1 Game Theory and Literature: An Overview

1.1 Introduction

Fiction has been one of the most fertile grounds for humanistic applications of game theory. Novels, short stories, plays, narrative poems, and even the librettos of operas—all have been subject to game-theoretic exegesis, as have stories in the Bible. It is these applications, all of which involve noncooperative game theory, that I survey in this chapter.[1]

I will sketch but not present technical details of several models, in part because my primary purpose is to emphasize literary themes amenable to game-theoretic treatment, and in part because I will model in depth the choices of characters in specific literary works later. This survey is meant to be reasonably comprehensive, but it is only a survey: Those interested in the modeling details will need to consult the sources cited, which I hope to encourage by stimulating interest in literature as a fruitful source of ideas worthy of strategic analysis.[2]

Game theory provides a parsimonious framework and an important set of tools for the literary theorist. Although there are no rigorous tests to determine what the "right" interpretation of a work of fiction is, some interpretations are clearly more tenable than others. Game theory

---

1. This chapter is adapted from Brams 1994a with permission. Some specialized terms (e.g., *noncooperative game theory* in the present sentence) often will not be defined in this chapter but can be found in the glossary and will be discussed later in the book.

2. I have excluded from this survey work that originally appeared in movies and television. This is not because I think that writing for these popular media is superficial or otherwise unworthy of game-theoretic treatment. On the contrary, some of the plots, such as for the popular TV series *The Sopranos*, reflect game-theoretic reasoning par excellence. But this enormous body of work, in my opinion, requires its own book-length treatment.

explicates the strategic choices of characters by illuminating the linkage between their motives and their actions. It also offers insight into certain interpretive questions, such as whether the ordinary calculations of literary characters can explain their extraordinary actions in some of literature's great tragedies.

My review of applications of game theory to literature has both a critical and a historical dimension. In an attempt to gain an understanding of how and why the applications evolved as they did, I asked several people who have applied game theory to fiction three questions, which are given in section 1.2. I also give in this section a chronological listing (table 1.1) of the literary works and opera librettos to which game theory has been applied.

In later sections, I make use of the respondents' answers to see what inspired them to tackle a particular literary work and what they think a game-theoretic perspective brings to the understanding and interpretation of that work. I also asked them whether this work, in turn, stimulated them to probe new theoretical questions.

In discussing fictional works analyzed by game theorists, I begin in section 1.3 by showing how two authors (Arthur Conan Doyle and Edgar Allan Poe), instead of confronting the consequences of the so-called minimax theorem in their fiction, sidestepped them. I then present in some detail an application that illustrates how one writer (William Faulkner) captured the spirit of the theorem, even invoking a fictitious "Player" to make seemingly random choices, which, according to the minimax theorem, are optimal under certain conditions.

In section 1.4, problems of coalition formation in zero-sum games take center stage in a play (by Harold Pinter); they also pervade a political novel (by C. P. Snow), arguably to the detriment of character development. One analyst, in fact, contends that emotions tend to be submerged when there are clear-cut winners and losers, whereas ambivalence is better expressed in literary plots with nonzero-sum elements.

Several works of fiction that may be interpreted as nonzero-sum games are reviewed in sections 1.5 and 1.6—some quite critically, because of what I believe are some misuses of game theory in these applications. Contrary to the views of some, I argue in section 1.5 that great tragedies—such as Shakespeare's *Othello*, Puccini's *Tosca*, and Shakespeare's

Richard III (usually classified as a history but certainly a tragedy for many of its characters)—can be well understood in rational-choice terms: Their high drama is less a product of irrational behavior than a train of events, and rational choices in response to them, that spiral out of control. Joseph Heller's *Catch-22* illustrates the frustration that may burden characters who see no way out of their predicaments.

What start out as rather mundane calculations may become anything but routine in their consequences for the players. These include a classic coordination problem (in an O. Henry story), compounded by incomplete information that also plagues Portia's suitors in a game they play with her father in Shakespeare's *The Merchant of Venice*. But incomplete information also creates opportunities for signaling and credible commitments, which are prominent in works by Homer, Shakespeare, Joseph Conrad, George V. Higgins, and Richard Wagner that are briefly discussed in section 1.6.

