Strategies, Behaviors, and Discounting in Radio Resource Sharing Games

Lars Berlemann
ComNets Aachen University
52074 Aachen, Germany
ber@comnets.rwth-aachen.de

Stefan Mangold
Philips Research Laboratories
Briarcliff Manor NY 10510, USA
stefan.mangold@philips.com

Bernhard Walke
ComNets Aachen University
52074 Aachen, Germany
walke@comnets.rwth-aachen.de

***Abstract*—In hot spot scenarios, Wireless Local Area Networks (WLANs) often have to share a common radio channel. These WLANs are here modeled as competing players that support Quality of Service in the presence of other players. The competition is modeled with a stage-based game structure. In this paper, we model existing games by means of repeated stages. This is referred to as Multi Stage Game. During the course of a Multi Stage Game, players adapt behaviors, i.e. modify their protocol parameters, by taking into account the history of past achieved utilities (the payoffs per stage). We show that players attempting to maximize their payoff are able to improve their payoffs through dynamic strategies. Strategies define what behavior a player selects under consideration of a potential interaction. Further on, a discounting-based decision process for determining what behavior to select is introduced. Simulation results indicate that setting the player's discounting factors based on the quality of service requirements leads to predictable outcomes in many competition scenarios.**

***Keywords*—Behaviors in Radio Resource Competition; Discounting in Multi Stage Games; Quality of Service as Utility; Strategies for Radio Resource Sharing**

I. INTRODUCTION

As *Wireless Local Area Networks* (WLANs) operate in unlicensed frequency bands, they often have to deal with interference from other WLANs. We discuss in this paper multiple WLANs, each represented as a player, that operates at the same frequency channel (therefore referred to as overlapping WLANs). The known coexistence problem arises if overlapping WLANs attempt to support *Quality of Service* (QoS) under the assumption that the overall utilization of the resource is at its limit [4].

Players' QoS requirements define an individual *utility* function. This utility function is used to map QoS requirements into behaviors, and achieved QoS into levels of satisfaction (i.e. outcome, payoff). The achieved QoS can differ from the required QoS, because the players' resource allocations influence each other. The mutual influences and the corresponding utility degradation leads to the *payoff* of a player. Players may benefit (i.e. players may achieve a higher payoff) from a dynamic interaction, by adapting behaviors to the environment and the behaviors of other players [3].

Players evaluate their individual expectations of future outcomes of a *Multi Stage Game* (MSG) based on so-called discounted payoffs from the *Single Stage Games* (SSG).

The paper is outlined as follows. Section II summarizes the problem description and the SSG model of [1]-[4]. To facilitate the understanding of the terms used in this paper we illustrate the concepts with the help of the *Universal Modeling Language* (UML), see Fig. 1. In Section III, we describe the principles for a dynamic interaction based on repeated SSGs which form an MSG. The possibility to establish cooperation through punishment within MSGs is discussed in Section IV, together with a presentation of simulation results. The paper ends with a summary and outlook on possible applications of the concepts introduced in this paper.

II. SINGLE STAGE GAME

An SSG is based on the resource allocation timing of the medium access control in a WLAN, and consists of three phases. The three phases define thus a stage. (1) The players decide about their action, which means they *demand* resource allocation times and durations (this is assumed an instant of time, and occurs at the beginning of a SSG, i.e. at the beginning of a stage). (2) The allocation process during the SSG may result in resource allocation delays and even collisions of allocation attempts. Hence, the *observed* allocation points may differ from the demanded allocations (this is assumed to require some time, for example *100ms* or equivalently the duration between two broadcasted management frames, i.e. two beacons). (3) After the allocation process, players calculate the outcomes with the help of individually defined utility functions (this again is assumed an instant of time and occurs at the end of a SSG, i.e. at the end of a stage). Each stage has the same duration.

A. Quality of Service Definitions

Two abstract and normalized representations of the QoS parameters are used: the normalized throughput $\Theta(n) \in [0,1]$ and the normalized delay $\Delta(n) \in [0,0.1]$. The delay variation $\Xi^i(n)$ can be derived from the delay and is thus not further considered here. The normalized throughput $\Theta^i(n)$ represents the share of capacity a player i demands in stage n of the game, and is defined in the following Eq. (1).

