Chapter 2.3

Application and Value of Deception

The two previous chapters stressed those aspects of deception that relate to human cognition. However, one can ask whether deception is a purely human phenomenon and question if deception can arise in the solution to mathematical games played by strictly rational players that are not affected by human emotions. In this Chapter, we argue that the answers to these questions are no and yes, respectively. We also illustrate how the framework of non-cooperative games provides answers to the questions of how to use deception optimally and how to guard against it.

In mathematical game theory there are several alternative notions for what is meant by the "optimal policy for a rational player." Some of the most common are as follows:

1. The *Stackelberg leader's policy* is based on the assumption that a player will select a course of actions and advertise it to the other players (followers), which will then select their policies. It implicitly assumes that the leader will not deviate from the advertised policy and that the followers will know this to be true.

2. The *Stackelberg follower's policy* is the best response to an advertised leader's policy.

3. *Security policies*, refer to a situation in which players choose a course of actions that maximizes their reward, assuming that other players are taking the least favorable courses of actions.

4. A *Nash set of policies* is a collection of policies, one for each player, chosen so that if a single player deviates from the assigned policy that player will do worse (or at least no better).

Stackelberg policies are inherently based on trust so they are clearly the wrong framework for deception—the first thought of a human Stackelberg leader would be to advertise the wrong policy. Security policies generally lead to very conservative decisions because players guard themselves against the opponents' worst case actions, even if these actions are also harmful for the opponents[1]. These policies can potentially lead to deception, but generally lead to poor rewards. This leaves us with Nash policies, which as we shall see can indeed exhibit deception and counter-deception.

[0]The author of this chapter would like to thank his former students Yusuf S. Ateşkan and Hüseyin H. Kızılocak for their contribution to this work.

[1]In some zero-sum games, security policies are also Nash policies, in which case these statements do not apply.

For simplicity, in this chapter we restrict our attention to zero-sum two-player games. In this case, Nash policies are also called saddle-point policies and are defined formally as follows: Consider a game between players A and B, and let $J(\alpha, \beta)$ denote the value of a criterion that player A would like to minimize and player B would like to maximize. The criteria depends on the policies α and β utilized by players A and B, respective. A particular pair of policies (α^*, β^*) is called a *saddle-point equilibrium* for the game if

$$J(\alpha^*, \beta) \leq J(\alpha^*, \beta^*) \leq J(\alpha, \beta^*), \qquad \forall \alpha, \beta. \tag{1}$$

The left inequality shows that β^* is the best policy for player B (which is the maximizer) in case player A decides to play α^*. Conversely, the right inequality shows that α^* is the best policy for player A (which is the minimizer) in case player B decides to play β^*. In view of this, the players have no incentive to deviate from their saddle-point policies, justifying the use of the word "equilibrium." This of course assumes that there is no collusion among players (i.e., that we are in a non-cooperative setting). The reader is referred to [1] for the formal definitions of other types of solutions to mathematical games.

Information, Computation, and Deception

Mathematical games are usually classified as either having full or partial information. In *full-information* games both players know the whole state of the game as they make their decisions. By state, we mean all information that is needed to completely describe the future evolution of the game, when the decision rules used by both players are known. Examples of full information games include chess, checkers, and Go. *Partial-information* games differ from these in that at least one of the players does not know the whole state. Poker, bridge, and hearts are examples of such games since each player does not know the hand of the others.

From a formal point of view, full- and partial-information games differ in what types of policies are acceptable. In full-information games, a policy should be understood as a function that selects one particular action for each possible state of the game. In partial-information games, policies are also functions, but now they select one action for each set of (past) observations collected by the player. The definition of saddle-point equilibrium in Eq. 1 is valid for both types of games, provided that we restrict our attention to the appropriate sets of policies.

In full information games, as players plan their next moves, they only need to hypothesize about their own and their opponent's future moves to predict the possible outcomes of the game. This is key to using dynamic programming [2] to solve full-information games. Partial information games are especially challenging because this reasoning may fail. In many partial information games, to predict the possible outcomes of the game, players must hypothesize not only about the future moves of both players, but also about the past moves of their opponent. Note that when the state of the game is known, it is pointless to hypothesize over the opponents past moves, but when the state is not known the past actions of the opponent

are key to constructing an estimate of the current state and of the eventual outcome of the game, based on past observations.

The need to take the past into account when trying to decide on future actions leads to a tremendous increase in complexity. In general, partial information games are poorly understood and the literature is relatively sparse. Notable exceptions are games with lack of information for one of the players [3, 4] and games with particular structures such as the Duel game [5], the Rabbit and Hunter game [6], the Searchlight game [7, 8], etc.

The existence of Nash equilibrium in behavioral (i.e., feedback) policies for finite game was established in [9], but the procedure used to construct the Nash policies suffers from an exponential complexity growth in the size of the game. Chapters 2.4 in this book presents a more efficient construction for saddle-point policies based on dynamic programming, which can be used when a zero-sum finite game has a *nested information structure*, i.e., when one of the players has access to all observations of the opponent (and perhaps some additional observations). This provides a worse case scenario for the player with information inferiority, allowing one to investigate how to avoid potential deceptions. Chapter 3.1 discusses yet another technique to find solutions for finite-games: a tree search. This technique is applicable to deterministic games of sequential moves, i.e., when players alternate in taking actions. Tree searches for partial information games also suffer from an exponential complexity growth, but the approaches discussed in Chapter 3.1 can alleviate this problem to some extent.

