, called the *coutility function* of $\mathcal{G}$, and
- a function $E_{\mathcal{G}} : X \times (Y \to R) \to \mathscr{P}\Sigma_{\mathcal{G}}$, called the *equilibrium function* of $\mathcal{G}$.

We sometimes write $\mathcal{G} : (X, S) \xrightarrow{\Sigma} (Y, R)$ to make the set of strategies explicit. Intuitively, the set X contains the possible states or histories of the game, Y the moves, R the utilities and S the coutilities. The set $\Sigma_{\mathcal{G}}$ contains the strategies we are trying to pick an optimal one from. The play function $P_{\mathcal{G}}$ selects a move given a strategy and a state, while the coutility function $C_{\mathcal{G}}$ chooses the coutility to extrude from the game, given a strategy, state and utility. Finally, if $\sigma \in E_{\mathcal{G}}\, x\, k$, then σ is an optimal strategy in state x and with utility given by $k : Y \to R$. Permitting arbitrary equilibrium functions, as opposed to only considering fixed ones, is one of the key steps for achieving compositionality.

		Player 2	
		C	D
Player 1	C	$-1, -1$	$-5, 0$
	D	$0, -5$	$-3, -3$

Figure 1: Payoff matrix for the Prisoner's Dilemma.

Example 2. In the *Prisoner's Dilemma*, two suspects in a major crime are held in separate cells. Each of them can be convicted of a minor offence, but unless one of them acts as an informer against the other, there is not enough evidence for a major conviction. Each prisoner must decide to either *cooperate* with the other prisoner and stay quiet, or to *defect* and betray the other. If both stay quiet, each will spend one year in prison, while if one and only one defects, she will walk away free and be used as a witness against the other prisoner, who will spend five years in prison. However if they both defect, they will each spend three years in prison. The situation is summarised in Figure 1. The Prisoner's Dilemma is represented as an open game

$$\mathcal{PD} : (\mathbf{1}, \mathbb{R} \times \mathbb{R}) \to (\{C, D\} \times \{C, D\}, \mathbb{R} \times \mathbb{R})$$

as follows: The history of the game is trivial (so $X = \mathbf{1}$, a singleton set), and the moves of the game consists of a choice to cooperate or defect for each player (so $Y = \{C, D\} \times \{C, D\}$). Both utility and coutility is represented by the number of years lost in prison for each player (so $R = S = \mathbb{R} \times \mathbb{R}$). A strategy simply consists of choosing a move, so $\Sigma_{\mathcal{PD}} = \{C, D\} \times \{C, D\}$, with play function $P_{\mathcal{PD}}(x, m) = m$. Coutility is given by $C_{\mathcal{PD}}(m, x, r) = r$, while

$$\begin{aligned}(m_1, m_2) \in E_{\mathcal{PD}}(x, k) \Leftrightarrow\ & m_1 \in \arg\max(\pi_1 \circ k(-, m_2)) \text{ and} \\ & m_2 \in \arg\max(\pi_2 \circ k(m_1, -))\end{aligned}$$

We see that $(m_1, m_2) \in E_{\mathcal{PD}}(x, k)$ exactly when (m_1, m_2) is a Nash equilibrium. No particular utility function is hardcoded into this game; we get the concrete Prisoner's Dilemma by specialising to the utility function k corresponding to the payoff matrix in Figure 1, in which case (D, D) is the only equilibrium.

The main result of Ghani et al. (2018) can be stated as follows:

Theorem 3. *The collection of pairs of sets, with open games* $\mathcal{G} : (X, S) \to$

(Y, R) *as morphisms, forms a symmetric monoidal category* Open.[1]

Proof. The composition of $\mathcal{G}$ and $\mathcal{H}$ is given by the game with strategies $\Sigma_{\mathcal{H}\circ\mathcal{G}} = \Sigma_{\mathcal{G}} \times \Sigma_{\mathcal{H}}$, play function the composition of the respective play functions from $\mathcal{H}$ and $\mathcal{G}$, and coutility function the composition in reverse of the coutility functions from $\mathcal{H}$ and $\mathcal{G}$, using the play function of $\mathcal{G}$ to produce a state for $\mathcal{H}$. Finally $(\sigma_1, \sigma_2) \in E_{\mathcal{H}\circ\mathcal{G}}\, x\, k$ if and only if $\sigma_1 \in E_{\mathcal{G}}\, x\, k'$, where $k'\, y = C_{\mathcal{H}}\, \sigma_2\, y\, (k\, (P_{\mathcal{H}}\, \sigma_2\, y))$, and $\sigma_2 \in E_{\mathcal{H}}\, (P_{\mathcal{G}}\, \sigma_1\, x)\, k$.

The monoidal product is given by Cartesian product in the category of sets, with componentwise action on the strategies, play functions and coutility functions of open games, and $(\sigma_1, \sigma_2) \in E_{\mathcal{G}\otimes H}\, (x_1, x_2)\, k$ if and only if $\sigma_1 \in E_{\mathcal{G}}\, x_1\, ((\pi_1 \circ k)(\_, P_{\mathcal{H}}\, \sigma_2\, x_2))$ and $\sigma_2 \in E_{\mathcal{H}}\, x_2\, ((\pi_2 \circ k)(P_{\mathcal{G}}\, \sigma_1\, x_1, \_))$. The unit of this monoidal structure is $(\mathbf{1}, \mathbf{1})$, while the symmetry is inherited from the Cartesian product in Set. □

Returning to the Prisoner's Dilemma, we see that it can be composed from simpler games as follows:

Example 4. The Prisoner's Dilemma game $\mathcal{PD} : (\mathbf{1}, \mathbb{R} \times \mathbb{R}) \to (\{C, D\} \times \{C, D\}, \mathbb{R} \times \mathbb{R})$ from Example 2 arises as

$$\mathcal{PD} = \mathcal{G} \otimes \mathcal{G}$$

where $\mathcal{G} : (\mathbf{1}, \mathbb{R}) \xrightarrow{\{C,D\}} (\{C, D\}, \mathbb{R})$ is the game with play function $P_{\mathcal{G}}(x, m) = m$, coutility function $C_{\mathcal{G}}(x, m, r) = r$ and equilibrium function $E_{\mathcal{G}}(x, k) = \arg\max k$. We can further hardwire a specific utility function k by creating a *closed game* $\mathcal{U}_k \circ \mathcal{PD} : (\mathbf{1}, \mathbb{R} \times \mathbb{R}) \to (\mathbf{1}, \mathbf{1})$ by composing $\mathcal{PD}$ with the game $\mathcal{U}_k : (\{C, D\} \times \{C, D\}, \mathbb{R} \times \mathbb{R}) \xrightarrow{\mathbf{1}} (\mathbf{1}, \mathbf{1})$, where $\mathcal{U}_k$ has trivial play function, coutility function $C_{\mathcal{U}_k}(m, y, x) = k(y)$, and every strategy is an equilibrium. It is not hard to see that $(m_1, m_2) \in E_{\mathcal{PD}}(x, k)$ if and only if $(m_1, m_2) \in E_{\mathcal{U}_k\circ\mathcal{PD}}(x, !)$, where $! : \mathbf{1} \to \mathbf{1}$ is the unique function into $\mathbf{1}$. However, we prefer to work with open games without specified utility functions whenever possible, so that we can compose them into larger games as needed.

[1] Actually, one needs to quotient by the equivalence relation induced by isomorphism of strategies, but we simplify presentation here by dealing with representatives directly.

3. Subgame Perfection and Conditioning

Intuitively, we play two rounds of a game by composing the game with itself. However, this is not quite right: in the composite game $\Sigma_{\mathcal{H}\circ\mathcal{G}} = \Sigma_{\mathcal{H}} \times \Sigma_{\mathcal{G}}$, and thus the second game $\mathcal{H}$ cannot react to the moves played by the first game $\mathcal{G}$. This clearly does not match practice as any later play should be able to react differently to different previous plays. Further, an optimal strategy should react optimally to all previous plays, even if those previous plays are sub-optimal. In the game-theoretic literature, this is known as *subgame perfection* (see also Abramsky and Winschel (2017) for a coalgebraic treatment). Rather than introduce a new form of composition, we introduce a new operator $A \to (-)$ for conditioning a game to react to every possibility in some set A.

Definition 5. Let A be a set. Given a game $\mathcal{H} : (X,\, S) \xrightarrow{\Sigma} (Y, R)$, we define the game $A \to \mathcal{H} : (A \times X,\, S) \xrightarrow{A\to\Sigma} (A \times Y,\, R)$ with

- play function $P_{A\to\mathcal{H}}\,(a, x)\,(f : A \to \Sigma_{\mathcal{H}}) \;=\; (a,\, P_{\mathcal{H}}\, x\,(fa))$
- coutility function $C_{A\to\mathcal{H}}\,(a, x)\, f\, r \;=\; C_{\mathcal{H}}\, x\,(fa)\, r$
- equilibrium function

$$f \in E_{A\to\mathcal{H}}(a, x)\,(k : A \times Y \to R) \quad \text{iff} \quad (\forall a' \in A)\;\, fa' \;\in\; E_{\mathcal{H}}\, x\;\, k(a', \_)$$

Note how the description of subgame perfection in the previous paragraph is reflected mathematically by requiring a strategy in $A \to \mathcal{H}$ to be a set of strategies, one for each element of A, and that for a strategy f to be optimal in $A \to \mathcal{H}$, each of its components must be optimal in $\mathcal{H}$. Clearly we have:

Lemma 6. *The mappings $(X, S) \mapsto (A \times X, S)$ and $\mathcal{H} \mapsto (A \to \mathcal{H})$ define a functor $A \to (-) : \mathsf{Open} \to \mathsf{Open}$.* □

This functor is not only of interest for infinite iterations of games, but more generally allows any game to observe information and react accordingly.

4. 2-Cells and Coutility-Free Games

Given a game $\mathcal{G} : (X, S) \xrightarrow{\Sigma} (Y, R)$, we will construct its infinite iteration $\mathcal{G}_\omega$ compositionally as the final coalgebra of the functor $F_{\mathcal{G}}(\mathcal{H}) := (Y \to \mathcal{H})\circ\mathcal{G}$.

However, this means that games will acquire universal properties and thus we need a notion of morphism between games. Further, $\mathcal{G}_\omega$ will satisfy $\mathcal{G}_\omega \cong (Y \to \mathcal{G}_\omega) \circ \mathcal{G}$, and hence the equation

$$C_{\mathcal{G}_\omega}\, x\, \sigma\, r = C_{\mathcal{G}}\, x\, \sigma_0\, (C_{\mathcal{G}_\omega}\, x'\, \sigma_1\, r)$$

relating the coutilities of $\mathcal{G}_\omega$ and $\mathcal{G}$ must hold. Here, the strategy σ for $\mathcal{G}_\omega$ decomposes into σ_0 for the first round and σ_1 for later rounds, and x' is the state after completing the first round. This equation does not always have a unique solution — e.g. if $C_{\mathcal{G}}\, x\, \sigma\, r = r$. To recover uniqueness, we restrict to games $\mathcal{G}$ where $C_{\mathcal{G}}\, x\, \sigma\, r = r$ in this paper. This is not a severe restriction, as there is no coutility in traditional game theory. We will also only consider state-free games. We are currently working on removing both these restrictions. Next, for $F_{\mathcal{G}}(\mathcal{H})$ to type check, the utility and coutility sets of $\mathcal{G}$ must be the same, and thus we only consider games whose utility and coutility is a fixed set R. To summarise: in this paper we consider games $\mathcal{G} : (\mathbf{1}, R) \xrightarrow{\Sigma} (Y, R)$ with state $\mathbf{1}$, utility and coutility the set R, and coutility function $C_{\mathcal{G}}\, \sigma\, r = r$. We define morphisms between such games as follows:

Definition 7. Let R be a set. Given two games $\mathcal{G} : (\mathbf{1}, R) \xrightarrow{\Sigma} (Y, R)$ and $\mathcal{G}' : (\mathbf{1}, R) \xrightarrow{\Sigma'} (Y', R)$, a morphism $\alpha : \mathcal{G} \to \mathcal{G}'$ consists of a pair of functions $\alpha = (\alpha_Y : Y \to Y', \alpha_\Sigma : \Sigma \to \Sigma')$ such that

(i) $\alpha_Y(P\, \sigma) = P'\, (\alpha_\Sigma\, \sigma)$, and

(ii) for every $\sigma \in \Sigma$ and $k : Y' \to R$, if $\sigma \in E\, (k \circ \alpha_Y)$ then $\alpha_\Sigma(\sigma) \in E'\, k$.