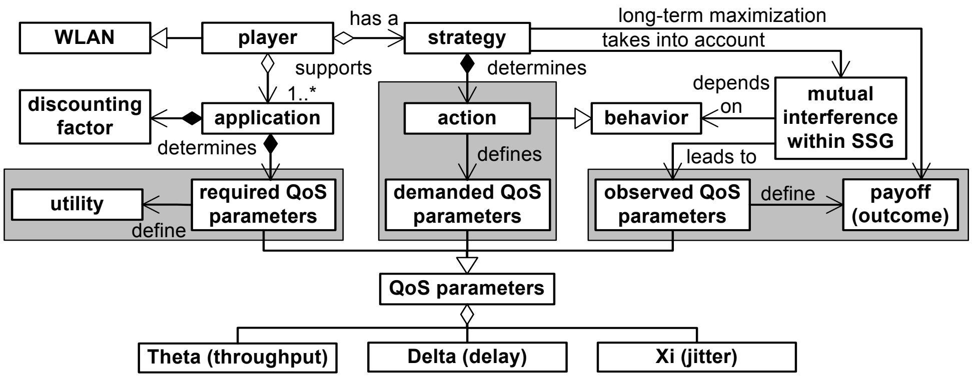


Fig. 1: UML representation of the Multi Stage Game. A WLAN is represented by a player, which has a strategy to determine what action to select. An action specifies a behavior. There are three types of QoS parameters, all in general constructed by Theta, Delta and Xi.

$$\Theta^i(n) = \frac{1}{SSGlength} \sum_{l=1}^{L^i(n)} d_l^i(n) \quad \in [0,1]. \qquad (1)$$

$L^i(n)$ is the number of allocations per stage n and $SSGlength$ the duration of the SSGs. The parameter $d_l^i(n)$ describes the duration of an allocation l, $l=1...L$, of player i in stage n. The normalized allocation delay $\Delta^i(n)$ specifies the maximum delay that the player i tolerates in stage n. In particular, this delay describes the observed maximum delay between two allocations:

$$\Delta^i(n) = \frac{1}{SSGlength} max\left[D_l^i(n)\right]_{l=1..L^i(n)-1} \quad \in [0,0.1]. \qquad (2)$$

$D_l^i(n)$ is the time between two successive allocations l and $l+1$ of player i in the stage n. Each player of the game structure has three different sets of QoS parameters: the 'required' (*req*), 'demanded' (*dem*) and 'observed' (*obs*) QoS parameters. Fig. 1 and Fig. 2 illustrate the interdependencies of these parameters in the context of a repeated SSG.

Player i's required QoS parameters Θ^i_{req} and Δ^i_{req} are defined by the QoS traffic the player is trying to support. At the beginning of each SSG, player i decides about demanded allocations corresponding to Θ^i_{dem} and Δ^i_{dem}. The demands are selected by players from stage to stage and determine the allocation point of times and lengths within a stage, i.e. within an SSG. In general, a player's observation deviates from the demand because of mutual influences in the allocation process. The observation determines the observed QoS parameters Θ^i_{obs} and Δ^i_{obs} as outcome of the SSG. The *action* $a^i(n)$ of player i is defined by the demand. An action of player i consists of the two demanded QoS parameters and is defined as

$$a^i := \left(\Theta^i_{dem}, \Delta^i_{dem}\right). \qquad (3)$$

We model a game consisting of two players, where an opponent player is referred to as player -*i*, according to the classical notation in game theory.

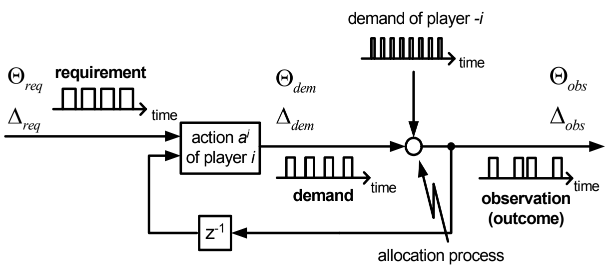


Fig. 2: The different QoS parameters of player i. Indicated are required resource allocations (left), demanded resource allocations (middle) that may deviate from the actual requirements if a player decides to improve the outcome. This outcome is defined by the observed resource allocations (right).