Infinite games with partial information are especially challenging. Chapter 3.2 considers one such zero-sum game, subject to linear dynamics and a quadratic cost on the state. The information structure is a variation on the nested case discussed above, in which one player has perfect access to the state of the game, whereas the observations of the opponent are partial and noisy. This scenario is also useful to study how the player with information inferiority can avoid potential deceptions.

Partial information games are particularly interesting because a player can obtain future rewards by either one of the two following mechanisms:

1. Choosing an action that will take the game to a more favorable state.

2. Choosing an action that will make the other player believe that the game is in a state other than the actual one. Moreover, this should lead the opponent to act in our own advantage.

The latter corresponds to a deception move that is only possible in partial information games because it relies on the fact that the perception of one player can be manipulated by the other one. Later in this chapter we will see that this type of moves can occur in saddle-point solutions to mathematical games.

Most of this chapter is focused on deception due to partial information. However, deception can also arise due to *limited computation* and is not restricted to partial-information games. To understand how, consider a deterministic game of full information and suppose

that player A has sufficient computational power to explore all possible game outcomes for all possible choices of the players' actions. However, assume that player B can only evaluate the possible evolution of the game for N steps into the future. In this scenario, the player with computational superiority (player A) can select actions that will prompt a response by player B that appears good when the preview horizon is limited to N steps, but that will prove damaging in the long run. This type of move also reveals deception, which is now caused by an asymmetry in the computational power available to the players. This asymmetry was widely used by humans players against early chess-playing computers, and gave the former a significant advantage until raw computing power and advanced board evaluation metrics tilted the balance in favor of computers.

Deception due to limited computation can reveal itself in numerous ways. In the scenario described above, we assumed that player B can only explore the game's decision tree N steps into the future. If instead the tree exploration is limited in terms of the total number of tree-nodes explored, one can imagine that player B may decide to explore a few tree branches more exhaustively than others, based on some heuristics. Player A can take advantage of this by creating "diversions" that bias player B's search away from the tree branch that player A ultimately plans to follow. This type of deception can be recognized in humans playing board games, which sometimes try to draw the attention of their opponents away from the areas in which they plan to make decisive attacks. In the context of games played by humans one should understand "limited computation" to mean "limited attention-span."

In military operations, deception due to limited computation (or more precisely, limited attention-span) is often combined with partial information to maximize its efficiency. A notable historical event was operation Overlord during the Second World War, which culminated with the D-day invasion of France in June 1944 by the allied forces. As discussed in Chapter 2.1, multiple diversions were used to attract the attention of the German command to the Pas de Calais area, making it believe that the Normandy incursion was a diversionary landing and that the true sea-borne assault was going to take place in Pas de Calais several days later. This led Hitler to delay sending several of his divisions to fight the allies in the Normandy coast, a move that eventually came too late for the Germans. This example reveals the tremendous benefits (and risks) that can arise when the effect of limited attention-span is amplified by limited access to information. The use of deception in the context of military operations has been documented and studied by several authors [10, 11, 12, 13, 14, 15]. For example, it is also widely recognized that, by the end of the cold war, the trend in Soviet naval electronic warfare was changing toward an independent type of combat action instead of a purely support role. This new role emphasized radio deception and misinformation at all levels of command [11]. The detection of false targets or decoys is now an important area of research in radar systems [12, 14].

The potential use of deception has been recognized in several other areas, such as price negotiation [16, 17], multi-object auctioning [18], pursuit-evasion [10, 19], human relations [20], and card games [21]. In [16, 17], the authors analyze a negotiation where the players do not know each other's payoffs, but receive estimates from their opponents. To increase their gain, each player may bias the estimate given. In [17], an advising scheme is proposed

to make deception mostly useless. In [18], it is analyzed how a bidder can use deception to lower the price of an item sold in a multi-object auction. A pursuit-evasion game is analyzed in [10], where the evader corrupts the information available to the opponent to gain an advantage. The authors assume that the evader can jam the pursuer's sensor and therefore induce measurement errors, produce false targets, or interrupt the observations. The use of deception in abstract games was also reported in [22, 23, 24]. In these papers the authors analyze games in which one of the players is forced to make a decision based on information that has been tampered with by the opponent.

Games of Deception

In the beginning of this chapter we claimed that deception can naturally arise in saddle-point solutions to mathematical games. We will now use a series of games to demonstrate this. In the process, we will also see how "rational" players deal with deception. The games considered are inspired by military air-strike operations. These games are sufficiently simple so that they can be solved in closed form, but they are still sufficiently rich to exhibit deception and counter-deception policies. These games were first introduced in [25]. We refer the reader to the technical report [26] for all the relevant mathematical derivations.

[Figure 1 about here.]

A Prototype Non-Cooperative Game

Consider the air-strike game represented schematically in Figure 1. The *attacker* must choose one of two possible targets (A or B) and the *defender* must decide how to better defend them. We assume that the defender has a finite number of assets available that can be used to protect the targets. To make these defense assets effective, they must be assigned to a particular target and the defender must choose how to distribute them among the targets. To raise the stakes, we assume that the defender only has three defense units and is faced with the decision of how to distribute them among the two targets. We start by assuming that both players make their decisions independently and execute them without knowing the choice of the other player. Although it is convenient to regard the players as "attacker" and "defender," this type of games also arise in non-military applications. For example, the "attacker" could be trying to penetrate a market that the "defender" currently dominates.

