Mathematical Biology

Reaction networks and evolutionary game theory

**Tomas Veloz · Pablo Razeto-Barry ·
Peter Dittrich · Alejandro Fajardo**

Abstract The powerful mathematical tools developed for the study of large scale reaction networks have given rise to applications of this framework beyond the scope of biochemistry. Recently, reaction networks have been suggested as an alternative way to model social phenomena. In this "socio-chemical metaphor" molecular species play the role of agents' decisions and their outcomes, and chemical reactions play the role of interactions among these decisions. From here, it is possible to study the dynamical properties of social systems using standard tools of biochemical modelling.

T. Veloz (✉)
Mathematics Department, University of British Columbia, Kelowna, BC V1V 1V7, Canada
e-mail: tomas.veloz@ubc.ca

T. Veloz · P. Razeto-Barry
Instituto de Filosofía y Ciencias de la Complejidad IFICC, Los Alerces 3024,
Ñuñoa, Santiago, Chile
e-mail: prazeto@ificc.cl

P. Razeto-Barry
Universidad Diego Portales, Vicerrectoría Académica, Manuel Rodríguez Sur 415,
Santiago, Chile

P. Dittrich
Friedrich-Schiller-University Jena, Institute of Computer Science Bio Systems
Analysis Group, 07743 Jena, Germany
e-mail: peter.dittrich@uni-jena.de

A. Fajardo
Departamento de Biología, Universidad de los Andes, Cra 1 N 18A-12, Bogotá, Colombia
e-mail: alfajardoa@gmail.com

In this work we show how to use reaction networks to model systems that are usually studied via evolutionary game theory. We first illustrate our framework by modeling the repeated prisoners' dilemma. The model is built from the payoff matrix together with assumptions of the agents' memory and recognizability capacities. The model provides consistent results concerning the performance of the agents, and allows for the examination of the steady states of the system in a simple manner. We further develop a model considering the interaction among Tit for Tat and Defector agents. We produce analytical results concerning the performance of the strategies in different situations of agents' memory and recognizability. This approach unites two important theories and may produce new insights in classical problems such as the evolution of cooperation in large scale systems.

1 Introduction

A reaction network consists of a set of possible units (species) that can exist in a system, and the interactions (reactions) among these species. Discrete or continuous dynamical laws are imposed to study the evolution of a reaction network, e.g. the changes in the species' concentrations with time (Feinberg and Horn 1974; Gillespie 2007). Since the seventies, a variety of mathematical tools grounded in dynamical systems, complexity theory, abstract algebra, and others, have been applied to the study of large scale biochemical systems where numerical methods and computational simulations do not suffice (Kacser and Burns 1973; Feinberg and Horn 1974; Reddy et al. 1993; Fontana and Buss 1994; Schilling and Palsson 1998; Schuster et al. 1999; Dittrich and Speroni di Fenizio 2008; Österlund et al. 2011).

The success of these applications has motivated the use of reaction networks as a framework to model systems in domains beyond biochemistry such as ecology, theoretical and applied computer science, among others (for a review see Dittrich 2009). Recently, a reaction network model of the political system has been developed by Dittrich and Winter (2007). They were able to identify the stable social structures of this particular system using tools from a novel reaction network scheme called chemical organization theory (Dittrich and Speroni di Fenizio 2008). The fundamental difference of this approach, with respect to other social system approaches, is that the unit of modeling are the agents' *communications* (e.g. decisions) instead of the agents themselves. Molecular species represent communications, and the interaction of two or more communications consume and produce new communications. A social stable structure in this setting corresponds to a set of communications that self-maintains in time, analogous to the biological notion of autopoietic structure (Maturana and Varela 1974; Razeto-Barry 2012). Interestingly, the first scholar arguing that *stable social structures are not formed by agents, but communications* (in this case represented by decisions) was Luhmann (1986) in the context of sociology, and Dittrich and Winter

(2007) formalized this idea using reaction networks. This suggests that the modelling of social systems by means of reaction networks may open a path for a promising theoretical treatment of large scale social and biological systems.

On the other hand, Game Theory has been one of the most important theoretical frameworks to analyze social problems in economics, political sciences, anthropology and biological and cultural evolution, under the label of evolutionary game theory (EGT) (Binmore 2007; Gintis 2008; Rasmusen 2007; Taylor 1995; Vincent and Brown 2007; Nowak 2006a; Bowles and Gintis 2011). Particularly, EGT has proved to be one of the most powerful conceptual and practical approaches to understand the classical problem of the evolution of cooperation (Maynard Smith 1982; Axelrod 1984, 1997). Nowak (2006b) reviews five mechanisms that may explain the evolution of cooperation, all of them based in EGT. The simple formulation of these mechanisms permit the development of analytical expressions which predict under what conditions cooperation may surge. Nevertheless, in most of cases analytical results are impossible or very difficult to obtain in social and biological problems. Models are then studied via simulations, even for situations where simple strategies are interacting (Nowak and May 1992; Nowak and Sigmund 1993; Axelrod 1997). Social and biological models based on Game Theory take agents as fundamental units (Axelrod 1997; Hofbauer and Sigmund 1998). The latter is a stringent condition that may be generalized using reaction networks.

Here we illustrate how to model the EGT systems using reaction networks. Our aim is to apply, for the first time as far as we know, analytical tools used in reaction networks to study Game Theory. Furthermore, we show the similitudes and differences between reaction networks and agent based approaches. Emphasis is made on the capacity of reaction networks to generate models that are not based in the agents as the fundamental unit, but in their communications, that require higher abstraction but are more general (Luhmann 1995; Dittrich and Winter 2007).

The paper is organized as follows: In Sect. 2 we introduce the standard definitions of reaction networks. In Sect. 3 we illustrate the socio-chemical metaphor, modelling the repeated prisoner's dilemma using reaction networks. In Sect. 4 we build a reaction network model for a system with agents following Tit for Tat and Defector strategies (Axelrod 1984) considering different situations of agents' recognizability and memory capacities, and we study the performance of the strategies for each case applying reaction networks analysis methods. We conclude and present a discussion of the methodology introduced in this work in Sect. 5.

2 Reaction networks

In a reaction network we deal with two types of objects: *molecular species* (from now on species) and *reactions*. Species are the elements of a *species set* $\mathcal{M} = \{m_1, \ldots, m_n\}$, and each reaction R is modeled by a pair $R = (A, B)$, where $A, B \in \mathscr{P}_M(\mathcal{M})$, and $\mathscr{P}_M(\mathcal{M})$ denotes the set of *multisets* formed by elements in $\mathcal{M}$. A multiset A is defined by a pair $(\mathcal{M}, \eta_A)$, where the function $\eta_A : \mathcal{M} \to \mathbb{N}_0$ states the number of *occurrences* $\eta_A(m)$ (or *multiplicity*) of m in A.

To be consistent with the notation used in chemical literature, we will denote the multiset A by $A = \sum_{m_i \in \mathcal{M}} a_i m_i$, i.e. each species m_i is preceded by its multiplicity $\eta_A(m_i)$ denoted by a_i. Moreover, we will refer the reaction $R = (A, B)$ by $R = A \rightarrow B$.

From now on, let $\mathcal{R} = \{R_1, \ldots, R_r\}$, where $R_i = A_i \rightarrow B_i$, $A_i = a_{i1}m_1 + \cdots a_{in}m_n$ and $B_i = b_{i1}m_1 + \cdots b_{in}m_n$, for $i = 1, \ldots, r$. Now we can define a *reaction network*, which captures the notion of a chemical system by specifying the species and reactions that may occur.

Definition 1 A reaction network is a pair $\langle \mathcal{M}, \mathcal{R} \rangle$.

2.1 Dynamical aspects

In a reaction network, the occurrence of reactions leads to a dynamical process of consumption and production of species. In order to represent this process, we introduce the *stoichiometric matrix* $\mathbf{S} = (s_{ij})$ associated to $\langle \mathcal{M}, \mathcal{R} \rangle$. $\mathbf{S}$ is a $n \times r$ matrix, where n is the number of species and r is the number of reactions in the reaction network, and $s_{ij} = b_{ji} - a_{ji}$ is the *stoichiometric* coefficient of species m_i in the reaction R_j. If s_{ij} is positive, then m_i is *produced* by reaction R_j. If s_{ij} is negative, then m_i is *consumed* by reaction R_j. Thus, the stoichiometric matrix represents how the species are consumed and produced by the reactions. The stoichiometric matrix is at the core of the reaction network analysis (Schuster et al. 1999; Schilling and Palsson 1998), and its properties have been extensively studied (Kacser and Burns 1973; Heinrich and Rapoport 1974).

To model the occurrence of reactions in a reaction network, we introduce a non-negative *flux vector* $\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_r)$. For each $i = 1, \ldots, r$, we have that $\mathbf{v}_i$ represents the *rate of occurrence* of reaction R_i in the reaction network. Thus, the application of $\mathbf{v}$ on the stoichiometric matrix $\mathbf{S}$ represents a reaction process where the rate of the reaction R_i in the system is given by $\mathbf{v}_i$ for $i = 1, \ldots, r$. Hence, it is natural to define the *production rate vector* by $\mathbf{f} = \mathbf{S}\mathbf{v}$. For $i = 1, \ldots, n$, we have that $\mathbf{f}_i$ is the rate of production of the species m_i in the reaction process determined by $\mathbf{v}$.