In section 1.7, I consider game-theoretic analyses of the devil in Johann Wolfgang von Goethe's *Faust* and of God in the Hebrew Bible. *Faust* is modeled as a differential game (described later), whereas in chapter 2 two stories in the Hebrew Bible are viewed as simple ordinal games, in which players can order or rank outcomes from best to worst but cannot attach cardinal utilities, or numerical values, to them. The latter games and a biblical story in chapter 5 are interconnected by the continuing presence of God, who exhibits an abiding interest in using threats to cement His reputation and thereby tries to deter future untoward actions, including some by His chosen people, the Israelites.

Sir Gawain and the Green Knight, a medieval narrative poem that has been explicitly modeled as a game of incomplete information, is discussed in section 1.8. In the model, reputation plays a prominent role in explaining the actions of the main characters. Also modeled is the dual character of Sir Gawain, who, in an intrapsychic game between his two natures, has only incomplete information about the Green Knight.

In section 1.9 I offer some observations on the state of the art—an apt phrase, because game theory, as applied to literature, is still more an art than a science. I also discuss possible new uses of the theory, such as the exploration of games played between the author and the reader that incorporate the expectations of each player. I conclude that game theory

offers a structure for clarifying strategic issues in plot design and character development that literary theories often ignore.

1.2   Method of Inquiry

Besides considering the merits of different applications, it is useful to inquire how game theory has gained the foothold that it has in literary analysis.[3] For this purpose, I wrote several game theorists who at some time had applied game theory to literature and asked them the following questions:

l. What inspired you to make the application(s) you did? Are there other humanistic works that you considered?

2. Does game theory offer unique insights into these works? Or does it offer more a framework for elucidating strategic conflict that these works illustrate?

3. Do these applications make a contribution to game theory, viewed as an applied field? What kind?

A number of respondents did not confine themselves just to these questions but went on to express wide-ranging views, replete with examples, of what benefits game theory can bring to the study of the humanities and vice versa.

To organize this rather open-ended information, I have grouped applications partly in terms of the theory (e.g., zero-sum games, games of coordination) and partly in terms of literary motifs (e.g., the role of emotions, the rational foundations of tragedy). At the same time I try to give a historical perspective to the applications by reporting what influenced game theorists, told mostly in their own words, to make the applications they did, and what they see as their benefits to both literature and game theory.

Because many readers will not be familiar with all the applications that have been made, I have included some information about the applications themselves, especially if they seemed representative or unique in

---

3. I use the word *foothold* with care: Game theory has hardly taken literary analysis by storm, perhaps in part because the theory is often misunderstood by humanists.

their approaches. In one instance, I describe in some detail how one writer (William Faulkner), in a grim tale of pursuit and mayhem, better captured the unpredictability of strategies in two-person zero-sum games without a saddlepoint—in which mixed or randomized strategies are optimal—than the usual authors cited on this subject.

The examples I discuss illustrate how game theory can enhance one's understanding of the strategic elements of fiction. The feedback may also go in the other direction, whereby a story, for example, may force the theorist to rethink how game theory may need to be extended or refined to mirror the strategic situation that it describes.

Before discussing some of the applications and looking at responses to the questions I posed, a chronological listing of literary works to which game theory has been applied is worth perusing (see table 1.1). Game-theoretic exegeses of these works range from a few sentences to lengthy articles. They also vary greatly in technical level, from relatively informal strategic descriptions to sophisticated mathematical analyses.

In applying game theory to literary works, it is useful to bear in mind the admonition of Howard (1971, 146) that "skillful authors often conceal certain essential motivations of their characters in order to reproduce the mystery we often feel in real life as to why people behave in the way they do." Game theory helps one unravel the mystery, at least in literary works in which there is a plot and the characters indicate reasons for acting the way they do. Plotless or surrealistic works, while they may have aesthetic appeal, are least amenable to this kind of analysis.

1.3   Avoidance and Acceptance of the Minimax Theorem

A number of conflicts in the literary works I assay can be viewed as zero-sum, in which what one player wins the other players lose. If there are only two players, the fundamental theorem of game theory or minimax theorem—proved in von Neumann (1928), sixteen years before the appearance of the first edition of his monumental treatise on game theory with Morgenstern (von Neumann and Morgenstern 1944/1953)—established that there is always a solution that guarantees the players at least a particular value, whatever the opponent does. However, it may be in mixed strategies, which are randomized choices designed to keep an opponent guessing.

Table 1.1
Literary works and operas to which game theory has been applied

The listing below is given in the order in which the first application was made. If I have quoted from a work in the text, the source used is given in the References. Not included in this listing are works analyzed by the literary scholars mentioned in note 16.

