# Backward Induction and Repeated Games With Strategy Constraints: An Inspiration From the Surprise Exam Paradox

Jiawei Li, *Member, IEEE*, Graham Kendall, *Senior Member, IEEE*, and Athanasios V. Vasilakos, *Senior Member, IEEE*

***Abstract*—Backward induction has led to some controversy in specific games, the surprise exam paradox and iterated prisoner's dilemma for example, despite its wide use in solving finitely repeated games with complete information. In this paper, a typical misuse of backward induction is revealed by analyzing the surprise exam paradox, and the reason why backward induction may fail is investigated. The surprise exam paradox represents a set of repeated games with strategy constraints and has not been fully investigated in game theory. The agents in real-world activities always face constraints in decision making, for example, a budget limitation. In a repeated game with strategy constraints, the players' choices in different stages are not independent and later choices depend on previous choices because of the strategy constraints. Backward induction cannot be applied in its normal use and it needs to be combined with Bayes' theorem in solving these kinds of problems. We also investigate how the strategy constraints influence the equilibrium and show how to solve repeated games with strategy constraints by analyzing a repeated battle of the sexes game with a budget constraint.**

***Index Terms*—Backward induction, Bayes' theorem, repeated game theory, surprise exam paradox.**

## I. Introduction

BACKWARD induction, also called retrograde analysis in game playing or backward reasoning in optimization and AI [32], [16], [5], has been a commonly used methodology for game tree search since the establishment of game theory [38]. When applying it, we first anticipate what choice the players would make in the last stage. Given the choice in the last stage, we then deduce what the players would choose in the second to last stage, and so on. It has been applied to solve complicated games like *Chess* [37], *Go* [12], and *Chinese Chess* [17]. Despite its wide use, there is still some controversy, especially as regards its use in the traveler's dilemma, the iterated prisoner's dilemma, and the surprise exam paradox.

Manuscript received December 29, 2012; revised April 15, 2013; accepted April 26, 2013. Date of publication May 06, 2013; date of current version September 11, 2013. This work was supported in part by the Engineering and Physical Science Research Council (EPSRC) under Grant EP/H000968/1 and in part by the KU Research Sector.

J. Li is with the School of Computer Science, University of Nottingham, Nottingham NG8 1BB, U.K. (e-mail: jiawei.li@nottingham.ac.uk).

G. Kendall is with the School of Computer Science, University of Nottingham, Semenyih, Selangor Darul Ehsan 43500, Malaysia (e-mail: graham.kendall@nottingham.edu.my).

A. V. Vasilakos is with the Computer Science Department, Kuwait University, Safat 13060, Kuwait (e-mail: th.vasilakos@gmail.com).

Digital Object Identifier 10.1109/TCIAIG.2013.2261894

Experimental economists have found that subjects in the traveler's dilemma do not make choices according to backward induction [7], [13]. In finitely iterated prisoner's dilemma experiments, cooperation between participants is not uncommon, although it is not an equilibrium choice [23], [15], [22], [2]. Cooperation can be achieved in the early stages if there are sufficient iterations. The rate of cooperation at the beginning of the game is significantly high, and it tends to be rare at the end of the game. Although theorists have attempted to rescue the argument for backward induction by means of bounded rationality [3], incomplete information [25], reputation [26], and other-regarding preferences [18], [9], [29], the debate is far from being settled [8], [31].

The surprise exam paradox, also known as the unexpected examination or the unexpected hanging paradox, has evoked research interest among mathematicians and philosophers for decades. A version of the surprise exam paradox is described as follows [14].

> "A teacher announces in class that an examination will be held on some day during the following week, and moreover that the examination will be a surprise. The students argue that a surprise exam cannot occur. For suppose the exam were on the last day of the week. Then on the previous night, the students would be able to predict that the exam would occur on the following day, and the exam would not be a surprise. So it is impossible for a surprise exam to occur on the last day. But then a surprise exam cannot occur on the penultimate day, either, for in that case the students, knowing that the last day is an impossible day for a surprise exam, would be able to predict on the night before the exam that the exam would occur on the following day. Similarly, the students argue that a surprise exam cannot occur on any other day of the week either. Confident in this conclusion, they are, of course, totally surprised when the exam occurs (on Wednesday, say). The announcement is vindicated after all. Where did the students' reasoning go wrong?"

The paradox has been thoroughly investigated as a pragmatic problem. The logical school of thought suggests that the statement of the teacher is self-contradictory and self-referencing so that the students' reasoning is based on an unsound assumption [33], [34], [11]. The epistemological school of thought agrees that the statement of the teacher is self-contradictory and tries to make the problem clearly defined. It explains that the students' assumptions about what they will know in the future are inconsistent [14], [36], [27], [20]. Although the paradox itself is considered to be resolved, there is still no consensus on the resolution.

The paradox is modeled as a repeated matching pennies game in [35]. With this model, the teacher and the students play a matching pennies game at each stage and the game ends once the exam has occurred (it looks like a centipede game in this sense). The Nash equilibrium of a three-stage surprise exam paradox is studied. It concludes that a surprise exam can be given if the teacher assigns a nonzero probability to giving the exam on the last day. This conforms to [10], in which the teacher's probability distribution to maximize the surprise is deduced. Another approach models the surprise exam paradox as an information-dependent game, in which a set of prediction profiles is added [21]. The set of predictions chosen by the players has an influence on the payoff matrix as well as the strategies. It shows that there is inconsistency of information in the process of game playing. Several more general models of the paradox are discussed in [19].

Backward induction cannot be applied in its normal way in the surprise exam paradox, given that the teacher must give an exam on someday during the following week as it is promised. This is a constraint on the teacher's strategy and both the teacher and the students know it. Because of the constraint, the choices of both sides at each stage are not independent and backward induction cannot be applied in its normal way.

In this paper, we analyze the surprise exam paradox by using a new game-theoretical framework, which we call repeated game with strategy constraints, and we show that backward induction should be used together with Bayes' theorem to solve this kind of repeated game.