B. *Utility and Payoff under Competition*

The definition of the individual utility function considers the main characteristics of QoS. The utility U^i is used to determine what player i gains from a specific action a^i [5]. The utility of player i depends on the two utility terms U^i_Θ and U^i_Δ. These two terms are related to the observed share of capacity and times of resource allocation by the following equation:

$$U^i\left(a^i\right) = U^i_\Theta\left(a^i\right) \cdot U^i_\Delta\left(a^i\right). \qquad (4)$$

Here U^i is a non negative real number normalized to values between *0* and *1*. Details on the utility function can be found in [3]. If the player exclusively utilizes the radio resource without competition, then the player's QoS observation equals its QoS demand. To evaluate the outcome of an SSG under competition, the opponent's action has to be considered. Therefore, the *payoff* V^i of player i is defined as

$$V^i\left(a^i, a^{-i}\right) \rightarrow U^i\left(a^i\right). \qquad (5)$$

The payoff, as outcome of the stage, completes the SSG and highlights the dependency of player i's payoff V^i on the opponent's action a^{-i}.

C. *Behavior in Single Stage Games*

The actions that are available to a player, i.e. all combinations of demanded QoS parameters Θ_{dem} and Δ_{dem}, are illustrated in the action portfolio of Fig. 3. An action in the area of 'Selfish' behavior leads to an aggressive allocation scheme: a selfish player allocates long resource allocations, which are not required according to its current QoS requirements. Hence, this behavior can be interpreted as blocking out the opponent completely in demanding a high share of capacity without regarding its own QoS requirements, and is therefore referred to as selfish behavior.

The behavior 'Cooperation' intents to gain highest payoffs in the case of game-wide cooperation. Initially, 'Cooperation' allows only the opponent player *-i* to meet its requirements. Player *i* gains from cooperation as well in the case of a cooperating opponent. Nevertheless, cooperation implies weakness against a non-cooperating opponent, because the cooperating player can easily be blocked out.

The compromise between 'Cooperation' and 'Selfish' behavior is the 'BestResponse'. This behavior selects the action that would result in the highest expected payoff, provided the estimation of the opponent's behavior is correct.

The evaluation of potential actions with the help of the action portfolio can be extended towards the aspect of mutual influence. All actions of the player that reduce the payoff of the opponent can be considered as *punishment*, because the player is aware of its influence on the opponent.

III. Multi Stage Games

The above introduced game structure of an SSG, including the behavior of a player, enables a further dynamic interaction. This potential interaction within MSGs is introduced in the following.

A. Discounting – the Role of the Future

Rational acting players assign present payoffs a higher value than potential uncertain payoffs in the future. A general approach to model this preference is to discount the payoffs for each stage of the game.

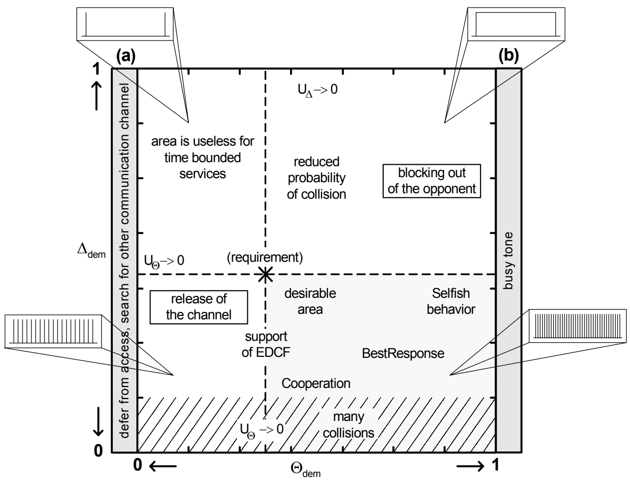


Fig. 3: Portfolio of available actions, the corresponding utilities and the resulting consequences on the opponents. Exemplary allocation schemes are depicted in each corner to illustrate the dependency of the allocations on the demanded QoS parameters. The case of (a) can be compared to a leaving of the MSG and (b) is an occupation of all resources for all time.

Therefore, a discounting factor δ^i , $0 < \delta^i < 1$, of player *i* is used, which reflects in the present stage the worth (or, the weight/importance) of the future payoffs of the next stages. The player's preference, i.e. δ^i , may be derived from corresponding QoS traffic types, which is here assumed to be constant over the MSG. Thus, also the discounting factor does not vary over the course of the MSG. A δ^i near one implies that future payoffs are considered similarly to the payoff of the current stage. Contrary, a player with a δ^i near zero only has its focus on the present payoff and completely neglects potential future payoffs.

