Draft:

# Agents, Games, and Evolution

## An Essay on Constructive Rationality

## Part I: Starters

Steven Orla Kimbrough

Draft: © 27 December 2006

$Id: age-master-partI.tex,v 2.2 2006/12/27 19:54:51 sok Exp $

Professor Steven Orla Kimbrough, University of Pennsylvania, The Wharton School, Philadelphia, PA 19104–6366. Tel: (215) 898-5133 or 610-667-8894. Fax: (215) 898-3664. Email: kimbrough@wharton.upenn.edu. Web:
http://opim-sky.wharton.upenn.edu/~sok/

# Contents

**Preface** **v**

## I Starters 1

**1 Contexts of Strategic Interaction 3**

1.1 Two Kinds of Decisions . . . . . . . . . . . . . . . . . . . . . . . . . . 4
1.2 Categorizing Games . . . . . . . . . . . . . . . . . . . . . . . . . . . 5
1.3 Game Theory and the Theory of Games . . . . . . . . . . . . . . . 9
1.4 Why Study Games? . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
1.5 Methods of Study . . . . . . . . . . . . . . . . . . . . . . . . . . . . 12

**2 Four Themes 17**

2.1 Problems of the Social Order . . . . . . . . . . . . . . . . . . . . . . 18
2.2 Self-Organization and Emergence . . . . . . . . . . . . . . . . . . . 22
2.3 Combinatorial Complexity . . . . . . . . . . . . . . . . . . . . . . . 24
2.4 Concepts of Rationality . . . . . . . . . . . . . . . . . . . . . . . . . 24
2.4.1 Type A Setup and Analysis . . . . . . . . . . . . . . . . . . . 28
2.4.2 Type B Setup and Analysis . . . . . . . . . . . . . . . . . . . 35
2.5 Discussion: Accessibility & Games . . . . . . . . . . . . . . . . . . 37
2.6 Bibliographic Notes . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
2.6.1 Social Order . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
2.6.2 Emergence . . . . . . . . . . . . . . . . . . . . . . . . . . . . 43
2.6.3 Rationality . . . . . . . . . . . . . . . . . . . . . . . . . . . . 44

**3 Games in the Wild 45**

3.1 War . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47
3.2 Trading and Investing . . . . . . . . . . . . . . . . . . . . . . . . . . 48

3.3 Athletic Contests . . . . . . . . . . . . . . . . . . . . . . . . . . 50
3.4 Gambling . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 51
3.5 Business Strategy . . . . . . . . . . . . . . . . . . . . . . . . . . 55
3.6 Negotiation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 57
3.7 Coordination, Symbiosis, Mutualism, Cooperation . . . . . . . . . 58
3.8 Conversation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 61
3.9 Games against Yourself . . . . . . . . . . . . . . . . . . . . . . . 62
3.10 Confidence Games . . . . . . . . . . . . . . . . . . . . . . . . . 63
3.11 Statesmanship . . . . . . . . . . . . . . . . . . . . . . . . . . . 64
3.12 Bibliographic Notes . . . . . . . . . . . . . . . . . . . . . . . . 64

**4 Games in the Abstract** **67**
4.1 Interrogation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 67
4.2 Prisoner's Dilemma . . . . . . . . . . . . . . . . . . . . . . . . . 72
4.3 Hawk-Dove . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 74
4.4 Stag Hunt or Assurance . . . . . . . . . . . . . . . . . . . . . . . 77
4.5 Chicken . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 79
4.6 Battle of the Sexes . . . . . . . . . . . . . . . . . . . . . . . . . 80
4.7 Inspector versus Evader . . . . . . . . . . . . . . . . . . . . . . . 82
4.8 A Zero-Sum Game . . . . . . . . . . . . . . . . . . . . . . . . . 83
4.9 PD Property Games: ##12, 47, 48 & 57 . . . . . . . . . . . . . . 84
4.10 Other Games, Other Forms . . . . . . . . . . . . . . . . . . . . . 87

**A Useful Mathematical Results** **89**
A.1 Miscellaneous Useful Facts . . . . . . . . . . . . . . . . . . . . . 89
A.2 Solving for Mixed Equilibria in 2×2 Games . . . . . . . . . . . . 89
A.3 Replicator Dynamic Equilibrium for 2×2 Games . . . . . . . . . 90
A.4 Pareto Optimal Mixed Equilibrium . . . . . . . . . . . . . . . . . 91

**B Strategic Treatments in Movies and Literature** **93**

**C Resources on the Web** **95**

**References** **99**

**Index** **104**

# Preface

Rational, decision-making agents optimize when they can, and meliorize when they must.

# Part I

# Starters

# Chapter 1

# Contexts of Strategic Interaction

Ideas have lives of their own. They arise at surprising times and come from surprising places. Ideas interact with other ideas. They become associated and these associations—or societies—lead to new social structures and to new ideas. These in turn may calve off in groups and form new societies. Of course, ideas live and have their being in human minds and cultural artifacts. Without us they wouldn't exist. Nor would most of us exist without the benefit of ideas that have created and sustained our civilization. Nor would there be animals without photosynthesizing plants, or photosynthesizing plants without photosynthesizing bacteria. Interdependence pervades.

The subject of this book is a certain society of ideas. I shall call it the AGE society (Agents, Games, and Evolution), without making any claim that the name ideally describes its subject. What name does? As with any interesting and reasonably complex society, AGE is a product, and continues to be a producer, of history. Its structure, its dynamics, even its constituents are often opaque and puzzling, and everything is in flux. Hence the need for study.

I shall not attempt to define the subject at hand. It likely can't be done and wouldn't be of much value even if it could. Hyperprecision would only be a distraction at this point. Rather, I will plunge in, immerse us at the center of AGE society, and explore from where we find ourselves. Better to start somewhere reasonable, then ask questions, attend to relevant problems and data, refine and recombine concepts and hypotheses, and build models and conduct experiments. All this is to be undertaken in an iterative, exploring, probing, nondeterministic search for sharper clarity, deeper understanding, and useful results. This shall be our mode throughout. If the process seems to be a sort of groping, adaptive muddling, so be it. The means are informed by the main results.

Let us begin, then, by discussing *contexts of strategic interaction* (CSIs), also known as *games.* Here is a—perhaps *the*—main theme in our AGE society of ideas.

## 1.1 Two Kinds of Decisions

Decision-making individuals, or *agents,* encounter two types of decisions. In *parametric* decisions, the outcome that an agent receives depends on its choices and states of nature, including states generated by random processes. In *strategic* decisions, the outcome or reward that the agent gets depends in addition on strategic decisions made by other agents. The theory of games and this book are about strategic decision making. A game, in this special sense, may be described as a context of strategic interaction or CSI.

This characterization of the parametric-strategic distinction is perhaps not entirely adequate. If the second and only other agent has decisions that matter to the first agent, but the second agent's decisions are known to the first agent before it decides, then we might best classify the decision context as parametric. And there are other issues. Even so, our characterization provides a sense good enough for present purposes of the difference between parametric decisions and decisions in games. The key thing in a game is that there are at least two players (agents) the values of whose decisions depend on the decisions of the other player(s) and these latter decisions are not known to the deciding players. Examples, rather than minutely-hewed definitions, will serve us best.

Here are some decision situations naturally modeled as parametric:

- A car buyer considers whether to purchase a hybrid automobile. On the negative side, the car is more expensive and less roomier than the non-hybrid under consideration. On the positive side, better mileage will reduce operating expenses and there is a moral benefit from having a smaller carbon footprint.

  What if what other people think of the buyer's decision matters to the buyer? If the buyer can predict their behavior, then it may be best to factor the consequences into the payoffs for the buyer. The decision would remain parametric. On the other hand, if the buyer and her neighbor are purchasing cars this afternoon, are in ignorance of each other's choices, and both gain or lose value, depending on the other's choice, then we have a strategic context.

- A patient is considering whether to have surgery or not. The surgery may or may not be successful, the condition may stabilize or may worsen rapidly without surgery. There are chances all around and the patient's values to take into account. Even though the actions of others matter—the surgeon may or may not do a good job—this is (best thought of as) a parametric decision. The surgeon has no interest in doing a bad job; the preferences and rewards of that agent are not material for the decision to hand.

Here are some other decision situations naturally modeled as strategic:

- A pitcher and a batter face each other in a baseball game. The count is 3 balls and 2 strikes. The pitcher (and catcher) must decide which pitch to throw (fastball? curveball?) and where to throw it (inside? outside? up? down?). The batter has to decide which pitch he thinks the pitcher will throw and prepare himself for that pitch. Unless the batter is "sitting on" a pitch, he has little chance of hitting it well, but can hope to foul it off if he guesses wrong.

- A deer knows several places that have been safe and where the eating may be good. Other deer have similar knowledge. In addition, transiting between eating places is risky. How long should the deer stay in its present patch, which contains other deer and is being depleted?

These should suffice for establishing a clear enough distinction between parametric and strategic contexts. We shall see many examples of games, and hence strategic situations, in what follows.

## 1.2 Categorizing Games

It is useful to distinguish various types of games. The generic $2\times 2$ game, and specifications thereof, will serve present purposes. See Figure 1.1, page 6, for the general, canonical strategic form for two players each having two strategies. In a $2\times 2$ game, there are two players, called Row and Column. The game is played once, and is said to be a *one-shot* game. Each player has two strategies among which it can chose. Row's strategies are labeled $R_1$ and $R_2$ and Column's $C_1$ and $C_2$ in the figure. The payoffs are given in the cells of the figure. If Row Plays $R_1$ and Column plays $C_1$, then Row's payoff is $r_1$ and Column's is $c_1$, and similarly for the other cells. Unless stated otherwise, it is assumed that the game is one-shot,

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | $r_1$, $c_1$ | $r_2$, $c_2$ |
| $R_2$ | $r_3$, $c_3$ | $r_4$, $c_4$ |

Figure 1.1: Canonical game matrix for the $2 \times 2$ game in strategic form

that Row and Column do not communicate with each other, cannot form binding agreements,[1] and in fact are not known to each other. Play is anonymous. Also, we assume that the payoffs are specified as numbers and that uniformly more is better.[2]

*The strategic form of a game*

The game form on display in Figure 1.1 is called the *strategic form* of the game.[3] There are other ways of representing games (such as *extensive form*). When we need them, we shall introduce them.

*Games of pure conflict*

First are *games of pure conflict*. In these games, the interests of the players (agents) are entirely opposed: one's gain is another's loss. One way (there are others) for a game in strategic form to be a game of pure conflict is for the payoffs in each cell to add to a common number (assuming payoffs are different in at least some cells). Figure 1.2 is an example.

*Inessential games*

Second, at the other extreme, in some games the interests of the players are perfectly and trivially aligned. Such games are said to be *no-conflict* games [56, page 22] or, better, *inessential* games. Figure 1.3 shows an example.

Observe that both players prefer that the strategy play $(R_1, C_1)$ obtain. If they both know the game and are fully rational, surely that's what will happen. We will not be much interested in such games.

*Coordination games*

In our third category of game, the interests of the players are largely in agreement, but there is more than one mutually agreeable outcome. These are called *coordination games*. Figure 1.4 shows an example.

[1] Our games, unless otherwise noted, are called *non-cooperative* games, in distinction to *cooperative* games.

[2] Some results in classical game theory assume only an ordering on the payoffs, other results assume the payoffs are given as utility values. This subtlety need not detain us at present.

[3] It is also called the *normal form* of the game, but this expression has fallen out of favor.

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 2, 3 | 4, 1 |
| $R_2$ | 1, 4 | 3, 2 |

Figure 1.2: Game #11-ORDINAL, a constant sum 2×2 game in strategic form [56]

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 4, 4 | 3, 3 |
| $R_2$ | 2, 2 | 1, 1 |

Figure 1.3: Game #1, an inessential 2×2 game in strategic form [56]

*Mixed-motive games*

Finally, our fourth category, which will attract most of our attention, is the category of *mixed-motive* games. In these games, the players' interests are in partial conflict. Figure 1.5, page 8, shows an example. Note its similarity with the coordination game.

The Stag Hunt game is an important mixed-motive game and one we will investigate at length. Figure 1.6, page 8, shows an example.

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 4, 4 | 1, 1 |
| $R_2$ | 1, 1 | 4, 4 |

Figure 1.4: A 2×2 coordination game in strategic form

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 4, 1 | 0, 0 |
| $R_2$ | 0, 0 | 1, 4 |

Figure 1.5: A Battle of the Sexes game in strategic form

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 4, 4 | 1, 3 |
| $R_2$ | 3, 1 | 2, 2 |

Figure 1.6: A Stag Hunt game in strategic form; game #61-ORDINAL in [56]

## 1.3 Game Theory and the Theory of Games

*TSI: theory of strategic interaction or theory of games*

The *theory of games,* or better the *theory of strategic interaction* (TSI), seeks to understand games. It aims to explain what happens, predict what will happen, and support intervention in contexts of strategic interaction. It is, in short, the science of strategic interaction.

*Orthodox game theory*

*Classical,* or *orthodox,* game theory (exemplified by standard texts, such as [4, 26, 38, 39, 44]) is part of the theory of games (TSI). Very roughly, but sufficient for present purposes, orthodox game theory is about investigating the mathematical properties of models of games, under the assumption of ideal (or at least very strong) rationality on the part of the players. This programme of research has been, and remains, remarkably productive, and it has generated indispensable insights into strategic interaction, insights that are invaluable for all other research programmes in the theory of games.

Game theory, as presently and foreseeably constituted, is not a complete theory of games. Its scope is much more limited. This is generally recognized, even among game theorists themselves. See the passage from Binmore's text *Fun and Games* quoted in Figure 1.7, page 10. Binmore is commendably frank and forthcoming, saying in effect that if you want to understand how to play a real-world game, don't look to game theory, instead "consult a psychologist." Why? Because orthodox game theory is about "rational" play under very strong, usually unrealistic, assumptions of what constitutes rationality. When these assumptions do not hold, game theory is silent.

These facts have led some game theorists to highly critical positions on the field. Gintis is especially blunt in print (but others say worse in conversation).

> Ironically, game theory is often hoisted on its own pétard: many of its most fundamental predictions—predictions that would have been too vague to test with any confidence in the pre-game-theoretic era—are *decisively and repeatedly disconfirmed,* in laboratory settings, with substantial agreement among experimenters, regardless of their theoretical priors. [27, page xxiv] (emphasis in original)

This is due, in large part, to hyper-strong assumptions about the rationality of players. In consequence,

> It is better to drop the term "rational" altogether, . . . .
>
> In the same vein, we do not follow classical game theory in asking how agents "learn" to play optimal strategies, because the cognitive

What is important here is that game theory does not pretend to tell you how to make judgments about the shortcomings of an opponent. In making such judgments, you would be better advised to consult a psychologist than a game theorist. Game theory is about what players will do when it is understood that both are rational in some sense. Sometimes, . . . this means that an orthodox game-theoretic analysis is not necessarily a very helpful guide on how to play against real people. . . .

Does this mean that game theory is useless? Obviously I do not think so or I would not devote my time to it. It is however true that, unless there are good reasons for supposing that the people involved will behave rationally, game theory cannot realistically be used in a naive way to make predictions about what real people will do. As a consequence, a player would often be unwise to use the strategy that a game theory book may label as "optimal" because this will usually only be optimal if *everyone* plans to play optimally. Of course, there are circumstances in which it *is* reasonable to work on the hypothesis that people will behave in a reasonably rational manner. Economics is somewhat shakily founded on the assumption that this will typically be the case in commercial and business transactions. However, it would be skating on very thin ice to use game theory for predictive purposes if none of the following criteria were satisfied:

- The game is simple.
- The players have played the game many times before[a] and hence have had much opportunity for trial-and-error learning.
- The incentives for playing well are adequate.

[a] Against different opponets each time. If you are to play a particular game against the same opponent many times, one much model the repeated situation as a single "supergame".

Figure 1.7: Ask a psychologist: Passage from *Fun and Games*, [4, pages 50–1]

> processes involved in "learning" are probably, under most conditions, much less important than the forms of imitation underlying the replicator dynamic...and cultural transmission.... In short, evolutionary game theory replaces the idea that games have "solutions" that agents "learn," with the idea that games are embedded in natural and social processes that produce agents who play effectively.
>
> Dispensing with the rationality postulate does not imply that people are *irrational* (whatever that means). The point is that the concept of "rationality" does not help us understand the world. [27, pages xxv-xxvi]

The position I take in this book is that rationality is a concept well worth saving, even if we need to develop new varieties of it. That, however, is a long story, one that will unfold as the book does. For now a working distinction will be enough to proceed. Details and elaboration will follow.

Orthodox game theory employs, as I said, a very strong notion of rationality. Players are assumed to choose so as to maximize the possible benefits to them. It is crucial to understand that by "maximize" game theory means "actually achieve the maximum" rather than, say, "aim for, or try to, achieve the maximum." To be rational is to maximize in this strong sense. Classical game theory asks What will happen in games in which all of the players are rational in this way? That is an important question and game theory has delivered admirably in addressing it.

*What this book is about*

This book investigates a different question: What will happen in games in which the players lack the ideal rationality of game theory, but are in possession of plausible learning regimes which are invoked during strategic interaction and are used to inform future actions?

To sloganize: game theory is about strategic situations in which the players *achieve maximization* for themselves, while agent-based game theory is about players who *seek improvement* for themselves.

We are only at the beginning. What follows in this book is a detailed elaboration of these points.

## 1.4 Why Study Games?

Why are we interested in contexts of strategic interaction (CSIs)? Games are interesting because games—war, diplomacy, poker, business strategy, and so on—are each interesting in themselves and each are examples from a larger pool of important phenomena meriting attention. Besides interdependent, interactive choice—the characterizing feature of games—we see the interplay of reasoning, calculation and reckoning, deception, skill, bluffing, power, adaptation, flexibility, cooperation, learning, arbitrage, coordination, norms, communication, markets, social organization, and much else that is pervasive in, and fundamental to our understanding of, the social order. (And belonging to the AGE society of ideas.) Games are interesting because they are vortices of many interesting phenomena. Social phenomena manifest themselves—play themselves out—in games.

What would we like to know about games? As in any field of science we seek to describe, explain, predict, and intervene. We wish to describe and classify games in the wild systematically. Chapter 3 is merely a hint at a much-needed natural history of games. We wish to explain and predict game outcomes. This often called *solving the game* in the classical literature. We also wish to understand—explain and predict—how it is outcomes are reached. How do agents of various sorts (experienced humans, naïve humans, monkeys, rats, lichens, organizations, artificial agents) find and implement their strategies of play? How does play unfold over time (and over space when geography is relevant)? Finally, we seek understanding of games in order to intervene in the world. We might hope to improve our own play in strategic contexts, or to design better social institutions (such as markets for electrical power that resist manipulation—"gaming"—as in the Enron affair [71]), or to field artificial agents that labor on our behalves (perhaps for negotiation or purchasing over the Internet) . The scope of potential investigation is both magnificent and beyond our means. We should be content with modest progress, while keeping ourselves reminded of the larger issues. That, at least, describes my aims in this essay.

## 1.5 Methods of Study

Games in the wild, we must always remind ourselves, are the primary phenomena that motivate study of contexts of strategic interaction. The games we make up or develop as abstract models are ultimately interesting only because they contribute to understanding games in the wild. How, in particular then, can we study strategic

interaction? Various ways are open to us:[4]

1. *a priori.* CSIs or games in the wild may be abstracted and reduced to formal models, then studied mathematically, typically upon assumption of axioms of rationality. From this perspective, the theory of games is a branch of mathematics. Much of classical game theory proceeds in this mode. Standard textbooks and reference works include [4, 26, 39, 44, 67].

2. *in vivo.* Games, or strategic situations, may be studied *in situ,* as they (more or less) naturally occur. This is an historical—"natural history"—mode of investigation, but of course the history may be contemporary and the means of study may use techniques from anthropology, sociology, or journalism. Pioneers of this approach include Thomas Schelling (e.g., [62, 60, 61, 63]), and Jon Elster (e.g., [13, 15]).

3. *in vitro.* We can study games by doing experiments with real ("wet") agents, including humans (e.g., [17, 36]), monkeys (e.g., [18]), even blue jays (e.g., [70]). And why not lichens and bacteria? The literature uses such names as *behavioral game theory* and *experimental economics* to refer to these kinds of investigations.

4. Algorithmic or *in silico.* There is much to be learned about games by representing agents as decision algorithms that choose their plays, and then studying the behavior of the resulting system. Fundamentally, this method of simulation or experimental mathematics is a variant of the *a priori* method. Let us call it *algorithmic game theory.* By allowing ourselves to use computational methods (instead of purely analytic mathematics) we may greatly extend the range, scope, and realism of models addressed, and concommitantly reduce the stringency of the assumptions required.

In what follows, we shall draw upon each of these methods. Our main focus methodologically, however, will be on algorithmic (or *in silico*) studies of artificial agents, using agent-based models (ABMs). Such studies may be, and have been, conducted from a variety of perspectives. Agents may be modeled as naked strategies (what I call *identity-centric* agents), possibly reactive or adaptive strategies, that play in tournaments (e.g., Axelrod's original and seminal study [2]) or that play in a populated ecology which evolves under the replicator dynamic (e.g.,

[4] I am grateful for the discussion in [50].

[2, 27, 68]) or that play in a differentiated geography (aka: spatial games; e.g. [16, 30]).

Again, we shall draw upon these and related studies but focus our efforts elsewhere. That focus has four main aspects:

1. *Finite, non-ideal* contexts of strategic interaction. Agents are finite beings. Their rationality, their abilities to reckon and foresee, are limited. The algorithms with which we model these agents must be computable[5] and computable without exorbitant use of resources. Play unfolds in a finite population, for a finite time, and in a finite space. A major theme will be to compare and contrast results under finite, non-ideal and infinite, ideal regimes of play. Classical game theory employs what philosophers call an *externalist* theory of rationality. Here we are asking different questions and shall be focusing on *internalist* notions of rationality. The upshot of this point will emerge as we proceed.