We trust the reader will not be confused by the fact that games are morphisms in Open but also have morphisms between them — this simply reflects inherent 2-categorical structure. The category whose objects are open games $\mathcal{G} : (\mathbf{1}, R) \xrightarrow{\Sigma} (Y, R)$ for some Σ, Y (and a fixed R), and whose morphisms are the morphisms between such open games is denoted $2\mathsf{Open}_R$. We are now in position to define the functor $F_{\mathcal{G}} : 2\mathsf{Open}_R \to 2\mathsf{Open}_R$ whose final coalgebra will be the infinite iteration of the game $\mathcal{G}$.

Theorem 8. *Let R be a set and $\mathcal{G} : (\mathbf{1}, R) \xrightarrow{\Sigma} (Y, R)$. The mapping $F_{\mathcal{G}}(\mathcal{H}) = (Y \to \mathcal{H}) \circ \mathcal{G}$ extends to a functor $F_{\mathcal{G}} : 2\mathsf{Open}_R \to 2\mathsf{Open}_R$.*

Proof. Given a morphism $\alpha : \mathcal{H} \to \mathcal{H}'$, define $F_{\mathcal{G}}(\alpha) : F_{\mathcal{G}}(\mathcal{H}) \to F_{\mathcal{G}}(\mathcal{H}')$ by

$$(F_{\mathcal{G}}(\alpha))_{\Sigma}\,(\sigma, f) = (\sigma, \alpha_{\Sigma} \circ f) \qquad\qquad (F_{\mathcal{G}}(\alpha))_{Y}\,(y, z) = (y, \alpha_{Y}\, z)$$

The play function and equilibrium preservation conditions are easily checked. □

5. The iterated game as a final coalgebra

From now on, let R be an arbitrary set, used as utility and coutility for all our games, and write $\mathsf{2Open}$ for $\mathsf{2Open}_R$.

5.1. Definition of the iterated game

Let us fix an arbitrary open game $\mathcal{G} : (1, R) \xrightarrow{\Sigma} (Y, R)$ that we want to iterate infinitely often via the final coalgebra of the functor $F_{\mathcal{G}} : \mathsf{2Open} \to \mathsf{2Open}$ from the previous section, mapping $\mathcal{H} : (1, R) \xrightarrow{\Sigma_{\mathcal{H}}} (Y_{\mathcal{H}}, R)$ to $(Y \to \mathcal{H}) \circ \mathcal{G} : (1, R) \xrightarrow{\Sigma \times (Y \to \Sigma_{\mathcal{H}})} (Y \times Y_{\mathcal{H}}, R)$. We first describe $F_{\mathcal{G}}$-coalgebras, then our candidate $\mathcal{G}_{\omega}$ for the final $F_{\mathcal{G}}$-coalgebra, and conclude with a proof that $\mathcal{G}_{\omega}$ really is final. As a first step we need to recall two endofunctors on the category of sets, and their final coalgebras.

Fact 9. *Given two sets I and O we let $D(I, O) : \mathsf{Set} \to \mathsf{Set}$ be the functor given by $D(I, O)X := O \times X^I$ and by $D(I, O)(f : X \to Y) := \mathrm{id}_O \times f^I$. Furthermore, for a set Y, we define the functor $S(Y) : \mathsf{Set} \to \mathsf{Set}$ by putting $S(Y)X = Y \times X$ and $S(Y)(f : X \to Y) := \mathrm{id}_Y \times f$. The final $D(I, O)$-coalgebra is*

$$(I^* \to O) \xrightarrow{\langle \mathrm{now}, \mathrm{ltr} \rangle} O \times (I^* \to O)^I$$

where I^* is the set of finite words over I, and $\mathrm{now}(f) := f(\epsilon)$ *and* $\mathrm{ltr}(f) = \lambda i.\lambda w.f(iw)$ *(cf. Rutten (2000, Ex. 9.5)*[2]*). The final $S(Y)$-coalgebra is*

$$Y^{\omega} \xrightarrow{\langle \mathrm{hd}, \mathrm{tl} \rangle} Y \times Y^{\omega}$$

where Y^{ω} is the set of infinite streams over Y, $\mathrm{hd}(y_0 y_1 \dots) := y_0$, *i.e.,* hd *maps a stream to its first element (its "head") and* $\mathrm{tl}(y_0 y_1 y_2 \dots) := y_1 y_2 \dots$, *i.e.,* tl *maps a stream to its tail (cf. Rutten (2000, Ex. 9.4)).*

[2] *Loc.cit.* proves this for $I = 2$ but the argument can be easily adapted for arbitrary I.

The above final coalgebras are fundamental for our representation of iterated games: The final $S(Y)$-coalgebra consists of all infinite sequences of moves of the one-round game, while the final $D(Y, \Sigma)$-coalgebra represents the set of strategies that map lists of moves — representing moves chosen in previous rounds — to a strategy for the next round. As notation, for $\sigma : Y^* \to \Sigma$ we abbreviate $\mathrm{now}(\sigma)$ to σ_0, $\mathrm{ltr}(\sigma)$ to σ', and use $(::) : Y \times Y^\omega \to Y^\omega$ to denote the cons-operator on streams. We now define the ω-iteration of $\mathcal{G}$.

Definition 10. The ω-iteration $\mathcal{G}_\omega : (1, R) \to (Y^\omega, R)$ of $\mathcal{G} : (1, R) \to (Y, R)$ has strategies $\Sigma_{\mathcal{G}_\omega}$ given by $\Sigma_{\mathcal{G}_\omega} := Y^* \to \Sigma_{\mathcal{G}}$, and play function $P_{\mathcal{G}_\omega}$ given by

$$P_{\mathcal{G}_\omega}\sigma = P_{\mathcal{G}}\sigma_0 :: P_{\mathcal{G}_\omega}(\lambda z.\sigma(P_{\mathcal{G}}\sigma_0 :: z))$$