1. Arthur Conan Doyle, *Sherlock Holmes* (several books in a series) (Morgenstern 1935; von Neumann and Morgenstern 1944/1953; Vorob'ev 1968)—mystery

2. William Shakespeare, *The Merchant of Venice* (Williams 1954/1966)—play

3. William Shakespeare, *Othello* (Rapoport 1960; Teodorescu-Brinzeu 1977)—play

4. William Shakespeare, *Measure for Measure* (Schelling 1960)—play

5. O. Henry (William Sidney Porter), "The Gift of the Magi" (Rapoport 1960; Vorob'ev 1968; Rasmusen 1989)—short story

6. Giacomo Puccini, *Tosca* (Rapoport 1962)—opera

7. William Shakespeare, *Henry V* (Schelling 1966; Dixit and Nalebuff 1991)—play

8. Joseph Conrad, *The Secret Agent* (Schelling 1966)—novel

9. Alexander Pushkin, *Eugene Onegin* (Vorob'ev 1968)—novel

10. William Shakespeare, *Hamlet* (Vorob'ev 1968; Brams 1994; Howard 1996)—play

11. Edgar Allan Poe, "The Purloined Letter" (Davis 1970)—short story

12. Harold Pinter, *The Caretaker* (Howard 1971)—play

13. William Shakespeare, *Richard III* (Lalu 1977)—play

14. Agatha Christie, *The Mousetrap* (Steriadi-Bogdan 1977)—play

15. Homer, *The Odyssey* (Elster 1979; Mehlmann 2000)—mythology

16. *Bible* (Brams 1980/2003; Dixit and Nalebuff 2008; Brams and Kilgour 2009)—religious work

17. C. P. Snow, *The Masters* (Riker 1986)—novel

18. Boris Pasternak, *Dr. Zhivago* (Howard 1988)—novel

19. Johann Wolfgang von Goethe, *Faust* (Mehlmann 1990, 2000)—play

20. *Sir Gawain and the Green Knight* (anonymous) (O'Neill 1991)—medieval poem

21. *The Feast of Bricriu* (anonymous) (O'Neill 1991)—medieval tale

22. William Faulkner, *Light in August* (Brams 1994a, 1994b)—novel

23. William Shakespeare, *Hamlet* (Brams 1994b)—play

24. William Shakespeare, *King Lear* (Chami 1996; Dixit and Nalebuff 2008)—play

25. Aristophanes, *Lysistrata* (Brams 1997a)—play

26. William Shakespeare, *Macbeth* (Brams 1997a)—play

27. Joseph Heller, *Catch-22* (Brams and Jones 1999; Dixit and Nalebuff 2008)—novel

28. William Goldman, *The Princess Bride* (Dixit and Nalebuff 2008)—novel

29. Giuseppe Verdi, *Rigoletto* (Dixit and Nalebuff 2008)—opera

30. Friedrich von Schiller, *Wallenstein* (Holler and Klose-Ullman 2008)—play

31. Richard Wagner, *Lohengrin* (Huck 2008)—opera

32. Richard Wagner, *Tannhäuser* (Harmgart, Huck, and Müller 2008, 2009)—opera

33. William Shakespeare, *Much Ado about Nothing* (Chwe 2009)—play

34. Richard Wright, *Black Boy* (Chwe 2009)—novel

Mixed strategies introduce an element of uncertainty into the play of a game and turn a certain guarantee into a guarantee of an average amount, or an expected value. In *Sherlock Holmes*, Conan Doyle portrayed this element in the difficult choice that he gave to Sherlock Holmes, pursued by the notorious Professor Moriarty, of whether to get off his train at Dover or at Canterbury, an intermediate stop. He chose Canterbury, anticipating that Moriarty would take a special faster train to Dover to try to catch him if he got off there. Holmes's anticipation was correct, but Morgenstern (1935, 174) asks the critical question: "But what if Moriarty had been still more clever, had estimated Holmes's mental abilities better and had foreseen his actions accordingly?"