2. *Identity-centric* more so than strategy-centric agents. Humans, and indeed monkeys and blue jays, in contexts of strategic interaction may be said to *have* strategies (rather than to *be* strategies), and to be capable of changing them in response to experience. These players, and most of the agents we shall consider, may meaningfully be said to have identities distinct from the strategies they employ at any given time. They are more than naked strategies. In particular, they are

3. *Exploring, probing* agents, not merely reactive agents. Humans, monkeys, blue jays, and most of our agents face the exploration–exploitation dilemma/-tradeoff, addressed throughout the machine learning literature.

   Finally, the strategic contexts we will focus on will be

4. *Chronic* and *social* more so than acute and singular, *ongoing* and *widespread* more so than unique. The games may be *repeated* (played many times) or *iterated* (played many times by the same plaers) or be like other games that will be played, rather than being unique, non-repeatable events.

A word of elaboration and justification for this last aspect of our focus. Contexts of decision, or choice, may be distinguished into *strategic* (or game-theoretic, the principal subject of this book) and *parametric* (not strategic, the principal subject

[5] Technically, effectively computable, which I assume to be coterminus with the partial recursive functions. We are only interested in algorithms that implement partially recursive functions.

of the field of decision analysis).[6] Further, contexts of decision or choice may be distinguished into *acute* ("one-shot" or once-only) and *chronic*. Herrnstein's name for chronic choices [34]—*distributed*—is an apt description, and I shall use it. Chronic decisions or choices are distributed, usually in time. We may think of habits as chronic decisions that are, or become, more or less settled. The following table, then, summarizes this framework:

| | Acute (one-shot) | Chronic (distributed) |
|---|---|---|
| Parametric | decision analysis | decision analysis |
| Strategic | classical game theory | agent-based game theory |

Table 1.1: Framework categorizing decision/choice contexts. This is an essay on agent-based game theory.

Further, as I have noted, decision contexts may be distinguished into *individual* and *social*, although this distinction is perhaps more applicable to strategic than to parametric contexts. In any event, there will be much emphasis in what is to follow on *social* aspects of distributed strategic choice.

We often *are* inclined to think of games in terms of acute, dramatic points of decision. This is captured in the penultimate stanza of "Casey at the Bat" (Ernest L. Thayer, alias Phin, page 4 of the San Francisco *Daily Examiner,* June 3, 1888).

> The sneer has fled from Casey's lip, the teeth are clenched in hate.
> He pounds, with cruel violence, his bat upon the plate.
> And now the pitcher holds the ball, and now he lets it go,
> and now the air is shattered by the force of Casey's blow.

Of course the final stanza is

> Oh, somewhere in this favored land the sun is shining bright.
> The band is playing somewhere, and somewhere hearts are light.
> And, somewhere men are laughing, and little children shout,
> but there is no joy in Mudville – mighty Casey has struck out.

[6]See Paul Bloom's *Descartes' Baby* [5] for an accessible presentation of evidence that human mentality innately is organized in recognition of the distinction between parametric and strategic decisions.

But first, many games, many contexts of strategic interaction, *are* distributed or chronic, or approximately so. Agents do business with a particular merchant, doctor, lawyer, restauranteur repeatedly.[7] Agents have friends, partners, lovers, spouses, colleagues they encounter more than once. Agents have competitors in the market for more than a day. Agents are embedded in societies. Very often indeed, strategic contexts cannot be separated from the future or the past.

And second, is Casey's situation really unique, even for Casey? True enough, Casey is in a zero-sum game in the sense that only one team can win. It is also true that in any given at-bat the pitcher in baseball has the advantage; anyone can strike out. Most likely, however, there will be another game tomorrow or the next day. Casey's interest lies in maximizing the expectation of his future contributions to the team. Getting angry, pounding the bat, focusing exclusively on this game and this moment is, perhaps, not the wisest of moves on Casey's part. Better to take the long view. Better to have the pitcher strike you out than for you to strike yourself out. Perhaps the long view can inform the acute. Perhaps, at least sometimes, learned policies of play in the chronic case should drive or at least inform play in the acute case. What follows has among its aims the investigation of such conjectures and their ramifications.

$Id: intro-age.tex,v 1.36 2006/12/27 01:20:14 sok Exp $

[7] When play is repeated between or among the same agents, I shall say that it is, in addition to being repeated, *iterated.*

# Chapter 2

# Four Themes

My purpose of this chapter is to introduce four topics that figure prominently, four themes that recur frequently, in the remainder of this book. Very briefly and quite approximately they are:

1. The Social Order

   The social order is the regular, systematic patterns of behavior evidenced by a particular society (or collection of individuals). Trust is one such behavior that will engage us. It is commonly described as the "cement of society," the glue that makes social order possible. For example, among a group of individuals conducting business, handshakes and verbal commitments will often suffice instead of written, witnessed agreements closely supervised by legal authorities. What accounts for this? When are more stringent measures required and why?

2. Self-Organization and Emergence

   We often find in complex systems *self-organizing* behavior: *spontaneous order* [4, page 393] emerges without centralized direction or control. Any property (including behavior) of a larger object is an emergent property if it arises through a decentralized process, "from the bottom up" (rather than centrally, "from the top down"), conducted by smaller, constituent objects. For example, the theory of evolution by natural selection views species (larger objects) as "emerging" from competition in the "struggle for life" among individuals (smaller objects). The process is decentralized and not, according to the theory, designed and directed from above.

3. Combinatorial Complexity

   From a small alphabet we may combine letters to form a large number of words. By combining in different ways a modest number of words (just a few thousand) we can compose not only millions of meaningful sentences, but millions of coherent books. The possible ways of combining a number of distinct objects is called a *combinatorial space* for those objects. In investigating agents, games, and evolution, we routinely encounter combinatorial spaces of enormous size (aka: complexity) and must face the problem of understanding how agents and evolution can search and navigate these spaces.

4. Rationality

   As we shall see, rationality is a contested concept. Superficially at least, *rational* and *rationality* commonly refer to several rather different ideas. For example, a rational belief is often said to be a belief for which the believer has good reasons. But also, a rational choice is often said to be one that is consistent with a body of preferences. We shall need a concept of rationality applicable to agents in games, agents that need not be possessed of human intelligence, agents whose rationality is real even if less than ideal.

The remaining sections of this chapter elaborate more thoroughly, but still quite incompletely, upon these topics. This serves as a start for further elaboration, which is continued throughout the book.

## 2.1 Problems of the Social Order

Societies are collections of individuals, sustained over time and exhibiting systematic behavior. They offer even the casual observer an inexhaustible supply of interesting questions and puzzles. Why do animals live in groups? What explains the sizes and behavioral orderings of these groups? Why do animals sometimes not live in groups? Among humans, why and when do voluntary associations form? What maintains them and what destroys them? Why are territories established and what explains the degrees of effort made to defend them? When does cooperation arise? How is it destroyed? How is cooperation to be distinguished from mutual parasitism and when does one occur rather than the other? Is trust to be distinguished from cooperation? If so, what are its conditions and dynamics? Can the unrestrained actions of purely selfish individuals yield a productive and

stable social order? If so, under what conditions? Under what conditions does genuine altruism arise? How is the social order affected by incentives given the participating agents? How can under- and over-incentivizing be recognized? How can optimal levels be determined? How do intelligence and memory affect strategic outcomes and social structure? How stable are the social structures we see about us? What is most likely to make them unstable and what forms of organization would succeed them?

These are just a few examples of the important problems of the social order. More generally, we can ask: for a given society, Why are things the way they are? and What would happen if a particular intervention were to occur? These and the questions above are questions and issues that arise in attempting to understand the formation and behavior of collections of individual agents, be the agents human or not, be the collections communities, societies, entire ecologies or any other sustained form of interaction among agents exhibiting at least minimal forms of behavior. These problems are addressed throughout the social sciences, in much of biology (behavioral ecology is entirely devoted to these problems, cf., [37]), and in much of philosophy (especially, but not exclusively, political and moral philosophy). Each of these disciplines recognizes, to some degree or other, the relevance of strategic decision making and strategic interaction (games) for understanding the social order.

Certain questions, themes, and modes of explanation have figured prominently in the science of the social order, regardless of academic discipline. We are well advised to begin among them. Exhibit 1 is a famous passage from Adam Smith, *The Wealth of Nations*, Book I Chapter II (1776).

> It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our necessities but of their advantages.

As in many a folk saying—think: “It ain’t over ’til it’s over”—Smith is stating a truism for the sake of making a point. What is the point? We can safely avoid scholarly exegesis in this case. Two ideas lurk. The first is that individually self-interested behavior can lead to socially good results. To get one’s dinner one appeals to the self-interest of the butcher and company by paying them. As the folk saying says, “Every trade has two sides.” The suggestion in both cases—Smith and the folk saying—is that these are transactions in which all parties benefit. Acting from self-interest is sufficient to explain both obtaining dinner and

sustaining the butcher, the brewer, and the baker in their livelihoods. To some it will be tempting to strengthen this first idea to the claim that myopically self-interested behavior always or nearly always leads to socially good results. While the weaker claim seems unassailable as an observation, let us leave this second claim as something to be decided by scientific investigation. We shall address it often in the sequel.

The second idea lurking within Smith's passage has the modern terms *emergence* and *self-organization* associated with it. The notion is that (some) macro-phenomena appear as a result of micro-processes that neither resemble the macro-phenomena nor involve intentions to produce the macro-phenomena. The suggestion to hand is that the orderly conduct of business and the sustaining of commercial activity is *emergent.* It is a by-product of a self-organizing system in which large numbers of small transactions are conducted only with self-interest in mind (whether myopic or not). It is not the product of centralized planning and direction. Smith's name for this idea is a name we continue to use: *the invisible hand*. Here is the single passage in *The Wealth of Nations* (Chapter II of Book IV) in which Smith uses the term.

> As every individual, therefore, endeavours as much as he can both to employ his capital in the support of domestic industry, and so to direct that industry that its produce may be of the greatest value; every individual necessarily labours to render the annual revenue of the society as great as he can. He generally, indeed, neither intends to promote the public interest, nor knows how much he is promoting it. By preferring the support of domestic to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse for the society that it was no part of it. By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it.

Note that, contrary to many popular accounts, Smith is *not* claiming that emergent properties, produced by the invisible hand, are always or even usually socially, or even individually, good. Often, yes, but he offers no stronger claim, certainly nothing like the "Greed is good" ideology articulated in the movie *Wall Street* (1987, directed by Oliver Stone) and on many blogs and pundit shows these

days. Smith is not espousing social Darwinism. Nor is Smith holding a brief for outsourcing or globalization. It is "support of domestic to that of foreign industry" that promotes the interest of society.

Adam Smith's account of economic social order is generally one of good cheer. Self-interested individuals engage in transactions that make them better off and social gains emerge as well. The pattern is far from universal, yet occurs often enough. Although Smith was an acute observer of emergent phenomena, he was hardly the first. Thomas Hobbes was among the earliest and remains important. Here is the signature passage from his *Leviathan* (1651, CHAPTER XIII OF THE NATURAL CONDITION OF MANKIND AS CONCERNING THEIR FELICITY AND MISERY).

> Hereby it is manifest that during the time men live without a common power to keep them all in awe, they are in that condition which is called war; and such a war as is of every man against every man. For war consisteth not in battle only, or the act of fighting, but in a tract of time, wherein the will to contend by battle is sufficiently known: and therefore the notion of time is to be considered in the nature of war, as it is in the nature of weather. For as the nature of foul weather lieth not in a shower or two of rain, but in an inclination thereto of many days together: so the nature of war consisteth not in actual fighting, but in the known disposition thereto during all the time there is no assurance to the contrary. All other time is peace.
>
> Whatsoever therefore is consequent to a time of war, where every man is enemy to every man, the same consequent to the time wherein men live without other security than what their own strength and their own invention shall furnish them withal. In such condition there is no place for industry, because the fruit thereof is uncertain: and consequently no culture of the earth; no navigation, nor use of the commodities that may be imported by sea; no commodious building; no instruments of moving and removing such things as require much force; no knowledge of the face of the earth; no account of time; no arts; no letters; no society; and which is worst of all, continual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish, and short.

Hobbes's view was that in the absence of a controlling power—the Leviathan—the war of all against all would prevail and nothing would emerge except "con-

tinual fear, and danger of violent death; and the life of man, solitary, poor, nasty, brutish, and short."

We can fairly say that subsequent experience has confirmed both Hobbes and Smith to a degree. Trade flourishes, life is often sweet, yet wars, horrific cruelty and destruction continue. Whether we see peace as the interim between wars or war as the interim between periods of peace, we have to admit the continued presence of both. (Hobbes wrote at a time of horrific civil wars and Smith at the beginning of the Industrial Revolution.) It is a problem of the social order to explain and understand why. Note in particular that in Hobbes's world, "where every man is enemy to every man," there is no trust or cooperation. Note further that in Smith's more felicitous world trust and cooperation are required, at least oftentimes. Explaining when and why trust and cooperation will arise and be sustained is a central concern in understanding the social order.

## 2.2 Self-Organization and Emergence

It was Hobbes's view that trust and cooperation are not, and for the most part cannot be, emergent properties. They cannot arise by self-organization, emerging socially "from the bottom up"—or in Hobbes's phrase from "the state of nature"—as a result of decentralized, microbehavior by numbers of individuals. Instead, according to Hobbes, trust and cooperation—as macroproperties of a society—will only occur among a group of people if there is "a common power to keep them all in awe." Hobbes had in mind a strong monarch. Deferring for the moment the question of whether, and under what conditions, Hobbes was right about the (lack of) emergence of trust and cooperation, there are plenty of credible examples of emergence.

Darwin ends *On the Origin of Species* (first edition, 1959) with this wonderful and famous paragraph, summarizing the general mechanism by which species emerge by a self-organizing (decentralized, bottom-up) process :

> It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us. These laws, taken in the largest sense, being Growth with Reproduction; Inheritance which is almost implied by reproduction; Variability from the

> indirect and direct action of the external conditions of life, and from use and disuse; a Ratio of Increase so high as to lead to a Struggle for Life, and as a consequence to Natural Selection, entailing Divergence of Character and the Extinction of less-improved forms. Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.

To what extent can social order also be explained as emergent from self-organizing, microbehavior? What about failures of the social order, in particular what are called social dilemmas? As Robyn Dawes write in his thoughtful paper, "Social Dilemmas," these

> . . . are defined by two simple properties: (a) each individual receives a higher payoff for a socially defecting choice (e.g. having additional children, using all the energy available, polluting his or her neighbors) than for a socially cooperative choice, no matter what the other individuals in society do, but (b) all individuals are better off if all cooperate than if all defect. [10]

These are pervasive and

> While many thinkers have simply pointed out that our most pressing societal problems result from such dilemmas, most have addressed themselves to the question of how to get people to cooperate. Answers have ranged from imposition of a dictatorship (Leviation) to "mutual coercion mutually agreed upon," to appeals to conscience.

What is the scope of behavior for emergence and self-organization in the context of social dilemmas? Is centralized control needed? If so, under what conditions? What are the risks and pathologies of the various kinds of structures for relieving social dilemmas?

## 2.3 Combinatorial Complexity

To a first approximation, it may be said that the task of micro-evolution is to find genetic codings for felicitous proteins, which are used to catalyze chemical reactions. Proteins are linear sequences of amino acids (again, I'm approximating). Life forms on earth use 20 amino acids. So there are $20^n$ possible proteins of length $n$. Naturally occurring proteins may easily be 1000 amino acids long. The number of possibilities—$20^{1000}$—is vastly greater than can be searched exhaustively. The estimated total number of atomic particles in the universe is only on the order of $10^{80}$. Most of the $20^{1000}$ possibilities are thought to be useless. How then is successful search by evolution possible?

The number of species used for food by we humans is small compared to the total number of species, yet the number of possible recipes—the combinatorial space of recipes—is enormous. Restricting ourselves only to herbs and spices, Harold McGee in his magnificent work, *On Food and Cooking,* [48] lists more than 100 herbs and spices commonly used in cooking. Many of these come in multiple varieties (for example, parsley has just one entry). Neglecting varieties (Italian or small leaf parsley?) and quantities (How much parsley?), and counting only whether a spice or herb is present or not in a recipe, there are thus more than $2^{100}$ or about $10^{30}$ possible combinations of spices. With fewer than $32 \times 10^6$ seconds per year, trying one recipe/combination per second would take more than $10^{21}$ years. But there are fewer than $10^{11}$ years since the Big Bang. Why, we might wonder, are chefs able to produce new and tasty recipes year and after year? Are happy combinations easily found by purely random search, or is something else going on? Cloves with cumin and ginger anyone?

The number of strategies possible in iterated games is similarly explosive. Consider a $2 \times 2$ game played repeatedly from player $R$'s perspective. Played once, there are 2 possible strategies of play. Played twice, there are $2 \times 2^4$ possible (pure) strategies. Played $n$ times the number (of pure strategies) is $2^{4^0} \times 2^{4^1} \times 2^{4^2} \times \ldots \times 2^{4^{n-1}}$. For even modest $n$, we're better off cooking. How is it possible for players to arrive at sensible, well-founded strategies of play in iterated $2 \times 2$ games? What about larger games?

## 2.4 Concepts of Rationality

We use the terms *rational, rationality* and their ilk in ordinary language, in informal but carefully considered theories, and in formal, rigorously articulated

theories. Consequently, several distinct senses of these words can be found. In ordinary language, in a commonsense but somewhat vague sense, to be rational is roughly to pursue one's goals efficiently, sensibly, intelligently. To be irrational, or not rational, is to be possessed of self-destructive, or at least self-frustrating, impulses that one cannot control.

WordNet, the electronic dictionary, finds four senses of "rational" as an adjective (`http://wordnet.princeton.edu/`):

- S: (adj) rational (consistent with or based on or using reason) "rational behavior"; "a process of rational inference"; "rational thought"
- S: (adj) intellectual, rational, noetic (of or associated with or requiring the use of the mind) "intellectual problems"; "the triumph of the rational over the animal side of man"
- S: (adj) rational (capable of being expressed as a quotient of integers) "rational numbers"
- S: (adj) rational (having its source in or being guided by the intellect (distinguished from experience or emotion)) "a rational analysis"

and two for "rationality" as a noun:

- S: (n) rationality, reason, reasonableness (the state of having good sense and sound judgment) "his rationality may have been impaired"; "he had to rely less on reason than on rousing their emotions"
- S: (n) rationality, rationalness (the quality of being consistent with or based on logic)

No doubt correct, and quite consistent with other sources, WordNet's definitions nonetheless tell us little about what constitutes rationality. The Wikipedia (`http://en.wikipedia.org/wiki/Rationality`, accessed 22 August 2005) is also explicit in recognizing different senses of, and concepts for, the word rationality, some of which point usefully towards operational theories, e.g., "A logical argument is sometimes described as rational if it is logically valid," and "In economics, sociology, and political science, a decision or situation is often called rational if it is in some sense optimal, and individuals or organizations are often called rational if they tend to act somehow optimally in pursuit of their goals," and

"Rationality is a central principle in artificial intelligence, where a rational agent is specifically defined as an agent which always chooses the action which maximises its expected performance, given all of the knowledge it currently possesses."

Others might be added. Many philosophers, for example, would say that someone—say a believer—is rational regarding something—say a belief—if the one has good reasons for the thing. Notoriously, there are many people who often find (correctly or not) that good reasons (and that type of rationality) conflict with self interest (and that type of rationality).

We could go on. We could add new senses, we could refine and more deeply articulate existing senses. Others have; we needn't and won't.

Given such a variety of concepts of rationality, and corresponding senses of the word rational, it is temping to ask which is right. What is it *really* to be rational? These are questions I wish deliberately and forthrightly to evade. I see no good reason why there should be just one concept that we, in ordinary speech or even in carefully wrought theories, attach to the word rational. Let a 100 senses bloom. There may be many ways of being rational, ways whose discovery and application should not be blocked definitionally. This risks confusion, of course, if one sense is meant and another understood. But forewarned is forearmed. Many words have multiple senses. If it matters, speech and communication may be clarified. For similar reasons, Amartya Sen has embraced vagueness with regard to rationality [65, page 4]:

> Rationality is interpreted here, broadly, as the discipline of subjecting one's choices—of actions as well as of objectives, values and priorities—to reasoned scrutiny. Rather than defining rationality in terms of some formulaic conditions that have been proposed in the literature (such as satisfying some prespecified axioms of "internal consistency of choice," or being in conformity with "intelligent pursuit of self-interest," or being some variant of maximizing behavior), rationality is seen here in much more general terms as the need to subject one's choices to the demands of reason.

Present purposes call for a concept of rationality appropriate for decision making in contexts of strategic interaction (games). It is useful to distinguish two such concepts. The first—type A—is well articulated and explored in classical game theory. The second—type B—is the kind of strategic rationality that is the focus of this book. We can get a sense, good enough for present purposes, of what these two concepts are and how they differ by illustrating their application to four games.

Our first game is the well-known Standard Prisoners' Dilemma (SPD), presented in Figure 2.1 in strategic form.

| | $c_1$ | $c_2$ |
|---|---|---|
| $r_1$ | (3, 3) | (0, 5) |
| $r_2$ | (5, 0) | (1, 1) |

Figure 2.1: Standard Prisoners' Dilemma (SPD). Player $R$ chooses between strategies $r_1$ and $r_2$. Player $C$ chooses between $c_1$ and $c_2$.

The interpretation of a game in strategic form is straightforward. There are two players,[1] the row player, $R$, and the column player, $C$. The row player must choose one of its available strategies, either $r_1$ or $r_2$ in figure 2.1, and column must choose one of its strategies, either $c_1$ or $c_2$ in the present case. When, as in the Standard Prisoners' Dilemma game, there are two players each with two strategies, we say the game is a $2 \times 2$ game: two players, two strategies each. For games in strategic form, we stipulate that each player picks its strategy without observing the other player's choice. Similarly, the players can make no enforceable agreement about which strategies to pick. Once the strategies are picked, the payoffs to the players are determined, as shown by the cells in the figure. If, for example, row chooses $r_2$ and column chooses $c_1$, then row's payoff is 5 and column's is 0. In a $2 \times 2$ game, there are four possible outcomes—$(r_1, c_1), (r_2, c_1), (r_1, c_2), (r_2, c_2)$—and each outcome has a payoff for each of the players. The payoffs for SPD are given as the entries in figure 2.1.