To define the equilibrium function $E_{\mathcal{G}_\omega} : (Y^\omega \to R) \to \mathscr{P}\Sigma_{\mathcal{G}_\omega}$, we first define an operator $\Phi : (\mathscr{P}\Sigma_{\mathcal{G}_\omega})^{(Y^\omega \to R)} \to (\mathscr{P}\Sigma_{\mathcal{G}_\omega})^{(Y^\omega \to R)}$ by putting

$$\sigma \in \Phi\,\Gamma\,k \quad \text{if} \quad \sigma_0 \in E_{\mathcal{G}}(\lambda y.k(y :: P_{\mathcal{G}_\omega}(\sigma' y))) \tag{1}$$
$$\text{and} \quad \forall y' \in Y.\ \sigma' y' \in \Gamma(\lambda z.k(y' :: z)) \tag{2}$$

Clearly $(\mathscr{P}\Sigma_{\mathcal{G}_\omega})^{(Y^\omega \to R)}$ forms a complete lattice by lifting the complete lattice structure of $\mathscr{P}\Sigma_{\mathcal{G}_\omega}$ pointwise to the function space. Furthermore, Φ is obviously a monotone operator on that complete lattice and therefore has a smallest and a greatest fixpoint. We define $E_{\mathcal{G}_\omega}$ to be the greatest fixpoint of Φ.

Notice that the above approach means we do not have to fix a particular utility function $Y^\omega \to R$ in advance by some arbitrary form of discounting, but rather work with all possible utility functions, allowing the user maximum flexibility. The definition of $E_{\mathcal{G}_\omega}$ contains a coinduction principle which we will use (i) in this section to characterise $\mathcal{G}_\omega$ as a final coalgebra; and (ii) in Section 6 to prove properties about equilibria in $\mathcal{G}_\omega$. The coinduction principle is

$$(\forall \quad \Gamma \in (\mathscr{P}\Sigma_{\mathcal{G}_\omega})^{(Y^\omega \to R)}) \quad (\Gamma \leq \Phi(\Gamma) \implies \Gamma \leq E_{\mathcal{G}_\omega})$$

where $\leq$ denotes the pointwise inclusion order on $(\mathscr{P}\Sigma_{\mathcal{G}_\omega})^{(Y^\omega \to R)}$ given by $\Gamma \leq \Delta$ if $\Gamma(k) \subseteq \Delta(k)$ for all $k : Y^\omega \to R$. This principle is sound, as the following lemma shows:

Lemma 11. *Let $\sigma \in \Sigma_{\mathcal{G}_\omega}$. Then* *(i) for all utility functions $k : Y^\omega \to R$ we have $\sigma \in E_{\mathcal{G}_\omega}(k)$ if and only if $\sigma \in \Phi(E_{\mathcal{G}_\omega})(k)$, and*

(ii) for every $\Gamma \in (\mathscr{P}\Sigma_{\mathcal{G}_\omega})^{(Y^\omega \to R)}$, if $\Gamma \leq \Phi(\Gamma)$ then also $\Gamma \leq E_{\mathcal{G}_\omega}$.

Proof. The first item follows since $E_{\mathcal{G}_\omega}$ is a fixpoint of Φ, the second because it is the greatest such, thus also the greatest post-fixpoint with respect to the order $\leq$. □

5.2. Proof of finality

In this section we are going to show that $\mathcal{G}_\omega$ is a final coalgebra of the functor $F_\mathcal{G} = (Y \to (-)) \circ \mathcal{G} : 2\mathsf{Open} \to 2\mathsf{Open}$. We have two things to show:

(i) $\mathcal{G}_\omega$ is an $F_\mathcal{G}$-coalgebra, and

(ii) for any other $F_\mathcal{G}$-coalgebra $\gamma : \mathcal{H} \to F_\mathcal{G}(\mathcal{H})$, there exists a *unique* $F_\mathcal{G}$-coalgebra morphism $(\mathrm{unf}_\Sigma, \mathrm{unf}_Y) : \mathcal{H} \to \mathcal{G}_\omega$.

The first item is formulated in the following proposition, and follows straightforwardly using Lemma 11:

Proposition 12. *The ω-iteration $\mathcal{G}_\omega$ of $\mathcal{G}$ is an $F_\mathcal{G}$-coalgebra with coalgebra map $\alpha = (\langle \mathrm{now}, \mathrm{ltr}\rangle, \langle \mathrm{hd}, \mathrm{tl}\rangle) : \mathcal{G}_\omega \to F_\mathcal{G}(\mathcal{G}_\omega)$.* □

We are now ready to prove that $\mathcal{G}_\omega$ indeed is the final $F_\mathcal{G}$-coalgebra. To this end we consider an arbitrary $F_\mathcal{G}$-coalgebra $\mathcal{H}$ with coalgebra map $(\langle \mathrm{now}_\mathcal{H}, \mathrm{ltr}_\mathcal{H}\rangle, \langle \mathrm{hd}_\mathcal{H}, \mathrm{tl}_\mathcal{H}\rangle)$. We have to prove that there is a morphism $\langle \mathrm{unf}_\Sigma, \mathrm{unf}_Y\rangle : \mathcal{H} \to \mathcal{G}_\omega$ such that the following diagram commutes:

$$\begin{array}{ccc} F_\mathcal{G}\mathcal{H} & \xrightarrow{F_\mathcal{G}(\mathrm{unf}_\Sigma, \mathrm{unf}_Y)} & F_\mathcal{G}\mathcal{G}_\omega \\ {\scriptstyle (\langle \mathrm{now}_\mathcal{H}, \mathrm{ltr}_\mathcal{H}\rangle, \langle \mathrm{hd}_\mathcal{H}, \mathrm{tl}_\mathcal{H}\rangle)}\uparrow & & \uparrow{\scriptstyle (\langle \mathrm{now}, \mathrm{ltr}\rangle, \langle \mathrm{hd}, \mathrm{tl}\rangle)} \\ \mathcal{H} & \xrightarrow[(\mathrm{unf}_\Sigma, \mathrm{unf}_Y)]{} & \mathcal{G}_\omega \end{array}$$