The game described above can be played as a zero-sum game with the following criterion, which the attacker tries to minimize and the defender tries to maximize:

$$J := \begin{cases} c_0 & \text{no units defending the target attacked} \\ c_1 & \text{one unit defending the target attacked} \\ c_2 & \text{two units defending the target attacked} \\ c_3 & \text{three units defending the target attacked.} \end{cases}$$

This cost is shown schematically in Figure 2.

[Figure 2 about here.]

Without loss of generality we can normalize the constants c_i to have $c_0 = 0$ and $c_3 = 1$. The values for the constants c_1 and c_2 are domain specific, subject to the reasonable constraint that $0 < c_1 \leq c_2 < 1$. Implicit in the above cost is the assumption that both targets have the same strategic value. We only make this assumption for simplicity of presentation.

As formulated above, the attacker has two possible choices (attack target A or attack target B) and the defender has a total of four possible ways of distributing the defensive units among the two targets. Each choice available to a player is called a *pure policy* for that player. We will denote the pure policies for the attacker by α_i, $i \in \{1, 2\}$, and the pure policies for the defender by δ_j, $j \in \{1, 2, 3, 4\}$. These policies are enumerated in Tables 1(a) and 1(b), respectively. In Table 1(b), each "$\circ$" represents one defensive unit. The defender polices δ_1 and δ_2 will be called *3-0 configurations* because they correspond to situations in which the defender assigns all the units to a single target. The policies δ_3 and δ_4 will be called *2-1 configurations* and correspond to situations in which 2 units are assigned to one target and one unit to the other.

[Table 1 about here.]

The game under consideration can be represented in its *extensive form* by associating each policy of the attacker and the defender with a row and column, respectively, of a 2×4 matrix G. The entry g_{ij}, $i \in \{1, 2\}$, $j \in \{1, 2, 3, 4\}$ of G corresponds to the cost J when the attacker chooses policy α_i and the defender chooses policy δ_j. The matrix G for this game is given by

$$G := \begin{array}{c} \begin{matrix} \delta_1 & \delta_2 & \delta_3 & \delta_4 \end{matrix} \\ \begin{bmatrix} 1 & 0 & c_2 & c_1 \\ 0 & 1 & c_1 & c_2 \end{bmatrix} \end{array} \begin{matrix} \\ \alpha_1 \\ \alpha_2 \end{matrix} \tag{2}$$

We are interested in saddle-point policies for the game. A *saddle-point equilibrium* in pure policies is a pair of policies $\{\alpha_{i^*}, \delta_{j^*}\}$, one for each player, for which

$$g_{i^*j} \leq g_{i^*j^*} \leq g_{ij^*}, \qquad \forall i, j.$$

As mentioned before, saddle-point policies are chosen by rational players since they guarantee a cost no worst than $g_{i^*j^*}$ for each player, *no matter what the other player decides to do.* As a consequence, playing at a saddle-point is "safe" even if the opponent discovers our policy of choice. They are also reasonable choices since a player will never do better by unilaterally deviating from the equilibrium. However, there are no saddle-points in pure policies for the game described by Eq. 2. In fact, all the pure policies violate the "safety" condition mentioned above. Suppose, for example, that the attacker plays policy α_1 (i.e., it always attacks target A). This choice is not safe in the sense that, if the defender guesses it, this player can choose the policy δ_1 (i.e., it assigns all units to target A), which subjects the attacker to the highest possible cost. Similarly, α_2 is not safe and therefore cannot also be in a saddle-point policy.

To obtain a saddle-point, one needs to enlarge the policy space by allowing each player to randomize among the available pure policies. In particular, suppose that the attacker chooses policy α_i, $i \in \{1, 2\}$ with probability a_i and the defender chooses policy δ_j, $j \in \{1, 2, 3, 4\}$ with probability d_j. When the game is played repeatedly, the expected value of the cost is then given by

$$\mathrm{E}[J] = \sum_{i,j} a_i g_{ij} d_j = a' G d.$$

Each 2-vector $a := \{a_i\}$ of probabilities is called a *mixed policy for the attacker*, whereas each 4-vector $d := \{d_j\}$ of probabilities is called a *mixed policy for the defender*. It is well know that at least one saddle-point equilibrium in mixed policies always exists for finite matrix games (cf. Minimax Theorem [1, p. 27]). In particular, there always exists a pair of mixed policies $\{a^*, d^*\}$ for which

$$a^{*\prime} G d \leq a^{*\prime} G d^* \leq a' G d^*, \qquad \forall a, d.$$

Assuming that both players play at the saddle-point the cost will then be equal to $a^{*\prime} G d^*$, which is called the *value of the game*. For this game, the unique saddle-point equilibrium for the matrix G in Eq. 2 is given by

$$a^* := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}', \tag{3}$$

$$d^* := \begin{cases} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \end{bmatrix}' & c_1 + c_2 \leq 1 \\ \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}' & c_1 + c_2 > 1 \end{cases} \tag{4}$$

with value equal to

$$a^{*\prime} G d^* = \max\left\{ \frac{c_1 + c_2}{2}, \frac{1}{2} \right\}.$$

This equilibrium corresponds to the intuitive solution that the attacker should randomize between attacking targets A or B with equal probability [i.e., randomize between α_1 and α_2, according to Eq. 3], and the defender should randomize between placing most of the defensive units next to A or next to B also with equal probability (according to Eq. 4). The optimal choice between 3-0 or 2-1 configurations (policies δ_1/δ_2 versus δ_3/δ_4) depends on the parameters c_1 and c_2. From Eq. 4 we conclude that 3-0 configurations are optimal when $c_1 + c_2 \leq 1$, otherwise the 2-1 configurations are preferable. Figure 3 shows the saddle-point policies when $c_1 + c_2 \leq 1$.