In the radio resource sharing games, players do not know when the interaction with other players ends, i.e., which stage is the last stage. This is typically modeled by infinite games: players take decisions under the assumption that the game continues ad infinitum. Player *i*'s multi-stage payoff V^i of an infinite game is given as the sum over its payoffs V_n^i of stage *n* discounted with δ^i :

$$V^i = \sum_{n=0}^{\infty} \left(\delta^i\right)^n V_n^i = \frac{1}{1-\delta^i} V_n^i, \text{ if } V_n^i = const. \qquad (6)$$

B. Strategies – the Decision Taking

A *strategy* describes the alternatives a player has for an action under consideration of a potential interaction with the influenced players. In our game structure the strategy of a player is the decision-making process about the own action. Following Osborne and Rubinstein, strategies are "steady social norms that support mutually desirable payoffs" [6].

We distinguish between static and dynamic trigger strategies. Static strategies are based on behaviors, which have been introduced above. A static strategy is the continuous selection of a behavior throughout the course of the whole game. Dynamic trigger strategies consider the strategy of the opponent. In the scenarios we investigate here it is not possible for players to identify the opponent's strategy because players do not exchange information about their strategies directly. Nevertheless, players are able to classify the opponent's behavior as introduced in the next section. Consequently, a player may react on the opponent's action in following a trigger strategy based on this classification.

Trigger strategies lead to simple interactions: the opponent's behavior of the last stage determines a specific action chosen through the player's trigger strategy.

1) Classifying the Opponent's Behavior

A player classifies the opponent's behavior to estimate its intention while reducing the complexity of its strategies. All players are aware of the opponent's influence on their own payoff and are able to classify these actions under consideration of the corresponding opponent's behavior: whether it is *cooperation* or *deviation* based on the action portfolio. We define 'deviation' to be equivalent to the 'BestResponse' behavior. With this behavior, a player attempts to maximizes its payoff, independently to what the effect on the opponent's payoff may be.

Tab. 1: Payoff table for an MSG of two players.

Pl.*1*↓ Pl.*2*→	DEV	COOP
DEV	V^1_{DD}, V^2_{DD}	V^1_{DC}, V^2_{DC}
COOP	V^1_{CD}, V^2_{CD}	V^1_{CC}, V^2_{CC}

2) Payoff Table

For analysis, Tab. 1 defines the payoffs of two players depending on the classification of cooperation (C, or COOP) or deviation (D, or DEV).

An example scenario of two players representing overlapping WLANs is evaluated in the following. The QoS requirements for normalized throughput and normalized delay are defined as

$$\left(\begin{pmatrix}\Theta^1_{req}\\ \Delta^1_{req}\end{pmatrix},\begin{pmatrix}\Theta^2_{req}\\ \Delta^2_{req}\end{pmatrix}\right)=\left(\begin{pmatrix}0.4\\ 0.051\end{pmatrix},\begin{pmatrix}0.4\\ 0.042\end{pmatrix}\right). \quad (7)$$

Analytic results of this paper are evaluated with the help of the Matlab simulator YouShi, introduced in [1]. The payoff table in Tab. 2 shows the results taken from SSGs evaluated with YouShi, when using the QoS requirements of Eq. (7).

3) Static Strategies

Static strategies are the continuous selection of one behavior without regarding the opponent's strategy. seethe details can be found in [3]. We classify available behaviors as discussed earlier in this paper, such that the set of static strategies is reduced to two strategies labeled as 'COOP' and 'DEV'. The cooperation strategy (COOP) is characterized through cooperating.

Cooperation is selected at any stage, independently from the opponent's behavior. The deviation strategy (DEV) consists of the behavior of permanent deviation. The player maximizes its own payoff, independently from the opponent, while reducing the opponent's payoff. Fig. 4 (a,b) illustrates the CCOP and DEV strategy as a state machine, respectively.

4) Trigger Strategy TitForTat

Trigger strategies have been analyzed for the first time by Friedman [7]. We use the known *TitForTat* (TFT) trigger strategy in the following as an example to discuss dynamic strategies. The TFT strategy implies cooperation as long as the opponent is cooperating, with cooperation in the initial stage.