It is easy to see that such a $F_\mathcal{G}$-coalgebra morphism — if it exists — must be unique because commutativity of the above diagram implies commutativity of the following two diagrams in the category of sets:

$$\begin{array}{ccc} \Sigma_{\mathcal{G}} \times \Sigma_{\mathcal{H}}^{Y} & \xrightarrow{D(Y,\Sigma_{\mathcal{G}})(\mathrm{unf}_\Sigma)} & \Sigma_{\mathcal{G}} \times \Sigma_{\mathcal{G}_\omega}^{Y} \\ \uparrow{\scriptstyle \langle \mathrm{now}_{\mathcal{H}},\mathrm{ltr}_{\mathcal{H}}\rangle} & & \uparrow{\scriptstyle \langle \mathrm{now},\mathrm{ltr}\rangle} \\ \Sigma_{\mathcal{H}} & \xrightarrow[\mathrm{unf}_\Sigma]{} & \Sigma_{\mathcal{G}_\omega} \end{array} \qquad \begin{array}{ccc} Y \times Y_{\mathcal{H}} & \xrightarrow{\mathrm{id}_Y \times (\mathrm{unf}_Y)} & Y \times Y^{\omega} \\ \uparrow{\scriptstyle \langle \mathrm{hd}_{\mathcal{H}},\mathrm{tl}_{\mathcal{H}}\rangle} & & \uparrow{\scriptstyle \langle \mathrm{hd},\mathrm{tl}\rangle} \\ Y_{\mathcal{H}} & \xrightarrow[\mathrm{unf}_Y]{} & Y^{\omega} \end{array} \tag{3}$$

In other words unf_Σ and unf_Y have to be $D(Y, \Sigma_{\mathcal{G}})$- and $S(Y)$-coalgebra morphisms, respectively, and these are uniquely determined by the fact that their codomains are the respective final coalgebras.

This means that to show that $\mathcal{G}_\omega$ is a final $F_{\mathcal{G}}$-coalgebra, we have to prove that the pair of functions $\langle \mathrm{unf}_\Sigma, \mathrm{unf}_Y\rangle$ defined via the diagrams in (3) is a $F_{\mathcal{G}}$-coalgebra morphism. We need several lemmas.

Lemma 13. *For every $\sigma \in \Sigma_{\mathcal{H}}$ we have $\mathrm{unf}_Y(P_{\mathcal{H}}(\sigma)) = P_{\mathcal{G}_\omega}(\mathrm{unf}_\Sigma(\sigma))$.*

Proof. To see this, we define the relation

$$Q := \{(\mathrm{unf}_Y(P_{\mathcal{H}}(\sigma)), P_{\mathcal{G}_\omega}(\mathrm{unf}_\Sigma(\sigma))) \mid \sigma \in \Sigma_{\mathcal{H}}\} \subseteq Y^\omega \times Y^\omega \ .$$

Using Prop. 12 and (3), it is not hard to prove that Q is a $S(Y)$-bisimulation, i.e., that for each $(\tau_1, \tau_2) \in Q$ we have $\mathrm{hd}(\tau_1) = \mathrm{hd}(\tau_2)$ and $(\mathrm{tl}(\tau_1), \mathrm{tl}(\tau_2)) \in Q$. From the coinduction principle it follows that any two streams related by Q are equal, which implies the lemma. □

We now turn to the verification of the equilibrium condition for $(\mathrm{unf}_\Sigma, \mathrm{unf}_Y)$. First we use $(\mathrm{unf}_\Sigma, \mathrm{unf}_Y)$ to define an indexed predicate on $\Sigma_{\mathcal{G}_\omega}$ (which can be thought of as the image of $E_{\mathcal{H}}$ under $(\mathrm{unf}_\Sigma, \mathrm{unf}_Y)$). This predicate will be a post-fixpoint of Φ which will then imply the desired equilibrium condition.

Definition 14. We define an indexed predicate $\hat{E}_{\mathcal{H}} : (Y^\omega \to R) \to \mathscr{P}\Sigma_{\mathcal{G}_\omega}$ by putting $\sigma \in \hat{E}_{\mathcal{H}}k$ if $\exists \sigma' \in \Sigma_{\mathcal{H}}$ s.t. $\mathrm{unf}_\Sigma(\sigma') = \sigma$ and $\sigma' \in E_{\mathcal{H}}(k \circ \mathrm{unf}_Y)$.

Definition 15. We define a map $(\_)^* : (Y^\omega \to R) \to (Y \times Y_{\mathcal{H}} \to R)$ by putting $k^* = \lambda y.\lambda z.k(y :: \mathrm{unf}_Y(z))$.

Lemma 16. *For $k : Y^\omega \to R$ and $\sigma' \in \Sigma_{\mathcal{H}}$, if $\sigma' \in E_{\mathcal{H}}(k \circ \mathrm{unf}_Y)$, then*

(i) $\mathrm{now}_{\mathcal{H}}(\sigma') \in E_{\mathcal{G}}(\lambda y.k^*(y, P_{\mathcal{H}}(\mathrm{ltr}_{\mathcal{H}}(\sigma')(y))))$*, and*

(ii) for all $y' \in Y$ we have $\mathrm{ltr}(\sigma')(y') \in E_{\mathcal{H}}(\lambda z.k^*(y', z))$.

Proof. Suppose $\sigma' \in E_{\mathcal{H}}(k \circ \mathrm{unf}_Y)$. Observe that $k \circ \mathrm{unf}_Y = k^* \circ \langle \mathrm{hd}_{\mathcal{H}}, \mathrm{tl}_{\mathcal{H}} \rangle$, so this is equivalent to $\sigma' \in E_{\mathcal{H}}(k^* \circ \langle \mathrm{hd}_{\mathcal{H}}, \mathrm{tl}_{\mathcal{H}} \rangle)$ and — as $\langle \mathrm{hd}_{\mathcal{H}}, \mathrm{tl}_{\mathcal{H}} \rangle$ is a morphism of open games — we obtain $\langle \mathrm{now}_{\mathcal{H}}, \mathrm{ltr}_{\mathcal{H}} \rangle (\sigma') \in E_{F_{\mathcal{G}}\mathcal{H}}(k^*)$. The lemma now follows by spelling out the definition of $E_{F_{\mathcal{G}}\mathcal{H}}(k^*)$. □

We are now ready to prove the key fact that $\hat{E}_{\mathcal{H}}$ is a post-fixpoint of Φ.