The usual manner to describe the dynamics of the species concentrations $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, is by applying the mass-action kinetics law. This law assumes that for each $i = 1, \ldots, r$, the ith coordinate $\mathbf{v}_i$ of the flux vector depends on the concentration of the species and a non-negative vector $\mathbf{k} = (\mathbf{k}_1, \ldots, \mathbf{k}_r)$ of reaction rates as follows

$$\mathbf{v}_i = \mathbf{k}_i \prod_{j=1}^{r} \mathbf{x}_j^{a_{ji}} \tag{1}$$

Thus, the dynamics of the species' concentration is described by the following system of ODEs

$$\dot{\mathbf{x}} = \mathbf{S}\mathbf{v}(\mathbf{x}, \mathbf{k}), \tag{2}$$

We call the system (2) a *chemical reaction system*.

Remark The dynamical aspects of a reaction network can also be defined in terms of deterministic or stochastic discrete processes. We will stick here to the continuous schema for illustrative purposes. However, the discrete and continuous schemas are complementary (Heiner et al. 2008).

3 Modelling evolutionary games using reaction networks

3.1 The chemical metaphor in social systems

Luhmann introduced the notion of communication as the basis of societies' structuring and ordering (Luhmann 1986; Razeto-Barry and Cienfuegos 2011). Luhmann's concept of communication is defined as the flow produced by the exchange of *social-symbols*. These symbols depend on the system. For example, for simple economical, legal and political systems, the communication flow is done through *money*, *justice* and *power* respectively. In a general case, all these systems have some degree of entanglement, and hence, communications in one system may affect the others.

Luhmann's communication concept has been formalized and applied by Dittrich and Winter (2007) in a *toy-model* of the political system. They define 13 communications, e.g. *social movement demands* (s_d), *social movement members* (s_m), *potential collective binding decisions* (p_{cbd}), and a set of 20 reactions to model the interactions among these communications. For example, the reaction $R = s_m + p_{cbd} \rightarrow s_d$ models that *social movement members decisions concerning a potential collective binding decision implies social movement demand decisions*. The molecule p_{cbd} might correspond to a potential law such as *increase the tax*, s_d corresponds to the communications that the social movement members discuss or spread (for example within social networks), and s_m corresponds to a social movement demand decision (e.g *do not increase the tax*), that may be expressed as a protest, or by other actions.

In chemical systems, species are determined by their concentration. In these models, the concentration of species x represents the *visibility of the communication x* in the communication system as a whole. To illustrate this clearly, Dittrich and Winter's social model also considers the reaction $s_d \rightarrow p_{cbd}$, which means that potential collective binding decisions can be generated from (and proportional to) the social movement demands, and add decaying reactions for each communication, to model that communications become less influential with time.

This *communication species based* model is essentially different to any agent based model. It neglects the particular actions that each agent takes, and focuses on the visibility of the communications (decisions) that are present in the system. This approach entails the discovery of the most visible communications, and the groups of communications that are able to self-replicate by the occurrence of the reactions, and hence become stable in time. These sets of self-replicable communications may represent higher level communication structures such as *rebel* or *repressive* social tendencies. Hence, the socio-chemical metaphor provides an elegant framework to study *emergent* structures in social systems.

Note that in Dittrich's model, each type of molecular species represents a decision. In principle however, molecular species can represent other kinds of elements required

for the occurrence of the interactions. In this paper we will consider two types of such elements: Decision and payoff species, and we will refer in general to such elements as *communications*.

3.2 Agents, memory and recognizability

One important difference between the communication species based model presented here and the agent based models resides at the cognitive capacities that the unit of modeling hold in each paradigm. In the agent based models, the unit of modeling (the agent) may possess memory, recognize common phenotypes (e.g. green beard effect, Gardner and West 2009), or hold other complex cognitive abilities associated to its interaction process (Bowles and Gintis 2011; Hamilton 1964). These abilities are the basis of important mechanisms of natural selection such as group or kin selection. In contrast, communication species have no cognitive abilities. Instead, this paradigm identifies the possible interactions that occur within agents that have these complex abilities, and then builds a list of "reactions" that establish the exchange of the elements involved in these interactions (see Fig. 1). Thus, social complexity is introduced in the reactions rather in the intrinsic properties of the interacting species. This is coherent with the Luhmannian framework which emphasizes that social structures can be represented only by the visible interactions among agents rather by the intrinsic properties of agents (Luhmann 1995).
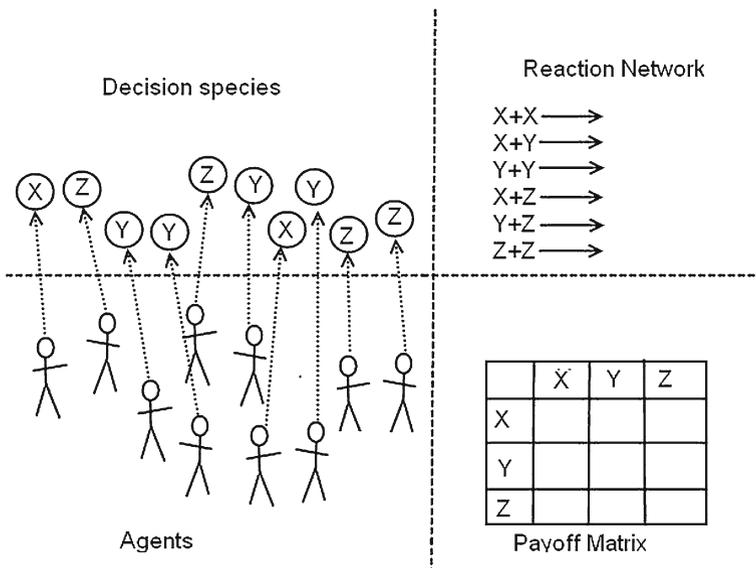
Fig. 1 Contrast between agent-based (*bottom*) and decision species (*top*) models. The interaction among agents corresponds to a vessel of decision species interacting, and the payoff matrix corresponds to a set of reactions which consumes a pair of decisions to produce the payoff of each decision and two new decisions determined by the strategies

We will show that when we are interested in population dynamics and long-term changes, memory and recognizability between agents can be represented by certain kinds interaction rules between communications. Therefore, the representation of the intrinsic properties of the agents can be substituted by the representation of such rules of communication.

The ability that agents have to recognize and remember each other implies that the interactions between two or more types of strategy (e.g. one type of strategy invading another) can be studied from the specific agent-agent interactions that occur among all the types of agent together (Axelrod 1984). Thus, the extent to which agents can recognize other agents is a determinant element in agent based modeling. This feature has been explored in EGT (Nowak 2006b) in the context of the evolution of cooperation, where five important mechanisms of recognizability have been formalized (among members of the same family, of the same group, of the same network, of the same society, of past interactions).

Situations of partial or null recognizability among agents are practically equivalent to situations when agents have failures in memory (e.g. an agent does not remember what its opponents did) or make involuntary mistakes (an agent makes a *move* when it tried to make another one) (Axelrod 1984). Moreover, it has been argued based on computer simulations and real life examples, that a mechanism called *generalized reciprocity* might be useful to explain the evolution of cooperation in small groups (Pfeiffer et al. 2005). In this mechanism agents do not necessarily recognize the agents they are interacting with, and base their actual decision only on the result of their last interaction (no matter with whom it was). Thus, generalized reciprocity does not assume advanced cognitive abilities of the agents. Hence, it is a powerful mechanism to explain natural behaviors in animals with little cognitive abilities.

In our framework, communication species are the fundamental units of modelling instead of agents. Hence, when decision species interact there is no identification of the agent that takes each decision. This is because our unit of modelling (decision species) does not possess memory. However, it is possible to simulate the long term effects of memory in a closed system (see Sect. 4.3), and establish the recognition among different types of agents (strategies) in our framework (see Sect. 4.4). Strategy recognition is an instance of recognition used in studies on downstream indirect reciprocity, such us reputation recognition, which require high level cognitive capacities, and by which probably is only applicable to humans (see West et al. 2011). An agent type is determined by the strategy that performs, and each strategy determines its decisions according to the possible interactions that occur in the agents' setting. To model the interaction of different types of agents, we label each decision according to two features: the strategy (type of agent) from which the decision comes from, and the type of decision that the species represents (e.g. cooperate or defect). From here, the species interactions (reactions) are built to represent the possible interactions among the decisions of each strategy (see Fig. 1).

Despite the differences between the agent and communication species approaches concerning recognizability, we will show that the initial motivation of EGT, i.e. to study the interactions among different types of agents, can be still grasped by our approach. We claim that the communication species approach is a complementary setup to the agent based approach. It provides the means to study the qualitative

dynamics of the strategies, i.e. analytic results concerning the asymptotic disappearing or growth of the concentration decisions in the system, as well as steady states where different decisions balance their concentrations. This complementarity is analogous to the mutual feedback between reaction networks analysis and *in silico* chemical simulations (Wiechert 2002; Österlund et al. 2011).

3.3 Modelling decisions and payoffs using molecular species

In EGT, each player (or agent) has a *strategy*, i.e. a set of rules that determines how the player is going to play. In EGT, strategies are not the result of a rational decision process by the players, as in classical Game Theory, but may be imposed by the genotype or cultural influence that each player has (Richerson and Boyd 2005; Bowles and Gintis 2011). EGT studies scenarios of frequency dependent selection: the players payoffs depend on the composition of the population. With an appropriate system of differential equations, it is possible to represent the changes of the populations of the different types of agents in time. Although there are many ways to construct this system, the replicator equation (Taylor and Jonker 1978) is widely used, essentially for its simplicity and because its captures our intuition: the proportion of the strategies that have higher payoff than the average grows, while the proportion of the strategies that perform under the average diminishes.

Note that the replicator equation implies a reaction network: it can be written as $\dot{x}_k = x_k(f_k(x) - \phi(x))$, where x_k is the proportion of players using strategy k, $f_k(x)$ is their fitness and $\phi(x) = \sum_{i=1}^{n} x_i f_i(x)$ the average fitness of the population (dilution flow in chemical reactions) represented by $x = (x_1, \ldots, x_n)$. The reaction network consists of decay reactions of the form $x_k \rightarrow \emptyset$ due to the dilution term $-x_k\phi(x)$.