Most games studied in game theory so far are games without strategy constraints, i.e., the players choose their strategies independently and respectively. However, many real-world activities can be modeled by games with strategy constraints when the agents have a budget limitation. As a typical example, income is a constraint for individual household consumption. In the classic battle of the sexes game, for example, the husband and the wife choose between attending an opera and watching a football match [28]. Consider the situation that the couple makes choices repeatedly and they have a budget limitation on the entertainment expenditure. The budget forms a constraint on their strategies if the two alternatives have different prices. Some problems in engineering have been studied as games with coupled constraints, for example, the constrained potential game [39], the network game [6], [40], the optical signal-to-noise ratio (OSNR) game [30], and general constraint games [1]. The above research focuses on the methodologies of computing Nash equilibrium of specific games, and, thus, is problem dependent. This paper studies a general form of repeated game with strategy constraints and focuses on the influence of constraints on the players' repeated choices.

The contribution of this study is twofold. First, it offers a methodological contribution to the application of backward induction. Although the use of backward induction is limited to repeated games with complete information, it may still be misused. The surprise exam paradox shows a typical misuse of backward induction. When there is a constraint on the strategies of the players, the choices of a player in multiple stages are coupled. Or, in other words, the choices in previous stages have an influence on later choices. This influence needs to be taken into consideration if backward induction is applied. Second, the surprise exam paradox is analyzed by using a new game-theoretic framework, repeated games with strategy constraints, and the reason why backward induction fails in the paradox is revealed. We show that backward induction can be used together with Bayes' theorem to compute the players' choices in solving repeated games with strategy constraints.

The rest of this paper is organized as follows. Section II analyzes the surprise exam paradox as a repeated game with strategy constraints. Section III discusses backward induction. Section IV introduces the finitely repeated games with strategy constraints and analyzes an example of a repeated battle of sexes game with a budget constraint. Section V concludes the paper.

| | | *The Teacher* | |
|---|---|---|---|
| | | Exam | No Exam |
| *The Students* | Predict Exam | *(x, -x)* | *(-y, y)* |
| | Predict No Exam | *(-x, x)* | *(0, 0)* |

Fig. 1. Payoff matrix of the stage game of the surprise exam paradox (where $x \geq y > 0$).

## II. The Surprise Exam Paradox

First, we represent the surprise exam paradox as a normal form game. The game is a two-player $n$-stage repeated game. At each stage, the teacher will choose between two alternatives: exam $(E)$ and no exam $(N)$, while the students will choose between two predictions: exam $(E)$ and no exam $(N)$ at the same time. If the students predict an exam successfully, they gain a payoff $x$ and the teacher receives $-x$. If the students fail to predict an exam, they receive $-x$ and the teacher receives $x$. If the students choose $E$ while the teacher chooses $N$, the teacher and the students receive payoffs of $y$ and $-y$, respectively. If the students choose $N$ while the teacher chooses $N$, both sides receive zero. The payoff matrix is shown in Fig. 1.

There is a constraint for the strategy of the teacher: the teacher must choose $E$ exactly once in the $n$ stages. This constraint is known to both players.

Because of the constraint, this game is different from traditional finitely repeated games, in which the players' choices at each stage are independent. The game cannot be analyzed by applying backward induction directly.

Consider a one shot game with the payoff matrix as shown in Fig. 1. Let us first consider the situation in which there is not any constraint on the strategies of the teacher and the students. Since no pure-strategy Nash equilibrium exists in this game, assume that both players choose a probability distribution. The teacher assigns probability $p$ to $E$ and the students assign probability $q$

to $E$. The expected payoffs to the teacher and the students, $P_T$ and $P_S$, can be expressed as

$$\begin{aligned} P_T &= pq(-x) + p(1-q)x + q(1-p)y \\ &= px + qy - pq(2x+y) \\ P_S &= -P_T = pq(2x+y) - px - qy. \end{aligned}$$

The Nash equilibrium is that the teacher chooses $p = y/(2x+y)$ and the students choose $q = x/(2x+y)$. The expected payoffs to the teacher and the students are $P_T = (xy)/(2x+y)$ and $P_S = (-xy)/(2x+y)$.

When the constraint is in force, the teacher has to choose $E$. Thus, the students will choose $E$, and the payoffs to the teachers and students are $-x$ and $x$, respectively. In this section, we start from a two-stage game, and then extend it to a three-stage, and then $n$-stage game.

### A. Two-Stage Game ($n = 2$)

Assume that the probabilities of the teacher choosing $E$ in the first and second stages are $p_1$ and $p_2$, respectively. We have $p_1 + p_2 = 1$ because the teacher must choose $E$ exactly once in two stages. Similarly, the probabilities of the students choosing $E$ in two stages are $q_1$ and $q_2$.

In the two-stage game, if the teacher has chosen $E$ in the first stage, she must choose $N$ in the second stage, and, thus, the students must choose $N$ in the second stage. On the other hand, if she has chosen $N$ in the first stage, she must choose $E$ in the second stage, and, thus, the students must choose $E$. Thus, the payoffs of the teacher and the students in the second stage are determined, given the teacher's choice in the first stage. The payoffs to the teacher and the students can then be computed as

$$\begin{aligned} P_T &= p_1 x + q_1 y - p_1 q_1 (2x+y) + p_2(-x) \\ &= 2p_1 x + q_1 y - p_1 q_1 (2x+y) - x \\ P_S &= -P_T. \end{aligned}$$

If the teacher chooses $p_1 = y/(2x+y)$, the payoffs $P_T$ will be $(-x(2x-y))/(2x+y)$, which is independent of the values of $q_1$ and $q_2$. On the other hand, if the students choose $q_1 = (2x)/(2x+y)$, the payoff $P_S$ will be $(x(2x-y))/(2x+y)$ no matter what value of $p_1$ and $p_2$ the teacher chooses. So this is the Nash equilibrium of this game.

The Nash equilibrium is that the teacher chooses $p_1 = y/(2x+y)$ and $p_2 = (2x)/(2x+y)$ and the students choose $q_1 = (2x)/(2x+y)$ and $q_2 = y/(2x+y)$. The expected payoffs to the teacher and the students are $P_T = (-x(2x-y))/(2x+y)$ and $P_S = (x(2x-y))/(2x+y)$.

Note that both $p_i$ and $q_i$ are *a priori* probabilities, that is, the teacher and the students assign these probabilities before the first stage of the game. After the first stage, the students may update their strategies (probability distribution) according to the outcome of the first stage. For example, if the teacher has chosen $N$ at the first stage, the students should choose $E$ at the second stage no matter what the value of $q_2$. Let $q'_2$ denote the students' posterior probability, given that no exam occurs in the previous stage. There $q'_2 = 1$. The relationship between $q_2$ and $q'_2$ is $q_2 = (1-q_1)q'_2$. Because *a priori* probabilities are not easy to compute directly, posterior probabilities are used in the situations with $n \geq 3$.