Lemma 17. *Let $\sigma \in \Sigma_{\mathcal{G}_\omega}$ be a strategy such that $\sigma \in \hat{E}_{\mathcal{H}} k$ for some $k : Y^\omega \to R$. Then $\sigma \in \Phi(\hat{E}_{\mathcal{H}})(k)$.*

Proof. The assumption $\sigma \in \hat{E}_{\mathcal{H}} k$ means that there is some $\sigma' \in \Sigma_{\mathcal{H}}$ such that $\mathrm{unf}_\Sigma(\sigma') = \sigma$ and $\sigma' \in E_{\mathcal{H}}(k \circ \mathrm{unf}_Y)$. We need to verify that

(a) $\mathrm{now}(\sigma) \in E_{\mathcal{G}}(\lambda y.k(y :: P_{\mathcal{G}_\omega} \mathrm{ltr}(\sigma)(y)))$, and

(b) for all $y' \in Y$ we have $\mathrm{ltr}(\sigma)(y') \in \hat{E}_{\mathcal{H}}(\lambda z.k(y' :: z))$.

For (a), note that by the diagram for strategies in (3) we have $\mathrm{now}(\sigma) = \mathrm{now}(\mathrm{unf}_\Sigma(\sigma')) = \mathrm{now}_{\mathcal{H}}(\sigma')$. Using the first item of Lemma 16, we obtain

$$\begin{aligned}
\mathrm{now}(\sigma) \quad &\in \quad E_{\mathcal{G}}(\lambda y.k^*(y, P_{\mathcal{H}}(\mathrm{ltr}_{\mathcal{H}}(\sigma')(y)))) \\
&= \quad E_{\mathcal{G}}(\lambda y.k(y :: \mathrm{unf}_Y(P_{\mathcal{H}}(\mathrm{ltr}_{\mathcal{H}}(\sigma')(y))))) \\
&\overset{\text{Lemma 13}}{=} \quad E_{\mathcal{G}}(\lambda y.k(y :: P_{\mathcal{G}_\omega}(\mathrm{unf}_\Sigma(\mathrm{ltr}_{\mathcal{H}}(\sigma')(y))))) \\
&\overset{(3)}{=} \quad E_{\mathcal{G}}(\lambda y.k(y :: P_{\mathcal{G}_\omega}(\mathrm{ltr}(\mathrm{unf}_\Sigma(\sigma'))(y)))) \\
&= \quad E_{\mathcal{G}}(\lambda y.k(y :: P_{\mathcal{G}_\omega}(\mathrm{ltr}(\sigma)(y))))
\end{aligned}$$

which establishes (a).

For (b), it suffices to define for each $y' \in Y$ a suitable strategy $\sigma'_{y'} \in \Sigma_{\mathcal{H}}$ such that $\mathrm{unf}_\Sigma \sigma'_{y'} = \mathrm{ltr}(\sigma)(y')$ and $\sigma'_{y'} \in E_{\mathcal{H}}(\lambda z.k(y' :: \mathrm{unf}_Y z))$. We claim that for an arbitrary $y' \in Y$ the strategy $\sigma'_{y'} := \mathrm{ltr}_{\mathcal{H}}(\sigma')(y')$ meets these conditions. The first condition is again an easy consequence of (3) and the fact that $\mathrm{unf}_\Sigma(\sigma') = \sigma$. For the second condition we note that $\sigma'_y \in E_{\mathcal{H}}(\lambda z.k^*(y', z))$ as a consequence of $\sigma' \in E_{\mathcal{H}}(k \circ \mathrm{unf}_Y)$ and the second item of Lemma 16. The claim follows now from $\sigma'_{y'} \in E_{\mathcal{H}}(\lambda z.k^*(y', z)) = E_{\mathcal{H}}(\lambda z.(y' :: \mathrm{unf}_Y(z)))$. □

We are now ready to prove the main theorem of this section.

Theorem 18. *Let $\mathcal{G} : (1, R) \to (Y, R)$ be an open game and let $\mathcal{G}_\omega$ be its ω-iteration. Then $\mathcal{G}_\omega$ is a final $F_{\mathcal{G}}$-coalgebra.*

Proof. By our discussion at the beginning of this subsection it suffices to show that for an arbitrary $F_{\mathcal{G}}$-coalgebra $(\mathcal{H}, (\langle \mathrm{now}_{\mathcal{H}}, \mathrm{ltr}_{\mathcal{H}} \rangle, \langle \mathrm{hd}_{\mathcal{H}}, \mathrm{tl}_{\mathcal{H}} \rangle))$ the map $(\mathrm{unf}_{\Sigma}, \mathrm{unf}_Y)$ consisting of the coalgebra morphisms in (3) is a morphism of open games. Lemma 13 shows that $(\mathrm{unf}_{\Sigma}, \mathrm{unf}_Y)$ satisfies the play function morphism condition. For checking the equilibrium condition, consider an arbitrary $\sigma' \in \Sigma_{\mathcal{H}}$ and a $k : Y^{\omega} \to R$ such that $\sigma' \in E_{\mathcal{H}}(k \circ \mathrm{unf}_Y)$. Then clearly we have $\mathrm{unf}_{\Sigma}(\sigma') \in \hat{E}_{\mathcal{H}}(k)$. As $\hat{E}_{\mathcal{H}}$ is a post-fixpoint of Φ by Lemma 17, we have $\hat{E}_{\mathcal{H}}(k) \subseteq E_{\mathcal{G}_{\omega}}(k)$, and thus $\mathrm{unf}_{\Sigma}(\sigma') \in E_{\mathcal{G}_{\omega}}(k)$ as required. □

6. Using coinduction to reason about infinite games

In this section we show how the coinduction principle inherent in the definition of the equilibrium $E_{\mathcal{G}_{\omega}}$ of the ω-iteration of a game $\mathcal{G}$ as a final coalgebra can be used to reason about equilibria in infinite games.