[Figure 3 about here.]

In the game described so far, deception is not possible since the players are forced to make a decision without any information. We will change that next.

Full Manipulation of Information

Suppose now that the game described above is played in two steps. First the defender decides how to distribute the defensive units. The defender may also disclose the position of some of the units to the attacker. On the second step the attacker decides which target to strike. To do this, the attacker may use the information provided by the defender. For now, we assume that this is the only information available to the attacker and therefore the defender completely controls the information that the attacker uses to make a decision. This game is represented schematically in Figure 4.

[Figure 4 about here.]

The rationale for the defender to voluntarily disclose the position of the defensive units is to deceive the attacker. Suppose, for example that the attacker uses two units to defend target A and only one to defend B (as in Figure 4). By disclosing that units were placed next to B, the defender may expect the opponent to attack A and, consequently, suffer a heavier cost.

In this new game, the number of admissible pure policies for each player is larger than before. The attacker now has 8 distinct pure policies available because, for each possible observation (no unit detected, unit detected defending target A, or unit detected defending target B), one has two possible choices (strike A or B). These policies are enumerated in Table 2(a).

[Table 2 about here.]

In policies α_1 and α_2 the attacker ignores any available information and always attacks target A or target B, respectively. These policies are therefore called *blind*. In policies α_3 and α_4, the attacker never selects the target where a defensive unit was detected. These policies are called *naive* because they reflect the belief that the target where defense units are visible is better defended. In policies α_5 and α_6, the attacker chooses the target where a defensive unit was detected. These policies are called *counter-deception* since they presume that a unit is being shown close to the least defended target. These policies are represented schematically in Figure 5(a).

The defender has ten distinct pure policies available, each one corresponding to a particular configuration of the defensive units and a particular choice of which units to disclose (if any). These are enumerated in Table 2(b), where "•" represents a defense unit whose position has been disclosed and "∘" a defense unit whose position has not been disclosed. Here, we are assuming that the defender will, at most, disclose the placement of one unit because more than that would never be advantageous. In policies δ_1 through δ_4 nothing is disclosed about the distribution of the units. These are called *no-information* policies. In policies δ_9 and δ_{10} the defender shows units placed next to the target that has fewer defenses. These are *deception* policies. Policies δ_5 through δ_8 are *disclosure* policies, in which the defender is showing a unit next to the target that is better defended. These policies are represented schematically in Figure 5(b).

[Figure 5 about here.]

This game can be represented in extensive form by the following 8×10 matrix

$$G := \begin{array}{c} \begin{matrix} \delta_1 & \delta_2 & \delta_3 & \delta_4 & \delta_5 & \delta_6 & \delta_7 & \delta_8 & \delta_9 & \delta_{10} \end{matrix} \\ \begin{bmatrix} 1 & 0 & c_2 & c_1 & 1 & 0 & c_2 & c_1 & c_2 & c_1 \\ 0 & 1 & c_1 & c_2 & 0 & 1 & c_1 & c_2 & c_1 & c_2 \\ 0 & 1 & c_1 & c_2 & 0 & 0 & c_1 & c_1 & c_2 & c_2 \\ 1 & 0 & c_2 & c_1 & 0 & 0 & c_1 & c_1 & c_2 & c_2 \\ 1 & 0 & c_2 & c_1 & 1 & 1 & c_2 & c_2 & c_1 & c_1 \\ 0 & 1 & c_1 & c_2 & 1 & 1 & c_2 & c_2 & c_1 & c_1 \\ 1 & 0 & c_2 & c_1 & 0 & 1 & c_1 & c_2 & c_1 & c_2 \\ 0 & 1 & c_1 & c_2 & 1 & 0 & c_2 & c_1 & c_2 & c_1 \end{bmatrix} \end{array} \begin{matrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \end{matrix} \tag{5}$$

Just as before, for this game to have saddle-point equilibria one needs to consider mixed policies. However, this particular game has multiple equilibria, one of them being

$$a^* := \left[\tfrac{1}{2} \quad \tfrac{1}{2} \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0\right]',$$
$$d^* := \begin{cases} \left[\tfrac{1}{2} \ \tfrac{1}{2} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0\right]' & c_1 + c_2 \leq 1 \\ \left[0 \ 0 \ \tfrac{1}{2} \ \tfrac{1}{2} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0\right]' & c_1 + c_2 > 1 \end{cases}$$

with value equal to

$$a^{*\prime} G d^* = \max\left\{\frac{c_1 + c_2}{2}, \frac{1}{2}\right\}.$$

This shows that

1. the attacker can ignore the information available and simply randomize among the two blind policies α_1 and α_2 with equal probability; and

2. the defender gains nothing from disclosing information and can therefore randomize among the no-information policies, also with equal probability.