The production term $x_k f_k(x)$ implies catalytic reactions of the form $x_k + i_1 + \ldots i_m \rightarrow 2x_k + i_1 + \ldots i_m$, where $i_1, \ldots i_m$ are those species that are necessary for k to replicate, i.e., $f(x)$ is positive if these species have positive concentrations in x. When the replicator equation is applied in EGT, a node of the network represents an agent type, i.e., a strategy, and x_k the amount of agents playing strategy k. This is fundamentally different to the approach followed here, where a node of the reaction network represents a decision or payoff species.

The canonical example of a two player game is the prisoner's dilemma. In this game, two prisoners are kept in different rooms and each is offered the same deal: if one testifies (defects) against the other and the other remains silent (cooperates), the former goes free while the latter receives a ten years sentence. If both remain silent, they get each a one year sentence, and if they both testify against the other, both prisoners get a seven year sentence. The game, in its more general form, can be represented using a payoff matrix (see Table 1), where b represents the benefit of

Table 1 Payoff table for the interaction of cooperative and defector decisions

Gain/Loss	D	C
D	$(0, 0)$	$(b, -c)$
C	$(-c, b)$	$(b - c, b - c)$

receiving a cooperative decision and c the cost of cooperating. In this setting, one may model a general interaction where two entities can cooperate or defect with each other. It is easy to note that in a *one shot game*, rational players will defect. From an evolutionary perspective, non rational defectors have a higher fitness than non rational cooperators and thus the former would be favored by natural selection (see Nowak 2006b for a complete discussion). It is not obvious then how cooperation may surge. The prisoner's dilemma has proved to be the best framework to study under what condition cooperation may be favored over defection (Rapoport and Chammah 1965; Axelrod 1984; Nowak 2006b).

In the following, we will represent the player's possible decisions (cooperate or defect), and the payoff they get (it could represent years or a certain good) by species. Let C, D be the species representing the cooperative and defector decisions respectively. The interaction of two players is equivalent to a chemical reaction where the reactants are two decision species. The reaction produces a payoff for each decision: we define the species G_C, G_D to represent the positive payoff for the cooperative and defector decisions respectively, analogously L_C, L_D represent the negative payoff of C and D. We consider two different types of species to model the positive and negative payoffs to avoid having negative concentrations on the species representing the payoff. Therefore, the set of species that models the prisoner's dilemma is $M = \{C, D, G_C, G_D, L_C, L_D\}$. We develop the details of the model in the next section.

3.4 The iterated prisoner's dilemma

We will build the set of reactions of the prisoner's dilemma based on the payoff table shown in Table 1. As stated previously, the reactants of each reaction correspond to a pair of decisions. Thus we have four possible reactions with the two decisions C, D: (C, C), (D, D), (C, D) and (D, C).

For simplicity, we will assume that the concentration of each type of decision C and D is fixed in the system. However, the reactions will generate species representing negative and positive profit. Therefore, in this case we have a set of decisions generating species representing positive or negative payoffs by their interactions. We are interested in determining which strategy performs best in time (i.e. gains a higher payoff) and thus would be favored, for example, by natural selection.

The interaction between two cooperative decisions is modeled by the reaction

$$R_1^* = 2C \rightarrow 2C + 2bG_C + 2cL_C, \tag{3}$$
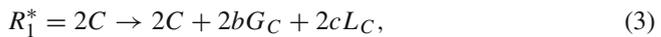
this means that when two cooperative decisions interact, $2b$ units of positive profit and $2c$ units of negative profit are generated. Analogously, the other three interactions are represented by the reactions

$$R_2' = D + C \rightarrow D + C + bG_D + cL_C, \tag{4}$$

$$R_2'' = C + D \rightarrow C + D + bG_D + cL_C \tag{5}$$

$$R_3^* = 2D \rightarrow 2D. \tag{6}$$

Note that in R_3^* the reactants are equal to the products. In chemistry this is called a *zero stoichiometry reaction*. These reactions may be excluded from the system, given that when they occur, the state of the system is not altered. Hence, the reaction network that models this system is $\langle \{C, D, G_C, G_D, L_C, L_D\}, \{R_1^*, R_2', R_2''\} \rangle$.

3.5 Dynamic analysis

To study the dynamics of the reaction network built above, we derive its chemical reaction system (2). Therefore, we have that the systems' dynamics is governed by the following system of differential equations:

$$\begin{aligned}
\dot{C} &= \dot{D} = \dot{L}_D = 0, \\
\dot{G}_C &= 2bk_1 C^2, \\
\dot{G}_D &= b(k_2' + k_2'')DC, \\
\dot{L}_C &= 2ck_1 C^2 + c(k_2' + k_2'')DC.
\end{aligned} \tag{7}$$

where k_1, k_2', k_2'' correspond to the reaction rates of R_1^* and R_2', R_2'' respectively. Recalling that the population of each decision does not vary in time, we will study which decision generates more profit depending on the values of b and c, the initial concentrations $C(0) = C_0$, $D(0) = D_0$, and the reaction rates k_1, k_2', k_2''.

We define the profit generated by one cooperative decision as $P_C = \frac{G_C - L_C}{C}$. Analogously, the profit generated by one defector decision is given by $P_D = \frac{G_D - L_D}{D}$. Solving the Eq. (7) we have that

$$\begin{aligned}
P_C(t) &= 2k_1(b - c)C_0 t - (k_2' + k_2'')c D_0 t \\
P_D(t) &= b(k_2' + k_2'')C_0 t
\end{aligned} \tag{8}$$

From (8) it is clear that in the absence of a mechanism that makes cooperative interactions more likely to occur, i.e. when $k_1 = k_2' = k_2''$, the Defector strategy is more profitable than the Cooperator strategy. Let us now assume that $k_2' = k_2'' = k_2$, supposing the symmetry of the reactions R_2' and R_2'' (a general assumption in EGT), and that k_1 and k_2 are not equal. We have that the *profit ratio* $P_r = \frac{P_C}{P_D}$ would favour the cooperation if and only if $P_r > 1$. This condition is equivalent to

$$r > \frac{c}{b}, \tag{9}$$

where $r = \frac{(k_1 - k_2)C_0}{k_1 C_0 + k_2 D_0}$. Therefore, $k_1 > k_2$ is a necessary (but not sufficient) condition for cooperative decisions being more profitable than defector decisions. The system's conditions that make $k_1 > k_2$ may be the consequence of kin or group selection in a standard EGT setting (Nowak 2006b). Remarkably, Eq. (9) recovers the same kind of relation for the evolution of cooperation, the comparison of a well defined quantity r and the benefit b and cost c of cooperating, that has been found using different approaches (Nowak et al. 2004). Moreover, Eq. (9) allows us to understand why the
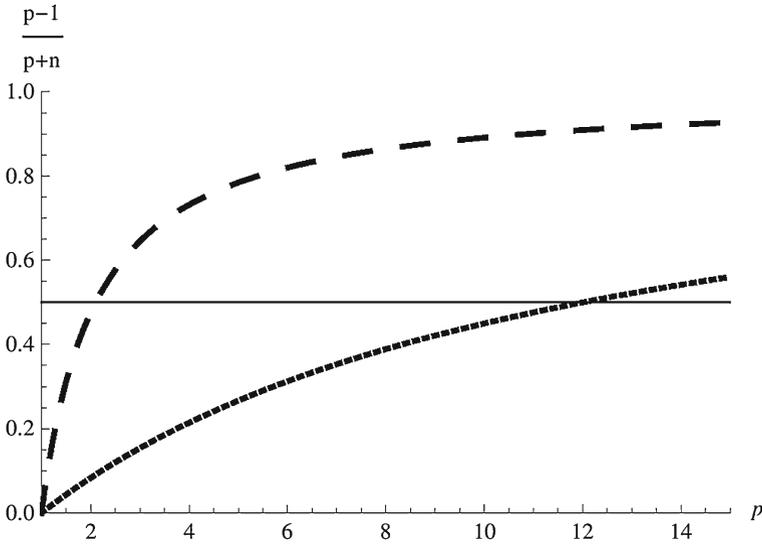
Fig. 2 Example of conditions that favor cooperation. The x-axis corresponds to the parameter $p = \frac{k_1}{k_2}$, the y-axis to the left-hand side of Eq. (10). *The large dashed curve* corresponds to the case $n = 0.1$, i.e. $D = 0.1C$, and *the short dashed curve* to $n = 10$, i.e. $0.1D = C$. The *solid horizontal line* at 0.5 is a reference for the case $\lambda = 0.5$. Hence, cooperation is more profitable for the values of p in which the curve is above 0.5. In these invasion settings, where the concentration of the invasor is 10 % of the invaded concentration, cooperative decisions resist the invasion of defectors when $p > 2$, and are able to invade a group of defectors when $p > 12$.

emergence of cooperation is non-trivial and requires specific mechanisms: First, let $k_1 = pk_2$ and $\lambda = \frac{c}{b}$, where $p > 1 > \lambda$ (and hence requiring $k_1 > k_2$ and $b > c$)

Secondly, by setting $D_0 = nC_0$ we have that $0 < n \ll 1$ corresponds to the situation when a small number of defector decisions *invades* a cooperators decision system, and $n \gg 1$ correspond to the invasion of cooperators to defectors. Under these conditions Eq. (9) becomes

$$\frac{p-1}{p+n} > \lambda. \tag{10}$$

When $n \ll 1$ (and thus when defector invades) we have that (10) can be approximated by $\frac{p-1}{p} > \lambda$. On the other hand, when $n \gg 1$ (and thus when cooperator invades) we have that (10) can be approximated by $\frac{p-1}{n} > \lambda$. Hence, the condition that allows cooperators to *resist* an invasion of defectors is less stringent than the condition in which cooperators invade defectors (see Fig. 2 for details). Thus, our approach can be used to study the changes in the profit relationships when the system is perturbed with a small concentration of new decisions. Hence, this analysis may serve as the basis of an analogous to the notion of evolutionarily stable strategies (ESS), an important concept in EGT (Weibull 1995).