Morgenstern originally posed this question in his first book (Morgenstern 1928), which coincidentally was published the same year as von Neumann's proof of the minimax theorem. Unaware of the minimax theorem, Morgenstern saw the Holmes-Moriarty story as an illustration of a paradox in which "an endless chain of reciprocally conjectural reactions and counter-reactions . . . can never be broken by an act of knowledge but always only through an arbitrary act—a resolution" (Morgenstern 1935, 174). Although prescient in recognizing the arbitrariness of the resolution, Morgenstern did not yet know its mixed-strategy form—which involves randomly choosing among pure strategies according to some probability distribution—that had actually been calculated for specific games before the minimax theorem was proved (Dimand and Dimand 1990).[4]

Conan Doyle's resolution, on the other hand, was to make Holmes one whit more clever than Moriarty, ignoring that Moriarty himself might have been able to make an anticipatory calculation similar to Holmes's. Moreover, the matter does not end there: Holmes could have anticipated Moriarty; Moriarty, Holmes; and so on, leading to Morgenstern's "endless chain of reasoning."

In the short story "The Purloined Letter," Edgar Allan Poe broke this chain by assuming that an extremely clever boy could always calculate exactly how far ahead his less clever opponents would reason. Then, in a game in which this boy guessed whether an opponent was concealing

---

4. But in von Neumann and Morgenstern 1944/1953, 177–178, the authors proposed a 2×2 payoff matrix for the Holmes-Moriarty game and found the mixed-strategy solution, pointing out that Holmes's "complete victory" is "somewhat misleading."

an odd or an even number of marbles in his hand, the clever guesser would be able to anticipate his opponent, and whether the opponent was a "simpleton" or someone of great cunning (but not greater than his own). Here is how the clever boy, according to Poe, was able to do this:

When I wish to find out how wise, or how stupid, or how good, or how wicked is any one, or what are his thoughts at the moment, I fashion the expression of my face, as accurately as possible, in accordance with the expression of his, and then wait to see what thoughts or sentiments arise in my mind or heart, as if to match or correspond with the expression. (quoted in Davis 1970, 26–27)

Labeling this reasoning "tortuous," Davis points out that "the adversary can undo all the boy's labor by simply randomizing, in which case it will take nothing short of the Delphic Oracle to gain an edge." This is also true in William Goldman's *The Princess Bride* (1973), in which the two antagonists try to outguess each other about which of two cups of wine has been poisoned. Goldman's clever twist is that both cups have been poisoned, but one person has made himself immune to the poison, so it doesn't matter which cup he takes (Dixit and Nalebuff 2008, 141–143).

To return to "The Purloined Letter," Davis wrote that he chose this example "because of the irony of Poe's comment: 'As poet and mathematician, he would reason well; as mere mathematician, he could not have reasoned at all'" (quoted in Davis 1970, 27). On the contrary, Davis argues, "as mathematician (using the minimax theorem) *he need not reason at all*—random play is sufficient to confound the boy" (italics in original).

Hence, it is the mathematician—who, according to Poe, "could not have reasoned at all"—who can play this game at least to a draw, even against an incredibly clever opponent.[5] By randomizing, the mathematician robs the opponent of any control over the outcome and so ensures the value of the game.

This is a fundamental insight of the minimax theorem that neither Conan Doyle nor Poe seems to have understood. (To be sure, the cunning

---

5. Of course, knowing *exactly* how clever an opponent is, the boy can always win, but this cleverness is better characterized as omniscience, which even the biblical God did not possess (Brams 1980/2003, 1983/2007). Moreover, it can lead to a "paradox of omniscience" (section 9.7), which hurts the omniscient player.

these writers attributed to their characters may make for better fiction than resolving each game with the flip of a coin.) But just the opponents' *knowledge* of this greater cunning would have been sufficient for them to even the score by choosing mixed strategies, because it protects them from being outsmarted. Apparently, however, they did not have even this knowledge—or, more accurately, the writers did not choose to give it to them.

Not all writers portray their characters in such a one-sided fashion. For example, knowledge is more shared, and calculations more even-handed, in the climactic scene of William Faulkner's novel *Light in August* (first published in 1932), in which Percy Grimm pursues Joe Christmas, a prisoner who has just escaped his captors. Though hand-cuffed in front, Christmas, like Grimm, has a gun. Grimm thinks, as the pursuit by bicycle and on foot nears its end, like a game theorist: "He can do two things. He can try for the ditch again, or he can dodge around the house until one of us gets a shot" (Faulkner 1950, 404).

Grimm runs for the ditch, but soon he realizes that "he had lost a point. That Christmas had been watching his legs all the time beneath the house. He said, 'Good man'" (Faulkner 1950, 405).