It should be noted that there are saddle-point equilibria that utilize different policies. For example, when $c_1 + c_2 > 1$, an alternative Nash equilibrium is

$$\bar{a}^* := \left[0 \quad 0 \quad \tfrac{1}{4} \quad \tfrac{1}{4} \quad \tfrac{1}{4} \quad \tfrac{1}{4} \quad 0 \quad 0\right]', \tag{6}$$
$$\bar{d}^* := \left[0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad \tfrac{1}{4} \quad \tfrac{1}{4} \quad \tfrac{1}{4} \quad \tfrac{1}{4}\right]'. \tag{7}$$

In this case, the defender randomizes between deception and disclosure policies with equal probability and the attacker between the naive and counter-deception policies. However, in zero-sum games all equilibria yield the same value, so the players have no incentive to choose this equilibrium that is more complex in terms of the decision rules. Finally, it should also

be noted that, because of the equilibrium interchangeability property for zero-sum games, the pairs $\{a^*, \bar{d}^*\}$ and $\{\bar{a}^*, d^*\}$ are also saddle-point equilibria [1, p. 28]. This means that one still gets an equilibrium, e.g., if the defender can play $\bar{d}^*$ (randomization between deception and disclosure) and the attacker plays a^* (randomization between blind policies).

We have just seen that, in this version of the game, the attacker gains nothing from using the observations available, even though these observations give precise information about the position of some of the defense units. At an intuitive level, this is because the information available to the attacker is completely controlled by the opponent. If the defender chooses to disclose the position of some of the defensive units, this is done solely to get an advantage. This can be seen, for example, in the equilibrium given by Eqs. 6–7. We shall consider next a version of the game where the defender no longer has complete control over the information available to the attacker. For the new game, the attacker may sometimes improve the cost by using the available information.

Partial Manipulation of Information

Suppose that when the defender decides to "show" one of the defensive units, the unit is simply not camouflaged, making it easy to find by the surveillance sensors used by the attacker. In the previous game, we assumed that not camouflaged units are always detected by the attacker and camouflaged ones are not. We will now deviate from this ideal situation and assume that (i) not camouflaged units may not be detected and, more importantly, (ii) camouflaged units may sometimes be detected by the attacker. We consider a generic probabilistic model for the attacker's surveillance, which is characterized by the conditional probability of detecting units next to a particular target, given a specific total number of units next to that target and how many of them are being shown:

$$\text{Prob}\left(\text{defenses detected near target A} \mid \mathbf{n}_A = n,\ \mathbf{s}_A = s\right) = \chi(n, s),$$

where $\chi(\cdot)$ is the *sensor characteristic function*, $\mathbf{n}_A \in \{0, 1, 2, 3\}$ is the total number of units defending target A, and $\mathbf{s}_A$ is the number of these that are shown. Since there is no incentive for the defender to show more than one unit, $\mathbf{s}_A$ is restricted to the set $\{0, 1\}$. For simplicity, we assume that the surveillance of target B is identical and independent, i.e.,

$$\text{Prob}\left(\text{defenses detected near target B} \mid \mathbf{n}_B = n,\ \mathbf{s}_B = s\right) = \chi(n, s),$$

where the symbols with the B subscript refer to the defense units assigned to target B.

For most of the discussion that follows, the sensor characteristic function $\chi(\cdot, \cdot)$ can be arbitrary, provided that it is monotone non-decreasing with respect to each of its arguments (when the other is held fixed). The monotonicity is quite reasonable since more units (shown or not) should always result in a higher probability of detection. A few particular cases should be considered:

- When $\chi(n, s) = 0$, $\forall\ n, s$ we have the first game considered in this chapter, since the defense units are never detected.

- When

$$\chi(n,s) = \begin{cases} 1, & s > 0 \\ 0, & s = 0 \end{cases} \qquad \forall n \in \{0,1,2,3\}, s \in \{0,1\}$$

the defense units are detected if and only if they are shown. This corresponds to the second game considered here, where the defender has full control over the information available to the attacker.

- Another interesting situation occurs when the probability of detecting two camouflaged units is still larger than the probability of detecting one non-camouflaged unit, i.e.,

$$\chi(2,0) \geq \chi(1,1). \tag{8}$$

When this happens we say that the attacker's sensors are *trustworthy*, because the probability of detecting camouflaged defense units near the better defended target is still larger than the probability of detecting a single non-camouflaged unit by the less defended target (cf. Figure 6).

[Figure 6 about here.]

We will see shortly that when the sensors are trustworthy the attacker should choose naive policies. A special case of trustworthy sensors arises when $\chi(n,s)$ is independent of s for all values of n. In this case we have sensors that cannot be manipulated by the defender since the detection is independent of the number of units "shown" be the defender.

In terms of the policies available, the game considered in this section is fairly similar to the one in the previous section. The only difference is that, in principle, the attacker may now detect defense units next to both targets. In practice, this means that Table 2(a) should have a forth column entitled "units detected at A and B," which would result in 16 distinct pure policies. It turns out that not detecting any unit or detecting units next to both targets is essentially the same. Therefore, we shall consider for this game only the 8 policies in Table 2(a), with the understanding that when units are detected next to both targets, the attacker acts as if no units were detected. It is not hard to show that this introduces no loss of generality. The defender's policies are the same as in the previous section, and are given in Table 2(b).