4 Tit for Tat versus Defector

Tit for Tat strategy was introduced by Anatol Rapoport, during the first computational tournament of strategies playing the iterated prisoner's dilemma (Axelrod 1984). A Tit for Tat agent cooperates always in its first interaction with another agent and afterwards imitates its last opponent's move, i.e. it cooperates when the opponent cooperates and defects when the opponent defects. This strategy is quite easy to implement because the agent only needs to remember the last play of the other agents. In this sense, Tit for Tat does not require advanced cognitive capacities (even guppies use this strategy, Dugatkin 1992). It is a cooperative strategy based in reciprocity. Under certain conditions, Tit for Tat is an ESS, and it has been proven empirically, via simulations, that it is strong against many other strategies when they are all interacting together (Axelrod 1984). We will apply our model to the interaction between Tit for Tat agents and agents behaving according to a strategy that always defect (Defector).

4.1 Decision species and initial concentrations

Recall that we will label decision species according to both the strategy and decision they represent. Thus, we have three types of decisions:

1. T_C, Tit for Tat cooperates,
2. T_D, Tit for Tat defects,
3. D, Defector defects.

Analogous to the previous section, we define molecules G_T, L_T, G_D, L_D to model the positive and negative profits of Tit for Tat and Defector strategies respectively, and we assume that no strategy has accumulated previous profit. Thus we have that $G_T(0) = L_T(0) = G_D(0) = L_D(0) = 0$. Note that Defector agents never cooperate, thus according to the payoff matrix (see Table 1) we have that $L_D(t) \equiv 0$. Moreover, as Tit for Tat strategy always cooperate in the first interaction, we have that $T_D(0) = 0$.

In the standard agent-based setup of this system, two Tit for Tat agents will always cooperate, and they are able to recognize a Defector agent during the rest of the game once they interact with it (they will never cooperate more than once with the same Defector). Settings of partial or null agent-recognizability deviate from this situation. For example, two Tit for Tat agents can mutually defect or they can cooperate twice in a row with the same Defector agent if they cannot recognize the other agent, or at least the other agent's strategy.

Depending on the cognitive capacities of the agents (memory, agent or strategy recognizability, etc.), the reactions will represent the different kinds of interactions that are necessary to model the system. Hence, the elements participating in the reaction (the decision species) may refer to different conditions required for the occurrence of such interactions (e.g. if the agent cooperates or defects having recognized the other agent). Moreover, the reaction rates will be used to calibrate the relative frequencies at which the interactions occur.

In the remaining of this section we will show how to build a reaction network (and its associated chemical reaction system) that models the interaction of Tit for Tat and Defector agents in three different situations of memory and recognizability, and we

will illustrate how to study dynamical properties in each case applying methods from reaction networks.

4.2 Non-recognizability

### 4.2.1 Building the reaction network

In this case agents cannot recognize other agents. This is an oversimplified situation. Indeed, a Tit for Tat agent, as it is defined in the literature (Axelrod 1984), needs to *recognize* its opponent and *remember* the decision the opponent took in their last interaction to decide its action. Hence, it is not technically correct to call this behaviour as an equivalent to the Tit for Tat strategy, but this kind of strategy may be associated to the simple upstream indirect reciprocity (Nowak and Rosch 2006), and illustrates our model in a simple scenario.

Note that Tit for Tat agents always cooperate in their first interaction, thus if two Tit for Tat agents encounter the first time they interact, we can model their interaction by the interaction of two Tit for Tat cooperative decisions, generating the payoff of two cooperative decisions. Moreover, as Tit for Tat strategy imitates its last opponent behaviour for the next interaction we have that this interaction is modeled by the reaction

$$R_1 = 2T_C \rightarrow 2T_C + 2bG_T + 2cL_T.$$

In general, when two cooperative Tit for Tat agents interact, reaction R_1 describes their interaction. It is important to emphasize that we are not assuming any recognition process here. The only assumption we made is that cooperative decisions interact cooperatively. Hence their interaction is independent of the recognition situation.

Analogously, if a Tit for Tat agent interacts for first time with a Defector we have that their interaction corresponds to the reactions

$$R_2' = T_C + D \rightarrow T_D + D + bG_D + cL_T, \quad \text{and} \quad R_2'' = D + T_C \rightarrow D + T_D + bG_D + cL_T.$$
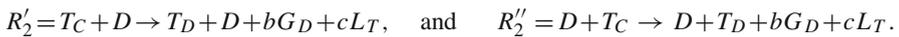
In R_2' and R_2'' the Tit for Tat agent cooperates, thus a cooperative decision T_C is part of the reactants, but it switches to T_D in the products because its opponent defects.

The next case is a Tit for Tat agent that has interacted only with cooperative Tit for Tat agents (T_C) and encounters a Tit for Tat agent that defects (T_D). This interaction is modeled by the reaction

$$R_3' = T_C + T_D \rightarrow T_D + T_C + cL_T + bG_T \quad \text{and} \quad R_3'' = T_D + T_C \rightarrow T_C + T_D + cL_T + bG_T.$$

Note that the decision reactants and products in R_3' and R_3'' are equal. This is because decision T_C switches to T_D and T_D switches to T_C (because Tit for Tat imitates the its last opponent's behaviour on the next interaction).

The three remaining cases entail zero stoichiometry reactions (as in reaction (6)): First, two Tit for Tat agents can interact in a non-cooperative way (the reactants are two T_D). Secondly, a defector Tit for Tat agent (decision T_D) may interact with a defector agent (D). Finally, two Defectors may interact.

Thus, the reaction network that models the interaction between a population of Tit for Tat agents and a population of Defectors in the non-recognizability case is given by

$$\langle \{D, G_D, T_C, T_D, G_T, L_T\}, \quad \{R_1, R_2', R_2'', R_3', R_3''\}\rangle.$$
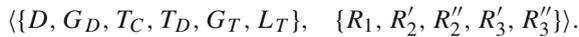
4.2.2 Dynamic analysis

In this case there is no assumption on the preference of certain interactions over others, and thus the three reactions rates will have the same constant value $k = k_1 = k_2 = k_3$. Applying mass action kinetics we obtain the following chemical reaction system:

$$\begin{aligned}
\dot{D} &= 0, \\
\dot{T_C} &= -2kT_C D, \\
\dot{T_D} &= 2kT_C D, \\
\dot{G_T} &= 2kb\left(T_C^2 + T_C T_D\right), \\
\dot{G_D} &= 2kbT_C D, \\
\dot{L_T} &= 2kc\left(T_C^2 + T_C T_D + T_C D\right).
\end{aligned} \tag{11}$$

The chemical reaction system (11) is simple enough to be solved analytically. However, before showing the solution of (11) we will show that a simple analysis of the topology of the network permits us to identify certain qualitative properties of the system. We say that a species m is consumed (produced) within $\mathscr{R}$ if and only if there exists a reaction $R \in \mathscr{R}$ such that the multiplicity of m as reactant in R is higher (smaller) than the multiplicity of m as product. Based on this, we say that a set of species $\mathbf{S} \subseteq \mathscr{M}$ is *semi-self-maintaining* if and only if every species $m \in \mathbf{S}$ that is *consumed* within $\mathscr{R}$ is also produced within $\mathscr{R}$. Semi-self-maintenance has been proven to be a necessary notion for the stability of a reaction network (Dittrich and Speroni di Fenizio 2008). In our example, note that the decision T_C is consumed by reaction R_3 but it is not produced by none of the reactions. Thus, T_C cannot belong to a semi-self-maintaining set. Hence, $T_C(t) \to 0$ as $t \to \infty$. As a consequence, we are able to deduce that in the long term only T_D and D will remain with concentration higher than zero (see Fig. 3). In addition, the interaction between these two species does not generate any payoff. Hence, we can conclude that the total accumulated profit of both strategies in this case is finite.[1]

Solving the chemical reaction system we have that the ratio P_r between the profit $\frac{P_T}{T_C(0)}$ of the Tit for Tat strategy and the profit $\frac{P_D}{D(0)}$ of the Defector strategy is given by:

---

[1] Technically the profit of both strategies is an asymptotic decreasing function, and the total profit calculated as its integral between zero an infinity could be unbounded. However, the solution of the profit for both strategies can be bounded by a decreasing exponential function, leading to a finite total profit.
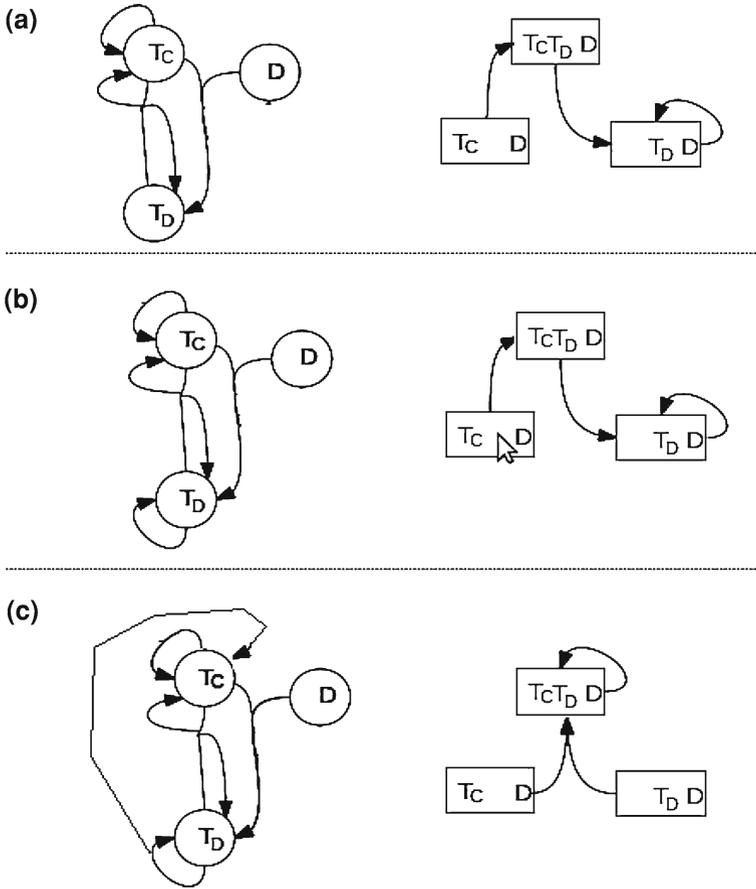
Fig. 3 Chemical reaction network (*left*) and the qualitative evolution of its set of decisions (*right*) for the systems modelled in Sects. 4.2, 4.3 and 4.4 (cases **a**, **b** and **c** respectively). For the sake of simplicity, the payoff species and some arrows are omitted in the *left diagrams*, and the trivially stable sets of decisions are omitted in the *right diagrams*. The only system that allows a steady state having the three decision species is **c**

$$P_r = 1 - \frac{c}{b}\left(\frac{D(0)}{T_C(0)} + 1\right)$$

Therefore, $P_r > 1$ has not a solution, and hence the Defector strategy collects more profit than Tit for Tat strategy.
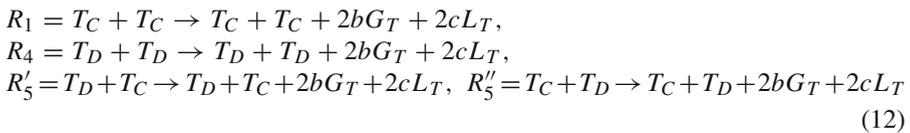
4.3 Simulating agent's memory

4.3.1 Building the reaction network

In the classic EGT situation where agents have memory, it is expected that each Tit for Tat agent cooperates only once with every Defector agent, and that Tit for Tat agents always cooperate between themselves. Moreover, a Tit for Tat agent
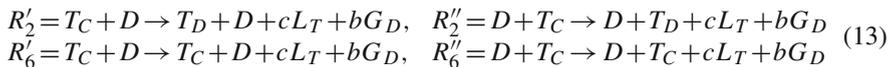
(with memory) will cooperate in its first interaction with every Defector agent and it will defect afterwards. This implies that, after a while, this Tit for Tat agent will have cooperated with all the Defectors in the system, and thus will not cooperate anymore with any of the Defectors. In order to model this system in our setup, we assume that from a long term analysis perspective, the fact that a Tit for tat agent that defects with the Defectors that has already encountered is equivalent to this agent cooperating with all the Defectors a number of times equals to the total number of Defectors in the system, and defecting with all of them afterwards. Hence, we assume in this section that T_D corresponds to the decision of a Tit for Tat agent that has interacted with all the Defectors (and thus will always defect with the Defectors), and T_C corresponds to the decision of a Tit for Tat agent that has cooperated with some but not all the Defectors, so it still needs to identify some of them (and thus can still cooperate with them).

The cooperation among Tit for Tat agents is modeled using the following reactions:

$$
\begin{aligned}
R_1 &= T_C + T_C \rightarrow T_C + T_C + 2bG_T + 2cL_T, \\
R_4 &= T_D + T_D \rightarrow T_D + T_D + 2bG_T + 2cL_T, \\
R'_5 &= T_D + T_C \rightarrow T_D + T_C + 2bG_T + 2cL_T, \quad R''_5 = T_C + T_D \rightarrow T_C + T_D + 2bG_T + 2cL_T
\end{aligned}
\tag{12}
$$

In these reactions, the two Tit for Tat decisions always lead to a cooperative payoff. Note that the reaction R_1 defined in this section is syntactically equivalent to the reaction R_1 of Sect. 4.2. However, the decision species of the Tit for Tat strategy (T_D and T_C) of Sect. 4.2 and of this section assume different features. Concretely, in Sect. 4.2 the decision species refer to pure decisions without any recognition or memory capacity, while in this the decisions represent a potentiality to cooperate or defect under the condition of having interacted with all the Defectors or not. In the remaining of the paper we will avoid renaming the reactions that are syntactically equivalent to reactions of previous sections.

We model the interaction between Tit for Tat and Defector agents with the following reactions:

$$
\begin{aligned}
R'_2 &= T_C + D \rightarrow T_D + D + cL_T + bG_D, \quad R''_2 = D + T_C \rightarrow D + T_D + cL_T + bG_D \\
R'_6 &= T_C + D \rightarrow T_C + D + cL_T + bG_D, \quad R''_6 = D + T_C \rightarrow D + T_C + cL_T + bG_D
\end{aligned}
\tag{13}
$$

Reactions R'_2, R''_2 in this case correspond to the situation where the Tit for Tat agent interacts *for the last time* with the Defectors in the system. Hence, after that last reaction, the cooperative Tit for Tat becomes a defector Tit for Tat that will not cooperate again with the Defectors in the system. Reactions R'_6 and R''_6 correspond to the situation when there are still some Defectors that the Tit for Tat agent has not encountered, and hence it cooperates, and stays as a cooperative decision for the next interaction. It is important to note that we can control the occurrence of these reactions with the rate constants because reactions R'_6 and R''_6 represent an interaction that does not depend on the number of Defectors in the system (and occurs at the same frequency than R_1, R_4 and R'_5), but R'_2 and R''_2 does depend on the number of Defectors, and hence occur at a different frequency with respect to the other reactions (much less frequent in this case).

To state this difference we need to set the reaction rates according to the time-scale that each reaction represents. First, note that R_1, R_4, R'_5, R''_5, R'_6 and R''_6 represent agent-agent interactions, and we will not assume preference of any of them over the others. Thus we set $k_1 = k_4 = k'_5 = k''_5 = k'_6 = k''_6 = k$. Secondly, reactions R'_2 and R''_2 represent the last interaction of the Tit for Tat agent with the defectors of the system. Hence, we set these reaction rates by $k'_2 = k''_2 = \frac{1}{\tau}k$, where τ is the expected time that a Tit for Tat agent takes to interact with all the Defectors in the system (the calculation of τ is included as an ESM appendix). Thus the reaction network that models this system is

$$\langle \{T_C, T_D, D, G_T, L_T, G_D\}, \quad \{R_1, R'_2, R''_2, R_4, R'_5, R''_5, R'_6, R''_6\}\rangle, \qquad (14)$$

and the chemical reaction system that corresponds to this system is

$$
\begin{aligned}
\dot{D} &= 0, \\
\dot{T}_C &= -\frac{k}{\tau}T_C D, \\
\dot{T}_D &= \frac{k}{\tau}T_C D, \\
\dot{G}_T &= 2kb\left(T_C^2 + 2T_C T_D + T_D^2\right) = 2kbT_C(0)^2, \\
\dot{G}_D &= 2kbT_C D\left(1 + \frac{1}{\tau}\right), \\
\dot{L}_T &= 2kc\left(T_C^2 + 2T_C T_D + T_D^2 + T_C D(1 + \frac{1}{\tau})\right) = 2kc\left(T_C(0)^2 + T_C D(1 + \frac{1}{\tau})\right).
\end{aligned}
\qquad (15)
$$

4.3.2 Dynamical analysis

In this case, the set $\{T_C, T_D, D\}$ is not a semi-self-maintaining set (see Sect. 4.2). Hence, as with the non-recognizability case, we have that $T_C(t) \to 0$ as $t \to \infty$ (see Fig. 3). However, in this case T_D decisions may cooperate with other Tit for Tat decisions. Thus in the asymptotic regime, P_D is constant and P_T is an increasing function. Therefore, Tit for Tat strategy collects more profit than the Defector strategy in the long term.

This case is also simple enough to be solved analytically. Solving (15) we have that the profit ratio is given by

$$P_r(t) = \alpha\frac{t}{1 - e^{-\gamma t}} - \beta \qquad (16)$$

where $\alpha = D(0)\frac{(b-c)k}{b(1+\tau)}$, $\beta = \frac{cD(0)}{bT_C(0)}$ and $\gamma = \frac{D(0)k}{\tau}$. Note that $\frac{t}{1-e^{-\gamma t}}$ is an increasing function when $\gamma > 0$. Hence, two cases are possible for this system: The first case is that the Tit for tat strategy is more profitable for all $t \geq 0$. The second case is that Defector strategy has a *bonanza* period (while Tit for tat agents learn who are the Defectors in the system), and after the bonanza period Tit for tat becomes more profitable. Applying L'Hopital rule to Eq. (16) we have that $P_r(0) = \frac{\alpha}{\gamma} - \beta$. Hence, $P_r(0) > 1$ is equivalent to

$$\frac{(b-c)}{b}\frac{\tau}{(1+\tau)} > 1 + \frac{cD(0)}{bT_C(0)} \qquad (17)$$
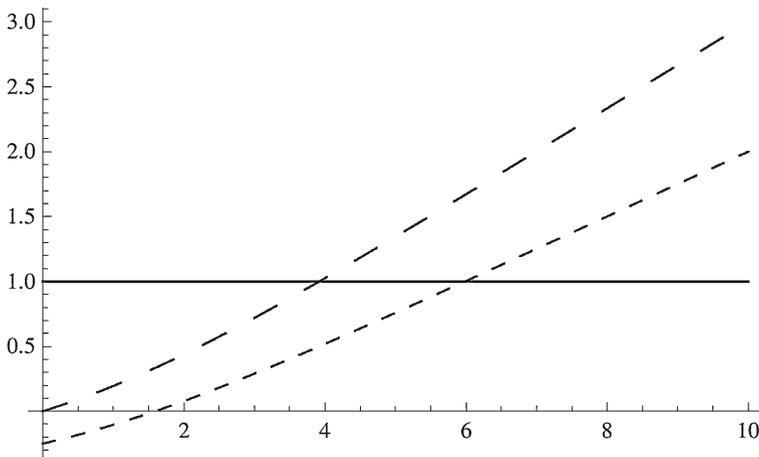
Fig. 4 Payoff ratio showing the bonanza period of Defectors. The x-axis corresponds to the time t, the y-axis to payoff ratio $P_r(t)$. We set $c = D(0) = T_C(0) = \tau = k = 1$. The *short dashing curve* corresponds to $P_r(t)$ when $b = 2$ and the *large dashing curve* to $b = 3$. The *solid horizontal line* equals to 1 is a reference to identify the bonanza period (for the values of t where $P_r(t) < 1$ Defectors collect more profit than Tit for tat agents). The bonanza period is approximately equal to 6 for $b = 2$ and to 4 for $b = 3$