Our second example game is a *constant sum* game: the total payoff is the same for every outcome. This makes the game one of pure conflict. $R$'s gain is $C$'s loss, and vice versa. Because the game happens to appear on page 90 of Federic Schick's *Making Choices* [64, page 90], I'll call this $2\times3$ game Schick90. See figure 2.2.

| | $c_1$ | $c_2$ | $c_3$ |
|---|---|---|---|
| $r_1$ | (3, 7) | (9, 1) | (1, 9) |
| $r_2$ | (5, 5) | (7, 3) | (6, 4) |
| $r_3$ | (4,6) | (2,8) | (8,2) |

Figure 2.2: Schick90: A game of pure conflict

[1] More general formulations are possible, but are not needed for present purposes.

Our third example game, Standard Stag Hunt (SSH), is also well-known. It is presented in figure 2.3. We will have occasion to discuss both Stag Hunt and Prisoners' Dilemma at length in the sequel.

| | $c_1$ | $c_2$ |
|---|---|---|
| $r_1$ | (3, 3) | (0, 2) |
| $r_2$ | (2, 0) | (1, 1) |

Figure 2.3: Standard Stag Hunt (SSH). Player $R$ chooses between strategies $r_1$ and $r_2$. Player $C$ chooses between $c_1$ and $c_2$.

Our fourth game is an ancient one. I'll call this version of it *One-Two-Twenty*. Two players take turns placing either 1 or 2 tokens on a table, starting from an empty table. The player placing the twentieth token on the table wins the game.

Here now are two approaches—type A and type B—to analyzing these games.

### 2.4.1 Type A Setup and Analysis

For type A analysis of a game (or a type A game), we need to specify the following items as constituting the setup of the game:

1. The players.

   For the examples to hand there are two players, $R$ and $C$ (think row or red or Robert, and column or cyan or Cynthia). In general there may be any finite number of players.

2. The pure strategy sets, $\Sigma^i$ for each player, $i$.

   A strategy (for player $i$) is a complete set of instructions for play of the game (by player $i$). In the Standard Prisoners' Dilemma game, $\Sigma^R = \{r_1, r_2\}$, and $\Sigma^C = \{c_1, c_2\}$. $\Sigma$ for a player is its set of *pure* strategies. When the game is presented in strategic form, as in figures 2.1, 2.2 and 2.3, the pure strategies for the row (column) player are the rows (columns) in the table. In addition to its pure strategies, each player also has *mixed strategies*. These are the probability-weighted combinations of the pure strategies. We denote mixed strategies with a tilde. For example, $\Sigma^C$ denotes player $C$'s set of pure strategies and $\tilde{\Sigma}^C$ denotes $C$'s mixed strategies. Since the pure strategies are a special case of probabilistic combination of pure strategies (one has a weight of 1, the others have 0), $\tilde{\Sigma}^C$ denotes all of $C$'s strategies.

3. For each outcome, a payoff vector giving payoffs in that outcome for each player.

   An *outcome* of a (type A) game is a strategy vector, giving the played strategy of each player.[2] Thus for Standard Prisoners' Dilemma, there are four possible outcomes: $(r_1, c_1)$, $(r_1, c_2)$, $(r_2, c_1)$, and $(r_2, c_2)$. The payoff vectors, $\omega(\cdot)$, for these outcomes are: $\omega(r_1, c_1) = (3, 3)$, $\omega(r_1, c_2) = (0, 5)$, $\omega(r_2, c_1) = (5, 0)$, and $\omega(r_2, c_2) = (1, 1)$. Our convention is that in payoff vector $(x, y)$, $R$ gets $x$ and $C$ gets $y$.

4. Rules of play for the game.

   Standardly, "The rules of a game must tell us *who* can do *what* and *when* they can do it. They must also indicate who gets *how much* when the game is over." [4, page 25]. I am handling the *how much* aspect of a game separately, as the payoff vectors $\omega$ (previous item).

   In the general case of a 2×2 game in strategic form, such as Standard Prisoners' Dilemma in figure 2.1, each player picks a strategy from its strategy set and this determines the outcome (and the payoffs). Players pick without observing each other's choices of strategy.

In addition to the game setup, type A games assume that the players have individual preference orders on the payoffs for the game. I'll illustrate with Standard Prisoners' Dilemma. The game has four possible payoff vectors: $\Omega = \{(3, 3), (0, 5), (5, 0), (1, 1)\}$. For player $R$ his *payoff* for a payoff vector $(x, y)$ is $x$ and because $R$ prefers higher payoffs to lower payoffs, $R$, let us assume, has the following *preference ordering* on $\Omega$: $(5, 0) \succ_R (3, 3) \succ_R (1, 1) \succ_R (0, 5)$. If an agent $a$ prefers $x$ to $y$ or is indifferent between $x$ and $y$, we write $x \succeq_a y$ (or just $x \succeq y$ if it is clear who the agent is). If an agent prefers $x$ to $y$ strictly (the agent is not indifferent between them), we write $x \succ y$. An agent is said to be *fundamentally rational* (with respect to a set of payoff vectors $\Omega$) if the agent has a preference ordering, $\succeq$, on $\Omega$ such that for every $a, b \in \Omega$:

fundamental rationality

1. The *totality* condition obtains:

   $a \succeq b$ or $b \succeq a$ (or both, in which case we say that the agent is indifferent between $a$ and $b$ and we write $a \sim b$), and

[2] My terminology here is nonstandard. Usually, *outcome* is used for what I am calling the payoff vector. Also, see below for the distinction between the *play* of a game and the *outcome* of play.

2. The *transitivity* condition obtains:

   If $a \succeq b$ and $b \succeq c$ then $a \succeq c$.

Clearly, $R$'s assumed preference ordering on the outcomes qualifies as fundamentally rational in this sense. For the sake of the example, we also assume that $R$'s counter-player, $C$, has a different preference ordering on the outcomes: $(0, 5) \succ_C (3, 3) \succ_C (1, 1) \succ_C (5, 0)$. It too is fundamentally rational.

Given a set of payoff vectors, we say that an agent's choice for or decision of $\omega \in \Omega$ is *rational* or *consistent* with respect to the rational preference ordering $\succeq$ on $\Omega$ if for all $\omega' \in \Omega$, $\omega \succeq \omega'$. In short, a choice is rational or consistent with regard to a fundamentally rational preference ordering, if there is no better choice available.

I am *not* suggesting that this is *the* definition of rationality, although economists and game theorists normally treat it that way (e.g., [57, page 19]). Because of its pervasive use and acceptance in this literature, I will honor the term. At times it will serve the purposes of clarity to refer explicitly to *fundamental rationality*. When confusion is unlikely, I will use *rational* and *rationality* either for this specific kind or in a general, unspecified sense (recall the passage from Sen, above) or for the specific kind under discussion.

Agents in games, however, do not get to choose payoff vectors, elements of $\Omega$, directly. Instead they get to choose strategies, elements of their $\tilde{\Sigma}^i$s. Type A analysis of games is about rational choice among strategies, assuming all players are fundamentally rational regarding the payoff vectors, $\Omega$. We thus need to define *type A rationality* in terms of choice among strategies. It is natural to define this sort of rationality much as we defined fundamental rationality with respect to $\Omega$. The complication, of course, is that in a game the strategy choices of all the players must be taken into account. A further complication is that we must allow for play of *mixed strategies*. We need to discuss this last complication first. A little notation, mostly. Nothing terribly complicated.

Each agent has its $\Sigma$, a set of basic or *pure strategies*. To repeat: for the Stag Hunt game, $R$'s pure strategies are $r_1$ (hunt stag) and $r_2$ (hunt hare), so we have $\Sigma^R = \{r_1, r_2\}$. In addition to the strategies in its $\Sigma$, an agent may also form a *mixed strategy* by probabilistic combination of its pure strategies. For the Standard Stag Hunt game, if $R$ plays $r_1$ with probability $\frac{1}{2}$ and $r_2$ with probability $\frac{1}{2}$ we can write $\tilde{r} = (r_1, \frac{1}{2}; r_2, \frac{1}{2})$. I'll denote a mixed strategy by using this tilde notation, e.g., $\tilde{r}$ is a mixed strategy. Note that even if $\Sigma$ is finite, $|\tilde{\Sigma}|$, the number of possible mixed strategies an agent can form from it is (uncountably) infinite.

Now the terminology. We need to distinguish the *play* of a game, from the *outcome* of (play of) a game, from the *payoff* resulting from the outcome of a game. The play of a game, *Play*, is the vector of strategies chosen by each player for playing the game. If every player plays a pure strategy, then the outcome of play, *Outcome*, is identical to the play. If at least one player plays a mixed strategy, however, chance must resolve the game into some pure strategy, and it is the resulting pure strategy that belongs to the outcome.

*Play* vs. *Outcome* vs. *Payoff*

The following example illustrates this simple framework. Let the game be Standard Stag Hunt (SSH, figure 2.3). Suppose that *Play* = $(\tilde{r}, c_2)$. That is, $R$ chooses to play the mixed strategy $\tilde{r} = (r_1, \frac{1}{2}; r_2, \frac{1}{2})$, defined above. $C$ decides to play her pure strategy, $c_2$. Given *Play*, chance now resolves it by instantiating any mixed strategies. Let us say that the coin is flipped and $R$'s $\tilde{r}$ gets resolved to $r_1$. Then the *Outcome* = $(r_1, c_2)$. Then the *Payoff* = $(0, 2)$, as indicated in figure 2.3, page 28. Note that $R$'s realized payoff is $\frac{1}{2} \cdot 0 = 0$. On average, however, $R$'s expected payoff (given that $C$ plays $c_2$) is $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 1 = \frac{1}{2}$. We can denote this compactly by writing E*Play*$(R, (\tilde{r}, c_2)) = \frac{1}{2}$. In words, the expected value, E, to $R$ of the play $(\tilde{r}, c_2)$ is $\frac{1}{2}$.

We are now in position to define rationality for type A games. Let $\tilde{s}^i \in \tilde{\Sigma}^i$, that is, $\tilde{s}^i$ denotes a (possibly mixed) strategy available to agent $i$. Numbering the players from 1 to $n$, we can denote the play of a game by *Play* = $(\tilde{s}^1, \tilde{s}^2, \ldots, \tilde{s}^i, \ldots, \tilde{s}^n)$. We say that agent $i$ is *individually economically rational* (IER) in the play of a game, if there is no strategy other than the one played by $i$ that would yield a superior payoff for $i$, assuming the play is otherwise unchanged. Formally, given *Play* = $(\tilde{s}^1, \tilde{s}^2, \ldots, \tilde{s}^i, \ldots, \tilde{s}^n)$, agent $i$ is IER if there is no $\tilde{t}^i \in \tilde{\Sigma}^i$ such that E*Play*$(i, (\tilde{s}^1, \tilde{s}^2, \ldots, \tilde{t}^i, \ldots, \tilde{s}^n) \succ$ E*Play*$(i, (\tilde{s}^1, \tilde{s}^2, \ldots, \tilde{s}^i, \ldots, \tilde{s}^n)$. Put otherwise, given $i$ played $\tilde{s}^i$, then $i$ is IER if $i$ has no better strategy, given how the other players played. Again, $i$ is IER if the strategy $i$ played is on average a *best response* to what the other players played.

IER: individually economically rational

An agent that is individually economically rational (IER) is also said to be *consistent* with its fundamental preference ordering on $\Omega$. Individually economically rational agents are said to maximize or optimize, given their fundamental preferences and the choices made by the counter-players in the game. In ordinary discourse, there is a sense of "Agent $i$ is maximizing" that is equivalent to "Agent $i$ is attempting to maximize" or "trying to maximize." This is *not* the sense employed in the present context. To the contrary, to say that an agent maximizes or is individually economically rational or is consistent with its fundamental preferences on $\Omega$ *means* that the agent actually succeeds in playing a strategy that is a best strategy, given the actual play by the other players. How the agent might

know which strategy to play, given that the strategy choices by the other players are hidden, is an entirely separate matter—in the type A setup—for the sake of determining whether the agent is individually economically rational or not.

For a given play, if (and only if) all of the players are individually economically rational (in this type A sense), we say that the play is a *Nash equilibrium* and that the outcome resulting from the play is *supported by* the Nash equilibrium play. Put equivalently, in a Nash equilibrium no player *individually* has incentive to change its chosen strategy. It is not precluded by the Nash equilibrium concept that two or more players might together change their strategies in such a way that both (or all) are better off. The concept of a Nash equilibrium is tied essentially to that of the possibilities for individuals acting alone, one at a time. Note again that individually economically rational outcomes and Nash equilibrium outcomes are defined without reference to how they may be arrived at or discovered. While analysts of a finite game may discern all of the solutions and test for individual rationality and Nash equilibrium, players in the game may or may not have sufficient means to hand to make these discoveries. Finally, we say that a *solution to a type A game setup* is a Nash equilibrium.

Nash equilibrium

With these notions to hand, we can now treat the example games from the type A perspective.

A type A story (or analysis) for the Prisoners' Dilemma game is especially straightforward. There is exactly one play that is individually economically rational for both players (and hence is a Nash equilibrium): $(r_2, c_2)$. Note that the associated payoff vector is $(1, 1)$ and that both players would do better if the play were $(r_1, r_1)$. Hence the dilemma. Type A analysis predicts the Nash equilibrium as the play of this game.

There is also an attractive line of reasoning that explains how the players might reach the Nash equilibrium by reasoning individually from their knowledge of the game. Player $R$ has two pure strategies. The second, $r_2$, is said to *dominate* the first because no matter which (mixed or pure) strategy the counter-player, $C$, plays, $R$ gets an outcome he (I'll use "she" for $C$) prefers more if he plays $r_2$ rather than $r_1$. So, if $R$ is to be consistent with his rational preference ordering on his $\Omega$, $R$ must choose to play strategy $r_2$. A completely analogous story applies to $C$. She must play $c_2$, her dominant strategy, if she is to be consistent with her preference ordering over her $\Omega$. Game solved.

Our second example game is Schick90, figure 2.2 page 27. Each player for this game of pure conflict has a plausibly attractive decision rule for finding a strategy: play a *maximin* strategy, a strategy that maximizes its minimum possible payoff. Taking row's perspective, playing $r_2$ guarantees $R$ a payoff of at least

5. Since the minimum payoff for $r_1$ is 1 and the minimum for $r_3$ is 2, $r_2$ is uniquely the strategy for $R$ that maximizes the player's minimum payoff. Similar reasoning identifies $c_1$ as $C$'s maximin strategy. Play of $(r_2, c_1)$ is in consequence individually economically rational for both players and is a—indeed the—Nash equilibrium for this game. Game solved.

Note a subtle difference between Standard Prisoners' Dilemma (SPD) and Schick90. In SPD each player is able to reason by elimination of dominated strategies and determine a uniquely attractive strategy of play completely independent of how the counter-player will play. No matter what row does, column is better off playing $c_2$. The situation is different in Schick90. When both players are IER each gets a payoff of 5. The reasoning and justification for $r_2$ (and $c_1$) relies on the fact that these are "safety strategies". No matter what the counter-player does, these strategies guarantee to their players a maximal minimum. No other strategy for each player is guaranteed to produce more than 5. If, however, one player knows, or has a reasonable amount of evidence, that the counter-player will not play its maximin strategy, then the player might be able to do better with a different strategy. For example, if $C$ is certain that $R$ will play $r_1$, then $C$ should play $c_3$ for a payoff of 9. Generally and in distinction to Prisoners' Dilemma, which strategy yields the highest payoff for a player now depends on which strategy is played by the counter-player. Reasoning by an individual player leading to a Nash equilibrium play will normally require the assumption of IER play by the counter-player(s).

The type A analysis of Stag Hunt is straightforward, but a bit problematic. Plays $(r_1, c_1)$ and $(r_2, c_2)$ are both Nash equilibria. There is in addition a third play, involving *mixed strategies*. If an agent plays $x$ with probability $p$ and $y$ with probability $(1 - p)$ we write this mixed strategy as $(x, p; y, 1 - p)$. For the Stag Hunt game in figure 2.3 the play in mixed strategies, $(\tilde{r} = (r_1, \frac{1}{2}; r_2, \frac{1}{2}), \tilde{c} = (c_1, \frac{1}{2}; c_2, \frac{1}{2}))$ is the third Nash equilibrium. Players playing at this equilibrium can expect a payoff of 1.5 each.

Two things make the Stag Hunt problematic in type A analysis. First, which play will prevail and what will be the distribution of plays when the game is surveyed across many plays? The Nash equilibrium concept by itself cannot discriminate among the three Nash equilibria. It is possible, of course, to single out one or another of the equilibria as favored in virtue of properties it has in addition to being a Nash equilibrium. Game theorists have tended to favor $(r_2, c_2)$ because it poses the least risk of a 0 payoff to any player, but there is not general agreement on this. There is a worry, moreover, that the properties so identified will not generalize to other games with multiple equilibria. Although principled selections may

be made in specific cases, the problem of multiple equilibria for type A analysis remains recalcitrant.

The second problem for type A analysis presented by the Stag Hung game is related to the first, and is perhaps but an aspect of it. This is the problem of specifying a procedure for arriving at an equilibrium. Sticking to just the two equilibria in pure strategies, $r_1$ is $R$'s strategy for one of the equilibria, while $r_2$ is the strategy for the other. But $R$ can't by himself pick one of the equilibria. If he prefers $(r_1, c_1)$ he can play $r_1$, but if $C$ prefers $(r_2, c_2)$ and she plays $c_2$ then the play is not a Nash equilibrium. Further, although $(\tilde{r}, \tilde{c})$ is a Nash equilibrium, plays combining mixed and pure strategies are not, e.g., $(\tilde{r}, c_2)$. If there is no procedure or path of reasoning by which the players coordinate on a single equilibrium, how is it that play is at an equilibrium?

The type A story for One-Two-Twenty is a bit complex, but not especially difficult. I will give it only in outline. A *state of the game*, $e = (i, n)$, is specified by which player, $i$, has the next play and how many stones, $n$, are on the board. Since there may be $0, 1, \ldots, 20$ stones on the board and at any time it may be either player's turn, there are 42 possible states. This number is reduced once we specify who goes first, but the details are not important for us. Let us say that $C$ goes first. It is easy to see that there is a winning strategy for $C$. Let the value of state $e$ to player $i$, $V^i(e)$, be 1 if once the game is in that state player $i$ can be guaranteed of a win. Clearly $V^C(C, 19) = 1 = V^C(C, 18)$, for $C$ can add either 1 or 2 stones to produce 20 on the board and thereby win the game. Consequently $V^C(R, 17) = 1$, since $R$ can then produce only states $(C, 18)$ and $(C, 19)$, which have a value to $C$ of 1. Continuing to reason backwards in a similar manner we find that $V^C(R, 14) = 1 = V^C(R, 11) = \ldots = V^C(R, 2)$. Consequently, $V^C(C, 0) = 1$. $C$ can begin the game by placing 2 stones on the table, thereby producing state $(R, 2)$. If $R$ produces state $(C, 3)$, then $C$ puts 2 stones down, producing state $(R, 5)$; otherwise, $R$ produces state $(C, 4)$ and $C$ puts 1 stone down, again producing state $(R, 5)$. Play continues in this fashion until $C$ wins the game. $C$'s strategy, combined with *any* play by $R$ is a Nash equilibrium. Note that if $C$ deviates from this strategy, say by producing state $(R, 6)$, then there is a strategy available to $R$ for winning the game. In the example, $R$ would produce $(C, 8)$ and be in position to win the game. Thus, in One-Two-Twenty rational (type A) play as specified by the Nash equilibrium accords well with what we would atheoretically expect rational players to do. The game favors whoever goes first. That player wins if the outcome is a Nash equilibrium.

### 2.4.2 Type B Setup and Analysis

The setup for, or description of, a type B game includes the following elements:

0. The supergames.

   A type B game consists of one or more supergames. Each supergame comprises many (2 or more) subgames. In a simple case, the subgame (aka: stage game) would be Standard Prisoners' Dilemma and the supergame would be 25 rounds of play of the subgame between two fixed players.

1. The players.

   Every game, including type B games, has a least 2 players. If there are only 2, I will continue to call them $R$ and $C$; otherwise they will be labeled $P_1, P_2, \ldots, P_i, \ldots$.

2. The policy sets, $\Pi^i$, for each player, $i$.

   A policy is a strategy (complete set of instructions) for playing a subgame. Policies are defined in such a way that a player may during the course of a supergame play under one policy for part of the supergame and under another policy for a different part. For example, if the supergame consists of 25 rounds of play of Standard Prisoners' Dilemma, a player might cooperate for rounds 1-11 and defect for rounds 12-25.

   A main difference between games of type A and games of type B, is that in the former we view players as choosing among strategies, while in the latter we view players as choosing (directly) among policies. They choose strategies only indirectly, as emerging from their policy choices.

3. For each outcome of every atomic subgame, a payoff vector specifying payoffs for that outcome for each (involved) player.

   A subgame is atomic if all of its outcomes are associated with elements of a relevant $\Omega$. A nonatomic subgame may be composed of atomic subgames or may have as payoffs the rewards from playing in other subgames.

4. The adaptation regime(s) used by the players, $\rho$ (or indexed , e.g., $\rho_i$ if there is more than one).

   Players play by following policies for play during a supergame. A particular action in a particular subgame is determined by the policy in effect for the player in question. The player's adaptation regime determines which policy

will be in effect at any given time. Necessary for games of type B, a game description of type A lacks this element entirely. I will discuss examples shortly.

5. Rules of play for the (super)game(s).

   As in type A game descriptions, the rules of play govern the sequencing and other conditions under which the players make their decisions.

The key differences, then, between a type B game setup (description, model) and one of type A are that (i) type B games are always supergames, consisting of multiple subgames, (ii) players directly choose policies rather than strategies, with the policies in turn determining play in subgames,[3] and (iii) players have adaptation regimes which produce their choices of policies. The salient feature of type B game setups is that players try policies, receive feedback from play of subgames, invoke their adaptation regimes, and either try new policies or continue on, depending on direction from their adaption regimes. This is a process of adapting (and perhaps learning) in policy space.