This game can also be represented in extensive form by an 8×10 matrix. The reader is referred to [26] on how to construct this matrix. As before, the game has saddle-point equilibria in mixed policies, but now the equilibrium policies depend on the sensor characteristic function $\chi(n,s)$. The saddle-point policies for this game were computed in [26] and are as follows:

1. When the sensors are trustworthy [i.e., when $\chi(2,0) \geq \chi(1,1)$], a saddle-point solution is given by

$$a^* := \begin{bmatrix} 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \end{bmatrix}',$$
$$d^* := \begin{cases} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}' & c_1 + c_2 \leq 1 - \chi(3,0) + \eta_1 \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}' & c_1 + c_2 > 1 - \chi(3,0) + \eta_1, \end{cases}$$

where $\eta_1 = (c_2 - c_1)\big(\chi(2,0) - \chi(1,1)\big)$. In this case, the attacker randomizes among the naive policies and the defender either randomizes among the deception policies or the 3-0 no-information configurations. The latter occurs when the attacker only incurs in significant cost when 3 units are in its path and therefore 2-1 configurations are not acceptable for the defender. The value of the game is

$$a^{*\prime} G d^* = \frac{\max\{1 - \chi(3,0), c_1 + c_2 - \eta_1\}}{2} \leq \frac{c_1 + c_2}{2},$$

which is smaller than the one obtained in the previous two games. This game is therefore more favorable to the attacker, which is now able to take advantage of the surveillance information.

2. When the sensors are not trustworthy $\big($i.e., $\chi(2,0) < \chi(1,1)\big)$ and $c_1 + c_2 \geq 1$, a saddle-point solution is given by

$$a^* := \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}',$$
$$d^* := \begin{bmatrix} 0 & 0 & \frac{\eta_2}{2} & \frac{\eta_2}{2} & 0 & 0 & 0 & 0 & \frac{1-\eta_2}{2} & \frac{1-\eta_2}{2} \end{bmatrix}',$$

where $\eta_2 := \frac{\chi(1,1) - \chi(2,0)}{\chi(1,1) - \chi(1,0)}$. In this case, the attacker randomizes among the blind policies and the defender randomizes between deception and no-information in 2-1 configurations. The probability distribution used by the defender is a function of the several parameters. However, the value of the game is always

$$a^{*\prime} G d^* = \frac{c_1 + c_2}{2},$$

which is the same as in the previous two games. This means that the surveillance sensors of the attacker are effectively rendered useless by the defender's policy. This happens because the sensors are not trustworthy and therefore the defender can significantly manipulate the information available to the attacker.

3. When the sensors are not trustworthy $\big($i.e., $\chi(2,0) < \chi(1,1)\big)$, but $c_1 + c_2 < 1$, a saddle-point solution is given by

$$a^* := \begin{bmatrix} \frac{1-\eta_3}{2} & \frac{1-\eta_3}{2} & \frac{\eta_3}{2} & \frac{\eta_3}{2} & 0 & 0 & 0 & 0 \end{bmatrix}',$$
$$d^* := \begin{bmatrix} \frac{1-\eta_4}{2} & \frac{1-\eta_4}{2} 0 & 0 & 0 & 0 & 0 & 0 & \frac{\eta_4}{2} & \frac{\eta_4}{2} \end{bmatrix}',$$

where $\eta_3 := \frac{1-c_1-c_2}{\chi(3,0)-\eta_1}$ and $\eta_4 := \frac{\chi(3,0)}{\chi(3,0)-\eta_1}$. In this case, the attacker randomizes between the blind and naive policies, whereas the defender randomizes between deception and no-information in 3-0 configurations. The value of the game is

$$a^{*\prime} G d^* = \frac{1}{2} - \frac{(1-c_1-c_2)\chi(3,0)}{2(\chi(3,0)-\eta_1)} \leq \frac{1}{2}.$$

which is smaller than the one obtained in the previous two games. Therefore, the attacker can attain a cost smaller than $\frac{1}{2}$, which would be obtained by only using blind policies.

Both in cases 2 and 3 the sensors are not trustworthy and the defender has sufficient power to manipulate the information available to the attacker so as to make it effectively useless. However, in the case 3, the 2-1 configurations required for deception are very costly to the defender and deception is no longer a very attractive alternative. Because of this fact, the defender will avoid it, giving an advantage to the attacker.

The Rational Side of Deception

The games considered above revealed several interesting facts about the rational use of deception in partial information games.

Deception can occur in saddle-point policies. For some parameters of the game with partial manipulation of information, the equilibrium policy for the defender attempts to make the attacker believe that the least defended target was the best defended one. This is achieved by disclosing the position of defense units near a poorly defended target and hiding the presence of other units near a better defended one. This confirms our original claim that deception can arise naturally in saddle-point solutions to mathematical games played by strictly rational players. We have thus shown that deception is not limited to the phenomena arising out of human cognitive limitations.

Deception provides a mechanism to remove information content from the observations. In the games with full and partial manipulation of information, the attacker had access to sensor data, which was not available in the initial game. Yet, this player was not able to fully utilize this data. In fact, in the game with full manipulation of information — as well as with partial manipulation of information, $c_1 + c_2 \geq 1$, and non-trustworthy sensors — the use of deception by the defender essentially forced the attacker to completely ignore the available observations. In both cases, the attackers could not improve their chances of success by using observations. This result is perhaps not surprising with full manipulation of information, in which case the defender completely controlled the information available to the attacker. However, even when the sensors allowed the attacker to "see through" the defender's camouflage, blind policies would still be optimal if the sensor's trustworthiness fell below a particular threshold.

The potential rewards collected by deception must be weighted against its cost. When $c_1+c_2 < 1$ in the game with partial manipulation of information, the attacker is biased towards the

naive policies and never uses the counter-deception ones. This means that the observations still provide useful information, even when the sensors can be substantially manipulated by the defender (including the case of non-trustworthy sensors). The explanation for this apparent paradox is that when $c_1 + c_2 < 1$ the use of 2-1 configurations does not provide sufficient defense, even for the target with two defense units. It is therefore costly for the defender (in terms of protecting the targets) to use the 2-1 configuration, which are the only ones for which deception is possible. The defender thus chooses to limit the use of the 2-0 deception policies to the extent that this player no longer removes all information content from the observations available to the attacker, even for non-trustworthy sensors.