From Eq. (17) we see that Defector strategy is more profitable than Tit for tat at $t = 0$ for any reasonable (positive) setting of the parameters. We can then estimate the bonanza period for Defectors from Eq. (16). We will not perform such analysis here, but point out that it can be developed by (algebraically or numerically) approximating the value of t that satisfies the transcendent equation (16) (see Fig. 4 for examples).
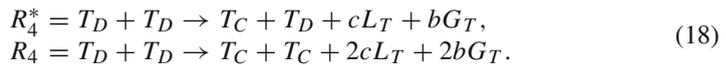
In this section we make use of a global property to define the reaction network. This global property refers to the long term capacity of Tit for tat agents to recognize the population of Defectors, i.e. Tit for tat agents recognize *the other's strategy* in the long term. Note that this global property is dependent on the assumption that the populations of agents in the system do not change in time (no mutation, no replication). While this assumption, and the use of a global property to define the reaction network, may be suspicious from an EGT perspective, the intention of this section is to illustrate that it is possible to consider different kinds of reactions to model an agent scenario. In particular, reactions may refer to different time-scales, which is reflected in the frequency at which they occur, and the frequencies of each reaction can be adjusted from statistical information of the system.

4.4 Strategy recognition

### 4.4.1 Building the reaction network

In this case, we will assume that agents do not have memory. However, agents can recognize other agent's strategy, i.e. agents may (or may not) recognize the strategy population to which the other agent belongs when they interact. The decision species in this case represent the decision when there is no recognition of the other's strategy,

i.e. a predisposition to cooperate or defect. However, if recognition occurs it may be the case that this predisposition is altered by the recognition, and hence a cooperative decision interacts as a defector decision and vice versa. Hence, the effect of the recognition may change the predisposition to cooperate or defect. This will be represented in the reactions we define. First, note that Defector agents will always defect, thus the recognition situations that we consider will not affect them in this analysis. Tit for Tat agents however, might change their decision after recognizing its opponent strategy. The products of each interaction will vary according to the decisions that are interacting, and whether or not the recognition occurs. To clarify this point let us consider the following case: two Tit for Tat defector decisions are interacting (the reactants of such reaction are $T_D + T_D$). When both agents do not recognize each other the products of this reaction correspond to the payoff of two defector decisions interacting (zero stoichiometry reaction). If only one agent recognizes that the other is a Tit for Tat, the products of the reaction will correspond to a defector–cooperator payoff. Finally if both agents recognize each other, the products will correspond to a cooperator–cooperator interaction. Thus the following reactions must be included in the system

$$
\begin{aligned}
R_4^* &= T_D + T_D \rightarrow T_C + T_D + cL_T + bG_T, \\
R_4 &= T_D + T_D \rightarrow T_C + T_C + 2cL_T + 2bG_T.
\end{aligned}
\tag{18}
$$

Note that in both reactions, according to the number of agents that recognizes other's strategy (one in R_4^* and two in R_4), one or two defector Tit for Tat decisions switch to a cooperative one because Tit for Tat strategy imitates the last opponent's play.

Furthermore, it can be the case that a defector Tit for Tat recognizes a cooperative Tit for Tat agent. This is modeled by

$$
R_5^* = T_C + T_D \rightarrow T_C + T_C + 2cL_T + 2bG_T, \quad R_5^{**} = T_D + T_C \rightarrow T_C + T_C + 2cL_T + 2bG_T
\tag{19}
$$
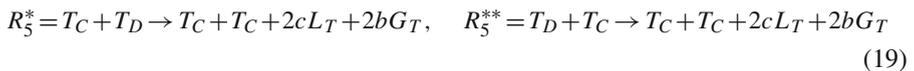
Finally we need to consider the interaction between a cooperative Tit for Tat and a Defector. If the Tit for Tat recognizes the Defector it will defect, so we have:

$$
R_2^* = T_C + D \rightarrow T_D + D, \; R_2^{**} = D + T_C \rightarrow D + T_D
\tag{20}
$$
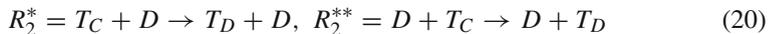
If the cooperative Tit for Tat agent does not recognize the Defector we have

$$
R_2' = T_C + D \rightarrow T_D + D + cL_T + bG_D, \quad R_2'' = D + T_C \rightarrow D + T_D + cL_T + bG_D
\tag{21}
$$
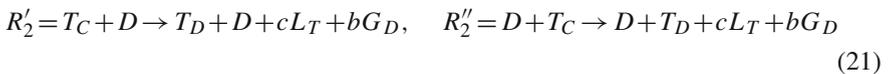
It is important to note that we do not consider the interaction where a defector Tit for Tat cooperates with a Defector agent because such case cannot be explained through the success or fail of a recognition process (but through a *confusion* process that we are not considering). If the defector Tit for Tat agent recognizes the defector, it will defect, and if it does not recognize it, it will defect as well. Thus, the reaction network

which models this situation is

$$\langle\{T_C, T_D, D, G_E, G_T, L_T\}, \quad \{R_1, R_2', R_2'', R_2^*, R_2^{**}, R_4, R_4^*, R_5^*, R_5^{**}\}\rangle. \quad (22)$$

4.4.2 Dynamical analysis

Let p be the probability that an agent recognizes other agent's strategy. We will modify the reaction rates according to the recognizability situation that occurs on each reaction, by multiplying the reaction rate k to the probability of occurrence of the reaction event. Note that in reaction R_1 both agents would cooperate, even if they do not recognize each other, thus $k_1 = k$. In R_4^* only one of the defector Tit for Tat agents recognizes the other thus $k_4^* = p(1-p)k$. In R_4, the two defector Tit for Tat recognize each other thus $k_4 = p^2 k$. In R_5^*, R_5^{**}, the defector Tit for tat agent recognizes the cooperator Tit for tat thus $k_5^* = k_5^{**} = pk$. Analogously $k_2' = k_2'' = (1-p)k$ and $k_2^* = k_2^{**} = pk$.

In this case the chemical reaction system is:

$$
\begin{aligned}
\dot{D} &= 0, \\
\dot{T}_C &= k\left(p(1+p)T_D^2 + 2pT_C T_D - 2T_C D\right), \\
\dot{T}_D &= -\dot{T}_C, \\
\dot{G}_T &= kb\left(2T_C^2 + p(1+p)T_D^2 + 2pT_D T_C + 2T_C T_D\right), \\
\dot{G}_D &= 2kb(1-p)T_C D, \\
\dot{L}_T &= \frac{c}{b}\dot{G}_T + 2kc(1-p)T_C D.
\end{aligned}
\quad (23)
$$

Note that by setting $p = 0$ we recover the case of non-recognizability shown in Sect. 4.2. By varying p from zero to one we can model different situations of recognition, and when $p = 1$, we have $G_D \equiv 0$. This means that Tit for Tat agents would never cooperate to Defectors (as expected because there is full recognition of other agent's strategy).

Considering that $T_C(t) + T_D(t) = T_C(0)$, we can obtain for both T_C and T_D a first order quadratic equation from (23). When the parameters k, b, c, p are set as specific values it is possible to find an analytic expression for the solution of the chemical reaction system (23). However, the solution of (23) considering general values for the parameters k, b, c, p is extremely complicated because different values on the parameters lead to qualitatively different solutions. Therefore, instead of determining the regions that lead to different solutions, then solving the equations in each parameter region, and finally looking at the relative profit relation as we did in the other cases, we will focus our analysis on conditions under which is possible to reach a steady state where all the decision species have concentration greater than zero, and study the payoff of the strategies under the steady condition. Differently to the two cases revised above, in this case T_C is produced by reactions R_4, R_4^*, R_5^* and R_5^{**}. Thus $\{T_C, T_D, D\}$ is a semi-self-maintaining set. For semi-self-maintaining sets, it is possible to verify whether they can reach a steady state by studying the stoichiometric matrix of the reaction network. Several methods such as flux balance analysis, elementary modes, among others have been developed for this purpose (Schilling and Palsson 1998; Schuster et al. 1999; Österlund et al. 2011). We will approach the steady state analysis of this system from a novel approach called Chemical Organization Theory

([Dittrich and Speroni di Fenizio 2008](#)) because it provides tools to connect a reaction network with the fixed points (and other higher-dimensional attractors) of its chemical reaction system ([Dittrich and Speroni di Fenizio 2008](#); [Peter and Dittrich 2011](#)).