Our example games can be used to illustrate type B game setups. Standard Prisoners' Dilemma (SPD) first. The supergame for this example consists of iterated play of SPD as a stage game. After each round of play the supergame halts with probability 0.02; otherwise another round is played. Players $R$ and $C$ have identical consideration sets of policies for play, $\Pi = \{$ALWAYS DEFECT, ALWAYS COOPERATE, TIT FOR TAT$\}$. Under the ALWAYS DEFECT policy, if the player is $R$ he plays $r_2$ whenever he has the policy in force and if the player is $C$ she plays $c_2$ whenever she has the policy in force. Similarly, they play $r_1$ and $c_1$ if ALWAYS COOPERATE is in force. Finally, under TIT FOR TAT, the player playing it cooperates ($r_1$ or $c_1$, depending on the player) in the first play for which the policy is in force. After that, so long as the policy is used, the player mimics the play (cooperative or not) of the counter-player in the previous round of play. Players independently pick policies and play them for a number of rounds of play. Each player keeps track of the performance of, the returns from, play with its policies, and uses this information when selecting a new policy for play. What will happen? Will ALWAYS DEFECT win out as the best of the available policies? Will a different policy win out or will there be no settling down of the process? Later we shall see.

[3] Policies need not be deterministic. They may involve randomized decisions. No special notation—such as use of the tilde in type A games, $\tilde{s}$—will be needed.

Our example model for Standard Stag Hunt is a *gridscape* model. Players are arrayed on a regular network. Think of a chessboard, but one that "wraps" around so that each cell has 8 neighbors. This is a model for a simple society. Agents are either stag hunters or hare hunters. Agents in parallel play all of their 8 neighbors and count up their points. Each agent then looks to its neighbors and sees if any have obtained more points by using a different strategy. If so, the agent adopts the strategy of a neighbor whose achieved points is highest and play continues. Figure 2.4 is a screen shot from the program `m1-symmetric-2x2.nlogo`, set up to run this model on a gridscape with 49,933 cells. The cells have been initialized randomly, with a strong bias towards hare hunting. Black cells are occupied by stag hunters, which are very much in the minority. How will the gridscape evolve when play commences? Will one or another policy (HUNT HARE, HUNT STAG) take over the gridscape or will a mixture be present indefinitely?

For either One-Two-Twenty or Schick90 (both of which are games of pure opposition) imagine again that two players $R$ and $C$ play the game repeatedly and that in the case of One-Two-Twenty $C$ is given the first move each time the game is played. Neither $R$ nor $C$, let us assume, have the capacity to analyze the game as we did above and to figure out an optimal strategy. Are there ways that simpler agents (simpler than us) might figure out the game? Clearly yes. An agent with a bit of memory and an elementary ability to reason backwards could surely learn by experience in iterated play of the game and achieve optimal, or at least high quality, play. In the case of One-Two-Twenty, the value of later states could be learned first by trial and error, followed by the value of neighboring earlier states. Eventually the entire game could be learned. How quickly might this be done? How well does the approach generalize to more complex games, such as chess and checkers?

## 2.5 Discussion: Accessibility & Games

A few terminological stipulations will facilitate discussion here and in the sequel.

Suppose that algorithm (or procedure or rule) $\alpha$ accepts inputs $\beta$ and produces $\gamma$. Let us then say that *$\gamma$ is accessible from $\beta$ via $\alpha$.* If needed, it is possible to give a more formal, rigorous definition of accessibility, but that requirement is not to hand. Examples can carry the burden of clarification: (1) $\gamma$ is $\neg P$, $\beta$ is $P \rightarrow Q, \neg Q$, and $\alpha$ is *modus tollens.* (2) $\gamma$ is 27, $\beta$ is $x = 3$, and $\alpha = x^3$. I deliberately leave open what sorts of things $\gamma$ and $\beta$ may be (e.g., numbers, statements, formulas, etc.). $\alpha$ is correspondingly open; it is any procedure—deterministic or

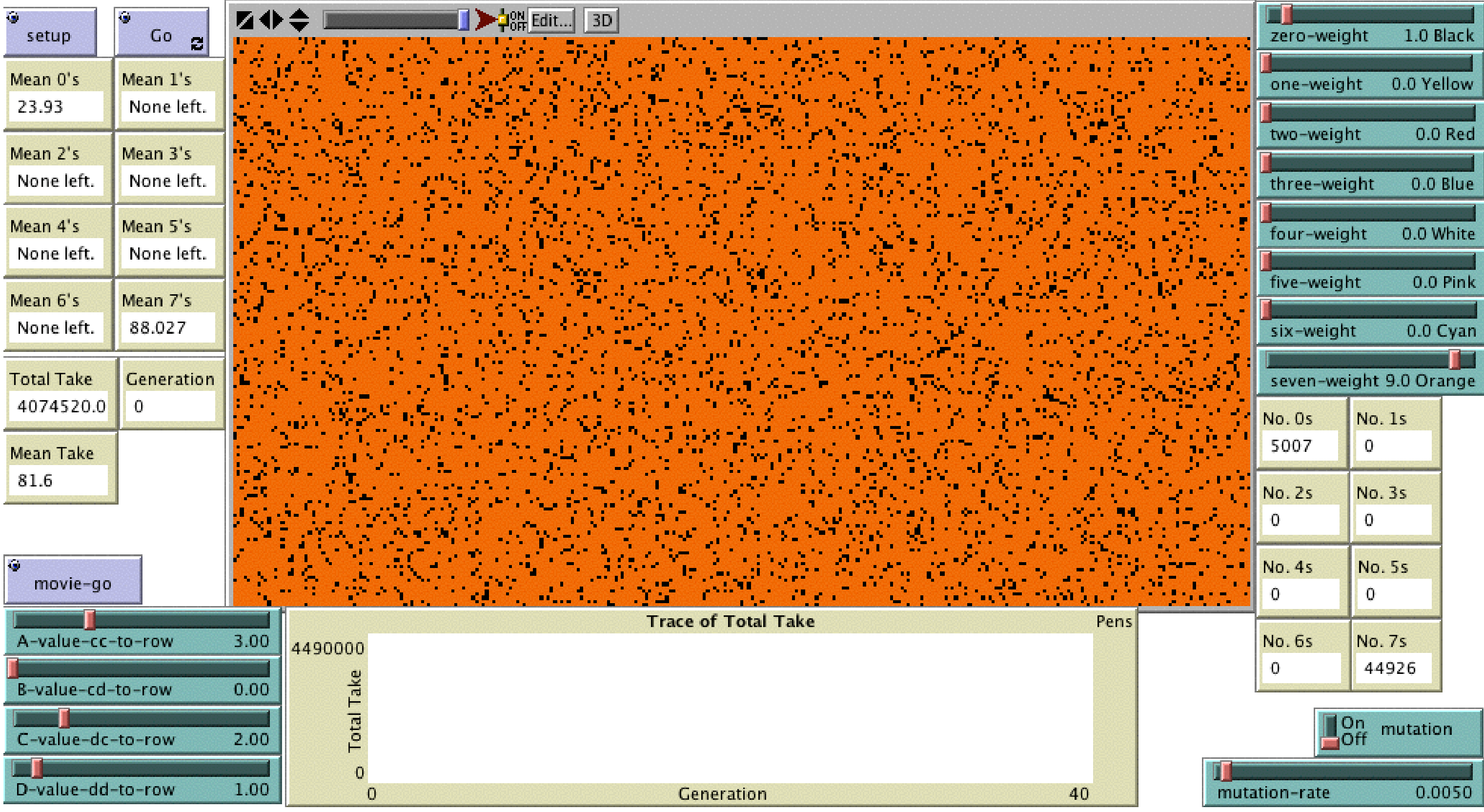


Figure 2.4: Setup for Iterated Play of Standard Stag Hunt on the Gridscape. Program: `m1-symmetric-2x2.nlogo`.

randomized—for producing $\gamma$ from $\beta$. $\gamma$ itself may be definite—*Rain tomorrow*—or probabilistic—*Chance of rain tomorrow greater than 0.7*.

Let us also say that $\gamma$ *is accessible from* $\beta$ if there is some (not necessarily specified) algorithm $\alpha$ such that $\gamma$ is accessible from $\beta$ via $\alpha$.

practicable accessibility

Under what conditions is something *not* accessible? Let us say that $\gamma$ *is practicably accessible from* $\beta$ *via* $\alpha$ if $\gamma$ is accessible without undue cost or delay from $\beta$ via $\alpha$. Quite clearly, much that is accessible (in principle) is not practicably accessible.

One way for absolute inaccessibility to occur is when applying $\alpha$ to $\beta$ yields more than one result. Then we have to say that any single result is not accessible from $\beta$ via $\alpha$. Equations with more than one root (solution) will perhaps be the most familiar example. Suppose that $y$, the position of an object, equals $t^2$ in minutes, with noon today set as $t = 0$. At what time is $y$ at 16? The equation allows two answers: 4 minutes before noon and 4 minutes after noon. So here we can say that $\gamma$ is *t = 4 minutes before noon or t = 4 minutes after noon* is accessible from the equation via mathematical solution. It would be incorrect to say that $\gamma$ is *t = 4 minutes before noon* is accessible from the equation via mathematical solution, even though $\gamma$ = *t is 4 minutes before noon* is consistent with the equation via mathematical solution. Accessibility resembles visibility under a microscope. A set (organelle) may be accessible (visible under the microscope) without its elements (components) being accessible (visible).

Back now to games. Let $\beta$ be a type A game with the assumption of fundamental rationality for all players. Let $\alpha$ be a procedure that selects player $R$'s strategy in every Nash equilibrium. Then in general neither $\gamma$ as $R$ *plays strategy* $r_1$ nor $\gamma$ as $R$ *does not play strategy* $r_1$ is accessible (from $\beta$ via $\alpha$). The Stag Hunt game illustrates the point. In some equilibria $R$ plays $r_1$, in some $r_2$, and in some a mixture. The Nash equilibrium concept does not reveal which equilibrium will obtain. Because player $R$ has more than one strategy involved in the equilibria, only a set (larger than 1) is accessible.

Practicable accessibility is also an issue for type A games. Under Zermelo's theorem (`http://www.ams.org/featurecolumn/archive/games3.html`, accessed 17 December 2005) any (type A) game that is finite, played by two players under perfect information (each player knows all the moves so far from the other player), and is strictly competitive is such that either the first player can force a win or a draw or the second player can force a win or a draw. Finding such an equilibrium, however, is another matter. Chess, checkers and many other board games qualify under the theorem, but their sizes and complexities preclude completion of the analysis. The equilibria of such a game are practicably inacces-

sible via the Nash equilibrium procedure (here, Zermelo's backwards induction procedure).

Some but not all type A games are not practicably accessible because of computational complexity. Some but not all are not accessible at all because of multiple equilibria. The situation is quite different for type B games. By hypothesis, the players each have adaptation regimes that select policies for play and these produce outcomes in atomic subgames. Convergence to a stable outcome or even stable distribution of outcomes may or may not occur, let alone convergence to an equilibrium. Because conditions of play may be stochastic, different runs of play may produce different results. We may think of $\gamma$ in the context of type B games ($\beta$) as a stream of outcomes and $\alpha$ as the adaptation regimes assigned to the several players. Conceived this way, something, some $\gamma$, is always accessible, practicably accessible, in type B games. They are designed to be that way.

At bottom, the distinction between type A and type B games is largely one of perspective or stance, of attitude we bring to the subject. In a type A model we are mainly interested in the Nash equilibria and are less concerned with accessibility issues.[4] The theory for type A games may be called *equilibrium game theory.* In a type B model we endow the players with policy spaces and adaptation regimes, which they deploy in conducting the game and from which their strategies emerge. We may call the theory for these games *effective game theory*. It investigates how game results emerge from the interplay of policy spaces and adaptation regimes.

effective game theory

effective rationality

Whatever rationality type B agents have is a *effective rationality*; play, $\gamma$, is produced by a policy in force, $\alpha$, typically relying on a history of play, $\beta$. I will have much to say later about effective rationality in games. For now, this observation. The perspective of effective game theory is likely to be apt under a number of conditions, including these:

1. When fundamental rationality cannot be assumed.

   Note that economic rationality is easily violated if agents have to rely on realistic perceptual mechanisms. Letting $a \sim b$ be interpreted as "$a$ is not perceptually distinguishable from $b$, " we can easily have $a \sim b$, $b \sim c$ and $a \succ c$. The relation $a \succ c$ might be used for such perceptual modalities as "is at least as hot as" and perhaps even "is worth at least as much as."

   Effective game theory (type B) representations may or may not assume fundamental rationality. In either case, the games will play out, agents will adapt and an evaluatable pattern of play will result.

[4]As usual, by "Nash equilibria" I mean Nash equilibria or refinements thereof.

2. When a type B game is a natural model of the system under examination, regardless of whether a type A representation is useful.

   The one-shot Prisoners' Dilemma is more naturally modeled as a type A game, while the indefinitely repeated version is usefully addressed from the perspective of effective game theory.

3. If at least some players are epistemically limited (nonhuman animals surely count) and cannot plausibly be assumed to be individually economically rational.

4. If the underlying game is too complex for meaningful type A (equilibrium) analysis.

   Equilibrium analysis for tic-tac-toe and One-Two-Twenty, constructivist analysis for chess and checkers.

5. If the underlying game has multiple equilibria after excluding implausible equilibria.

   The Folk Theorem (a real theorem; see standard texts such as [4, 38] for proofs and discussion) tells us that in indefinitely repeated games the equilibria proliferate.

6. If a procedure of play is sought by which the players can operate and with which explanations, predictions, and interventions may be made.

   There are special cases, such as when elimination of dominated alternatives produces a unique outcome, in which procedures are available for producing equilibria in type A games. With complexity or repetition, however, such cases are rare.

The main focus of this book is effective game theory (and effective rationality), in distinction to equilibrium theory (and economic rationality). I will throughout, however, honor and draw upon the concepts and insights of that approach. Before directly taking up effective analysis of games and their modeling with agent-based systems, a word on evaluation criteria. Details will be articulated during the analyses that follow.

In a effective (type B) game setup, we model players as agents with specific, albeit quite limited, powers of thought. Agents have a cognitive apparatus, $(\Pi, \rho)$, consisting of a consideration set of policies, $\Pi$, and an adaptation regime, $\rho$. Under equilibrium analysis there is only one criterion of evaluation: What are the

individually rational outcomes? Such parsimony is not available to us in the case of effective games. Instead, the list of criteria is best left open. Here is a list that serves to begin a fruitful discussion.

1. Performance against self.

   Does the apparatus do well against itself? If all agents use the apparatus, do the agents as a group prosper relatively well in the ambient environment?

2. Performance against others.

   Does the apparatus do well playing against other regimes that do well against themselves?

3. Exploitability.

   Is the apparatus catastrophically exploitable? Does it have weaknesses that may be discovered by another apparatus?

4. Robustness.

   Is the apparatus robust under perturbations of its parameters? Does the apparatus perform well against a field of others?

5. Learnability.

   Can the apparatus be parameterized in such a way that an agent can (easily) learn profitable, well-performing settings?

6. Computational cost.

   Is the apparatus computationally tractable? Is it simple or does it require excessive computational resources from the agent?

7. Informational requirements.

   Does the apparatus rely on plausibly available information? Or does it require information not likely to be available in the actual system being modeled?

The foregoing is enough to get us properly started. Let the games begin.

## 2.6 Bibliographic Notes

### 2.6.1 Social Order

Wikipedia has a useful entry for *social order*, `http://en.wikipedia.org/wiki/Social_order`. Here is an exerpt:

> Social order is a concept used in sociology, history and other social sciences.
>
> It refers to a set of linked social structures, social institutions and social practices which conserve, maintain and enforce "normal" ways or relating and behaving.
>
> Thus, a "social order" is a relatively stable system of institutions, pattern of interactions and customs, capable of continually reproducing at least those conditions essential for its own existence. The concept thus refers to all those facets of society which remain relatively constant over time.
>
> These conditions could include both property, exchange and power relations, but also cultural forms, communication relations and ideological systems of values.

My use here of *social order* is not tied to any particular doctrine or theory of society. Instead, I use the expression to refer generally to the regular, systematic patterns of behavior evidenced by a particular society (or collection of individuals).

Dawes [10], Hardin [32], and Elster [13, 14, 15] are excellent on connecting the study of strategic interaction with problems of the social order.

### 2.6.2 Emergence

The Wikipedia entry for emergence is quite good: `http://en.wikipedia.org/wiki/Emergence`.

NetLogo (`http://ccl.northwestern.edu/netlogo/`) comes with a library of models, many of which pertain to social order, emergence, or strategic interaction. The following NetLogo models, from the model library that comes with installation, are especially pertinent to the subject of emergence:

- `Flocking.nlogo`

- `Slime.nlogo`
- `Traffic Basic.nlogo`
- `Segregation.nlogo`
- `RumorMill.nlogo`

Thomas C. Schelling's work is seminal; see particularly [61]. Rodney Brooks on "situated agents" [8], Robert Frank's *Luxury Fever* [24], and John Holland's *Emergence* [35] all make stimulating reading. There is a terrific book on biology and emergence [9], and I would be remiss not to mention the delightful classic, *Vehicles* [6]. Grimm and Railsback [31] are very thoughtful on emergence and on agent-based models (the book uses the term *individual-based models* instead).

### 2.6.3 Rationality

Binmore [4] and Kreps [38] are good textbook treatments of economic rationality and game theory. Sen's essay, "Rational Fools," [66] is an indispensable antidote.

On the effective side, David Fogel's *Blondie24* [21] is a very readable account of his successful construction of a computer program that learned to play checkers at a high level.

# Chapter 3

# Games in the Wild

From the Wikipedia (`http://en.wikipedia.org/wiki/In_vivo`, accessed 2006-12-25):

> In vivo (Latin: (with)in the living) means that which takes place inside an organism. In science, in vivo refers to experimentation done in or on the living tissue of a whole, living organism as opposed to a partial or dead one. Animal testing and clinical trials are forms of in vivo research.

Games, or more descriptively *contexts of strategic interaction* (CSIs), are everywhere.[1] They pervade social situations and occur quite naturally (or appear "in the wild" as geneticists say of certain alleles). Two people play backgammon. They are in a game, or context of strategic interaction (CSI), because the reward (winning or losing) for each player depends at least in part on decisions made by the other player. One player cannot make a series of decisions that results in winning or losing, *independently of what the other player does.* The other player has to make decisions, too, and they matter. The context is interactive—two or more players are involved—and it is strategic because both players have interests, which they take into account in making their decisions.

[1]The term game is perhaps an unfortunate one for a number of reasons. It suggests a certain frivolity, also that only contexts of pure competition are of interest. More importantly, we need a distinction between a situation involving strategic interaction and a model of such a situation. *Game* gets used for both. When necessary to differentiate, I'll use game$_S$ for the situation, not always well defined with all vagueness left out, and game$_M$ for a model, presumably specified with great precision, of a game$_S$. Or, CSI for game$_S$ and game for game$_M$.

Backgammon is representative of many games in that it is purely competitive.[2] One player's win is the other player's loss. The interests of the players are, we may assume, entirely opposed. In other CSIs (or games) the players' interests are entirely coincident. These are what we call *games of pure coordination.* Two people are conversing by telephone when the connection is suddenly dropped. How should they attempt to resume the conversation? If both call back simultaneously both will get a busy signal or perhaps voice mail. They share a joint interest in mutually divining a decision that results in prompt and unfrustrated resumption of their conversation. Here we may assume the interests of the two agents are identical. Neither really cares who makes the new call, so long as it results in immediate resumption.

Lying between games of pure competition (e.g., backgammon) and games of pure coordination (e.g., resuming a broken phone call) are *mixed motive* games (or CSIs). A small group negotiates where to have dinner. No two people have identical preferences, but everyone agrees that failing to come to a congenial decision quickly is the worst outcome. Remarkably subtle moves will typically attend this familiar situation. Bluff, bluster, threat, compromise, accommodation, probing, retreating, appeal to norms, humor, and much else are routinely employed with facile skill by everyone who participates in such groups.

How are we to understand games? In particular, how are we to predict and explain both behavior and outcomes in games? This is a large and important question. I remind the reader that our mode here is to make some progress through an "iterative, probing, nondeterministic search for sharper clarity and deeper understanding." To this end, a rough characterization of our topic will be helpful:

> Games, or CSIs, essentially involve at least two *agents* (or players) who make *choices* and receive *rewards* (or payoffs).[3] The reward to an individual agent is based in part on its choices *and the choices made by the other agent(s)*, as well as the underlying structure of the situation.

Now consider a few representative, idiosyncratically-chosen examples of contexts of strategic interaction.

---

[2] At least approximately or often. Consider playing with a tyro and playing to lose for purposes of instruction.

[3] Later, it will be useful to distinguish *rewards,* which are received after each move in a strategic situation, and *returns,* which are the net of the rewards obtained in a multi-move strategic context. Unless otherwise noted, what I say about rewards applies to returns and *vice versa.*

## 3.1 War

Much more than a pure, brutal contest of strength, war has been recognized from the earliest writings as a field of interactive decision making. Deception especially has been and remains a primary theme; it is inherently a strategic concept. Think of the Trojan horse incident told in the *Iliad* and the story of the Cyclops in the *Odyssey.* Think of the elaborate obfuscations undertaken by the Allies in World War II concerning the time and place of D-Day. Sun Tzu, in *The Art of War* (http://www.chinapage.com/sunzi-e.html), the oldest known military treatise, wrote this:

> 18. All warfare is based on deception.
> 19. Hence, when able to attack, we must seem unable; when using our forces, we must seem inactive; when we are near, we must make the enemy believe we are far away; when far away, we must make him believe we are near.
> 20. Hold out baits to entice the enemy. Feign disorder, and crush him.
> 21. If he is secure at all points, be prepared for him. If he is in superior strength, evade him.
> 22. If your opponent is of choleric temper, seek to irritate him. Pretend to be weak, that he may grow arrogant.
> 23. If he is taking his ease, give him no rest. If his forces are united, separate them.
> 24. Attack him where he is unprepared, appear where you are not expected.
> 25. These military devices, leading to victory, must not be divulged beforehand.