At a qualitative level, the opening statements of the three paragraphs above seem intuitive and even "common sense." The role of the analysis presented in this chapter — based on the concept of saddle equilibrium — was to provide the specific policies for the defender that optimize the use of deception and the corresponding policies for the attacker that are most robust against manipulation of information. We believe that this is a key contribution that game theory has to offer to adversarial reasoning.

References

[1] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. No. 23 in Classics in Applied Mathematics, Philadelphia: SIAM, 2nd ed., 1999.

[2] R. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton University Press, 1957.

[3] S. Sorin and S. Zamir, ""Big Match" with lack of information on one side (III)," in Raghavan *et al.* [27], pp. 101–112.

[4] C. Melolidakis, "Stochastic games with lack of information on one side and positive stop probabilities," in Raghavan *et al.* [27], pp. 113–126.

[5] G. Kimeldorf, "Duels: An overview," in *Mathematics of Conflict* (M. Shubik, ed.), pp. 55–72, Amsterdam: North-Holland, 1983.

[6] P. Bernhard, A.-L. Colomb, and G. P. Papavassilopoulos, "Rabbit and hunter game: Two discrete stochastic formulations," *Comput. Math. Applic.*, vol. 13, no. 1–3, pp. 205–225, 1987.

[7] G. J. Olsder and G. P. Papavassilopoulos, "About when to use a searchlight," *J. Mathematical Anal. and Applications*, vol. 136, pp. 466–478, 1988.

[8] G. J. Olsder and G. P. Papavassilopoulos, "A Markov chain game with dynamic information," *J. Opt. Theory and Applications*, vol. 59, pp. 467–486, Dec. 1988.

[9] J. P. Hespanha and M. Prandini, "Nash equilibria in partial-information games on Markov chains," in *Proc. of the 40th Conf. on Decision and Contr.*, Dec. 2001.

[10] Y. Yavin, "Pursuit-evasion differential games with deception or interrupted observation," *Comput. Math. Applic.*, vol. 13, no. 1–3, pp. 191–203, 1987.

[11] M. Vego, "Soviet naval electronic warfare," *Signal*, vol. 44, no. 4, pp. 96–99, 1989.

[12] C. L. C. Oliveira, M. A. Grivet, and E. R. Pantoja, "Radar-ECM simulation system," in *Proc. of the Int. Microwave Conf.*, vol. 1, pp. 255–260, 1993.

[13] M. Tambe, "Recursive agent and agent-group tracking in real-time dynamic environment," in *Proc. of the Int. Conf. on Multi-agent Syst.*, pp. 368–375, 1995.

[14] S. Z. Yanglin, "Identification of false targets in bistatic radar system," in *Proc. of the IEEE National Aerospace and Electronics Conf.*, vol. 2, pp. 878–883, July 1997.

[15] R. W. Burns, "Deception, technology and the D-day invasion," *Engineering Science and Education J.*, vol. 4, pp. 81–88, Apr. 1995.

[16] I. Greenberg, "The effect of deception on optimal decisions," *Operations Research Lett.*, vol. 1, pp. 144–147, Sept. 1982.

[17] S. Matsubara and M. Yokoo, "Negotiations with inaccurate payoff values," in *Proc. of the Int. Conf. on Multi-agent Syst.*, pp. 449–50, 1998.

[18] D. B. Hausch, "Multi-object auctions: Sequential vs. simultaneous sales," *Management Science*, vol. 32, pp. 1599–1610, Dec. 1986.

[19] J. P. Hespanha, M. Prandini, and S. Sastry, "Probabilistic pursuit-evasion games: A one-step Nash approach," in *Proc. of the 39th Conf. on Decision and Contr.*, vol. 3, pp. 2272–2277, Dec. 2000.

[20] C. Reimold, "Games of deception," in *Proc. of the Int. Professional Communication Conf.*, pp. 96–101, 1989.

[21] D. Billings, D. Papp, J. Schaeffer, and D. Szafrone, "Poker as a testbed for machine intelligence research," in *Advances in Artificial Intell.: Proc. of the 12th Biennial Conf. of the Canadian Soc. for Computational Studies*, pp. 228–238, 1998.

[22] J. Spencer, "A deception game," *Amer. Mathematics Monthly*, pp. 416–417, 1973.

[23] V. J. Baston and F. A.Bostock, "Deception games," *Int. J. of Game Theory*, vol. 17, no. 2, pp. 129–134, 1988.

[24] T. Inohara, S. Takahashi, and B. Nakano, "Impossibility of deception in a conflict among subjects with interdependent preference," *Appl. Mathematics and Computation*, vol. 81, no. 2–3, pp. 221–244, 1997.

[25] J. P. Hespanha, Y. S. Ateşkan, and H. H. Kızılocak, "Deception in non-cooperative games with partial information," in *Proc. of the 2nd DARPA-JFACC Symp. on Advances in Enterprise Control*, July 2000.

[26] J. P. Hespanha, Y. S. Ateşkan, and H. H. Kızılocak, "Deception in non-cooperative games with partial information," tech. rep., EE–Systems, University of Southern California, Los Angeles, Feb. 2001. This report supersedes two previous reports with the same title, dated April 2000 and June 2000.