Tab. 2: Example with one steady operation point: (DEV|COOP) and a corresponding game outcome of (0.71|0.31). Neither player can gain a higher payoff in leaving (DEV|COOP).

Pl.*1*↓ Pl.2→	DEV	COOP
DEV	(0.25,0.05)	(0.71,0.31)
COOP	(0.24,0.78)	(0.40,0.56)

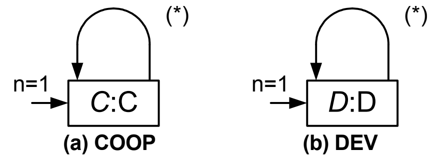


Fig. 4: State machines of the static strategies cooperation and deviation using the notation by Osborne and Rubinstein [6].

An opponent's deviation in stage *l* is punished with an deviation of the player in stage *l+1*, as illustrated in Fig. 5. The TitForTat strategy by Rapoport was the winning strategy of a tournament of a 200 times repeated 'Prisoner's Dilemma' [8]. The advantage of the TFT strategy is on the one hand the 'motivation' for the opponent to cooperate because of a potential punishment and on the other hand the robustness against non-cooperative strategies.

IV. ESTABLISHMENT OF COOPERATION

Player *i* only cooperates in an MSG if *i* gains a higher payoff than in deviating from cooperation. Therefore, Eq. (8) defines whether player *i* cooperates or not, resulting from Eq. (6). The payoff gain of deviation has to be compensated through punishment over one or several stages. Thus, the factor $n \in \mathbb{N}_0$ determines the number of stages of expected punishment through the opponent after a single deviation from cooperation. The left term of Eq. (8) contains the single deviation gain of the actual stage V^i_{DC} and the resulting stages of punishment leading to discounted payoffs of V^i_{CD}. The right side is the discounted payoff V^i_{CC} of game wide cooperation during the *n* stages.

$$V^i_{DC} + \sum_{k=1}^{n}\left(\delta^i\right)^k \cdot V^i_{CD} < \sum_{k=0}^{n}\left(\delta^i\right)^k \cdot V^i_{CC} \quad (8)$$

This equation enables to determine whether the player *i* can be convinced to cooperate or not because of the punishment threat, in isolating δ^i under the assumption of a game long punishment, i.e. $n \to \infty$:

$$\delta^i > \frac{V^i_{CC} - V^i_{DC}}{V^i_{CD} - V^i_{DC}} \quad (9)$$

Fig. 6 provides a basic understanding with the help of the exemplary player *1* of the example game from Tab. 2. In the figure, the MSG outcomes of player *i*, here player *1*, are shown.

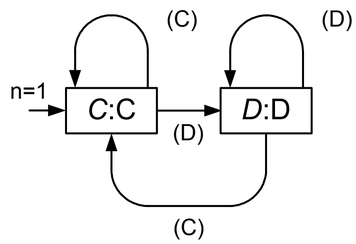
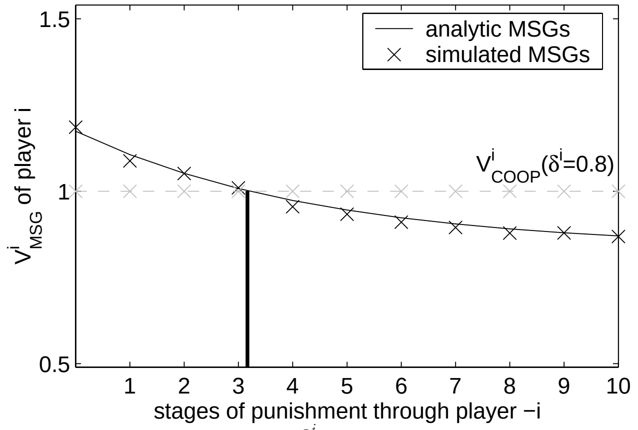
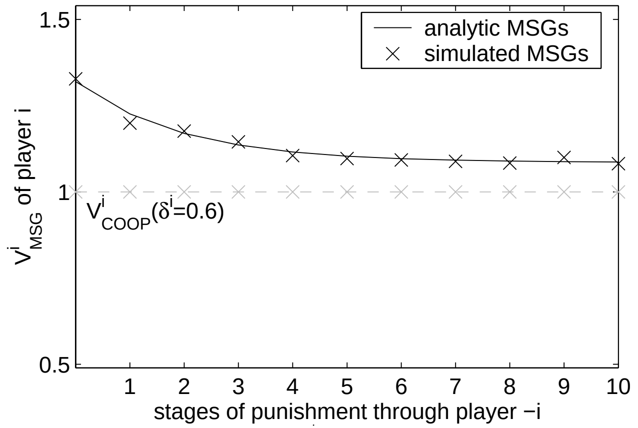