Other themes abound, but deception and surprise remain keystones to military strategy. Other works on the short list of classics in military strategy include: *On War,* by Karl von Clausewitz, *The Prince,* by Niccola Machiavelli, and *A Book of Five Rings,* Miyamoto Musashi.[4] Liddell Hart, e.g., [43] is an especially persuasive spokesman for the importance of military deception and surprise. Thomas Schelling is uniformly insightful and a joy to read, e.g., [62, 60, 61, 63]. The

[4] You may find these on-web at http://www.gametheory.net/html/books.html#classics.

*Memoirs* of Ulysses S. Grant are chock full of material to stimulate reflection on war and on interactive decision making in general. Here is my favorite passage. Grant is describing his first field command in the American Civil War.

> My sensations as we approached what I supposed might be a 'field of battle' were anything but agreeable. I had been in all the engagements in Mexico that it was possible for one person to be in; but not in command. If someone else had been colonel and I had been lieutenant-colonel I do not think I would have felt any trepidation. ...As we approached the brow of the hill from which it was expected we would see the enemy... my heart kept getting higher and higher until it felt as though it was in my throat. I would have given anything to have been back in Illinois, but I had not the moral courage to halt and consider what to do; I kept right on. When we reached a point from which the valley below was in full view I halted. The place where the Confederates had been encamped was still there but the troops were gone. My heart resumed its place. It occurred to me at once that [Colonel Thomas] Harris had been as much afraid of me as I had been of him. This was a view of the question I had never taken before; but it was one I never forgot afterwards. From that event to the close of the war, I never experienced trepidation upon confronting an enemy, though I always felt more or less anxiety. I never forgot that he had as much reason to fear my forces as I had his. The lesson was valuable.
>
> —Ulysses S. Grant, *Memoirs*

## 3.2 Trading and Investing

"Buy low, sell high" is great advice if (and only if) you know what to do. As the song says about that special form of trade and investment called love, "Nice work if you can get it, And you can get it if you try."

Examples of buying low or selling high? This is from *The Reader's Digest,* June 2003, pages 76–7:

> Customers at The Home Depot who overestimate how much paint they need return the unopened cans, which are stocked in the "Oops Paint" section. The "remnant" paint—perfect for bathrooms and other small projects—sells for $5 a gallon and $1 a quart (regular gallon

> prices are $21 to $25). "And it's not all chartreuse," says The Home Depot spokesperson Mandy Holton. "There are usually a lot of great neutrals." Best time to buy: Sundays and Mondays, because folks return unwanted paint over the weekend.

More generally, traders and investors are in the business of finding assets that are either under-valued or over-valued in the market. In other words, they seek opportunities for *risky arbitrage.* Risky because—unlike the paint at The Home Depot—the values of the assets in question are typically not known with much certainty. Arbitrage because the traders are looking to buy assets that are under-priced (and then resell them at their proper prices) or looking to sell ("unload") assets that are over-priced. In any event, the trick is to have and exploit knowledge that is superior to what is represented in the market. The nature of this knowledge and the means of getting it vary greatly. An investor in equities may look deeply and carefully at the fundamentals of the companies. Which are and which are not well managed, well positioned, in possession of new products and alliances? An investor may look at the "technical" data, the trends and other movements in prices. In the extreme, so-called day traders do this in real time, attempting to out-guess the market, that is out-do the other traders in discerning what is over-valued or under-valued. Note that investing on analysis of fundamentals would seem to have less strategic content than investing on technical grounds.

On-line, Internet-based examples are readily available for those who wish to trade or just to study and learn. Tradesports (http://www.tradesports.com) is a real-time, on-line trading market that affords an excellent case for study. Academic analogs—but with real money if you want—are available for elections at Iowa Electronic Markets (IEM; http://www.biz.uiowa.edu/iem/; http://www.biz.uiowa.edu/iem/markets/) and the University of British Columbia Election Stock Market (http://esm.ubc.ca/). A *BusinessWeek* article, "The 'Election Futures' Market: More Accurate than Polls?"[5] presents the case in a popular format that these markets predict election outcomes better than opinion polls. The Bush administration even toyed with creating a similar market for the purpose gauging intelligence in the Middle East (see "Betting on Terror: What Markets Can Reveal" by Floyd Norris in *The New York Times,* August 3, 2003). The idea was dropped after being exposed to public ridicule. Is it ridiculous? Consider: What would it take to "game" (distort for ulterior purposes) these markets? When would anyone want to do this? What might be done to prevent manipulation? Does it make sense to have an SEC for markets in international affairs?

[5] 1996; http://www.businessweek.com/1996/46/b3501116.htm.

There are always the public equity markets. Consider this comment on the bond market by a Salomon trader in the 1980s.

> The men on the trading floor [Salomon's bond trading area] may not have been to school, but they have Ph.D.'s in man's ignorance. In any market, as in any poker game, there is a fool. The astute investor Warren Buffett is fond of saying that any player unaware of the fool in the market probably *is* the fool in the market. In 1980, when the bond market emerged from a long dormancy, many investors and even Wall Street banks did not have a clue who was the fool in the new game. Salomon bond traders knew about fools because that was their job. Knowing about markets is knowing about other people's weaknesses. And a fool, they would say, was a person who was willing to sell a bond for less or buy a bond for more than it was worth. A bond was worth only as much as the person who valued it properly was willing to pay. And Salomon, to complete the circle, was the firm that valued the bonds properly.
>
> —*Liar's Poker,* Michael Lewis [40, page 35]

## 3.3 Athletic Contests

There are sports, called games in common parlance, that have little or no strategic content. They amount more or less to contests of skill. Among them are golf, bowling, darts, skiing, track and field events, and bob sledding. Still other competitive games, such as billiards, have strategic content only with fairly advanced play. These are not, for the most part, of interest as CSIs, contexts of strategic interaction, and will not concern us further.

Many other athletic contests most unambiguously count as CSIs. Baseball has given us a wonderful strategic slogan, entirely appropriate for war and other games: "Hit 'em where they ain't." Wee Willie Keeler hit .432 in 1897. Asked how a man of his diminutive size could put together such an average, Keeler responded: "Simple. I keep my eyes clear and I hit 'em where they ain't."[6] Deception—or the fake-out—plays as prominent a role in these athletic contests as it does in warfare. Think of the pitcher-batter duel in baseball, the fake-out moves in basketball, or the mixing of plays in American football.

[6] From http://www.baseballtips.com/slang.html.

Management of sports teams is as much a matter of strategic interaction as the play itself. In *Moneyball: The Art of Winning an Unfair Game,* Michael Lewis describes how the Oakland A's baseball team, with consistently small amounts spent on player salaries, is consistently able to contend in major league baseball and reach the playoffs [41]. In two words: risky arbitrage. The A's, and in particular their general manager Billy Bean, have identified predictive measures superior to those used by other teams, for example using on-base percentage instead of batting average to evaluate the worth of a batter. Better measures of value allow the A's to 'buy' (hire) under-priced players. They may not have the best team in baseball, but they have one of the best. Their efficiency in the sense of what it costs them to win a game is tops and they operate at a profit in a media market dominated by the San Francisco Giants baseball team.

## 3.4 Gambling

Many forms of gambling do not involve strategy or even much skill. Examples include playing slot machines and playing roulette. Not so with poker at the professional level. Poker is prototypical. It is to competitive games of strategy what robins are to birds: a standard, familiar, readily available example, displaying in typical form many of the characterizing features of the subject. Everyone over time gets roughly the same quality of hands, yet there is an enormous difference among players in their success rates. Everyone can count cards and figure the odds. What actually matters is bluffing, reading your opponents (discerning their "tells," behavior such as slamming down chips that indicates what is in their hands), and preventing your opponents from reading you. The following fine passage from a master poker player is well worth quoting at length:

> Let's take a quick glimpse at the high-stakes poker world, an enterprise that yields several of my friends over a million dollars a year! At this level, too, luck is a factor on any given day, week, or month, but what's different is that if you play better poker than your opponents do, pretty consistently, you'll find that over almost any *two*-month period your winnings have exceeded your losses. Furthermore, if you play better poker than your opponents over a *six*-month period, your results will have moved very solidly in the winning direction. Making a few well-timed bluffs each day will add up to a lot of money each year!

> In fact, if an inexperienced poker player were to sit down for a few hours with a group of world-class poker players, he would have virtually no chance to win over even an eight-hour period. This very fact is why five or six top pros might be willing to sit down in the same game with this fellow and each other: the money that even one amateur is likely to contribute makes it work their while to do battle with so many respected opponents.
>
> This is why so many of the top poker players today drive fine cars and live in palatial homes [the author of this passage lives with his family in Palo Alto]. Right now, as you're reading this book, there is a $600–$1,200-limit poker game at the Bellagio Casino in Las Vegas and a $400-$800-limit poker game at the Commerce Casino in Los Angeles. There is . . .
>
> If that's not enough action for you, four nights a week in Los Angeles, there is a $2,000-$4,000-limit Seven-Card Stud game a Larry Flynt's Hustler Club Casino, with Larry himself often playing. In the $400-$800-limit poker game it's easy to take a $25,000 swing in one hour. In the $2,000-$4,000-limit game, where movie stars, former governors, and billionaires play, it's not uncommon for someone to win or lose $250,000 in one night. In these "nosebleed" poker games (the term refers to the altitude of the stakes), strategy, discipline, calculation of the odds, and practiced observation contribute to a game that involves much more skill. Better play wins more hands in the long run.
>
> –*Play Poker Like the Pros* by Phil Hellmuth, Jr., 2003 [33, pages 4–5]

The society of poker players has given us an important concept—the *tell*—not only for poker but for CSIs generally. *Webster's Unabridged Dictionary* finds only two senses for *tell* as a noun. Quoting:

1. something that is told : TALK, TALE, ACCOUNT

   "have a tell with you —Eden Phillpott"

2. [Ar *tall*]

   : HILL, MOUND

   specif : an ancient mound in the Middle East composed of remains of successive settlements — compare TEPE

The *Oxford English Dictionary* is no more helpful. The Wikipedia gets it right. Quoting:

> Tell (poker)
>
> From Wikipedia, the free encyclopedia.
>
> In poker, a tell is a detectable change in a player's behavior that gives clues to that player's hand. Possible tells include leaning forward or back, placing chips with more or less force, fidgeting, changes in breathing or tone of voice, direction of gaze and actions with the cards, cigarettes, or drinks.
>
> For example, a player with a weak hand, hoping to bluff, may throw his chips into the pot forcefully and with a direct gaze at a player he hopes to discourage from calling.
>
> Tells may be common to a class of players or unique to a single player. A player gains an advantage if she observes another player's tell, particularly if that action is unconscious and reliable. However, better players may fake tells, hoping to lead their opponents into costly traps when they rely on the false information. So the observing, creating, and evaluating of tells can add another level to the play of poker.
>
> Mike Caro has published the most comprehensive information on tells; his *Book of Tells* (ISBN 0897461002) is now a standard reference on the subject.
>
> David Mamet's 1987 movie *House of Games* includes an interesting discussion and visual reference to tells as an essential part of the plot. The 1998 movie *Rounders* contains an even more subtle use of strategy: at one point, "Mike" discovers a tell in his opponent (that he eats cookies in a particular way after he has bet a very strong hand), and after using it once, he reveals to the opponent that he has this tell; although this eliminates the usefulness of the tell itself, it upsets his opponent so much that it affects his later play.
>
> –*Wikipedia,* http://www.wikipedia.org/wiki/Tell_(poker)

The Wikipedia is also better than the dictionaries on *tell* as a noun, not in the context of poker. Quoting:

> Tell
>
> From Wikipedia, the free encyclopedia. A tell (Arabic, or tel, Hebrew) is a mound site formed through successive human occupation over a very long timespan.
>
> The word is used as a term in archaeology, particularly Middle-Eastern archaeology. It is sometimes used in a toponym, that is, as part of a town or city name, the most well known example being the city of Tel Aviv. Often a modern city is located next to an ancient mound with a similar tell name, for example the city of Arad is a few kilometers (miles) away from an ancient mound called Tel Arad.
>
> This article is a stub. You can help Wikipedia by fixing it.
>
> External link
>
> * http://www.webref.org/anthropology/t/tell.htm
>
> Tell is an English verb meaning "to speak to" or "to talk to"; also "to give an order". For more information on what that is, see talking.
>
> Retrieved from "http://www.wikipedia.org/w/wiki.phtml?title=Tell" This page was last modified 01:47, 19 Jun 2003. All text is available under the terms of the GNU Free Documentation License.

Note the meaning associations and similarities between these two senses of *tell.* The strategic sense of *tell* is lucidly on display in the following passage about the great baseball player and base runner, Rickey Henderson.

> But Kennedy knew how devastating stealing could be: he had been with the San Francisco Giants in the 1989 World Series, when Henderson and the A's swept the Giants in four games and Henderson set a post-season record, with eleven stolen bases.
>
> Henderson agreed to give a demonstration, and there was a buzz as Goodman, Johnson, and the other players gathered around first base. Henderson stepped off the bag, spread his legs, and bent forward, wiggling his fingers. "The most important thing to being a good base stealer is you got to be fearless," he said. "You know they're all coming for you; everyone in the stadium knows they're coming for you. And you got to say to yourself, 'I don't give a dang. I'm gone.' " He said that every pitcher has the equivalent of a poker player's "tell," something that tips the runner off when he's going to throw home.

> Before a runner gets on base, he needs to identify that tell, so he can take advantage of it. “Sometimes a pitcher lifts a heel, or wiggles a shoulder, or cocks an elbow, or lifts his cap,” Henderson said, indicating each giveaway with a crisp gesture.
>
> Once you were on base, Henderson said, the next step was taking a lead. Most players, he explained, mistakenly assume that you need a big lead. “That’s one of Rickey’s theories: Rickey takes only three steps from the bag,” he said. “If you’re taking a big lead, you’re going to be all tense out there. Then everyone knows you’re going. Just like you read the pitcher, the pitcher and catcher have read you.”
>
> He spread his legs again and pretended to stare at the pitcher. “O.K., you’ve taken your lead; now you’re ready to find that one part of the pitcher’s body that you already know tells you he’s throwing home. The second you see the sign, then *boom*, you’re gone.” [28, page 58]

Hellmuth’s book has a great deal of information on Texas Hold ’em, which is generally the most popular form of poker in tournaments and is the variety of poker played at the World Series of Poker each year at Binion’s Horseshoe Hotel & Casino in Las Vegas.[7] *Positively Fifth Street* [49] by Jim McManus, a published poet, novelist, and professor, describes the 2001 World Series of Poker and the world around it. Strategic insight abounds in both works.

## 3.5 Business Strategy

The gambit, a term from chess, is a favorable trade, which the opponent may or may not realize is happening. The player offering the gambit offers a comparatively small loss in exchange for a larger gain in position or other form of resource. Here is something very like a gambit played big time in business.

> Analysts called it “Marlboro Friday”—Philip Morris announced on April 2, 1993 that it would reduce the U.S. price of its premium

[7] See http://conjelco.com/wsop.html and http://www.binions.com/home.asp. The Wikipedia has an introduction to Texas Hold ‘em: http://www.wikipedia.org/wiki/Poker/Texas_holdem. At the World Series of Poker, No Limit Texas Hold ’em is the game. “No Limit” means that the largest bet permitted is the size of the current wealth outside the pot of the poorest player still in the hand. Once a player has bet all of his or her chips, the player is said to be “all in,” since the player’s wealth is all in the pot. Once a player is all in for a particular hand, other players may call but may not raise.

> brand of cigarettes by 20%. The tobacco manufacturer also said it would increase the budget for its domestic advertising by a substantial amount. R.J. Reynolds, Philip Morris's biggest competitor, responded by matching the price cut on its own premium brands (Camel and Winston among them) and by pouring more money into its own domestic advertising.
>
> The pricing war that ensued cost both companies tens of millions of dollars. But was the domestic market share the real reason Philip Morris lowered the price of Marlboro cigarettes? Consider that just as R.J. Reynolds had depleted its cash resources trying to keep up with its chief opponent, Philip Morris was expanding aggressively into the Eastern European market, investing $800 million in Russia and other regions that were formerly part of the Soviet Union. R.J. Reynolds was in no position to fight back, having spent so much money to maintain its market share in the United States, and Philip Morris won the battle for Eastern European market share, hands down.
>
> –"Global Gamesmanship" by Ian C. MacMillan et al., *Harvard Business Review,* May 2003 [45]

Sometimes you can even make a profit on a gambit:

> One day earlier in his career [Robert] Dall was in the market to buy (borrow) fifty million dollars. He checked around and found the money market was 4 to 4.25 percent, which meant he could buy (borrow) at 4.25 percent or sell (lend) at 4 percent. When he actually tried to buy fifty million dollars at 4.25 percent, however, the market moved to 4.25 to 4.5 percent. The sellers were scared off by a large buyer. Dall bid 4.5. The market moved again, to 4.5 to 4.75 percent. He raised his bid several more times with the same result, then went to Bill Simon's office to tell him he couldn't buy money. All the sellers were running like chickens.
>
> "Then you be the seller," said Simon.
>
> So Dall became the seller, although he actually needed to buy. He sold fifty million dollars at 5.5 percent. He sold another fifty million dollars at 5.5 percent. Then, as Simon had guessed, the market collapsed. Everyone wanted to sell. There were no buyers. "Buy them back now," said Simon when the market reached 4 percent. So Dall

> not only got his fifty million dollars at 4 percent but took a profit on the money he had sold at higher rates. *That* was how a Salomon bond trader thought: He forgot whatever it was that he wanted to do for a minute and put his finger on the pulse of the market. If the market felt fidgety, if people were scared or desperate, he herded them like sheep into a corner, then made them pay for their uncertainty. He sat on the market until it puked gold coins. *Then* he worried about what he wanted to do.
>
> —*Liar's Poker,* Michael Lewis [40, page 88]

## 3.6 Negotiation

Negotiation exemplifies strategic interaction *par excellence.* After all, there is no point in negotiating if your counter-party's actions don't matter to you. Familiar as negotiation is to everyone, it is useful to be reminded that often negotiation is not explicit, at least not at first. Here is a description of this sort of encounter "in the wild."

> To begin to negotiate the environment does not, of course, mean that you enter the negotiation with a clear-cut goal in mind. A clear-cut goal is not needed even in purely human negotiations. Suppose you pass a stall in a market every week and notice an antique ornament for sale. At first it seems ugly, but as it grows familiar, you catch yourself wondering how it would look on your shelf. One day it rains while you are crossing the market and you take shelter in the stall. the ornament is still there; for something to do you ask its price. Even when a low price is mentioned you automatically snort in contempt, for you have no intention of buying. . . or have you? During the week that follows you decide that the price really was low and think of a friend who has a birthday soon and might like it. Next week you stop and begin to bargain.
>
> When did the negotiation begin? When you started to bargain? Or earlier, when you asked the price? Or earlier still, when you first noticed the ornament among an anonymous heap of others? Pointless to say, as pointless as to say where mind began.
>
> –*Language and Species* by Derek Bickerton [3, pages 234–5]

## 3.7 Coordination, Symbiosis, Mutualism, Cooperation

Contexts of strategic interaction are not all adversarial in the sense that one agent's gain is another's loss (so-called *constant-sum* or equivalently *zero-sum* games). In *coordination games* all players gain if they can arrive at a common outcome and lose if they fail. Think of the game of finding someone you have separated from during a shopping trip. You both wish to meet up again, but did not plan for the separation and have no easy means of communication. Schelling's early treatment of such games is masterful and well worth reading today [62].

Biologists have named and studied several kinds of interactive decision making that—in terms of game theory lingo—is not constant-sum. Symbiosis and mutualism are two of the most important for our purposes.

> Symbiosis (pl. symbioses) is an interaction between two organisms living together in more or less intimate association or even the merging of two dissimilar organisms.
>
> The term host is used for the larger of the two members of a symbiosis. The smaller member is called the symbiont.
>
> Symbiosis may be divided into two distinct categories: ectosymbiosis and endosymbiosis. In ectosymbiosis, the symbiont lives on the body surface of the host, including the inner surface of the digestive tract or the ducts of exocrine glands. In endosymbiosis, the symbiont lives in the intracellular space of the host.
>
> An example of mutual symbiosis is the relationship between anemonefishes of the genus Amphiprion (family, Pomacentridae) that dwell among the tentacles of tropical sea anemones. The territorial fish protects the anemone from anemone-eating fish, and in turn the stinging tentacles of the anemone protects the anemone fish from its predators (a special mucous on the anemone fish protects it from the stinging tentacles).
>
> The biologist Lynn Margulis, famous for the work on endosymbiosis, contends that symbiosis is a major driving force behind evolution. She considers Darwin's notion of evolution, driven by competition is incomplete, and claims evolution is strongly based on co-operation, interaction, and mutual dependence among organisms. According to

> Margulis and Sagan (1986), *Life did not take over the globe by combat, but by networking.*
>
> (From: http://www.wikipedia.org/wiki/Symbiosis.)

(See [46] for a recent treatment of this theme by Margulis and Sagan.)

> Mutualism is a interaction in which both organisms in a close relationship derive some degree of benefit. Mutualism is usually temporary or not obligatory.
>
> (From: http://www.wikipedia.org/wiki/Mutualism.)

Lichens—those familiar greenish splotches on trees and rocks—present a most striking example of symbiosis.