We will spell out a coinductive proof of the well-known fact that the strategy pair consisting of two grim trigger strategies is a subgame perfect equilibrium of the Iterated Prisoner's Dilemma for sufficiently large discount factors (McGillivray and Smith, 2000). To this aim consider the infinite iteration $\mathcal{PD}_{\omega}$ of the Prisoner's Dilemma game $\mathcal{PD}$ from Example 2. Recall that the one-round game

$$\mathcal{PD} : (1, \mathbb{R} \times \mathbb{R}) \xrightarrow{\Sigma_{\mathcal{PD}}} (Y, \mathbb{R} \times \mathbb{R})$$

has $\Sigma_{\mathcal{PD}} = Y = \{C, D\} \times \{C, D\}$, the coplay and play functions are identity functions, and the equilibrium $E_{\mathcal{PD}} : (Y \to \mathbb{R} \times \mathbb{R}) \to \Sigma_{\mathcal{PD}}$ is given by

$$\begin{aligned}(s_0, s_1) \in E_{\mathcal{PD}}(u : Y \to \mathbb{R} \times \mathbb{R}) \text{ if } & \pi_0 u(s_0, s_1) \geq \pi_0 u(t, s_1) \\ \text{and } & \pi_1 u(s_0, s_1) \geq \pi_1 u(s_0, t) \quad \text{for all } t \in \{C, D\}\end{aligned}$$

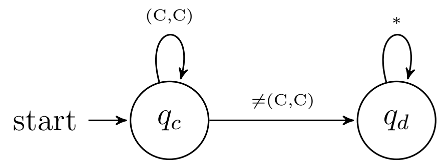
Furthermore we consider the utility function $u_{\mathcal{PD}}$ of the one-round Prisoner's Dilemma as specified in Figure 1 on page 4. The discounted utility for the Iterated Prisoner's Dilemma $k^{\beta}_{\mathcal{IPD}} : Y^{\omega} \to \mathbb{R}$ with discount factor $\beta \in [0, 1)$ is then given as

$$k^{\beta}_{\mathcal{IPD}}(\rho) = \sum_{i=0}^{\infty} \beta^i \cdot u_{\mathcal{PD}}(\rho(i))$$

where $\rho(i)$ for $i \geq 0$ denotes the ith element of a stream $\rho \in Y^{\omega}$.

A strategy in $\mathcal{PD}_\omega$ is an element of the final coalgebra of the functor $D(Y, \Sigma_{\mathcal{PD}})$. This can be represented as a finite automaton that reads in sequences of moves in $\mathcal{PD}$ and returns the strategy for the next round of the game (Rutten, 1998).

Definition 19. The grim trigger strategy is given by the following automaton:



where q_c is the initial state, the transitions are labelled by inputs (the $*$ denotes an arbitrary letter) and the outputs are given by $o(q_c) = (C, C)$ and $o(q_d) = (D, D)$.

We see that the grim trigger strategy stays in the 'cooperate state' q_c as long as both players cooperate, but moves to — and then stays forever in — the 'defect state' q_d as soon as one player defects. We will now prove that this strategy is an equilibrium of $\mathcal{PD}_\omega$, if the discount factor β is large enough. Recall that the equilibrium $E_{\mathcal{PD}_\omega}$ is defined as a greatest fixpoint of the operator

$$\Phi : (\mathscr{P}\Sigma_{\mathcal{PD}_\omega})^{(Y^\omega \to \mathbb{R}\times\mathbb{R})} \to (\mathscr{P}\Sigma_{\mathcal{PD}_\omega})^{(Y^\omega \to \mathbb{R}\times\mathbb{R})}$$

via conditions (1) and (2) on page 9. Therefore we will prove our claim by demonstrating that the grim trigger strategy is contained in a post-fixpoint of Φ; this is sufficient by Lemma 11. To ease presentation, we introduce the following notation:

Notation 20. Given a $D(Y, \Sigma_{\mathcal{PD}})$-coalgebra $\langle \text{now}, \text{ltr}\rangle : X \to \Sigma_{\mathcal{PD}} \times X^Y$, and a word $w \in Y^*$, we put for elements $x \in X$ and functions $k : Y^\omega \to \mathbb{R}\times\mathbb{R}$

$$\begin{aligned} x(w) &:= \widehat{\text{ltr}(x)}(w) : X \\ k_w &:= \lambda y.k(wy) : Y \to X \end{aligned}$$

Here $\widehat{\text{ltr}(x)}$ denotes the obvious extension of $\text{ltr}(x) : Y \to X$ to words.

Intuitively $x(w)$ is the state that is reached in an automaton when starting at position x and reading the input word w. Likewise, k_w should be thought of as calculating the payoff under the assumption that a finite sequence of moves w has been made already.

We now move to the definition of a post-fixpoint χ^β of Φ.

Definition 21. We define a function $\chi^\beta : (Y^\omega \to \mathbb{R} \times \mathbb{R}) \to \mathscr{P}\Sigma_{\mathcal{PD}_\omega}$ by

$$\chi^\beta(k) := \left\{ q_c(w) \mid w \in Y^*, k = (k^\beta_{\mathcal{IPD}})_w \right\}$$

where q_c denotes the initial state of the grim trigger strategy.

Since obviously $q_c \in \chi^\beta(k^\beta_{\mathcal{IPD}})$, in order to prove that the grim trigger strategy is an equilibrium (for large enough discount factor β), i.e. $q_c \in E_{\mathcal{PD}_\omega}(k^\beta_{\mathcal{IPD}})$, it suffices to prove the following lemma:

Lemma 22. *The indexed predicate χ^β as defined above is a post-fixpoint of Φ, i.e.,*

$$\chi^\beta(k) \subseteq \Phi(\chi^\beta)(k) \text{ for all } k : Y^\omega \to \mathbb{R} \times \mathbb{R},$$

if and only if $\beta \geq \frac{1}{3}$.

Proof. To prove this, we consider an arbitrary $k : Y^\omega \to \mathbb{R} \times \mathbb{R}$ and show that the inclusion holds. Let $q \in \chi^\beta(k)$. Then by definition of χ^β there is some $w \in Y^*$ such that $q = q_c(w)$ and $k = (k^\beta_{\mathcal{IPD}})_w$. We need to verify the following conditions:

1. $\mathrm{now}(q_c(w)) \in E_{\mathcal{PD}}(\lambda y.k_y(P_{\mathcal{PD}_\omega}(\mathrm{ltr}(q_c(w))(y))))$, and
2. $\forall y' \in Y.\ \mathrm{ltr}(q_c(w))(y') \in \chi^\beta(k_{y'})$.

Note that the second condition holds trivially by definition of χ^β, so we only need to check the first condition. We do this using a case distinction on $q_c(w)$.

Case $\boxed{q_c(w) = q_c}$ In this case we have $\mathrm{now}(q_c(w)) = (C, C)$ and it is easy to see that $w = (C, C) \ldots (C, C)$, $P_{\mathcal{PD}_\omega}(q_c(w(C, C))) = (C, C)^\omega$ and $P_{\mathcal{PD}_\omega}(q_c(wy)) = (D, D)^\omega$ for all $y \in \{(C, D), (D, C), (D, D)\}$.