### B. Three-Stage Game ($n = 3$)

Let $p_1$, $p_2$, and $p_3$ be the probabilities of the teacher choosing $E$ in the three stages and $q_1$, $q_2$, and $q_3$ the probabilities of the students choosing $E$. Let $p'_2$ be the posterior probability of the teacher choosing $E$ at the second stage, given that no exam occurs at the first stage, and $p'_3$ be the posterior probability of the teacher choosing $E$ at the third stage, given that no exam occurs in the first two stages. Similarly, $q'_2$ and $q'_3$ are the posterior probabilities of the students choosing $E$ at the second and third stages, respectively.

Obviously, $p'_3 = 1$ and $q'_3 = 1$. We also have $p'_2 = y/(2x+y)$ and $q'_2 = (2x)/(2x+y)$ according to the Nash equilibrium of the two-stage game. *A priori* probabilities can be computed by using posterior probabilities as follows:

$$\begin{aligned} p_2 &= (1-p_1)p'_2 \\ p_3 &= (1-p_1)(1-p'_2)p'_3 \\ q_2 &= (1-q_1)q'_2 \\ q_3 &= (1-q_1)(1-q'_2)q'_3. \end{aligned}$$

We just need to know the values of $p_1$ and $q_1$ in order to determine all *a priori* probabilities listed above.

The payoffs to the teacher and the students can be expressed as

$$\begin{aligned} P_T &= p_1 x + q_1 y - p_1 q_1 (2x+y) \\ &\quad + (1-p_1)(p'_2 x + q'_2 y - p'_2 q'_2 (2x+y)) \\ &\quad - (1-p_1)(1-p'_2)p'_3 x \\ P_S &= -P_T. \end{aligned}$$

This simplifies to

$$\begin{aligned} P_T &= p_1 \frac{4x^2}{2x+y} + q_1 y - p_1 q_1 (2x+y) + \frac{xy - 2x^2}{2x+y} \\ P_S &= -P_T. \end{aligned}$$

If the teacher chooses $p_1 = y/(2x+y)$, the payoff to the teacher $P_T$ is independent of the values of $q_i$. If the students choose $q_1 = (4x^2)/((2x+y)^2)$, the payoff to the students $P_S$ is independent of the values of $p_i$. Thus, the Nash equilibrium is that $p_1 = y/(2x+y)$, $p_2 = (2xy)/((2x+y)^2)$, $p_3 = (4x^2)/((2x+y)^2)$, and $q_1 = (4x^2)/((2x+y)^2)$, $q_2 = (2xy(4x+y))/((2x+y)^3)$, $q_3 = (y^2(4x+y))/((2x+y)^3)$. The payoffs to the teacher and the students are $P_T = (x(4xy - 4x^2 + y^2))/(2x+y)$ and $P_S = -P_T$.

### C. N-Stage Game ($n > 3$)

Let $p_i$ and $q_i$ ($i = 1, \ldots, n$) be *a priori* probabilities that the teacher and the students assign to choosing $E$ at the $i$th stage, and $p'_i$ and $q'_i$ ($i = 2, \ldots, n$) the posterior probabilities that the teacher and the students choose $E$ at the $i$th stage. We have $\sum_{i=1}^{n} p_i = 1$.

We just need to compute the values of $p'_i$, $q'_i$, $p_1$, and $q_1$. For the convenience of expression, let us define that $p'_1 = p_1$ and

$q'_1 = q_1$. We are about to compute the posterior probabilities $p'_i$ and $q'_i$ $(i = 1, \ldots, n)$. The *a priori* probabilities could be determined by $p'_i$ and $q'_i$, according to the Bayes' theorem

$$p_i = p'_i \prod_{j=1}^{i-1} (1 - p'_j) \tag{1}$$

$$q_i = q'_i \prod_{j=1}^{i-1} (1 - q'_j). \tag{2}$$

Posterior probabilities should be subgame perfect. Given that no exam occurs at the first stage, $p'_2$ and $q'_2$ will be the optimal choices for the teacher and the students at the second stage. Given that no exam occurs at the first, second, ... $i$th stage, $p'_{i+1}$ and $q'_{i+1}$ will be the optimal choices at the $(i+1)$th stage. Thus, we could adopt backward induction to compute $p'_i$ and $q'_i$.

Let $P_T(i)$ denote the Nash equilibrium payoff to the teacher, given that no exam occurs in the previous $i-1$ stages $(1 \leq i \leq n)$. We have

$$P_T(i) = p'_i x + q'_i y - p'_i q'_i (2x + y) + (1 - p'_i) P_T(i+1) \tag{3}$$

where $i = 1, \ldots n-1$. Note that $p'_i$, $q'_i$, and $P_T(i-1)$ are independent, and $p'_n = 1$ and $q'_n = 1$.

It is easy to deduce that

$$p'_i = \frac{y}{2x + y}, \qquad i = 1, \ldots n-1 \tag{4}$$

and $P_T(i) = (xy)/(2x+y) + (2x)/(2x+y) P_T(i+1)$. So, we have

$$P_T(i) = \frac{xy}{2x+y}\left(1 + \left(\frac{2x}{2x+y}\right)^2 + \cdots + \left(\frac{2x}{2x+y}\right)^{n-i-1}\right) + \left(\frac{2x}{2x+y}\right)^{n-i} P_T(n).$$

Since $P_T(n) = -x$, we have

$$P_T(i) = x - 2x\left(\frac{2x}{2x+y}\right)^{n-i}, \qquad i = 1, \ldots n-1.$$

Substitute $x - 2x((2x)/(2x+y))^{n-i-1}$ for $P_T(i+1)$ in (3), then we have

$$q'_i = \frac{1}{x}\left(\frac{2x}{2x+y}\right)^{n-i}, \qquad i = 1, \ldots n-1. \tag{5}$$

The probabilities (both *a priori* and *a posteriori*) that the teacher and the students assign to choosing $E$ at the $i$th stage could be computed by means of (1)–(5) for any value of $n \geq 3$.

The Nash equilibrium payoffs to the teacher and the students are

$$P_T = x - 2x\left(\frac{2x}{2x+y}\right)^{n-1} \tag{6}$$

$$P_S = -x + 2x\left(\frac{2x}{2x+y}\right)^{n-1}. \tag{7}$$

In Fig. 2, the strategies and payoffs of the teacher and the students are expressed as functions of $y/x$, given $n = 5$. Fig. 2(a)–(e) shows the difference between *a priori* probabilities and *a posteriori* probabilities.