> Lichens have been described as "dual organisms" because they are symbiotic associations between two (or sometimes more) entirely different types of microorganism -
>
> - a fungus (termed the mycobiont)
> - a green alga or a cyanobacterium (termed the photobiont).[8]
>
> There are many examples of symbiosis in nature, but lichens are unique because they look and behave quite differently from their component organisms. So, lichens are regarded as organisms in their own right and are given generic and species names. However, for taxonomic purposes the names are actually fungal names: lichens are regarded as a special group of fungi - the lichenised fungi.
>
> There are an estimated 13,500 to 17,000 species of lichens, extending from the tropics to the polar regions. Some of them grow on the bark of temperate trees or as epiphytes on the leaves of trees in tropical

[8] Author's note: From the Wikipedia, www.wikipedia.org/wiki/Cyanobacteria: "Cyanobacteria or blue-green bacteria are a group of aquatic bacteria that obtain their energy through photosynthesis. They are often referred to as blue-green algae, even though it is now known that they are not related to any of the other algal groups, which are all eukaryotes. Nonetheless, the description is still sometimes used to reflect their appearance and ecological role. Fossil traces of cyanobacteria have been found from around 3800 million years ago, making cyanobacteria some of the earliest living things known." Implied but not said, the cyanobacteria are prokaryotes, of ancient origin and lacking a cell membrane. It is thought by some that photosynthesizing plants acquired or incorporated the genomes of photosynthesizing bacteria.

> rain forests. Others occupy some of the most inhospitable environments on earth, growing on cooled lava flows and bare rock surfaces, where they help in the process of soil formation, and on desert sands where they help to stabilise the surface and enrich it with nutrients (see Cyanobacteria [cf., footnote 8, page 59]). Some other types of lichen grow abundantly on tundra soils, providing a vital winter food source for animals (including reindeer and caribou) in arctic and subarctic regions.Yet other lichens grow on or in the perennial leaves of some economically important tropical crop plants such as coffee, cacao and rubber, where they are regarded as parasites.
>
> All these features make lichens interesting and significant in environmental terms. But lichens also pose challenging scientific problems - how do two or more microorganisms interact at the cellular, genetical and biochemical levels to produce a unique, hybrid organism?
>
> (From: http://helios.bto.ed.ac.uk/bto/microbes/lichen.htm.)

A form of emergence occurs with lichens. Surprisingly, what appears to be, and in many ways is, a single individual is actually composed of, arises through the interactions of, individuals from two distinct biological kingdoms. It is even surprising who first noticed this underlying, symbiotic structure.

> Lichens are unusual creatures. A lichen is not a single organism the way most other living things are, but rather it is a combination of two organisms which live together intimately. Most of the lichen is composed of fungal filaments, but living among the filaments are algal cells, usually from a green alga or a cyanobacterium.
>
> In many cases the fungus and the alga which together make the lichen may each be found living in nature without its partner, but many other lichens include a fungus which cannot survive on its own – it has become dependent on its algal partner for survival. In all cases though, the appearance of the fungus in the lichen is quite different from its morphology as a separately growing individual.
>
> The true identity of lichens as symbiotic associations of two different organisms was first proposed by Beatrix Potter, who is best remembered for her children's books about Peter Rabbit. In addition to her books, she spent time studying and drawing lichens. Her illustrations are still appreciated for their detailed and accurate portrayal of the delicate beauty of these bizarre organisms.

(From http://www.ucmp.berkeley.edu/fungi/lichens/lichens.html.)

(Searching Google's image base on "lichens" turns up an excellent collection of images.)

Next, cooperation is—in its prototypical sense—a human social phenomenon, one that has been much noticed and remarked upon by social scientists, including game theorists. Cooperation, or roughly non-greedy behavior, has been called "the cement of society" [15] (by analogy with causation, which Hume called "the cement of the universe"). Without it, in the pungent phrasing of Thomas Hobbes, there would be

> no place for industry, because the fruit thereof is uncertain; and consequently no culture of the earth; no navigation, nor use of the commodities that may be imported by Sea; no commodious Building; no Instruments of moving and removing such things as require much force; no Knowledge of the face of the Earth; no account of Time; no Arts; no Letters; and which is worst of all, continuall feare, and danger of violent death; And the life of man, solitary, poore, nasty, brutish, and short. (Hobbes, *Leviathon*)

(See, e.g., http://plato.stanford.edu/entries/hobbes-moral/, http://www.philosophypages.com/ph/hobb.htm.) Without cooperation we are lost. How, then, does it arise and how might it be sustained? Hobbes thought that realistically it was necessary to turn power over to a sovereign (king or powerful government)—a leviathon—who would enforce cooperation on society. Others have thought that perhaps cooperation could emerge and be sustained naturally, without a central authority, much as, say, lichens emerge and are sustained naturally. Is this possible? If so, what is required of the games and the players?

## 3.8 Conversation

When we speak we have in mind how others will react to what we say and what we do not say. In this regard, a representative news story—"Official's comments set off euro's surge. U.S. Treasury's Snow said a weaker dollar would help U.S. exports. The dollar fell against the euro." by David McHugh—appeared in *The Philadelphia Inquirer* on May 13, 2003. The first sentence says it all: "The U.S. dollar fell to another four-year low against the euro yesterday, inching closer to its all-time low, after U.S. Treasury Secretary John Snow said a weaker dollar would

help U.S. exports." Secretary Snow never said he favored letting the dollar fall, but what he did say, as he no doubt understood, led the markets to infer that he favored a decline in the dollar. This form of strategic interaction is rife in linguistic communication and even has a special name: conversational implicature. Examples abound. A sign at Big Sur Lodge, Pfeiffer State Park, near a food counter:

Stressed?
Spelled
Backwards
Is
Desserts

Translation: Buy a dessert from us; it'll make you feel good. Or the concluding line in Hitchcock's movie, "Frenzy": "Mr. Rusk, you're not wearing your tie." Translation: You're the necktie murderer and I'm placing you under arrest. Or the use of irony, as in "Rick, Major Strasse is one of the reasons the German Reich enjoys the reputation it has today," from the movie "Casablanca." Translations: (to Strasse) The Reich is an impressive accomplishment and you are a big part of it; (to Rick) Watch out, this guy Strasse is a very bad man. Or the dialog-less eating scene in the movie "Tom Jones" with Albert Finney. Translation: This is just foreplay foreplay; the best is yet to come. See Paul Grice's "Logic and Conversation" [29] for the original treatment, still worth reading.

## 3.9 Games against Yourself

The long and justly celebrated story from the *Odyssey* of Ulysses and the Sirens continues to enchant and inform us. (Jon Elster has even written an entire book relating the story to modern social science [13].) From the perspective of strategic interactions, the story may be interpreted as a game played by Ulysses at one time against Ulysses at another time. At $t_0$, before approaching within earshot of the Sirens, Ulysses foresees that Ulysses at $t_1$, within earshot, will have preferences and inclinations quite at variance from Ulysses at $t_0$ and from Ulysses at $t_2$, post the encounter with the Sirens (if he should live that long). So Ulysses at $t_0$ cleverly prevents Ulysses at $t_1$ from acting as Ulysses at $t_1$ would prefer. He hears the Sirens and lives to tell the tale.

Robert Louis Stevenson's familiar story, *Dr. Jekyll and Mr. Hyde,* carries a similar theme. Thomas Schelling tells a fable about a man who is struggling to quit

smoking. A friend who smokes arrives at his house, converses, and leaves without incident. The friend, however, forgets his jacket and our protagonist notices the jacket contains a package of cigarettes. Not having an immediate compulsion to smoke and knowing the friend will return tomorrow, he puts the jacket away. Later, upon reflection, he recovers the jacket, removes the cigarettes, and destroys them.

## 3.10 Confidence Games

The con man (or woman) first gets your trust, your confidence, and then abuses it for profit. "Take the money and run" is the operating creed. Confidence rackets are celebrated in literature, theater, and film. Examples include Herman Melville's novel *The Confidence Man,* Thomas Mann's novel *Confessions of Felix Krull, Confidence Man: The Early Years,* Sinclair Lewis's novel *Elmer Gantry* (movie with Burt Lancaster and Jean Simmons), Jim Thompson's novella *The Grifters* (movie with Anjelica Huston, John Cusack, and Annette Bening), Guy Owen's short story "The Flim-Flam Man" (now published as *The Ballad of the Flim-Flam Man,* Coastal Carolina Press, May 2000; movie with George C. Scott and Sue Lyon), N. Richard Nash's play *The Rainmaker* (movie with Burt Lancaster and Katherine Hepburn), David Mamet's movie "House of Games" (with Lindsay Crouse and Joe Mantegna), and Meredith Wilson's Broadway musical *The Music Man* (movie with Robert Preston and Shirley Jones). This is from a Penn Web site, August 2003:[9]

```
6:30 pm - 8:30 pm     Confidence Games at the GSC

The GSC shows films about con artists:
Catch Me If You Can on 7/31;
The Thomas Crown Affair on 8/7;
The Spanish Prisoner on 8/14; and
The Grifters on 8/21.

Location:  Graduate Student Center, 3615 Locust Walk
Category:  Film
More info:
```

[9]See http://www.gametheory.net/ for yet another list of game-related movies, as well as lots of useful material on game theory.

http://www.upenn.edu/gsc/programs/film.htm#con

Con games lie at the core of much detective fiction and fact, as well as recently popular email scams. There is a confidence business, indeed an industry, with its own lessons and skills. (This takes us beyond the scope of the book. Those wishing to go further might consult such works as *How to Become a Professional Con Artist,* by Dennis M. Marlock.)

## 3.11 Statesmanship

Ending this list on a less cynical note, George Washington is understood to have been a politically ambitious man throughout his life. He actively, deliberately sought and schemed for the power, influence, and adulation he ultimately received. Washington notoriously wore his military uniform during the deliberations on the Declaration of Independence, just to remind the other delegates of his availability for command. In pursuing his ambitions Washington consistently and consciously followed a strategy of seeking rewards by actually deserving to get them. Resigning from the army at the end of the Revolution, an unexampled act, was a move calculated to make him fit for political leadership in a democracy. Declining to run for a third term as president was a move calculated to secure the success of the new country and of Washington's legacy.

Napoleon on his deathbed and in prison lamented that "They expected me to be another Washington."

## 3.12 Bibliographic Notes

The definitive natural history of games, of games in the wild, is yet to be written. Until—and even after—it is, Dixit and Nalebuff's very readable, *Thinking Strategically* [12], will be useful to anyone with a general interest in strategic interaction. Because games can make good stories, a number of popular books, aimed at the general reader, have appeared, offering informed and realistic accounts of particular sorts of strategic contexts. Brandenburger and Nalebuff's *Co-opetition* [7] addresses business strategy. Michael Lewis's *Liar's Poker* [40] is an often hilarious memoir of a bond trader. William Poundstone's *Fortune's Formula* [53] is a fascinating and insightful take on investing and its relation to gambling and the world of gambling. Lewis's *Moneyball: The Art of Winning an Unfair Game* [41] applies the investor's concept of arbitrage to the game of

baseball. James McManus's *Positively Fifth Street: Murderers, Cheetahs, and Binion's World Series of Poker* [49] tells a lurid, amusing, and ultimately insightful story about the Texas Hold 'Em variety of poker and its milieu. Poundstone's *Prisoner's Dilemma: John von Neumann, Game Theory and the Puzzle of the Bomb* [52] is an intellectual history for the general reader. It describes the people and the ideas (game theory, computation, and nuclear weapons) present and in play at the creation of game theory. (See also `http://en.wikipedia.org/wiki/John_von_Neumann`.) Jared Diamond's *Collapse: How Societies Choose to Fail or Succeed* [11] pushes the envelope of historical method in examining entire societies and their strategic choices. Robert M. Sapolsky is also a biologist who writings often shed light on contexts of strategic interaction, for example "A Natural History of Peace" [59] and *Monkeyluv: And other Essays on Our Lives as Animals* [58]. Robert H. Frank regularly produces readable and insightful books on strategic interaction from an economic perspective, for example *Passions within Reason: The Strategic Role of Emotions* [22], *The Winner-Take-All Society: Why the Few at the Top Get So Much More Than the Rest of Us* [23], *Luxury Fever: Why Money Fails to Satisfy in an Era of Excess* [24], and *What Price the Moral High Ground? Ethical Dilemmas in Competitive Environments* [25]. Histories are often rich in strategic detail. Allison's *Essence of Decision: Explaining the Cuban Missile Crisis* [1] is a notable example. Finally, business literature is suffused with works on strategy. Michael E. Porter's 1980 book, *Competitive Strategy: Techniques for Analyzing Industries and Competitors,* [51] is recognized as a classic of the genre.

Mitchell Zuckoff's article, "The Perfect Mark: How a Massachusetts Psychotherapist fell for a Nigerian E-mail Scam" [72] and David Lewis's "The Ballad of Big Mike" [42] are two eminently readable accounts of successful fraud by e-mail spam and of strategic aspects of American football, respectively. In addition, Lewis's story is a touching and uplifting account of the rise of an improbable player.

# Chapter 4

# Games in the Abstract

## 4.1 Interrogation

The scene is Sam Spade's apartment. Gutman and Cairo are partners in search of the Maltese Falcon. They have the guns. Spade has the information. Spade is being difficult.

> Cairo, his face and body twitching with excitement, exclaimed: "You seem to forget that you are not in a position to insist on anything."
>
> Spade laughed, a harsh derisive snort.
>
> Gutman said, in a voice that tried to make firmness ingratiating: "Come now, gentlemen, let's keep our discussion on a friendly basis; but there certainly is"—he was addressing Spade—"something in what Mr. Cairo says. You must take into consideration the—"
>
> "Like hell I must." Spade flung his words out with a brutal sort of carelessness that gave them more weight than they could have got from dramatic emphasis or from loudness. "If you kill me, how are you going to get the bird? If I know you can't afford to kill me till you have it, how are you going to scare me into giving it to you?"
>
> Gutman cocked his head to the left and considered these questions. His eyes twinkled between puckered lids. Presently he gave his genial answer: "Well, sir, there are other means of persuasion besides killing and threatening to kill."
>
> "Sure," Spade agreed, "but they're not much good unless the threat of death is behind them to hold the victim down. See what I mean? If

> you try anything I don't like I won't stand for it. I'll make it a matter of your having to call it off or kill me, knowing you can't afford to kill me."
>
> "I see what you mean." Gutman chuckled. "That is an attitude, sir, that calls for the most delicate judgment on both sides, because, as you know, sir, men are likely to forget in the heat of action where their best interest lies and let their emotions carry them away."
>
> Spade too was all smiling blandness. "That's the trick, from my side," he said, "to make my play strong enough that it ties you up, but yet not make you mad enough to bump me off against your better judgment."
>
> Gutman said fondly: "By Gad, sir, you are a character!"
>
> –From "The Fall-Guy" in *The Maltese Falcon* by Dashiell Hammett

The passage describes a context of strategic interaction. How might we model this game in the wild? A number of *game forms* are available. For the present we will use just one, the *strategic form*, which is well-suited to games with two players. See Figure 4.1 for the general, canonical strategic form for two players each having two strategies.

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | $r_1$, $c_1$ | $r_2$, $c_2$ |
| $R_2$ | $r_3$, $c_3$ | $r_4$, $c_4$ |

Figure 4.1: Canonical game matrix for the $2 \times 2$ game in strategic form

The interpretation is straightforward. There are two players: Row, who choses $R_1$ or $R_2$, and Column, who choses $C_1$ or $C_2$. We say Row has available the *strategies* $R_1$ and $R_2$, and similarly Column has strategies $C_1$ and $C_2$. The form is called *strategic form* because the players' strategies are laid out so plainly. If players have more than two strategies, we add rows or columns to the *game matrix* as necessary. If there are more than two players we can use a game cube or hypercube (for more than 3 players) if needed.

Choosing simultaneously, or at least in ignorance of each other's choices, there are four possible outcomes and associated *rewards* for the players. In terms of the canonical game matrix, Figure 4.1:

| Outcome | $R$'s reward | $C$'s reward |
|---|---|---|
| $R_1C_1$ | $r_1$ | $c_1$ |
| $R_1C_2$ | $r_2$ | $c_2$ |
| $R_2C_1$ | $r_3$ | $c_3$ |
| $R_2C_2$ | $r_4$ | $c_4$ |

Returning now to our scene of interrogation in Sam Spade's apartment, the players are Spade (let us say Row) and Gutman (Column). Spade may either Blab (B) regarding the whereabouts of the bird, or Keep Silent (K). Gutman will interrogate. He will either Press (P) or Extreme Press (E), the latter killing Spade in the process, or at least severely disabling him. Spade's preference ordering is $KP > BP > KE > BE$. Gutman's is $BP > BE > KP > KE$. That is, Spade prefers most that Row play K and Column play P; Gutman prefers most that Row play B and Column play P. Converting these rankings to convenient numbers gives us Figure 4.2. The numbers range from 1 to 4, more being better. Here they simply record the ranking of the rewards or outcomes for the players in question.

| | Press (P) | Extreme Press (E) |
|---|---|---|
| Blab (B) | 3, 4 | 1, 3 |
| Keep Silent (K) | 4, 2 | 2, 1 |

Figure 4.2: Interrogation: Spade is Row, Gutman is Column

What will happen? Gutman is surely right that the situation "calls for the most delicate judgment on both sides." Putting that aside for the present and looking at the game as abstracted in Figure 4.2, we can see that no matter which strategy Gutman pursues, Press or Extreme Press, Spade will prefer—is rewarded more by—Keeping Silent. If Gutman Presses, Spade gets 4 for Keeping Silent and 3 for Blabbing. If Gutman Extreme Presses, Spade gets 2 from Keeping Silent and 1 from Blabbing. Either way, Spade does better by Keeping Silent. Notice that

the actual values of the rewards to Spade matter little. We could replace 4 by $A$, 3 by $B$, 2 by $C$ and 1 by $D$ and this conclusion would follow for any numbers assigned to the letters, so long as higher numbers represent preferred rewards and $A > B > C > D$.

So we predict Spade will Keep Silent. What will Gutman do? Gutman is a smart fellow and he will presumably see the reasoning in the situation—as he appears to in the dialog—and conclude that Spade will choose to Keep Silent. Given that, Gutman's best strategy is to Press, since choosing Extreme Press won't get him the Maltese Falcon and burdens him with Spade's demise. Gutman's "By Gad, sir, you are a character!" is an admission of defeat.

Some terminology and associated concepts that we will need throughout: a dominant strategy, a Nash equilibrium, and a Pareto optimal outcome. We say Spade's (Row's) Keep Silent strategy *dominates* his Blab strategy, because on every alternative—here, either Press or Extreme Press—Spade's Keep Silent yields a higher reward to Spade than does his Blab. The typographic conventions we are using for game matrices were designed to make it easy to spot dominating or dominated strategies. See Figure 4.2. Gutman (Column) also has an absolutely dominate strategy: Press beats Extreme Press.

The *principle of dominance* enjoins us—or at least predicts of all rational players—never to choose a dominated strategy. If Spade and Gutman are rational in this rather minimal sense, KP (Gutman Presses and Spade Keeps Silent) will be the outcome. In this game the dominance principle is sufficient to predict a unique outcome. Is this in fact what will always happen? Gutman's remark that "men are likely to forget in the heat of action where their best interest lies and let their emotions carry them away" is surely apt. We should ask ourselves whether other rational causes might lead to violation of the principle. This is a matter for the sequel.

We say that KP is a *Nash equilibrium* (or NE) outcome because no player could change its strategy *unilaterally* and do better.[1] KP is an NE because for each player, given the play by the other player(s), its strategy is best. Specifically, KP is an NE because *given that Gutman chooses Press*, Spade can do no better with Blab, and *given that Spade chooses Keep Silent*, Gutman can do no better with Extreme Press. A Nash equilibrium outcome is a strategic standoff: no player can do better, given what the other players have played. One might, and classical

[1] Named in honor of John Nash, 1928–, who invented and developed the concept. His home page is http://www.math.princeton.edu/jfnj/. See http://www.nobel.se/economics/laureates/1994/nash-autobio.html for his biography for his The Bank of Sweden Prize in Economic Sciences in Memory of Alfred Nobel, 1994, prize.

game theory does, predict that game outcomes among rational players will be Nash equilibria. Again, we should ask ourselves whether rational causes might lead to violation of this *Nash equilibrium principle.* Again, this is a matter for the sequel.

We say that an outcome is *Pareto optimal* if there is no other outcome for which every player does better.[2] If an outcome is not Pareto optimal, we say it is dominated. (Notice the difference between a dominated strategy and a dominated outcome.) In our Interrogation game, Figure 4.2, BE is a dominated outcome; both Spade and Gutman do better with BP. Similarly, KE is dominated, this time by KP. Both KP and BP are Pareto optimal. We say that the set of Pareto optimal outcomes constitutes the *Pareto frontier*. The *Pareto frontier principle* has it that among rational players game outcomes will always be in (or on) the Pareto frontier. Once again, we should ask ourselves whether rational causes might lead to violation of this, the Pareto frontier principle. And what if the principles fail to apply or even even conflict?

Summing up, we can label the outcomes in the game matrix, [N] for Nash equilibrium, [P] for Pareto optimal, [NP] for Nash and Pareto, and unlabeled for none of the above. Here is Interrogation with labeling.

[2]Named in honor of Vilfredo Pareto, 1848-1923, who invented and developed the concept. See http://cepa.newschool.edu/het/profiles/pareto.htm.

| | Press (P) | Extreme Press (E) |
|---|---|---|
| Blab (B) | 4<br>[P]<br>3 | 3<br><br>1 |
| Keep Silent (K) | 2<br>[NP]<br>4 | 1<br><br>2 |

Figure 4.3: Interrogation with Labeling

A last word on the Interrogation game. Dashiell Hammett, the author of *The Maltese Falcon,* actually worked for a time as a private detective. He knew whereof he wrote. *The Maltese Falcon* was published on February 14, 1930, well before the flowering of game theory.

## 4.2 Prisoner's Dilemma

The Prisoner's Dilemma is the most famous and the most studied of games in the abstract. Invented about 1950 (and attributed to A.W. Tucker), experimentation began on it immediately [19, 20] and continues to this day. Entire books have been devoted to it, including [2, 52, 55]. Luce and Raiffa give the standard interpretation [44, page 95]. There is no need to revise it:

> Two suspects are taken into custody and separated. The district attorney is certain that they are guilty of a specific crime, but he does not have adequate evidence to convict them at trial. He points out to each prisoner that each has two alternatives: to confess to the crime the police are sure they have done, or not to confess. If they both do not confess, then the district attorney states he will book them on some very minor trumped-up charge such as petty larceny and illegal possession of a weapon, and they will both receive minor punishment; if they both confess they will be prosecuted, but he will recommend less than the most severe sentence; but if one confesses and the other does not, then the confessor will receive lenient treatment for turning state's evidence whereas the latter will get "the book" slapped on him.