Definition 2 Let $O \subseteq \mathcal{M}$ and $\mathcal{R}_O$ the set of reactions whose reactants are in O, i.e. the set of reactions that are *firable* within the species available in O. We say O is a *chemical organization* if and only if it holds the following properties:

– *Closure*: Species produced within $\mathcal{R}_O$ are contained in O.
– *Self-maintenance*: There exist a flux vector $\mathbf{v} \in \mathbb{R}_{>0}^{|\mathcal{R}_O|}$ such that $\mathbf{S}_O \mathbf{v} \geq \mathbf{0}$, where $\mathbb{R}_{>0}$ is the set of strictly positive real numbers, $|\mathcal{R}_O|$ is the cardinality of $\mathcal{R}_O$, and $\mathbf{S}_O$ is the stoichiometric matrix associated to the reaction network $\langle O, \mathcal{R}_O \rangle$.

Organizations are sets of species which do not produce novel species within their reactions (closed), and are able to sustain or increase the concentration of its species by a reaction process where each reaction occurs at a positive rate (self-maintaining). Although the notions of organization and elementary mode ([Schuster et al. 1999](#)) are similar in that both comprise a condition that relates the flux vector to the steady states of the chemical reaction system ([Kaleta et al. 2006](#)), organization entails a concept of *independent sustainability* within a reaction network because of the closure condition. Indeed, the set of species with concentration higher than zero of any fixed point of a chemical reaction system is an organization of the underlying reaction network (see theorem 42 in ([Dittrich and Speroni di Fenizio 2008](#))). We will analyze whether or not Tit for Tat and Defector strategies can be part of the same steady state, and in this sense if they can *coexist*. To do so we verify the conditions under which both strategies together form an organization. Note that in the cases studied in Sects. [4.2](#) and [4.3](#) we found that $T_C(t) \to 0$ as $t \to \infty$, and hence both strategies were not able to coexist in the long term.

We first need to point out that from $X = \{T_C, T_D, D\}$ we can obtain the rest of the (payoff) species by applying the reactions, i.e. the *closure* of X is $\mathcal{M}$. Moreover, the species representing payoff are produced but not consumed by the set of reactions. Hence, if X is a self-maintaining set, then $\mathcal{M}$ is a self-maintaining set, and thus an organization. Therefore, $\mathcal{M}$ would be a candidate to represent a fixed point of the chemical reaction system.

To prove the self-maintenance of X we need to verify if there exists a flux vector that produces species in an equal or higher quantity than its consumption. As the reactions which vary the number of species in X are R_4, R_4^*, R_5^*, R_5^{**}, R_2^*, R_2^{**}, R_2', R_2'', we can verify the self-maintaining condition reducing the stoichiometric matrix to only these reactions. The system of equations is given by:

$$\begin{pmatrix} 2 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ -2 & -1 & -1 & -1 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{v}_{sm} = \mathbf{f} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (24)$$

Solving this equation we have that there exist a solution only for the case $\mathbf{f} = \mathbf{0}$, and such case is fulfilled by the flux vector $\mathbf{v}_{sm} = (\mathbf{v}_4, \mathbf{v}_4^*, \mathbf{v}_5^*, \mathbf{v}_5^{**}, \mathbf{v}_2^*, \mathbf{v}_2^{**}, \mathbf{v}_2', \mathbf{v}_2'')$ satisfying the following relation

$$2\mathbf{v}_4 + \mathbf{v}_4^* + \mathbf{v}_5^* + \mathbf{v}_5^{**} = \mathbf{v}_2^* + \mathbf{v}_2^{**} + \mathbf{v}_2' + \mathbf{v}_2'' \tag{25}$$

We have found a relation among the fluxes that correspond to a steady production regime (fixed point) for all the decision species in the system. Note that we can rewrite Eq. (25) replacing $\mathbf{v}_i$ by the expression defined in (1), which assumes that our system is dynamically governed by the law of mass action kinetics, i.e. we can link the stoichiometric condition given by (25) with the corresponding chemical reaction system (Peter et al. 2010). In this case, the fluxes of $\mathbf{v}_{sm}$ can be obtained applying (1) to the reactions R_4, R_4^*, R_5^*, R_5^{**}, R_2^*, R_2^{**}, R_2' and R_2'' respectively. If we do such replacement, Eq. (25) becomes in the following equation

$$kp(1 + p)T_D^2 + 2kpT_C T_D = 2kT_C D. \tag{26}$$

Thus, recalling that $D(t) = D(0)$ and that $T_C(t) + T_D(t) = T_C(0)$ for all t, we have that a stable concentration for species $T_D = T_D^*$ is given by

$$T_D^* = \frac{pT_C(0) + D(0)}{p(1 - p)} \mp \frac{\sqrt{pT_C(0) + D(0)^2 + 2p(1 - p)T_C(0)D(0)}}{p(1 - p)}, \tag{27}$$

Hence, we have obtained a condition for steady concentrations of each decision species. Note that the steady concentration does not depend on k, and that by analyzing the squared root of (27) under the condition $T_D^* > 0$ we can determine whether there is zero, one or two steady concentrations.

We now assume the steady state condition (26), and study which strategy is more profitable. To do so, we need to inspect the production of the molecules G_T, L_T, G_D for the flux vectors that verify the self-maintenance of $\mathscr{M}$. As we already know that the production of species T_C, T_D and D is zero, we are only concerned with the production of the species G_T, L_T, G_D, i.e we construct the stoichiometric matrix corresponding to reactions $\{R_1, R_2', R_2'', R_4, R_4^*, R_5^*, R_5^{**}\}$ and compute the production of the payoff species under the steady decision production.

$$\begin{pmatrix} 2b & 0 & 0 & 2b & b & 2b & 2b \\ 2c & c & c & 2c & c & 2c & 2c \\ 0 & b & b & 0 & 0 & 0 & 0 \end{pmatrix} \mathbf{v}_P = \begin{pmatrix} \mathbf{f}_4 \\ \mathbf{f}_5 \\ \mathbf{f}_6 \end{pmatrix} \tag{28}$$

where $\mathbf{v}_P = (\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_2'', \mathbf{v}_4, \mathbf{v}_4^*, \mathbf{v}_5^*, \mathbf{v}_5^{**})$. Note that we have that $P_T = \frac{\mathbf{f}_4 - \mathbf{f}_5}{T_C(0)}$ and $P_D = \frac{\mathbf{f}_6}{D(0)}$. Hence P_r is defined by

$$P_r = \left(\frac{D(0)}{T_C(0)} \right) \left(\frac{b - c}{c} \right) \left(\frac{2\mathbf{v}_1 + 2\mathbf{v}_4 + \mathbf{v}_4^* + 2\mathbf{v}_5^* + \mathbf{v}_5^{**}}{\mathbf{v}_2' + \mathbf{v}_2''} \right). \tag{29}$$

From here, applying the condition (25) and replacing the elements of $\mathbf{v}_P$ by their kinetic expressions defined in (1), we obtain that $P_r \geq 1$ if and only if

$$1 + 2k \left(T_C^* + pT_D^* + pD(0) \right) > \frac{2k(1 - p)c}{b - c} T_C(0) \tag{30}$$

where T_D^* is obtained from (27) and $T_C^* = T_C(0) - T_D^*$. Thus we have found an algebraic expression that indicates the conditions under which strategy is more profitable in the steady decision production case.

5 Discussion

After introducing the reaction network framework, we developed the socio-chemical metaphor in a classical EGT scenario. We discussed how to model evolutionary games using reaction networks. Remarkably, the units of modelling are the agents' decisions and the payoffs obtained by the interaction of such decisions, instead of the agents themselves. This framework is fundamentally different to standard agent-based models used in EGT because we neglect the role of agents and study the system at the *decision level*. We model the evolution of the concentration of the species that represent decisions and payoffs of the strategies in the system. The consequences of memory, strategy recognition and other properties of agents are replaced by specific rules of interaction. This is based on the idea that in social systems visible communications are the key properties when population dynamics and long term changes are studied (Luhmann 1986). The concentration of a decision species is interpreted as its *visibility*, and the concentration of the payoff species is interpreted as the strategy performance within the system (see Fig. 1). As a first illustrative case, we studied the repeated prisoner's dilemma.

We were able to reproduce the results known from the literature concerning this problem, and we also provided an *invasion* analysis that explains why the emergence of cooperation is paradoxical from a payoff-based fitness standpoint (see Fig. 2). We then studied a system consisting of a group of Tit for Tat agents interacting with a group of Defectors in three situations where agents have different capacities to recognize and remember other agents. In each case, the reaction network is built from the information given by the payoff matrix together with the concrete assumptions on the agents' recognizability and memory capabilities. These capabilities play an important role in determining the meaning of the decision species, the possible reactions, and the kinetic rates of these reactions. For example, in Sect. 4.2 the species T_D represents a defector decision of a Tit for Tat agent and R_2 represents an (agent-agent) interaction between a Tit for Tat agent and a Defector, where both agents have neither memory nor recognizability capacity. However, in Sect. 4.3 the species T_D represents the decision of a Tit for Tat agent that can remember a group of Defectors, and R_2 represents that a Tit for Tat agent has interacted with all the Defectors in the system. In particular, we developed a reaction network model for the interaction of Tit for Tat and Defector agents when there is non recognizability (Sect. 4.2), when agents have memory capacity (Sect. 4.3), and when agents have a probability to recognize other agents' strategy (Sect. 4.4).