In the situation that $x = 1.0$, $y = 0.5$ and $n = 5$, for example, the optimal strategies of the teacher and the students are $p'_i = (1/5, 1/5, 1/5, 1/5, 1)$ and $q'_i = (256/625, 64/125, 16/25, 4/5, 1)$, or $p_i = (0.2, 0.16, 0.128, 0.102, 0.41)$ and $q_i = (0.41, 0.302, 0.184, 0.083, 0.021)$. The teacher and the students will receive $0.181$ and $-0.181$, respectively.

In another situation that $x = y = 1.0$ and $n = 5$, for example, we have $p'_i = (1/3, 1/3, 1/3, 1/3, 1)$ and $q'_i = (16/81, 8/27, 4/9, 2/3, 1)$. The Nash equilibrium is that the teacher assigns probability distribution $p_i = (0.333, 0.222, 0.148, 0.099, 0.198)$ and the students choose $q_i = (0.198, 0.238, 0.251, 0.209, 0.105)$ and the payoffs to the teacher and the students are $0.605$ and $-0.605$, respectively.

We have analyzed the surprise exam paradox as an $n$-stage repeated game and deduced the Nash equilibrium solution. With the Nash equilibrium solution, the teacher assigns the probabilities of choosing $E$ at each stage. Especially, there is $p_n \neq 0$. As shown in Fig. 3, both the teacher and the students should assign a nonzero probability to choosing $E$ at the last stage.

As $n$ increases, the expected payoff of the teacher $P_T$ increases and the expected payoff of the students $P_S$ decreases. If $n \to \infty$, there will be $P_T \to\to x$ and $P_S \to y$. This means that the probability that the teacher successfully gives a surprise exam is 1 if the semester is infinitely long.

## III. Backward Induction

The application of backward induction is limited to repeated games (or one-shot games with iterated decisions) as long as the games contain a finite game tree and they are with complete information. However, backward induction may still be misused in some cases. In this section, we discuss why backward induction "fails" in the surprise exam paradox and how to modify its usage to solve this, and similar problems.

It is clear what is wrong with the students' logic in the surprise exam paradox following the analysis in the previous section. They start with the statement "If the exam has not occurred before Friday, we know that it must be on Friday," which means that $p'_5 = q'_5 = 1$. This statement is true. But they deduce that "The exam cannot occur on Friday," which means that $p_5 = 0$. The second statement cannot be deduced from the first statement and is thus incorrect. According to (1), we know that $p_n$ depends on $p'_i$ $(i = 1, \ldots, n)$ and it is not solely determined by $p'_5$.

Given that "If the exam has not occurred after Wednesday, we know that it must be on Thursday," we cannot deduce the statement "The exam cannot occur on Thursday." The former statement is equivalent to $p'_4 = p'_5 = 1$ while the latter one is equivalent to $p_4 = 0$. $p_4 = 0$ does not hold even if $p'_4 = p'_5 = 1$ is true.