The game is standardly taken to hinge on cooperation. The prisoners (Row and Column) may behave cooperatively (C) by *refusing to confess* or they may behave

uncooperatively and defect (D) on each other. If both prisoners cooperate (refuse to confess) each receives a reward, R, for cooperation. R=3 in much of the literature and in what I will call the *Default* Prisoner's Dilemma (PD) game. Think of 3 as the number of years the prisoner gets to live out of jail during the next 5 years. More is better. If both prisoners defect (confess), each receives a reward of P, the penalty for mutual defection. P=1 in the Default PD. If one player cooperates (C) and the other player defects (D), the cooperator gets S, the sucker's payoff and the defector gets T, the temptation to defect. In the Default PD, S=0 and T=5. These assumptions are recorded in the left-hand game matrix of Figure 4.4. Typically,

| | D | C |
|---|---|---|
| D | 1<br>[N]<br>1 | 0<br>[P]<br>5 |
| C | 5<br>[P]<br>0 | 3<br>[P]<br>3 |

| | D | C |
|---|---|---|
| D | P<br>[N]<br>P | S<br>[P]<br>T |
| C | T<br>[P]<br>S | R<br>[P]<br>R |

Figure 4.4: Default and Canonical Prisoner's Dilemma: $T > R > P > S$ and $2R > (T + S)$

and in what I call the *Canonical* Prisoner's Dilemma (right-hand game matrix in Figure 4.4), the payoffs to each player are symmetric in the sense that T for row chooser equals T for column chooser, and so on. This is not strictly required for PD and wasn't true in the original PD experiments [19, 20].[3] Prisoner's Dilemma strictly requires that $T_i > R_i > P_i > S_i$ for each player $i$. Further, it is usually, although not always, required that $2R_i > (T_i + S_i)$ for each player $i$. Both the Default and the Canonical PDs meet both of these conditions.

The PD game is a dilemma because our rationality principles are in conflict. The principle of dominance advises both players to defect, D, since it is a dominant strategy for each of them. The Nash equilibrium principle concurs. DD is the only Nash equilibrium outcome of this game. DD, however, is not Pareto optimal: *both* players would do better if CC were the outcome. The Pareto frontier in this game is: $\{CD, CC, DC\}$. On one side we have the dominance principle and the

[3]The game matrix used was:

| (-1, 2) | ($\frac{1}{2}$, 1) |
|---|---|
| (0, $\frac{1}{2}$) | (1, -1) |

Nash equilibrium principle and on the other we have the Pareto frontier principle. The two sides are, in this game, directly in conflict.

## 4.3 Hawk-Dove

Two players confront each other over a resource whose full value is $V$ to either of them. Each player may play one of two strategies: H (Hawk) or D (Dove). Doves signal that they wish to share the resource equally. Hawks signal they are willing to fight to get the resource. When two Doves meet, each gives the characteristic sharing signal and the resource is divided equally, or, perhaps, a fair coin is tossed and the winner gets all. In any case, the expected return to each of the two Doves is $V/2$. When a Hawk meets a Dove, the Hawk (as it always does) signals fight, the Dove (as it always does) signals share, then the Dove retreats and the Hawk takes the entire resource. Finally, when two Hawks meet, each signal fight, neither retreats, both fight at a cost of $C$. In the end, the resource is shared equally, minus the cost, or, perhaps, half the time one Hawk gets the entire resource and half the time the other Hawk gets it. In any case, the expected return to each of the Hawks is $\frac{1}{2}(V-C)$. We can summarize the Hawk-Dove game in the following strategic form representation. Let us assume, sensibly and without loss of generality, that

| | H | D |
|---|---|---|
| H | $\frac{1}{2}(V-C)$<br>$\frac{1}{2}(V-C)$ | 0<br>[P]<br>V |
| D | V<br>[P]<br>0 | $V/2$<br>[P]<br>$V/2$ |

Figure 4.5: Hawk-Dove Game: $C > V > 0$

$V > 0$ and $C > 0$. Notice that the outcomes HD, DD, and DH are Pareto optimal: in each case it is impossible to find another outcome that will not make at least one of the players worse off. Note especially what happens when $C > V$ and when $V > C$. DD is not a Nash equilibrium in either case, since Row would prefer HD and Column would prefer DH. If $V > C > 0$, then HD is not Nash because Column would prefer H; similarly for DH, and HH is a Nash equilibrium. Conversely, if $C > H > 0$ fighting is very damaging. Row would refer DH to HH

and Column would prefer HD to HH, so HH is not a Nash equilibrium, while HD and DH are.

| | H $y$ | D $(1-y)$ |
|---|---|---|
| H $x$ | $\frac{1}{2}(V-C)$ (Row), $\frac{1}{2}(V-C)$ (Column) | V (Row), 0 (Column) |
| D $(1-x)$ | 0 (Row), V (Column) | $V/2$ (Row), $V/2$ (Column) |

Figure 4.6: Hawk-Dove Game with Mixed Strategies

What about *mixed equilibria*, in which strategies are played according to a probability distribution? Let Row play H with probably $x$ and D with probability $(1-x)$. Similarly, let Column play H with probability $y$ and D with probability $(1-y)$. See Figure 4.6, in which these probabilities are recorded in the margins. The expected return to Row, or the game value for Row, $G_R$ in this regime is

$$G_R = xy(\frac{1}{2}(V-C)) + x(1-y)V + 0 + (1-x)(1-y)(\frac{V}{2}) \quad (4.1)$$

$$= x[y(\frac{1}{2}(V-C)) + (1-y)V - (1-y)(\frac{V}{2})] + (1-y)(\frac{V}{2}) \quad (4.2)$$

Taking a derivative with respect to $x$ (see the Addendum to this chapter, §A.2):

$$\frac{dG_R}{dx} = [y(\frac{1}{2}(V-C)) + (1-y)V - (1-y)(\frac{V}{2})] \quad (4.3)$$

Setting this to 0, solving for $y$, and simplifying we get:

$$y = \frac{V}{C} \quad (4.4)$$

Making the same calculation for $G_C$, the value of the game to Column, we also get:

$$x = \frac{V}{C} \quad (4.5)$$

It is an equilibrium (indeed a stable one) for Row to play H with probability $x = \frac{V}{C}$ and for Column to play H with probability $y = \frac{V}{C}$. Note that this assumes $V < C$,

which is what we'll assume in our subsequent studies of the Hawk-Dove game. Note further that when $V - C$ is positive, Hawk-Dove is a degenerate Prisoner's Dilemma: $T = V, R = V/2, P = (V - C)/2, S = 0$ violates the $2R > T + S$ condition.

The method we used immediately above to find the mixed equilibria of a game in strategic form is entirely general, and especially tractable in a $2 \times 2$ game. We will see it again. There is another, very fruitful perspective we can take on "solving" the game. Suppose now that the regime of play is repeated or iterated. We have an infinite population of players, some of whom play H and some of whom play D. We draw them at random from the population and have them play each other. We record the returns the two players get and we adjust their frequencies accordingly in the next generation. See Figure 4.7 for a representation relevant to our current regime of play, which is called a *replicator dynamic.* Figure 4.7 is a special version of Table 4.6, with $x = y$ and the perspective of the row player.

| | H $x$ | D $(1-x)$ |
|---|---|---|
| H | $\frac{1}{2}(V-C)$ <br> $\frac{1}{2}(V-C)$ | 0 <br> V |
| D | V <br> 0 | $V/2$ <br> $V/2$ |

Figure 4.7: Hawk-Dove Game with a Mixed Population

What about equilibrium? Think of it this way. Suppose you could pick which strategy to play, H or D, knowing the current value of $x$. At what value of $x$ would you be indifferent between playing H and playing D? You would be indifferent when the expected return from playing H, $E(H)$, equaled the expected return from playing D, $E(D)$.

$$E(H) = \frac{1}{2}(V - C)x + V(1 - x) \quad (4.6)$$

$$E(D) = 0x + \frac{V}{2}(1 - x) \quad (4.7)$$

Setting them equal

$$\frac{1}{2}(V - C)x + V(1 - x) = \frac{V}{2}(1 - x) \quad (4.8)$$

and solving for $x$ we again get

$$x = \frac{V}{C} \tag{4.9}$$

Thus, the equilibrium reached by the replicator dynamic is a Nash equilibrium. This *replicator dynamic equilibrium* is also stable.[4] A replicator dynamic equilibrium is a special case of a Nash equilibrium; it is an equilibrium reached by an infinite population of strategies under the regime of the replicator dynamic. Finite populations over finite times, driven by evolution and natural selection, may approximate it. The replicator dynamic equilibrium, like the Nash equilibrium, will be a useful benchmark.

## 4.4 Stag Hunt or Assurance

Two agents go hunting and take up their places in a blind, which hides them both from each other and from any stags that happen by. Together they can expect to bag a stag, which will feed them each for 3 days. If, however, one of the players reneges and goes hunting for hare, that player can expect to bag two hares, enough to feed him for two days. The other player will receive nothing, neither stag nor hare. If both players renege, each can expect to bag one hare, a day's worth of food. The game, presented in strategic form in Figure 4.8 is so named in honor of a passage in Rousseau's *A Discourse on Inequality*:

> If it was a matter of hunting a deer, everyone well realized that he must remain faithful to his post; but if a hare happened to pass within reach of one of them, we cannot doubt that he would have gone off in pursuit of it without scruple...

The Stag Hunt is also called the Assurance game. What assurance does a player have that the other player won't renege? The game has been used to model arms races. To see why, relabel. For the row player, change Hunt stag to Refrain from deploying missile defense and Hunt hare to Fully deploy missile defense. For the column player change Hunt stag to Refrain from deploying missile defense penetration system and Hunt hare to Fully deploy missile defense penetration system. Hunting stag (or its strategic equivalent) is a cooperative play, as chasing hare is uncooperative. The Stag Hunt game is thus another kind of strategic context in which issues of cooperation arise.

[4] And it qualifies as an ESS (Evolutionary Stable Strategy). See [47] for the seminal introduction of the Hawk-Dove game and the concept of an ESS.

| | Hunt stag (S) | Chase hare (H) |
|---|---|---|
| Hunt stag (S) | 3 [NP] 3 | 2 0 |
| Chase hare (H) | 0 2 | 1 [N] 1 |

Figure 4.8: Stag Hunt (aka: Assurance game)

Note that there are two equilibria: SS (both Hunt stag) and HH (both Chase hare), only one of which is Pareto optimal, SS. What about a mixed equilibrium? Symbolizing for the general case, with allusion to Prisoner's Dilemma we have the game matrix of Figure 4.9:

| | Hunt stag (S) | Chase hare (H) |
|---|---|---|
| Hunt stag (S) | R [NP] R | T S |
| Chase hare (H) | S T | P [N] P |

Figure 4.9: Generic Stag Hunt (aka: Assurance game): $R > T > P > S$

Calculating the replicator dynamic equilibrium we have

$$E(S) = Rx + (1-x)S \tag{4.10}$$
$$E(H) = Tx + (1-x)P \tag{4.11}$$

Equating and solving for $x$ gives us

$$x = \frac{P+S}{R+P-T-S} \tag{4.12}$$

For the specific game in Figure 4.8

$$x = \frac{1+0}{3+1-2-0} = \frac{1}{2} \tag{4.13}$$

The expected return received by a player of the game at this equilibrium is

$$G_{x=0.5} = \frac{1}{4}(3 + 0 + 2 + 1) = 1.5$$

The single Pareto optimal outcome continues to dominate.

See "The Stag Hunt" by Brian Skyrms [69] for a thoughtful discussion of the significance of this game.

## 4.5 Chicken

Each player drives a car, racing it directly on a collision course with the other player's car. If both players Drive Straight they will crash into one another with dire consequences for each. If one player swerves, leaving the road clear for the other, that player is "chicken" and the non-swerving player gets accolades for bravery. If both players swerve, both are "chicken" and both dishonored, but less so than being the sole "chicken." Figure 4.10 abstracts this strategic context as a strategic form game. In our game, the Pareto frontier consists of SS (both swerve),

| | Swerve | Drive Straight |
|---|---|---|
| Swerve | 2<br>[P]<br>2 | 3<br>[NP]<br>1 |
| Drive Straight | 1<br>[NP]<br>3 | 0<br><br>0 |

Figure 4.10: Chicken

SD (row swerves, column drives straight), and DS. SD and DS are both Nash equilibria. Symbolizing for the general case, with allusion to Prisoner's Dilemma gives us Figure 4.11. Calculating the replicator dynamic equilibrium we have:

$$E(S) = Rx + S(1-x) \quad (4.14)$$

$$E(D) = Tx + P(1-x) \quad (4.15)$$

Equating and solving for $x$ gives us

$$x = \frac{P - S}{R - S - T + P} \quad (4.16)$$

| | Swerve | Drive Straight |
|---|---|---|
| Swerve | R [P] R | T [NP] S |
| Drive Straight | S [NP] T | P P |

Figure 4.11: Generic Chicken: $T > R > S > P$

In our specific case

$$x = \frac{0-1}{2-1-3+0} = \frac{1}{2} \tag{4.17}$$

This game apparently was originated in the movie, "Rebel without a Cause," starring James Dean and Natalie Wood. In the movie's "Chickie Run" scene, the two players race stolen cars towards a cliff overlooking the ocean. Whoever bails out first is "chicken." One of the players is unsuccessful in abandoning his car before it lurches over the cliff. James Dean survives and gets the girl.

## 4.6 Battle of the Sexes

We meet Della, loving wife of loving Jim in O. Henry's short story "The Gift of the Magi"[5] Della has a problem. They are a young couple, times are tough, and Jim's salary has been cut. They are poor.

> Della finished her cry and attended to her cheeks with the powder rag. She stood by the window and looked out dully at a gray cat walking a gray fence in a gray backyard. Tomorrow would be Christmas Day, and she had only $1.87 with which to buy Jim a present. She had been saving every penny she could for months, with this result. Twenty dollars a week doesn't go far. Expenses had been greater than she had calculated. They always are. Only $1.87 to buy a present for Jim. Her Jim. Many a happy hour she had spent planning for something nice

[5]Freely available at http://www.auburn.edu/~vestmon/Gift_of_the_Magi.html thanks to Project Gutenburg.

> for him. Something fine and rare and sterling—something just a little bit near to being worthy of the honor of being owned by Jim.

How is she to get money for his Christmas present?

> Now, there were two possessions of the James Dillingham Youngs in which they both took a mighty pride. One was Jim's gold watch that had been his father's and his grandfather's. The other was Della's hair. Had the queen of Sheba lived in the flat across the airshaft, Della would have let her hair hang out the window some day to dry just to depreciate Her Majesty's jewels and gifts. Had King Solomon been the janitor, with all his treasures piled up in the basement, Jim would have pulled out his watch every time he passed, just to see him pluck at his beard from envy.

So Della sells her hair and buys a gold chain for Jim's watch. Jim, of course, comes home with an expensive set of combs for Della's hair. He has sold his watch to buy them.

Abstracting this story we get the Battle of the Sexes as it is called in the game theory literature. Its story is slightly different. He (say Row) would like to go to the fight (F, boxing match). She (Column) would like to go to the opera (O). They both would prefer to attend the same event and they each have to commit to an event without communicating with the other (cell phones have been lost, time is short, etc.). Figure 4.12 gives a reasonable abstraction. (B=best return; S=second best return; 0=worst.) In pure strategies, there are two Nash equilibria and two

| | F | O |
|---|---|---|
| F | 1<br>[NP]<br>3 | 0<br><br>0 |
| O | 0<br><br>0 | 3<br>[NP]<br>1 |

| | F $y$ | O $\overline{y}$ |
|---|---|---|
| F<br>$x$ | S<br>[NP]<br>B | 0<br><br>0 |
| O<br>$\overline{x}$ | 0<br><br>0 | B<br>[NP]<br>S |

Figure 4.12: Battle of the sexes: Specific and Symbolic ($\overline{x} = (1 - x)$)

Pareto optimal outcomes: FF and OO. There is a mixed Nash equilibrium at

$$x = \frac{B}{B+S} \tag{4.18}$$

$$y = \frac{S}{B+S} \tag{4.19}$$

Because this is not a symmetric game (in a sense to be made clear in the sequel), the replicator dynamic equilibrium is not (yet) well defined.

## 4.7 Inspector versus Evader

A two-period game is played between the Inspector (Column) and the Evader (Row). Evader might be a drug smuggler, Inspector the Coast Guard, or Evader might be a country bent on developing nuclear weapons, Inspector the United Nations, and so on. Inspector can only inspect during one of the two periods. The evader has two strategies:

$E$: Evade the rules during the first period; evade during the second period if and only if Inspector inspects during the first period.

$\neg E$: Do not evade the rules during the first period; evade during the second period if and only if Inspector inspects during the first period.

Inspector has two strategies:

$I$: Inspect during the first period (and not during the second).

$\neg I$: Inspect during the second period (and not during the first).

In terms of outcomes, let us assume Evader's preferences are $E\neg I > \neg EI > \neg E\neg I > EI$ and Inspector's preferences are $\neg E\neg I > \neg EI > EI > E\neg I$. Let $0 < b, c < 1$, $0 < a, d$, and assign the outcome values as in Figure 4.13.

| | $I$ $y$ | $\neg I$ $(1-y)$ |
|---|---|---|
| $E$ $x$ | $-a$, $0$ | $1$, $-d$ [P] |
| $\neg E$ $(1-x)$ | $c$, $b$ | $0$, $1$ [P] |

Figure 4.13: Inspector versus evader game

The game has no Nash equilibrium in pure strategies. Using the technique of §A.2 we find that there is an equilibrium at

$$x = \frac{1-b}{1-b+d} \tag{4.20}$$

$$y = \frac{1}{1+a+c} \tag{4.21}$$

Checking further shows that this is a stable equilibrium. Notice that the larger $a$ and $c$ are the smaller $y$ is, and the larger $d$ is the smaller $y$ is. Is the rationality of this plain?

See [54, chapter 11] for a thorough discussion of this game. I have honored the specifics of his example.

## 4.8 A Zero-Sum Game

The game in Figure 4.14 has the special property that for any outcome, Row's gain is Column's loss, and *vice versa.* Such *zero-sum* games are contexts of pure

| | $C_1$ $y$ | $C_2$ $(1-y)$ |
|---|---|---|
| $R_1$<br>$x$ | $2$<br>[P]<br>$-2$ | $-4$<br>[P]<br>$4$ |
| $R_2$<br>$(1-x)$ | $-2$<br>[P]<br>$2$ | $1$<br>[P]<br>$-1$ |

Figure 4.14: A zero-sum game

opposition. Every outcome is on the Pareto frontier simply because if one player is relatively better off, the other player is relatively worse off. It happens in this particular game that there is no Nash equilibrium outcome, i.e., no Nash equilibrium in pure strategies. There is a Nash equilibrium in mixed strategies. Using the general results from appendix A, §A.2:

$$x = \frac{(1+2)}{(2+1)-(-4-2)} = \frac{1}{3} \tag{4.22}$$

$$y = \frac{-1-4}{(-2-1)-(4+2)} = \frac{5}{9} \tag{4.23}$$

## 4.9 PD Property Games: ##12, 47, 48 & 57

Recall the canonical game matrix for the $2 \times 2$ game (Figure 4.1, page 68), reprinted below:

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | $r_1$, $c_1$ | $r_2$, $c_2$ |
| $R_2$ | $r_3$, $c_3$ | $r_4$, $c_4$ |

Figure 4.15: Canonical game matrix for the $2 \times 2$ game in strategic form

If we accept limitations on the numerical values assigned to the rewards, the $r_i$s and the $c_j$s, then the number of $2 \times 2$ games can be restricted to a tractable size and the class studied systematically. If each reward value for an agent is unique and drawn from $\{1, 2, 3, 4\}$, then there are only 576 = $4! \times 4!$ distinct $2 \times 2$ games. Many of these are really equivalent, e.g., one can be transformed to another simply by exchanging rows, or one can be transformed to another by switching the rôles of the players: Row becomes Column, Column becomes Row. It turns out that there are exactly 78 unique $2 \times 2$ games under these assumptions. Anatol Rapoport and his co-workers have studied them all [56].[6]

Among these 78 unique 2×2 games, it turns out that exactly 4 have the *PD property*—first seen in the Prisoner's Dilemma—of having Pareto frontier outcomes distinct from Nash equilibrium outcomes. Figure 4.16 presents Rapoport's game #12.

[6] See [56, pages 14–7] for details on counting the number of $2 \times 2$ games.

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 2 [N] 2 | 1 [P] 4 |
| $R_2$ | 4 [P] 1 | 3 [P] 3 |

Figure 4.16: Game #12: Prisoner's Dilemma

Game #12 is a Prisoner's Dilemma: $T = 4, R = 3, P = 2, S = 1$. Figures 4.17, 4.18, and 4.19 present games #47, #48, and #57 respectively, with their [N] and [P] outcomes labeled.

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 3 [N] 2 | 1 [P] 4 |
| $R_2$ | 2 1 | 4 [P] 3 |

Figure 4.17: Game #47

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 2<br>[N]<br>2 | 1<br>[P]<br>4 |
| $R_2$ | 3<br><br>1 | 4<br>[P]<br>3 |

Figure 4.18: Game #48

| | $C_1$ | $C_2$ |
|---|---|---|
| $R_1$ | 3<br>[N]<br>2 | 2<br>[P]<br>4 |
| $R_2$ | 1<br><br>1 | 4<br>[P]<br>3 |

Figure 4.19: Game #57

Points arising:

1. Prisoner's Dilemma, as in game #12, is typically specified in a *symmetric* fashion. In the special case in which we have a *game in symmetric form* it happens that $V_R(i, j) = V_C(j, i)$: the value that the Row player receives if Row plays strategy $i$ and Column plays strategy $j$ is equal to the value to the Column player of Row playing $j$ and Column playing $i$. Here, e.g., $V_R(R_2, C_1) = 1 = V_C(R_1, C_2)$. Symmetric form games are more easily—hence more often—studied. In Part II we begin with symmetric form games, then move on to asymmetric games (aka: bimatrix games).