The pursuit continues until it reaches the house of Reverend Hightower, who, though knocked down and injured by Christmas when Christmas burst in, refuses to tell Grimm in which room Christmas has run to hide. But a fictitious "Player"—a literary device in the novel—guides Grimm. Grimm storms into the kitchen, where Christmas has overturned a table to protect himself, and Grimm fires his revolver. Before Christmas dies, Grimm castrates him with the butcher knife he finds in the kitchen.

This, the most gruesome scene in the novel, contrasts sharply with Grimm's pursuit of Christmas, which is all cool calculation. Faulkner seems to have invented Player to epitomize the calm and deliberate mind of the fanatic; Grimm, who is "moved," as in a parlor game, by Player, is utterly devoid of emotion, except when he explodes with savagery in the end. The beast in Grimm coexists with the cerebral Player, which is a juxtaposition that game theory normally does not entertain when it posits a player with one set of preferences.[6]

6. If more than one type of player is allowed, as in games of incomplete information, only one type is actually the true type—there are not different types embodied in a single player (e.g., with multiple personalities), though in section 1.8 I discuss intrapsychic games.

Grimm seems genuinely in the dark after he rushes into Hightower's house, repeatedly asking where Christmas went: "'Which room?' Grimm said, shaking him. 'Which room, old man?'" (Faulkner 1950, 406). After Grimm asks once again, Hightower attempts to exonerate Christmas for the alleged murder he committed, but Grimm "flung the old man aside and ran on" (at random?) into the kitchen (Ibid.).

Unlike Conan Doyle and Poe, Faulkner beautifully captures the uncertainty inherent in mixed strategies—and how to act in the face of this uncertainty. And act Grimm does: first, to his own disadvantage when he discovers that Christmas could follow his movements as he ran toward the ditch; second, to his advantage when, "waiting for Player to move him again" (Faulkner 1950, 406), he rushes into the kitchen. Faulkner has little to say about the motivations behind Christmas's choices, but it seems they were essentially arbitrary, as if Christmas, as well as Grimm/Player, was randomizing.

Faulkner does *not* assume that one player had superior calculational abilities. True, Grimm has Player on his side, so to speak, but this device, in my view, reinforces the desultory character of Grimm's choices. Calculated they may have been, but because Grimm, at each stage of the pursuit, has only imperfect information, he can never be sure what his best choice is. Grimm "won," finally, not because of sheer cleverness but because the game was unfair—the odds were heavily stacked against the fugitive, Christmas, whom Grimm so relentlessly hunted down.[7]

I have offered this analysis of a scene from *Light in August* to suggest that Faulkner is one fiction writer who had an astute if implicit understanding of mixed strategies in two-person zero-sum games of imperfect information. Doubtless, other examples could be found. While the scenes that Morgenstern and Davis discussed in *Sherlock Holmes* and "The Purloined Letter" have the earmarks of games in which mixed strategies are optimal, both Conan Doyle and Poe shrank from making their protagonists' opponents as smart as the protagonists themselves. They got tidy results that way, but the minimax solution in games of imperfect information shows that not all conflicts can be resolved by outguessing. Faulkner understood this, even if the formal calculations eluded him.

---

7. I assume that once Christmas is cornered, Grimm has the upper hand. Until then, however, the players seem evenly matched.

1.4 Are Zero-Sum Games Emotionless?

Zero-sum games with more than two players raise entirely new theoretical questions, chiefly related to what coalitions are likely to form and be stable. Nigel Howard (1990) reports that when he went to a performance of Harold Pinter's *The Caretaker*, he was struck by its similarity to the game of split-the-dollar—where a dollar (or better 99¢) is divided equally among three people unless at least two agree on another way of dividing it. This zero-sum game has no stable solution, because however the dollar is divided, there are always two players who can do better by agreeing on another split that excludes the third.

In the case of *The Caretaker*, there is a pecking order of respect, such that the least-respected character can always suggest to one of the other two a deal in which they give each other greater mutual respect at the expense of the third. Each of the play's three acts deals with the formation of one of the three two-person coalitions.

Howard (1971) describes the formation and disintegration of each coalition in the three acts, involving two brothers who share a house and a third man who might become their caretaker. The play ends with "no relationships," but with the possibility that new relationships will form once again, "causing the three acts to be repeated in sequence again and again" (Howard 1971, 145). Although *The Caretaker* "is almost classically austere and simple from a game-theoretic point of view" (Ibid.), Howard argues that "Pinter's view is however interesting in that at least he has risen to the level of dramatizing a three-player interaction" (Howard 1990).[8]

In analyzing C. P. Snow's *The Masters*, William H. Riker (1986) examines the more complex interactions of thirteen fellows in a Cambridge (UK) college, who must vote on a new master of their college in a zero-sum game (there are two candidates among the thirteen fellows, and only one can win). The novel is about the election campaign, in which "pride and ambition and humiliation and failure are displayed against a background of political bitterness" (Riker 1986, 52).