[27] T. E. S. Raghavan, T. S. Ferguson, and T. Parthasarathy, eds., *Stochastic Games and Related Topics: In Honor of Professor L. S. Shapley*, vol. 7 of *Theory and Decision Library, Series C, Game Theory, Mathematical Programming and Operations Research*. Dordrecht: Kluwer Academic Publishers, 1991.

List of Figures

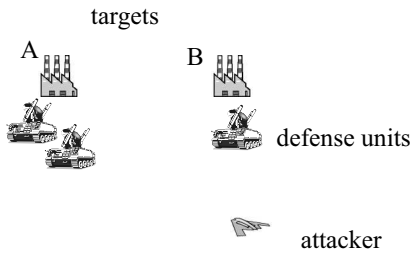


Figure 1: The air-strike game.

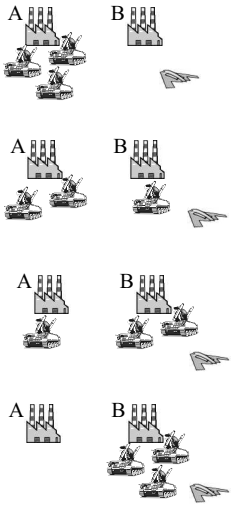


Figure 2: Cost structure for the air-strike game.

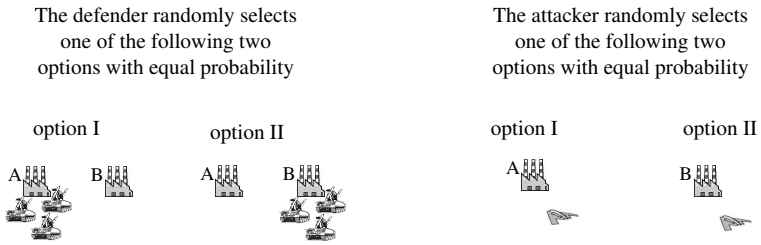


Figure 3: Saddle-point equilibrium in mixed policies for the game defined by Eq. 2 with $c_1 + c_2 \leq 1$. In this case, two units do not provide adequate defense so 3-0 configurations are always chosen.

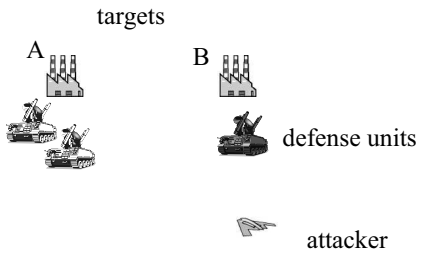
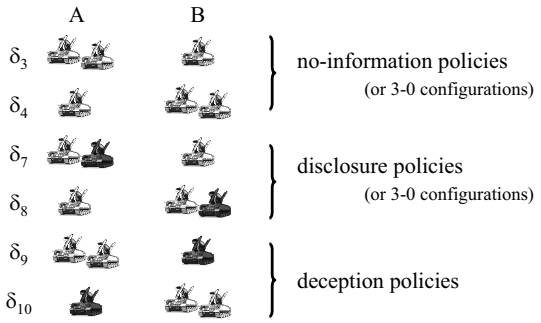
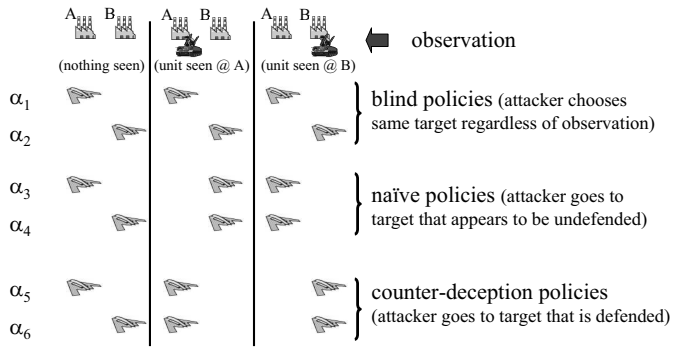


Figure 4: The air-strike game with manipulation of information. The two "transparent" defense units near target A are assumed invisible to the attacker, whereas the "solid" unit near target B can be seen by the attacker.



(a) Attacker policies



(b) Defender policies

Figure 5: Qualitative description of the pure policies in Table 2

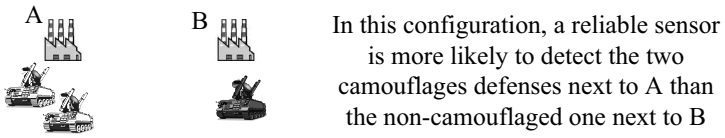


Figure 6: Trustworthy sensors.

List of Tables

policy	target assigned
α_1	A
α_2	B

(a) Attacker policies

policy	target A	target B
δ_1	○ ○ ○	
δ_2		○ ○ ○
δ_3	○○	○
δ_4	○	○○

(b) Defender policies

Table 1: Pure policies

policy	target assigned when ...		
	no obs.	unit detected at A	unit detected at B
α_1	A	A	A
α_2	B	B	B
α_3	B	B	A
α_4	A	B	A
α_5	A	A	B
α_6	B	A	B
α_7	A	B	B
α_8	B	A	A

(a) Attacker policies

policy	target A	target B
δ_1	○ ○ ○	
δ_2		○ ○ ○
δ_3	○○	○
δ_4	○	○○
δ_5	○ ○ ●	
δ_6		○ ○ ●
δ_7	○●	○
δ_8	○	○●
δ_9	○○	●
δ_{10}	●	○○

(b) Defender policies

Table 2: Pure policies for the air-strike game with manipulation of information.