Moreover, we built a chemical reaction system (ODE system) from each reaction network to study its dynamics (see Eq. 2). We were able to solve the chemical reaction system in the non recognizability and memory situations. Hence, we found analytical conditions for the performance of each strategy in such cases. From these performance conditions it is possible to infer if a strategy may invade the other by properly setting the

Table 2 Table showing the meaning attributed to the decision species and the usage of the kinetic constant to model the frequency of the reactions on the three cases studied in Sect. 4

Situation	Decision species' meaning in the reactions	Use of kinetic constant
Non-recognition	Pure decision	None
Memory	Decision having (or not) interacted with all the Defectors	Expected time to identify all the Defectors
Strategy recognition	Decision having (or not) recognized other's agent strategy	Recognition frequency

initial concentration values, and observe the influence of group selection by setting the reaction rates. Moreover, a simple topological analysis of the network (see the semi-self-maintenance condition defined in Sect. 4.2) makes possible to study asymptotic properties of the system. In the strategy recognition case (Sect. 4.4), the analytical solution of its chemical reaction system is extremely complicated. Instead of solving the chemical reaction system or applying numerical methods, we applied a novel method called Chemical Organization Theory (Dittrich and Speroni di Fenizio 2008) to study its dynamical properties. We found a steady state condition for the Tit for Tat and Defector groups combining stoichiometric and kinetic information of the reaction network (see Eq. 27), and we obtained an analytical expression for the performance ratio between the two strategies under the steady state condition (see Eq. 30).

We suggest that this approach may be promising for the modelling of complex EGT scenarios. Here we focus on the capacity of this framework to show the possible kinds of interactions that may exist depending on what strategies are present and the cognitive capacities of the agents (see Table 2). However, a more elaborated dynamical analysis and a methodology to generate models in complex situations are required. Concerning the dynamical analysis, we propose three key elements for such development. First, we expect to extend the model to situations where strategies change in relative proportion by including reactions that generate and destroy decisions. This would provide a dynamical (evolutive) process for the decisions and strategies, as it is usually modeled by the replicator equation in EGT. Secondly, it is necessary to investigate if the mass-action kinetics is the right dynamical law to build the differential equations that govern the evolution of the system. In the examples we have discussed it seems to be a reasonable assumption. However, network and other effects often occur in realistic human interactions (Nowak 2006a; Nowak and May 1992; Lieberman et al. 2005). These effects change the space homogeneity assumption underlying the mass-action law. Once these two issues are addressed, it is necessary to clarify which methods are applicable to the equations we would obtain in order to study their stability. We claim that COT is an interesting tool because provides a basic landscape of the steady states in the system. However, it does not provide concrete results concerning stability. Therefore, dynamical systems' methods need to be investigated in conjunction with recent extensions of COT (Strogatz 2000; Peter et al. 2010). Moreover, we propose to apply other methods that have been developed to study *mechanisms for stability* of reaction networks such as elementary modes (Schuster et al. 1999) and Petri net based analysis (Heiner et al. 2008), to analyze the *invasion* of strategies on these complex scenarios, and study whether or not different strategies are ESS, a crucial concept in EGT (Weibull 1995).

Concerning the modeling of more complex situations, it is necessary to develop a methodology that specifies the kinds of interactions and the kinds of meaning associated with the decision species given the specification of the agents' cognitive abilities and the strategies that are present in the system. We have illustrated how to model memory and strategy recognizability in a simple agent setting. However, in a more complex situation these abilities may not be modeled in the same way as we did here. Note that in this work the decision species refer to a strategy, and thus a strategy is defined as a set of decisions. In a more realistic setting, strategies are not predefined but emerge from the interactions. Our approach allows the specification of a set of interactions without attributing a specific strategy to each decision. Hence, we can use network analysis techniques such as elementary modes or COT to identify the decision pathways (analogous to metabolic pathways) that tend to self reproduce. This would indicate the emergence of higher order structures of decisions that can represent strategies or other behavioural tendencies.

We believe that the reaction network framework applied to EGT might be an important step towards a formal theory which explains the formation, interaction and evolution of societies.

References

Axelrod R (1984) The evolution of cooperation. Basic Books, Inc., New York

Axelrod R (1997) The complexity of cooperation. Agent-based models of competition and collaboration. Princeton University Press, Princeton

Binmore K (2007) Game theory. A very short introduction. Oxford University Press, Oxford

Bowles S, Gintis H (2011) A cooperative species: human reciprocity and its evolution. Princeton University Press, Princeton

Dittrich P (2009) Artificial chemistry. In: Meyers B et al (eds) Encyclopedia of complexity and system science, pp 113–136

Dittrich P, Speroni di Fenizio P (2008) Chemical organization theory. Bull Math Biol 69:1199–1231

Dittrich P, Winter L (2007) Chemical organizations in a toy model of the political system. Adv Complex Syst 1(4):609–627

Dugatkin L (1992) Tendency to inspect predators predicts mortality risk in the guppy (*Poecilia reticulata*). Behav Ecol 3(2):124–127

Feinberg M, Horn F (1974) Dynamics of open chemical systems and the algebraic structure of the underlying reaction network. Chem Eng Sci 29(3):775–787

Fontana W, Buss L (1994) 'The arrival of the fittest': toward a theory of biological organization. Bull Math Biol 56(1):1–64

Gardner A, West S (2009) Greenbeards. Evolution 64:25–38

Gillespie D (2007) Stochastic simulation of chemical kinetics. Annu Rev Phys Chem 58(3):35–55

Gintis H (2008) The bounds of reason: game theory and the unification of behavioral sciences. Princeton University Press, Princeton

Hamilton W (1964) The genetical evolution of social behaviour, I, II. J Theor Biol 7:1–16

Heiner M, Gilbert D, Donaldson R (2008) Petri nets for systems and synthetic biology. In: Proceedings of the formal methods for the design of computer, communication, and software systems. 8th international conference on formal methods for computational systems biology. Springer, Berlin, pp 215–264

Heinrich R, Rapoport T (1974) A linear steady-state treatment of enzymatic chains. Eur J Biochem 42:89–95

Hofbauer J, Sigmund K (1998) Evolutionary games and population dynamics. Cambridge University Press, Cambridge

Kacser H, Burns J (1973) The control of flux. Symp Soc Exp Biol 27:65–104

Kaleta C, Centler F, Dittrich P (2006) Analyzing molecular reaction networks: from pathways to chemical organizations. In: Proceedings of the 7th conference on protein expression in animal cells (PEACe), vol 34, pp 117–124

Lieberman E, Hauert C, Nowak M (2005) Evolutionary dynamics on graphs. Nature 433:312–316

Luhmann N (1986) The autopoiesis of social systems. In: Sociocybernetic paradoxes. Sage, London

Luhmann N (1995) Social systems. Stanford University Press, Stanford

Maturana H, Varela F (1974) Autopoiesis: the organization of living systems, its characterization and a model. Biosystems 5:187–1196

Maynard Smith J (1982) Evolution and the theory of games. Cambridge University Press, Cambridge

Nowak M (2006a) Evolutionary dynamics. Harvard University Press, Cambridge

Nowak M (2006b) Five rules for the evolution of cooperation. Science 8:1560–1563

Nowak M, May R (1992) Evolutionary games and spatial chaos. Nature 358:826–829

Nowak M, Rosch S (2006) Upstream reciprocity and the evolution of gratitude. Proc R Soc Lond B: Biol Sci 274:8605–609

Nowak M, Sigmund K (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. Nature 366:56–58

Nowak MA, Sasaki A, Taylor C, Fudenberg D (2004) Emergence of cooperation and evolutionary stability in finite populations. Nature 428(6983):646–650. doi:10.1038/nature02414

Österlund T, Nookaew I, Nielsen J (2011) Fifteen years of large scale metabolic modeling of yeast: developments and impacts. Biotech Adv 3:1–10

Peter S, Dittrich P (2011) On the relation between organizations and limit sets in chemical reaction systems. Adv Complex Syst 14(1):77–96

Peter S, Veloz T, Dittrich P (2010) Feasibility of organizations: a refinement of chemical organization theory. In: Proceedings of the eleventh international conference on membrane computing, pp 369–382

Pfeiffer T, Rutte C, Killingback T, Taborsky M, Bonhoeffer S (2005) Evolution of cooperation by generalized reciprocity. Proc R Soc B 272:1115–1120

Rapoport A, Chammah A (1965) Prisoner's dilemma. University of Michigan Press, Ann Arbor

Rasmusen E (2007) Games and information. An introduction to game theory. Basil Blackwell, Oxford

Razeto-Barry P (2012) What is autopoiesis? 40 years later. Orig Life Evol Biosph (in press)

Razeto-Barry P, Cienfuegos J (2011) La paradoja de la probabilidad de lo improbable y el pensamiento evolutivo de Niklas Luhmann. Convergencia 57:13–38

Reddy V, Mavrovouniotis M, Liebman M (1993) Petri net representations in metabolic pathways. In: Proceedings of the 1st international conference on intelligent systems for molecular biology. AAAI Press, pp 328–336

Richerson P, Boyd R (2005) Not by genes alone: how culture transformed human evolution. The University of Chicago Press, Chicago

Schilling C, Palsson B (1998) The underlying pathway structure of biochemical reaction networks. Proc Natl Acad Sci 95:4193–4198

Schuster S, Dandekar T, Fell D (1999a) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. Trends Biotechnol 17:53–60

Strogatz S (2000) Nonlinear dynamics and chaos. Westview Press, Cambridge

Taylor A (1995) Mathematics and politics: strategy, voting, power and proof. Springer, Berlin

Taylor P, Jonker L (1978) Evolutionary stable strategies and game dynamics. Math Biosci 40:145–156

Vincent T, Brown J (2007) Evolutionary game theory, natural selection and Darwinian dynamics. Cambridge University Press, Cambridge

Weibull J (1995) Evolutionary game theory. MIT Press, Cambridge

West S, Mouden CE, Gardner A (2011) Sixteen common misconceptions about the evolution of cooperation in humans. Evol Hum Behav 32:231–262

Wiechert W (2002) Modeling and simulation: tools for metabolic engineering. J Biotechnol 94(1):37–63