2. Games ##47, 48, and 57 are asymmetric games. They are further distinguished from game #12 and indeed all Prisoner's Dilemma games by having a Pareto dominated outcome that is *not* a Nash equilibrium. Thus is a nice setup arrived at to test our three principles (so far) of rationality. The principle of (strategy) dominance and the Nash equilibrium principle predict the

outcome $R_1C_1$ for all three games. The Pareto optimality principle predicts $\{R_1C_2, R_2C_2\}$. No principle predicts $R_2C_1$. Who is right and under what conditions?

3. None of these four games has a Nash equilibrium in properly mixed strategies.[7]

4. It is possible to have a Pareto outcome in mixed strategies, for example a mixture of the elements in $\{R_1C_2, R_2C_2\}$. Can it actually happen? If so, under what conditions?

## 4.10 Other Games, Other Forms

Games in the wild are natural phenomena and as such need to be represented or modeled for purposes of investigation. To do so, we abstract the phenomena into simpler, more tractable representations, which afford our inquiries. This is a general point, and it applies straightforwardly in contexts of strategic interaction.

My purpose in this chapter has been to provide an inventory of abstract games, linked insofar as possible to naturally-occurring games. We draw upon this stock in the sequel. The discussion has proceeded and will proceed by example, raising concepts and terminology as they are needed.

Each of the games discussed in this chapter has been a 2×2 game in strategic form. There are other important 2×2 games, there are plenty of important games that are not 2×2 games, and there are other game forms than the strategic. We shall see examples of each of these in the course of our discussion. We begin with the 2×2 game because it is an excellent place to begin our algorithmic, constructivist—"from the ground up"—study of contexts of strategic interaction.

> [I]f there is any hope of eventually constructing scientific theories of human behavior, we must first learn to perform controlled experiments with a view of drawing inferences from them that at least have *apparent* relevance to human motivations, learning, decisions—above all, to interactions. Gaming experiments include all these features, and experiments on $2 \times 2$ games are the simplest and most tractable that include the most important of them. The value of such experiments is that they can teach us not necessarily how people behave

[7] **Note JDL: Use Selten** *Models of Strategic Rationality,* **"A Note on ESS in Asymmetric Animal Conflicts". Agent normal form.**

in real life but how we can study certain aspects of characteristically human behavior systematically, from the ground up, as it were.

–*The* $2 \times 2$ *Game,* Rapoport & Guyer [56, page 13, underline added]

# Appendix A

# Useful Mathematical Results

## A.1 Miscellaneous Useful Facts

$\sum_{i=0}^{\infty} w^i = (1-w)^{-1}$ for $0 < w < 1$

PV factor =

$$\sum_{i=0}^{\infty} \frac{1}{(1+r)^i}$$

$r$ = interest rate. Axelrod's discount rate, $w$ = $(1+r)^{-1}$

Comment

- $r$ and $w$ are inversely related
- When the interest (hurdle) rate is small, the discount rate is large.
- When the interest rate is small, we are more interested in looking to the future.

## A.2 Solving for Mixed Equilibria in 2×2 Games

Recall Figure 4.1, our canonical game matrix for the $2 \times 2$ game in strategic form, which is reprinted below with a bit of relabeling and additional information: Row plays $R_1$ with probability $x$ and $R_2$ with probability $(1-x)$. Similarly, Column plays $C_1$ with probability $y$ and $C_2$ with probability $(1-y)$. The expected return for Row, $G_R$, is

$$G_R = xya_r + x(1-y)b_r + (1-x)yc_r + (1-x)(1-y)d_r \qquad \text{(A.1)}$$

| | $C_1$ $y$ | $C_2$ $(1-y)$ |
|---|---|---|
| $R_1$<br>$x$ | $a_c$<br>$a_r$ | $b_c$<br>$b_r$ |
| $R_2$<br>$(1-x)$ | $c_c$<br>$c_r$ | $d_c$<br>$d_r$ |

Figure A.1: Canonical game matrix for the $2 \times 2$ game in strategic form

The equilibrium values of $x, y$, $0 < x, y < 1$ if they exists, are found by taking the partial derivatives $\partial G_R/\partial x$ and $\partial G_C/\partial y$, setting the results to 0, and solving. We have

$$\frac{\partial G_R}{\partial x} = ya_r + (1-y)b_r - yc_r - d_r + yd_r = 0 \tag{A.2}$$

Solving for $y$ gives us:

$$y = \frac{d_r - b_r}{(a_r + d_r) - (b_r + c_r)} \tag{A.3}$$

The analogous calculation for $G_C$ yields

$$x = \frac{d_c - c_c}{(a_c + d_c) - (b_c + c_c)} \tag{A.4}$$

If (and only if) the resulting values for $x$ and $y$ are legitimate probabilities, we have found a mixed equilibrium for the $2\times2$ game.

## A.3 Replicator Dynamic Equilibrium for $2\times2$ Games

Given our standard $2\times2$ game matrix: the replicator dynamic equilibrium may be found by solving

$$Ax + B(1-x) = Cx + D(1-x) \tag{A.5}$$

which gives

$$x = \frac{(D-B)}{(D-B) + (A-C)} \tag{A.6}$$

| | $S_1$ $x$ | $S_2$ $\overline{x}$ |
|---|---|---|
| $S_1$ $x$ | $A$ / $A$ | $B$ / $C$ |
| $S_2$ $\overline{x}$ | $C$ / $B$ | $D$ / $D$ |

Figure A.2: Canonical game matrix for the symmetric $2 \times 2$ game in strategic form

# A.4 Pareto Optimal Mixed Equilibrium

Consider a canonical $2{\times}2$ *symmetric* game under the replicator dynamic, with $x$ the proportion of $C$ players and $\overline{x} = (1 - x)$ the proportion of $D$ players:

| | $S_1$ $x$ | $S_2$ $\overline{x}$ |
|---|---|---|
| $S_1$ $x$ | $A$ / $A$ | $B$ / $C$ |
| $S_2$ $\overline{x}$ | $C$ / $B$ | $D$ / $D$ |

Figure A.3: Canonical game matrix for the symmetric $2 \times 2$ game in strategic form

Although the labeling here resembles that often used for the general Prisoner's Dilemma, the case at hand should be taken generally. The expected value of the game, $G$, for a player is

$$G = Ax^2 + Bx(1-x) + C(1-x)x + D(1-x)(1-x) \tag{A.7}$$

$$= Ax^2 + Bx - Bx^2 + Cx - Cx^2 + D - 2Dx + Dx^2 \tag{A.8}$$

Differentiating

$$\frac{dG}{dx} = 2Ax + B - 2Bx + C - 2Cx - 2D + 2Dx \tag{A.9}$$

$$= \quad 2(A - B - C + D)x + (B + C - 2D) \tag{A.10}$$

Setting to 0 and rearranging:

$$x = \frac{(B + C - 2D)}{2((B + C) - (A + D))} \tag{A.11}$$

which will be maximal when

$$\frac{d^2G}{x^2} = (A + D - C - B) < 0 \tag{A.12}$$

# Appendix B

# Strategic Treatments in Movies and Literature

*Rebel without a Cause*, the "Chickie Run" scene, origin of the game of chicken?

*Wall Street* "Greed is good" speech and encomium to myopic opportunism. Does it work?

# Appendix C

# Resources on the Web

- Swarm and generally for information on agent-based (and individual-based modeling):

  http://www.swarm.org/wiki/Main_Page

- Home page of NetLogo:

  http://ccl.northwestern.edu/netlogo/

- Agent-Based Computational Economics: Growing Economies from the Bottom Up

  Group at Iowa, led by Leigh Tesfatsion.

  http://www.econ.iastate.edu/tesfatsi/ace.htm

- *Emergence* on-line zine:

  http://www.emergence.org/

- MIT Media Lab site on emergence:

  http://llk.media.mit.edu/projects/emergence/contents.html

- Stanford Encyclopedia of Philosophy, Game Theory

  http://plato.stanford.edu/entries/game-theory/

  Nice section on modular rationality.

- Russell Sage Foundation, Evolution and the Capacity for Commitment

  http://www.russellsage.org/publications/titles/pdf_files/

Randolph M. Nesse, editor

http://www.russellsage.org/publications/titles/evolutioncapacit

- Al Roth's game theory and experimental economics page

  http://www.economics.harvard.edu/~aroth/alroth.html

- Learning and Evolution in Games

  Drew Fudenberg and David K. Levine

  June 26, 1996; revised July 30, 1996

  http://levine.sscnet.ucla.edu/Papers/Essay/ESSAY7.htm

- Who's who and what's what in the history of rational choice theory, broadly defined

  http://faculty.fuqua.duke.edu/~rnau/choice/whoswho.htm

  http://faculty.fuqua.duke.edu/~rnau/choice/

  http://faculty.fuqua.duke.edu/~rnau/choice/513read.html

- William James site

  http://www.emory.edu/EDUCATION/mfp/james.html

- "Complete Classes of Strategies for the Classical Iterated Prisoner's Dilemma," Bruno BEAUFILS - Jean-Paul DELAHAYE - Philippe MATHIEU

  http://www.lifl.fr/IPD/references/from_lifl/ep98/html/

- No Preference: A survey of Cognitive Responses to Rational Choice Theory by: Adam Stone

  http://ist-socrates.berkeley.edu:4050/ratchoice.html

- SMAC's Bibliography on the Iterated Prisoner's Dilemma

  http://www.lifl.fr/IPD/bib.html

- Useful bibliography.

  http://merlin.fae.ua.es/fvega/bib.PDF

- Smart Mobs – interesting set of links

  http://www.smartmobs.com/book/book_bib_ch_3_0_30.html

- ON THE EARLY HISTORY OF EXPERIMENTAL ECONOMICS BY Alvin E. Roth

  http://www.economics.harvard.edu/˜aroth/history.html

- ”Using Adaptive Agent-Based Simulation Models to Assist Planners in Policy Development: The Case of Rent Control,” by Robert N. Bernard

  http://www.santafe.edu/sfi/publications/wpabstract/199907052

- ”AGENT-BASED SIMULATION OF DYNAMIC ONLINE AUCTIONS,” Hideyuki Mizuta and Ken Steiglitz

  http://www.cs.princeton.edu/˜ken/dynamicauction.pdf

- The Emergence of Parties: An Agent-Based Simulation Darren Schreiber University of California, Los Angeles dschreib@ucla.edu

  http://www.bol.ucla.edu/˜dschreib/EmergingParties/dschreibMPSA00.pdf

- *Journal of Artificial Societies and Social Simulation*

  http://www.soc.surrey.ac.uk/JASSS

# Bibliography

[1] Graham T. Allison and Philip Zelikow. *Essence of Decision: Explaining the Cuban Missile Crisis*. Longman, second edition, 1999.

[2] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, Inc., New York, NY, 1984.

[3] Derek Bickerton. *Language and Species*. The University of Chicago Press, Chicago, IL, USA, 1990.

[4] Ken Binmore. *Fun and Games: A Text on Game Theory*. D.H. Heath and Company, Lexington, MA, 1992.

[5] Paul Bloom. *Descartes' Baby: How the Science of Child Development Explains what Makes Us Human*. Basic Books, New York, NY, 2004.

[6] Valentino Braitenberg. *Vehicles: Experiments in Synthetic Psychology*. MIT Press, Cambridge, MA, 1986.

[7] Adam M. Brandenburger and Barry J. Nalebuff. *Co-opetition*. Currency Doubleday, New York, NY, 1996.

[8] Rodney A. Brooks. *Flesh and Machines: How Robots Will Change Us*. Pantheon Books, New York, NY, 2002.

[9] Scott Camazine, Jean-Louis Deneubourg, Nigel R. Franks, James Sneyd, Guy Theraulaz, and Eric Bonabeau. *Self-Organization in Biological Systems*. Princeton University Press, Princeton, NJ, 2001.

[10] Robyn M. Dawes. Social dilemmas. *Annual Review of Psychology*, 31:169–193, 1980.

[11] Jared Diamond. *Collapse: How Societies Choose to Fail or Succeed*. Viking, New York, NY, 2004.

[12] Avinash Dixit and Barry Nalebuff. *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*. W.W. Norton & Company, New York, NY, 1991.

[13] Jon Elster. *Ulysses and the Sirens : Studies in Rationality and Irrationality*. Cambridge University Press, Cambridge, UK, 1984.

[14] Jon Elster, editor. *Rational Choice*. New York University Press, New York, NY, 1986.

[15] Jon Elster. *The Cement of Society: A Study of Social Order*. Studies in rationality and social change. Cambridge University Press, Cambridge, UK, 1989.

[16] Joshua M. Epstein and Robert Axtell. *Growing Artificial Societies: Social Science from the Bottom Up*. The MIT Press, Cambridge, MA, 1996.

[17] Ido Erev and Alvin E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4):848–881, 1998.

[18] Christopher D. Fiorillo, Philippe N. Tobler, and Wolfram Schultz. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, 299:1898–1902, 2003.

[19] Merrill M. Flood. Some experimental games. Research Memorandum RM-789, RAND Corporation, Santa Monica, CA, 1952.

[20] Merrill M. Flood. Some experimental games. *Management Science*, 5(1):5–26, 1958.

[21] David B. Fogel. *Blondie24: Playing at the Edge of AI*. Morgan Kaufmann, San Francisco, CA, 2002.

[22] Robert H. Frank. *Passions within Reason: The Strategic Role of Emotions*. W. W. Norton & Co. Inc., New York, NY, 1988.

[23] Robert H. Frank. *The Winner-Take-All Society: Why the Few at the Top Get So Much More Than the Rest of Us*. Penguin, New York, NY, 1996.

[24] Robert H. Frank. *Luxury Fever: Why Money Fails to Satisfy in an Era of Excess*. The Free Press, New York, NY, 1999.

[25] Robert H. Frank. *What Price the Moral High Ground? Ethical Dilemmas in Competitive Environments*. Princeton University Press, Princeton, NJ, 2004.

[26] Drew Fudenberg and Jean Tirole. *Game Theory*. The MIT Press, Cambridge, MA, 1991.

[27] Herbert Gintis. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, Princeton, NJ, 2000.

[28] David Grann. Stealing time: What makes rickey henderson run? *The New Yorker*, pages 52–9, 12 September 2005.

[29] Paul Grice. *Studies in the Way of Words*, chapter Logic and Conversation, pages 22–40. Harvard University Press, Cambridge, MA, 1989 (originally, 1967). ISBN: 0-674-85270-2.

[30] Patrick Grim, Gary Mar, and Paul St. Denis. *The Philosophical Computer: Exploratory Essays in Philosophical Computer Modeling*. The MIT Press, Cambridge, MA, 1998.

[31] Volker Grimm and Steven F. Railsback. *Individual-based Modeling and Ecology*. Princeton Series in Theoretical and Computational Biology. Princeton University Press, Princeton, NJ, 2005.

[32] Garrett Hardin. The tragedy of the commons. *Science*, 162(3859):1243–1248, 13 December 1968.

[33] Phil Hellmuth, Jr. *Play Poker Like the Pros*. HarperCollins Publishers, Inc., New York, NY, 2003.

[34] Richard J. Herrnstein. *The Matching Law: Papers in Psychology and Economics*. Harvard University Press, Cambridge, MA, 1997.

[35] John H. Holland. *Emergence: From Chaos to Order*. Perseus Books, Reading, MA, 1998.

[36] John H. Kagel and Alvin E. Roth, editors. *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ, 1995.

[37] J. R. Krebs and N. B. Davies. *An Introduction to Behavioural Ecology*. Blackwell Scientific Publications, Oxford, UK, third edition, 1993.

[38] David M. Kreps. *A Course in Microeconomic Theory*. Princeton University Press, Princeton, NJ, 1990.

[39] David M. Kreps. *Game Theory and Economic Modeling*. Clarendon Press, Oxford, England, 1990.

[40] Michael Lewis. *Liar's Poker: Rising Through the Wreckage on Wall Street*. W.W. Norton & Company, New York, NY, 1989.

[41] Michael Lewis. *Moneyball: The Art of Winning an Unfair Game*. W.W. Norton & Company, New York, NY, 2003.

[42] Michael Lewis. The ballad of Big Mike. *The New York Times*, 24 September 2006. `http://select.nytimes.com/gst/abstract.html?res=F60B14F83B550C778EDDA00894DE404482`.

[43] B. H. Liddell Hart. *Strategy*. New American Library, 1988.

[44] R. Duncan Luce and Howard Raiffa. *Games and Decisions*. John Wiley, New York, NY, 1957. Reprinted by Dover Books, 1989.

[45] Ian C. MacMillan, Alexander B. van Putten, and Rita Gunther McGrath. Global gamesmanship. *Harvard Business Review*, May 2003. Reprint R0305D.

[46] Lynn Margulis and Dorion Sagan. *Acquiring Genomes: A Theory of the Origins of Species*. Basic Books, Inc., New York, NY, 2002.

[47] John Maynard Smith. *Evolution and the Theory of Games*. Cambridge Univesity Press, New York, NY, 1982.

[48] Harold McGee. *On Food and Cooking: The Science and Lore of the Kitchen*. Scribner, New York, NY, 2004.

[49] James McManus. *Positively Fifth Street: Murderers, Cheetahs, and Binion's World Series of Poker*. Farrar, Straus, Giroux, New York, NY, USA, 2003.

[50] Nancy J. Nersessian. The cognitive basis of model-based reasoning. In Peter Carruthers, Stephen Stich, and Michael Siegal, editors, *The Cognitive Basis of Science*, pages 133–153. Cambridge University Press, Cambridge, UK, 2003.

[51] Michael E. Porter. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. The Free Press, New York, NY, 1980/1998.

[52] William Poundstone. *Prisoner's Dilemma: John von Neumann, Game Theory and the Puzzle of the Bomb*. Anchor Books, Doubleday, New York, NY, 1992.

[53] William Poundstone. *Fortune's Formula: The Untold Story of the Scientific Betting System that Beat the Casinos and Wall Street*. Hill and Wang, New York, NY, 2005.

[54] Anatol Rapoport. *Two-Person Game Theory*. The University of Michigan Press, Ann Arbor, MI, 1966. Republished by Dover Publications, Inc., Mineola, New York, 1999.

[55] Anatol Rapoport and Albert M. Chammah. *Prisoner's Dilemma: A Study in Conflict and Cooperation*. The University of Michigan Press, Ann Arbor, MI, 1965.

[56] Anatol Rapoport, Melvin J. Guyer, and David G. Gordon. *The 2×2 Game*. The University of Michigan Press, Ann Arbor, MI, 1976.

[57] Alvin E. Roth and Marilda A. Oliveira Sotomayor. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press, Cambridge, United Kingdom, 1990.

[58] Robert M. Sapolsky. *Monkeyluv: And other Essays on Our Lives as Animals*. Scribner, New York, NY, 2005.

[59] Robert M. Sapolsky. A natural history of peach. *Foreign Affairs*, 85(1):`http://www.foreignaffairs.org/20060101faessay85110/robert--m--sapolsky/a--natural--history--of--peace.html`, January-February 2006.

[60] Thomas C. Schelling. *Arms and Influence*. Yale Univ Pr, New Haven, CT, 1967.

[61] Thomas C. Schelling. *Micromotives and Macrobehavior*. W.W. Norton & Company, New York, NY, 1978.

[62] Thomas C. Schelling. *The Strategy of Conflict*. Harvard Univ Pr, Cambridge, MA, 1980 (originally published 1960).

[63] Thomas C. Schelling. *Choice and Consequence*. Harvard Univ Pr, Cambridge, MA, 1985.

[64] Frederic Schick. *Making Choices: A Recasting of Decision Theory*. Cambridge University Press, Cambridge, United Kingdom, 1997.

[65] Amartya Sen. *Rationality and Freedon*, chapter Introduction: Rationality and Freedom, pages 3–64. Harvard University Press, Cambridge, MA, 2002.

[66] Amartya K. Sen. Rational fools: A critique of the behavioural foundations of economic theory. *Philosophy and Public Affairs*, 6:317–344, 1977.

[67] Martin Shubik. *Game Theory in the Social Sciences*. The MIT Press, Cambridge, MA, 1982.

[68] Brian Skyrms. *Evolution of the Social Contract*. Cambridge University Press, Cambridge, UK, 1996.

[69] Brian Skyrms. The stag hunt. World Wide Web, 2001. Proceedings and Addresses of the American Philosophical Association. http://www.lps.uci.edu/home/fac-staff/faculty/skyrms/StagHunt.pdf. Accessed September 2004.

[70] D. W. Stephens, C. M. McLinn, and J. R. Stevens. Discounting and reciprocity in an iterated prisoner's dilemma. *Science*, 298:2216–2218, 2002.

[71] Steven Stoft. *Power System Economics: Designing Markets for Electricity*. John Wiley & Sons, Inc., New York, NY, 2002.

[72] Mitchell Zuckoff. The perfect mark: How a Massachusetts psychotherapist fell for a Nigerian e-mail scam. *The New Yorker*, 15 May 2006. `http://www.newyorker.com/fact/content/articles/060515fa_fact`.

# Index

Big Bang, 24
bimatrix games, 86

choice
- acute, 15
- chronic, 15
- distributed, 15
- parametric, 15

decision
- acute, 15
- chronic, 15
- distributed, 15
- parametric, 15

effective rationality, 40
emergence, 60
equilibrium
- replicator dynamic, 77

Evolutionarily Stable Strategy, 77

game
- zero-sum, 83

game form, 68
game matrix, 68
games
- Battle of the Sexes, 7
- Stag Hunt, 7

Henderson, Rickey, 54

McGee, Harold, 24

movies
- *Wall Street*, 21

Nash equilibrium, 71
Nash equilibrium principle, 71

one-shot game, 5

Pareto frontier, 71
Pareto frontier principle, 71
Pareto optimal outcome, 71
PD property, 84
principle of dominance, 70

rationality
- effective, 40
- Nash equilibrium principle, 71
- Pareto frontier principle, 71
- principle of dominance, 70

replicator dynamic, 76
replicator dynamic equilibrium, 77

strategy, 68
symmetric form games, 86

tell, 54

zero-sum game, 83