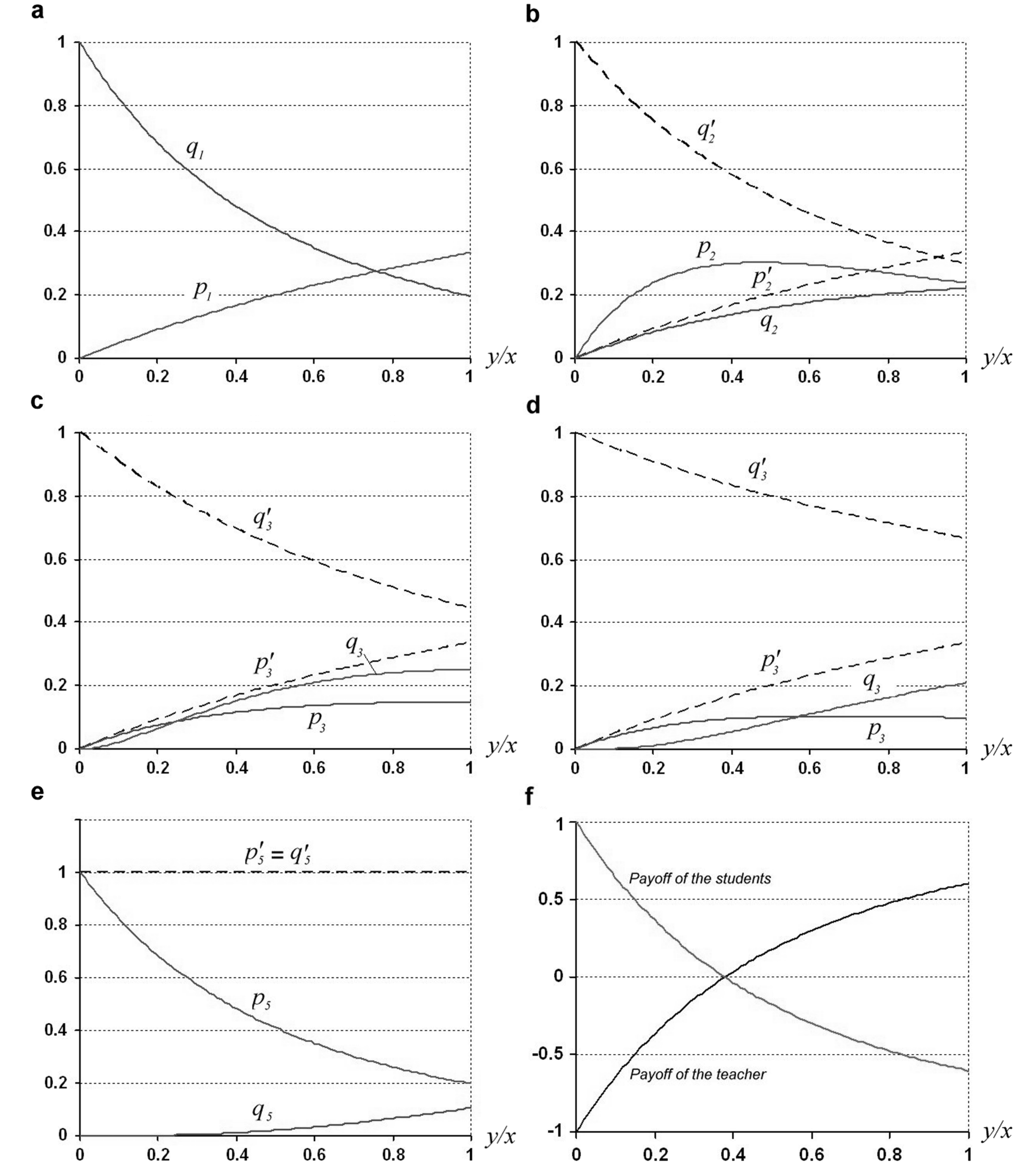


Fig. 2. The surprise exam paradox is analyzed as a five-stage game with strategy constraint. (a)–(e) The probabilities (both *a priori* and *a posteriori*) are expressed as functions of $y/x$. (f) The payoffs of the teacher and the students. (a) Probability of choosing $E$ on Monday. (b) Probability of choosing $E$ on Tuesday. (c) Probability of choosing $E$ on Wednesday. (d) Probability of choosing $E$ on Thursday. (e) Probability of choosing $E$ on Friday. (f) Payoff.

Similarly, $p_i = 0$ does not hold even if we have $p'_i = p'_{i+1} = \cdots p'_n = 1$. We express the students' logic in Fig. 4. Each odd step corresponds to an incorrect deduction.

It is obvious that step 1 is not the only mistake in the students' logic. Given $p_5 = 0$, the teacher can still find a strategy to give a surprise exam. In the situation that $x = y = 1$ and $n = 5$, for example, the teacher could assign $p' = (1/3, 1/3, 1/3, 1, 1)$ or $p = (1/3, 2/9, 4/27, 8/27, 0)$ and the expected payoff to the teacher would be 0.407 if the students choose their best strategy so far $q' = (8/27, 4/9, 2/3, 1, 1)$. Thus, the teacher can exclude

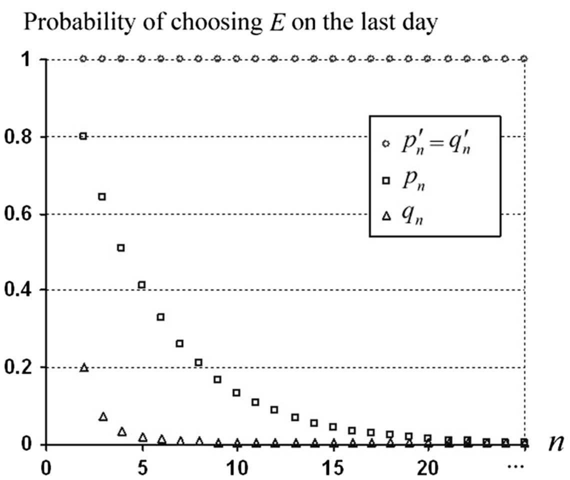


Fig. 3. The probabilities of choosing $E$ on the last day are expressed as the functions of $n$, given that $y/x = 0.5$.

$$\begin{array}{ccccccccccccccc}
p'_5 = 1 & & p_5 = 0 & & p'_4 = 1 & & p_4 = 0 & & p'_3 = 1 & & \cdots & & p'_2 = 1 & & p_1 = 0 \\
 & \Rightarrow & & \Rightarrow & & \Rightarrow & & \Rightarrow & & \Rightarrow & & \Rightarrow & & \Rightarrow & \\
q'_5 = 1 & & q_5 = 0 & & q'_4 = 1 & & q_4 = 0 & & q'_3 = 1 & & \cdots & & q'_2 = 1 & & q_1 = 0 \\
\text{Step 1} & & \text{Step 2} & & \text{Step 3} & & \text{Step 4} & & & & \cdots & & \text{Step 7} & & \text{Step 8}
\end{array}$$

Fig. 4. The students' logic in the surprise exam paradox. They start from $p'_5 = q'_5 = 1$ that is a correct statement. However, $p_5 = q_5 = 0$ cannot be deduced because *a priori* probabilities depend on the choices made in previous stages. Each odd step corresponds to an incorrect deduction.

the last day as a possible exam day. She should not exclude the last day purely because it leads to a higher probability of giving the surprise exam.

The surprise exam paradox shows a typical misuse of backward induction. When there is a constraint on the strategies of the players, the choices of a player in multiple stages are coupled. Or, in other words, the choices in previous stages have an influence on later choices. This influence needs to be taken into consideration if backward induction is applied.

We have shown how to use backward induction to solve an $n$-stage surprise exam paradox in Section II. Given $p'_n = q'_n = 1$, we first deduce the values of $p'_{n-1}$ and $q'_{n-1}$, and then $p'_{n-2}$ and $q'_{n-2}$, until $p'_2$ and $q'_2$. Backward induction is used to deduce posterior probabilities (or conditional probabilities). Given posterior probabilities, *a priori* probabilities could be computed by means of Bayes' theorem

$$P(A \mid B) = \frac{p(B \mid A)p(A)}{p(B)} \tag{8}$$

where $A$ and $B$ are two events, $p(A)$ and $p(B)$ are the (*a priori*) probabilities of $A$ and $B$, and $p(A \mid B)$ and $p(B \mid A)$ are the conditional probabilities of $A$, given $B$, and $B$, given $A$.

The traditional usage of backward induction in repeated games assumes that the choices in different stages are independent. These are $p(A \mid B) = p(A)$ and $p(B \mid A) = p(B)$, if the events $A$ and $B$ are independent. Only in this situation could *a priori* probabilities be equivalent to posterior probabilities.

In summary, backward induction could be used in its normal usage only if the choices of the players in different stages were independent. Otherwise, we need Bayes' theorem together with backward induction to compute the players' choices.

*The husband*

| *The wife* | Opera | Football |
|---|---|---|
| Opera | *(3, 2)* | *(0, 0)* |
| Football | *(0, 0)* | *(2, 3)* |

Fig. 5. Payoff matrix of the battle of the sexes game.

## IV. Repeated Games With Strategy Constraints

Let $G = \langle N, (S_i), (u_i) \rangle$ be an $n$-player normal form game. Each player has a finite strategy (action) space $S_i$, and a corresponding payoff function $u_i : S \to \mathbb{R}$, where $S = \Delta(S_i)$ is the set of player $i$'s mixed strategies. This stage game is finitely repeated at each discrete time period $t = 1, 2, \ldots, T$. Let $s^t \equiv (s_1^t, s_2^t, \ldots, s_n^t)$ be the strategy profile chosen in period $t$, and $s_{-i}$ a strategy profile of all players except for player $i$.