In order to check whether $\mathrm{now}(q_c(w)) \in E_{\mathcal{PD}}(\lambda y.k_y(P_{\mathcal{PD}_\omega}(q_c(wy))))$ we have to ensure that

$$\begin{aligned} \pi_0 k_{(C,C)}((C, C)^\omega) &\geq \pi_0 \left(k_{(D,C)}(D, D)^\omega\right) \\ \pi_1 k_{(C,C)}((C, C)^\omega) &\geq \pi_1 \left(k_{(C,D)}(D, D)^\omega\right) \end{aligned}$$

We use n to denote the length of the word $w = (C,C)\dots(C,C)$ (where $n = 0$ means that w is the empty word). Then

$$\begin{aligned}
k_{(C,C)}((C,C)^\omega) &= (-\sum_{i=0}^{\infty}\beta^i, -\sum_{i=0}^{\infty}\beta^i) \\
k_{(C,D)}((D,D)^\omega) &= k^{\beta}_{\mathcal{IPD}}(w(C,D)(D,D)^\omega) \\
&= (-\sum_{i=0}^{n-1}\beta^i - 3\sum_{i=n+1}^{\infty}\beta^i, -\sum_{i=0}^{n-1}\beta^i - 5\beta^n - 3\sum_{i=n+1}^{\infty}\beta^i) \\
k_{(D,C)}((D,D)^\omega) &= (-\sum_{i=0}^{n-1}\beta^i - 5\beta^n - 3\sum_{i=n+1}^{\infty}\beta^i, -\sum_{i=0}^{n-1}\beta^i - 3\sum_{i=n+1}^{\infty}\beta^i)
\end{aligned}$$

Due to symmetry, the equilibrium condition amounts to the following inequality:

$$-\sum_{i=0}^{\infty}\beta^i \geq -\sum_{i=0}^{n-1}\beta^i - 3\sum_{i=n+1}^{\infty}\beta^i$$

Cancelling the common history and using the formula $\sum_{i=n}^{\infty}\beta^i = \frac{\beta^n}{1-\beta}$, we see that this holds if and only if $\beta \geq \frac{1}{3}$.

Case $\boxed{q_c(w) = q_d}$ The argument works analogously. We now have $(q_d(w)) = (D,D)$ for all histories w and that $P_{\mathcal{PD}_\omega}((q_d(wy))) = (D,D)^\omega$ for all $y \in Y$.

To check that $q_d(w) \in E_{\mathcal{PD}}(\lambda y.k_y(P_{\mathcal{PD}_\omega}(q_c(wy))))$ we need to show that

$$\begin{aligned}
\pi_0 k_{(D,D)}((D,D)^\omega) &\geq \pi_0\left(k_{(C,D)}(D,D)^\omega\right) \\
\pi_1 k_{(D,D)}((D,D)^\omega) &\geq \pi_1\left(k_{(D,C)}(D,D)^\omega\right)
\end{aligned}$$

As in the previous case, the utility produced from the history w will cancel, and we will be left with the inequality

$$-3\sum_{i=0}^{\infty}\beta^i \geq -5 - 3\sum_{i=1}^{\infty}\beta^i$$

reducing to the true statement $-3 \geq -5$, independent of the value of β. $\square$

In summary, we have shown:

Proposition 23. *The grim trigger strategy is an equilibrium of the Iterated Prisoner's Dilemma game with discount factor β, i.e. $q_c \in E_{\mathcal{PD}_\omega}(k^{\beta}_{\mathcal{IPD}})$, if $\beta \geq \frac{1}{3}$.* □

7. Conclusions and Future Work

The main contributions of this paper are on the one hand a notion of morphism between open games and — based on this notion — the representation of the infinite iteration of a given game as a final coalgebra. This provides a first extension of the compositionality results from Ghani et al. (2018) to infinitely repeated games. Nevertheless a number of challenges remain: firstly, we need to extend our construction to state-full games and to games with non-trivial coutility function. The former seems straightforward. Of course we want to use this framework to provide new reasoning tools for such games based on coinduction and coalgebraic logics.

Acknowledgements. This work was supported by EPSRC grants EP/M016951/1, EP/K023837/1, and EP/N015843/1. We thank Jules Hedges, Viktor Winschel, Philipp Zahn, Alexander Kurz, and the participants at the Open Games workshop in Glasgow 2017 for interesting discussions about open games and their iteration.

References

Abramsky, S., Winschel, V., 2017. Coalgebraic analysis of subgame-perfect equilibria in infinite games without discounting. Mathematical Structures in Computer Science 27 (5), 751–761.

Axelrod, R., Dion, D., 1988. The further evolution of cooperation. Science 242 (4884), 1385–1390.

Ghani, N., Hedges, J., Winschel, V., Zahn, P., 2018. Compositional game theory. To appear at LICS 2018. arXiv:1603.04641.

Lescanne, P., 2012. Rationality and escalation in infinite extensive games. arXiv:1112.1185.

Mailath, G., Samuelson, L., 2006. Repeated Games and Reputations: Long-Run Relationships. Oxford University Press.

McGillivray, F., Smith, A., 2000. Trust and cooperation through agent-specific punishments. International Organization 54 (4), 809–824.

Nash, J., 1951. Non-cooperative games. Annals of Mathematics 54 (2), 286–295.

Oliva, P., Powell, T., 2015. A constructive interpretation of Ramsey's theorem via the product of selection functions. Mathematical Structures in Computer Science 25 (8), 1755–1778.

Rutten, J., 2000. Universal coalgebra: a theory of systems. Theoretical Computer Science 249 (1), 3–80.

Rutten, J., Turi, D., 1994. Initial algebra and final coalgebra semantics for concurrency. In: de Bakker, J. W., de Roever, W. P., Rozenberg, G. (Eds.), A Decade of Concurrency Reflections and Perspectives. Springer, pp. 530–582.

Rutten, J. J. M. M., 1998. Automata and coinduction (an exercise in coalgebra). In: Sangiorgi, D., de Simone, R. (Eds.), CONCUR'98. Springer, pp. 194–218.

Selten, R., 1965. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit: Teil I: Bestimmung des Dynamischen Preisgleichgewichts. Zeitschrift für die gesamte Staatswissenschaft 121 (2), 301–324.

Shubik, M., 1984. Game Theory In The Social Sciences. The MIT Press.