Fig. 5: State machines of the TitForTat (TFT) trigger strategy. It implies deviation, if the opponent deviates, and cooperation in the case of a cooperating opponent.



(a) $\delta^i = 0.8$



(b) $\delta^i = 0.6$

Fig. 6: To V^i_{COOP} normalized MSG outcomes of player *i*. Player *i* deviates for a single stage and is consequently punished by the opponent for varied stages.

The lines mark the pre-calculated payoffs, determined at the time of the player's decision which strategy to choose. The crosses show the real discounted observed payoffs from simulated MSGs with YouShi. The third player, representing the EDCF traffic, influences the payoffs in the simulated MSGs, while it is not considered in the analysis. Player *i* deviates for a single stage and is consequently punished by the opponent. Depending on the intensity of the punishment, i.e. the number of stages with punishment, the players discounted payoff from this single deviation is higher than the payoff because of pure cooperation. If the payoff from cooperation $V^i_{COOP}(\delta^i)$ is higher than the deviation payoff, cooperation can be established through the credible threat of punishment sustained by the opponent. For example in the case of $\delta^i = 0.8$, player *1* has to expect a punishment of four times $(n = 4)$ that it remains in cooperation: Eq. (8) is invalid for $n \geq 4$, see Fig. 6 (a).

For small values of δ^i the player *i* gives the short term payoff gain (which is achieved by deviation) a higher value than the long term gain (achieved by cooperation), see Fig. 6 (b) for $\delta^i = 0.6$. Thus cooperation cannot be established when this discounting factor is selected. Here, the player can be forced to cooperate for discounting factors between $0.672 < \delta^i < 1$, as can be calculated with Eq. (9). For values of δ^i below 0.672, no cooperation can be enforced, even if the punishment has an infinite duration.

V. Conclusion and Future Work

This paper introduces a stage-based game structure to solve the coexistence problem of overlapping wireless LANs and outlines a potential interaction under consideration of the players preference of future payoffs. Based on the game structure, proven concepts taken from economics, such as the Nash Equilibrium or Pareto efficiency can be used to evaluate the outcomes of MSGs.

The equilibrium analysis leads to predictable steady MSG outcomes with a determinable level of QoS. The consideration of Pareto efficiency and subgame perfection will help to judge games with several equilibria. Additionally, strategies and a learning in games, as for example the adaptation of the own strategy to the opponent ones, may improve the ability of the players to support successful QoS on a satisfying level.

References

[1] Mangold, S. and Berlemann, L. and Hiertz, G. "QoS Support as Utility for Coexisting Wireless LANs," in Proc. of *International Workshop in IP Based Cellular Networks*, Paris France, 23-26 April 2002.

[2] Mangold, S. and Berlemann, L. and Walke, B. "Radio Resource Sharing Model for Coexisting IEEE 802.11e Wireless LANs," in Proc. of *International Conference on Communication Technology*, Beijing China, 9-11 April 2003.

[3] Mangold, S. and Berlemann, L. and Walke, B. "Equilibrium Analysis of Coexisting IEEE 802.11e Wireless LANs," in Proc. of *International Symposium on Personal, Indoor and Mobile Radio Communications*, Beijing China, 7-11 September 2003.

[4] Mangold, S. "Analysis of IEEE 802.11e and Application of Game Models for Support of Quality-of-Service in Coexisting Wireless Networks," PhD Thesis, Aachen University ComNets, 2003.

[5] Fudenberg, D. and Tirole, J. "Game Theory," MIT Press, 1998.

[6] Osborne, M.J. and Rubinstein, A. "A Course in Game Theory," Cambridge Massachusetts USA, London England: The MIT Press, 1994.

[7] Friedman, J. W. "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies 38*, 1971.

[8] Axelrod, R. "The Evolution of Cooperation." New York: Basic Books, 1984.