Assume that each action $s_i$ will lead to a constant consumption $C_i$ and the player faces a budget constraint $B_i$ on total consumption in $T$ periods. The optimal strategy of each player in $T$ period is a sequence of actions $(s_i^1, s_i^2, \ldots, s_i^t)$, which is the solution of the following programming problem:

$$\max \sum_{t=1, s_i^t \in S_i}^{t=T} u_i^t \left(s_i^t, s_{-i}^t\right)$$

subject to:

$$\forall s_j^t, \qquad \bar{s}_j^t \in S_j, \qquad s_j^t \in S_{-i}, \qquad \bar{s}_j^t \neq s_j^t$$

$$\sum_{t=1}^{T} u_j^t \left(\bar{s}_j^t, s_{-j}^t\right) \le \sum_{t=1}^{T} u_j^t \left(s_j^t, s_{-j}^t\right)$$

$$\text{and } \sum_{t=1}^{T} C_i^t \le B_i.$$

The second constraint reflects the limitation on player $i$'s strategy. Because of the constraint, the set of possible strategies of player $i$ shrinks. However, the player is not necessarily worse off, compared with the situation without the strategy constraint. We show this by analyzing a repeated battle of the sexes game.

The battle of the sexes is a two-player coordination game. Two players, say the husband and the wife, choose between two alternatives "opera" and "football" independently. The payoff matrix is shown in Fig. 5. The values in brackets are the payoffs to the wife and husband, respectively. There are two pure strategy Nash equilibria (both players choose "opera" and both players choose "football") and one mixed strategy Nash equilibrium (both players choose their preferred choices with probability 0.6) in this game. If there is no communication between the two players, mixed strategy Nash equilibrium is very likely to be the players' choices. The players will coordinate with probability 0.48, leaving each player with an expected payoff of 1.2 (less than the payoff one would receive from choosing one's less favored event if coordination could be achieved).

Now consider a repeated game in which two players play the battle of the sexes game twice. Assume that the choice of "football" means "watching a live football match in a pub," and attending an opera is more expensive than the other choice. The

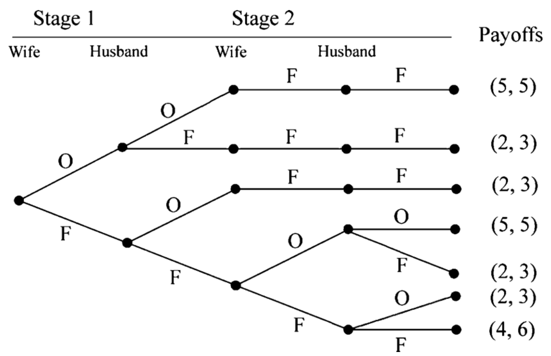


Fig. 6. Strategy tree of the two-stage battle of the sexes game.

strategy constraint is: Each player can choose "opera" no more than once in two stages.

Let us first consider the pure strategy. Because of the strategy constraint, a player cannot choose "opera" twice. Each player actually has three alternatives in two stages: (opera, football), (football, football), and (football, opera). The strategy tree can be expressed as shown in Fig. 6. Notice that some strategy combinations are not feasible. When the wife has chosen "opera" at the first stage, for example, the husband will not choose "opera" at the second stage because he knows that she must choose "football" at the second stage.

There are three pure strategy Nash equilibria in this game:

1) both choose "opera" at the first stage and both choose "football" at the second stage;
2) both choose "football" at the first stage and both choose "football" at the second stage;
3) both choose "football" at the first stage and both choose "opera" at the second stage.

Whichever strategy both players choose, they coordinate at least once (both choosing "football") in this game. The minimum payoffs to the wife and husband are 2.0 and 3.0, respectively. In the situation where there is no strategy constraint, however, miscoordination in both stages is possible.

Now let us consider the mixed strategy. Assume that the wife and the husband assign probabilities $p_1$ and $q_1$ to choosing "opera" at the first stage, respectively. Let $p_2'$ and $q_2'$ be the posterior probabilities of the wife and the husband choosing "opera" at the second stage, given that no one has chosen "opera" at the first stage. $p_2'$ and $q_2'$ have three possible values since there are three Nash equilibria in the one shot game (two pure strategy and one mixed strategy). Here, we take the mixed strategy and neglect the other two. The situation that both pure strategy and mixed strategy are adopted will be discussed later in this section. Thus, $p_2' = 0.6$ and $q_2' = 0.4$. The expected payoffs to the wife and husband $E_w$ and $E_h$ can be computed as

$$\begin{aligned}E_w &= 3p_1q_1 + 2(1-p_1)(1-q_1) + 2(1-(1-p_1)(1-q_1)) \\ &\quad + (1-p_1)(1-q_1)(3p_2'q_2' + 2(1-p_2')(1-q_2')) \\ E_h &= 2p_1q_1 + 3(1-p_1)(1-q_1) + 3(1-(1-p_1)(1-q_1)) \\ &\quad + (1-p_1)(1-q_1)(2p_2'q_2' + 3(1-p_2')(1-q_2')).\end{aligned}$$

These simplify to

$$\begin{aligned}E_w &= \frac{21}{5}p_1q_1 - \frac{6}{5}p_1 - \frac{6}{5}q_1 + \frac{16}{5} \\ E_h &= \frac{16}{5}p_1q_1 - \frac{6}{5}p_1 - \frac{6}{5}q_1 + \frac{21}{5}.\end{aligned}$$

There is a mixed strategy Nash equilibrium if $p_1 = 3/8$ and $q_1 = 2/7$. The expected payoffs are $P_w = 20/7$ and $P_h = 15/4$.

The above equilibrium is only one of the mixed strategy Nash equilibria, and there are several others in this game. For example, if we set $p_2' = q_2' = 0$, we have another mixed strategy Nash equilibrium $p_1 = 3/5$, $q_1 = 2/5$, $P_w = 16/5$, and $P_h = 21/5$. Among all Nash equilibria, the equilibrium with $p_1 = 3/8$ and $q_1 = 2/7$ is the least efficient because the players receive the lowest expected payoffs.

Compared with the situation where there is no strategy constraint (the expected payoff of each player is 2.4), it can be seen that both players are better off in the game with the strategy constraint. The strategy constraint helps the game converge to a Nash equilibrium. In the games with multiple Nash equilibria, the players could be confused about which Nash equilibrium to choose. Without communication between players, miscoordination is inevitable, even in repeated interactions. The strategy constraint provides extra information for the players in the process of game playing, so that they know more about the strategies of others. This is the reason why it helps the game converge to pure strategy Nash equilibrium in the coordination game.

## V. Conclusion

We have introduced a typical misuse of backward induction in repeated games by analyzing the surprise exam paradox. In the surprise exam paradox, the strategies of the teacher and the students in different stages are coupled so that backward induction cannot be applied in its normal use. The surprise exam paradox represents a special type of repeated game: a repeated game with strategy constraints. When the players' choices in a later stage are conditional and they depend on the choices in previous stages, backward induction should be used together with Bayes' theorem in order not to confuse *a priori* probabilities with the corresponding conditional probabilities.

In some repeated games, the players involved in strategy constraints are not necessarily worse off, compared to the situation without strategy constraints. In the repeated battle of the sexes game, for example, both players could be better off under a budget constraint. This result reminds us that the players in some repeated games may actively exert a constraint on their strategies if they can be better off by doing so. If these behaviors are possible, then we should be very careful in using backward induction to solve finitely repeated games. The tit-for-tat strategy in the iterated prisoner's dilemma could be an example of this [4], [24]. By adopting tit-for-tat, the player links his/her strategy to the strategy of the other player. If the other side realizes that they are better off choosing to cooperate, then mutual cooperation could be established. When the players adopt strategies like tit-for-tat in repeated games, backward induction cannot be applied in the normal way. This will be the direction of our future research.

## REFERENCES


[1] E. Altman and E. Solan, "Constrained games: The impact of the attitude to adversary's constraints," *IEEE Trans. Autom. Control*, vol. 54, no. 10, pp. 2435–2440, Oct. 2009.

[2] M. Assem, D. Dolder, and R. Thaler, "Split or steal? Cooperative behavior when the stakes are large," *Manage. Sci.*, vol. 58, pp. 2–20, 2012.

[3] R. Aumann, "Rationality and bounded rationality," *Games Econ. Behav.*, vol. 21, pp. 2–14, 1997.

[4] R. Axelrod*, The Evolution of Cooperation*. New York, NY, USA: Basic Books, 1984.

[5] H. Bal and V. Allis, "Parallel retrograde analysis on a distributed system," in *Proc. IEEE/ACM Supercomput. Conf.*, 1995, pp. 73–73.

[6] T. Basar and R. Srikant, "A Stackelberg network game with a large number of followers," *J. Optim. Theory Appl.*, vol. 115, no. 3, pp. 479–490, 2002.

[7] K. Basu, "The traveler's dilemma: Paradoxes of rationality in game theory," *Amer. Econ. Rev.*, vol. 84, no. 2, pp. 391–395, 1994.

[8] K. Binmore, "Rationality and backward induction," *J. Econ. Methodol.*, vol. 4, no. 1, pp. 23–41, 1997.

[9] G. Bolton and A. Ockenfels, "ERC: A theory of equity, reciprocity and competition," *Amer. Econ. Rev.*, vol. 90, pp. 166–193, 2000.

[10] D. Borwein, J. Borwein, and P. Marchal, "Surprise maximization," *Amer. Math. Monthly*, vol. 107, no. 6, pp. 517–527, 2000.

[11] J. Bosch, "The examination paradox and formal prediction," *Logique et Analyse*, vol. 15, pp. 505–525, 1972.

[12] B. Bouzy, "Go patterns generated by retrograde analysis," *Evaluation*, vol. 1, no. 3, p. 9, 2001.

[13] M. Capra, J. Goeree, R. Gomez, and C. Holt, "Anomalous behavior in a travelers dilemma?," *Amer. Econ. Rev.*, vol. 89, no. 3, pp. 678–690, 1999.

[14] T. Chow, "The surprise examination or unexpected hanging paradox," *Amer. Math. Monthly*, vol. 105, no. 1, pp. 41–51, 1998.

[15] R. Cooper, D. Jong, R. Forsythe, and T. Ross, "Cooperation without reputation: Experimental evidence from prisoner's dilemma games," *Games Econ. Behav.*, vol. 12, no. 2, pp. 187–218, 1996.

[16] H. Evrard and A. Craig, "Retrograde analysis of the cerebellar projections to the posteroventral part of the ventral lateral thalamic nucleus in the macaque monkey," *J. Comparat. Neurol.*, vol. 508, no. 2, pp. 286–314, 2008.

[17] H. Fang, T. Hsu, and S. Hsu, "Construction of Chinese Chess endgame databases by retrograde analysis," in *Computers and Games*, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer-Verlag, 2001, vol. 2063, pp. 96–114.

[18] E. Fehr and K. Schmidt, "A theory of fairness, competition and cooperation," *Quart. J. Econ.*, vol. 114, no. 3, pp. 817–868, 1999.

[19] J. Ferreira and J. Bonilla, "The surprise exam paradox, rationality, and pragmatics: A simple game theoretic analysis," *J. Econ. Methodol.*, vol. 15, pp. 285–299, 2008.

[20] J. Gerbrandy, "The surprise examination in dynamic epistemic logic," *Synthese*, vol. 155, no. 1, pp. 21–33, 2007.

[21] I. Gilboa and D. Schmeidler, "Information dependent games: Can common sense be common knowledge?," *Econ. Lett.*, vol. 27, pp. 215–221, 1988.

[22] E. Johnson, F. Colin, S. Camerer, and T. Talia, "Detecting failures of backward induction: Monitoring information search in sequential bargaining," *J. Econ. Theory*, vol. 104, no. 1, pp. 16–47, 2002.

[23] L. Kahn and J. Murnighan, "Conjecture, uncertainty, and cooperation in prisoners' dilemma games: Some experimental evidence," *J. Econ. Behav. Organisms*, vol. 22, pp. 91–117, 1993.

[24] G. Kendall, X. Yao, and S. Chong, "The iterated prisoners' dilemma 20 years on," in *Advances in Natural Computation 4*. Singapore: World Scientific, 2007.

[25] D. Kreps, P. Milgrom, J. Robert, and R. Wilson, "Rational cooperation in the prisoner's dilemma," *J. Econ. Theory*, vol. 27, no. 2, pp. 245–252, 1982.

[26] D. Kreps and R. Wilson, "Reputation and imperfect information," *J. Econ. Theory*, vol. 27, no. 2, pp. 253–279, 1982.

[27] D. Lehmann, "Knowledge, common knowledge and related puzzles (extended summary)," in *Proc. 3rd Annu. ACM Symp. Principles Distrib. Comput.*, 1984, pp. 62–67.

[28] R. Luce and H. Raiffa*, Games and Decisions: An Introduction and Critical Survey*. New York, NY, USA: Wiley, 1957, 1957.

[29] J. Oechssler, "Finitely repeated games with social preferences," *Exp. Econ.*, vol. 16, no. 2, pp. 222–231, Jun. 2013.

[30] Y. Pan and L. Pavel, "Global convergence of an iterative gradient algorithm for the Nash equilibrium in an extended OSNR game," in *Proc. IEEE Int. Conf. Comput. Commun.*, 2007, pp. 206–212.

[31] P. Reny, "Common knowledge and games with perfect information," in *Proc. Philosophy Sci. Assoc.*, 1989, pp. 363–369.

[32] J. Romein and H. Bal, "Solving Awari with parallel retrograde analysis," *Computer*, vol. 36, no. 10, pp. 26–33, 2003.

[33] R. Shaw, "The paradox of the unexpected examination," *Mind*, vol. 67, pp. 382–384, 1958.

[34] R. Smullyan*, Forever Undecided: A Puzzle Guide to Göedel*. New York, NY, USA: Knopf, 1987.

[35] E. Sober, "To give a surprise exam, use game theory," *Synthese*, vol. 115, no. 3, pp. 355–373, 1998.

[36] R. Sorensen, "Conditional blindspots and the knowledge squeeze: A solution to the prediction paradox," *Australasian J. Philosophy*, vol. 62, pp. 126–135, 1984.

[37] K. Thompson, "Retrograde analysis of certain endgames," *Int. Comput. Chess Assoc. J.*, vol. 9, no. 3, pp. 131–139, 1986.

[38] J. von Neumann and O. Morgenstern*, Theory of Games and Economic Behaviour*. Princeton, NJ, USA: Princeton Univ. Press, 1944.

[39] Q. Zhu, "A Lagrangian approach to constrained potential games: Theory and examples," in *Proc. IEEE Conf. Decision Control*, 2008, pp. 2420–2425.

[40] Q. Zhu, H. Tembine, and T. Basar, "A constrained evolutionary Gaussian multiple access channel game," in *Proc. Int. Conf. Game Theory Netw.*, 2009, pp. 403–410.


**Jiawei Li** (M'12) received the B.Sc. degree in ship engineering and the Ph.D. degree in engineering from the Harbin Engineering University, Harbin, China, in 1992 and 1998, respectively.

He is currently a Research Fellow at the School of Computer Science, University of Nottingham, Nottingham, U.K. His research interests include heuristic, adaptive learning, evolutionary game theory, and its applications.

**Graham Kendall** (M'03–SM'10) received the B.S. degree in computation (first class honors) from the University of Manchester Institute of Science and Technology, Manchester, U.K., in 1997 and the Ph.D. degree in computer science from the University of Nottingham, Nottingham, U.K., in 2001.

His previous experience includes almost 20 years in the information technology industry where he held both technical and managerial positions. He is the Dunford Professor of Computer Science at the University of Nottingham and is currently based at their Malaysia Campus where he holds the position of Vice-Provost (Research and Knowledge Transfer). He is a Director of three companies (EventMAP Ltd., Nottingham, U.K.; Aptia Solutions Ltd., Nottingham, U.K., and Nottingham MyRIAD Solutions Sdn Bhd, Malaysia) and is CEO of two companies. During his career, he has edited or authored ten books and published over 150 refereed papers. He has been awarded externally funded grants worth over \$6 million from a variety of sources including the Engineering and Physical Sciences Research Council (EPSRC) and commercial organizations. His research interests include adaptive learning (with an emphasis on evolving games), heuristic development (particularly hyperheuristics), optimization, scheduling (particularly sports), and artificial intelligence.

Prof. Kendall has guest edited special issues of journals including *Annals of Operations Research*, the *Journal of Scheduling*, the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, and the *International Computer Games Association Journal*. Prof. Kendall is a Member of the Automated Scheduling, Optimization, and Planning Research Group. He is a Fellow of the Operational Research Society. He is an Associate Editor of seven international journals, including two IEEE journals: the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION and the IEEE TRANSACTIONS ON COMPUTATIONAL INTELLIGENCE AND ARTIFICIAL INTELLIGENCE IN GAMES. He chaired the Multidisciplinary International Conference on Scheduling: Theory and Applications in 2003, 2005, 2007, 2009, and 2011, and has chaired several other international conferences, which included establishing the IEEE Symposium on Computational Intelligence and Games. He chaired this symposium in 2005 and 2006. He has been a Member of the program (or refereeing) committees of over 160 international conferences over the last few years.

**Athanasios V. Vasilakos** (M'00–SM'11) received the B.S. degree in electrical and computer engineering from the University of Thrace, Xanthi, Greece, in 1983, the M.S. degree in computer engineering from the University of Massachusetts, Amherst, MA, USA, in 1986, and the Ph.D. degree in computer engineering from the University of Patras, Patras, Greece, in 1988.

He is currently Professor at the Computer Science Department, Kuwait University, Safat, Kuwait. He has authored or coauthored over 250 technical papers in major international journals and conferences. He is the author/coauthor of five books and 20 book chapters in the areas of communications.

Prof. Vasilakos has served as General Chair, Technical Program Committee Chair for many international conferences. He served or is serving as an Editor or/and Guest Editor for many technical journals, such as the IEEE TRANSACTIONS ON NETWORK AND SERVICES MANAGEMENT, the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, the IEEE TRANSACTIONS ON COMPUTERS, *ACM Transactions on Autonomous and Adaptive Systems*, the IEEE COMMUNICATIONS MAGAZINE, *ACM/Springer Wireless Networks (WINET)*, and *ACM/Springer Mobile Networks and Applications (MONET)*. He is founding Editor-in-Chief of the *International Journal of Adaptive and Autonomous Communications Systems (IJAACS)* and the *International Journal of Arts and Technology (IJART)*. He is General Chair of the European Alliances for Innovation (funded by the European Union Parliament).