There are four switches in support for the two candidates as they vie for the votes of the eleven other fellows of the college. Riker shows how

---

8. It is worth noting that coalitional cycles of the kind that Howard identified can occur in nonzero-sum games; they are not exclusive to zero-sum games like split-the-dollar.

the maneuvers of one fellow, in particular, who abandons his original favorite the day before the election, ultimately succeed. Although Riker's analysis stresses social-choice theory rather than game theory (e.g., Riker shows that no logrolling is possible, based on the positions of the fellows on two dimensions), it is evident that the campaign is suffused with game-theoretic calculations.

The leaders of the two factions constantly plot to hold their coalitions together, and draw in new members, against opposition efforts to woo away potential defectors. Riker in fact explored this idea in an earlier game-theoretic model (Riker 1962); its best-known prediction—that only minimal winning coalitions will form under certain assumptions (the so-called size principle)—is exactly what happens in Snow's story.

Riker regards *The Masters* as uniquely political: It "is, so far as I know, the only one [novel] in which politics is not mere background but the very plot itself" (Riker 1986, 52); "all other novels concern character development, love affairs, hurried journeys, family history, etc." (Riker 1990). Riker admits that building coalitions is "hardly the stuff to release readers' adrenalin as do seductions, quarrels, or chases," but he believes "political ambition, and indeed political success, uniquely reveal tragic flaws in character," as demonstrated by Greek dramatists and Shakespeare (Riker 1986, 52).

To Howard (1990), by contrast, the most interesting conflicts are not zero-sum:

Such a zero-sum view is a common one, as shown by the frequent comparisons of politics or war with chess, poker, or football. I think it is unrealistic; all my experience with applying game theory leads me to think that people are both more clever than this (they don't see things as zero-sum when they aren't) and more stupid (the simplest game-theoretic model of their situation often shows them simple, gross, obvious things they have entirely failed to see).

For Howard, "Pinter's view is the bleak, cynical one obtained by supposing that adults do not grow out of the 'zero-sum' mentality of children," which he disdains:

In fact, I would think this mistaken view is a rare, sophisticated aberration of the 20th century elitism. In a two-person game, zero-sumness means absence of emotion, deceit, preference change, etc.—all the things that artists have traditionally been most interested in. (In the three-person case, zero-sumness no longer

excludes these—it merely means that they are necessarily exercised at the expense of a third party, as in *The Caretaker*). (Howard 1990)

In depicting *n*-person games of coalition formation, Pinter and Snow illustrated the fragility of coalitions in zero-sum games, which Howard (1971) and Riker (1986) explicated by showing how alliances may unravel. From a literary point of view, however, the question is not the stability of coalitions but whether such works are only plot and calculation—or something more. And if the latter, does the something more require that characters transcend their own rationality? I argue not: In an appropriate game, rationality—with respect to some plausible goals—perfectly well explains the choices of most characters we find compelling in literature.

In section 1.5, I turn to applications of game theory that have been made to nonzero-sum games. Whether game theory can illuminate emotions in such games, as Howard maintains, or is better suited to elucidating purely political plots and stories, as Riker maintains, is a question whose answer may shed light on the types of literature that have been selected for game-theoretic scrutiny. I pursue this question in later chapters with new applications, focusing on the emotions of frustration and anger in chapters 7 and 10.

1.5 The Rationality of Tragedy

The early use of game theory in literary exegesis includes Anatol Rapoport's interpretations of Shakespeare's *Othello* and Puccini's *Tosca* as nonzero-sum games (Rapoport 1960, 1962). In a two-person normal-form version of *Othello*, Othello may believe or not believe that Desdemona has been faithful, and Desdemona may deny or confess (falsely) her guilt; the tragedy occurs when Desdemona denies that she has given herself to Cassio, but Othello, with seeds of doubt planted by Iago, does not believe her.[9]

Rapoport also considers an extensive-form version of *Othello*, involving the four principals and "chance"; this game has fifty-five distinct

9. Neither does Hamlet believe that his uncle, Claudius, innocently acceded to the throne of his father after marrying his mother, Gertrude (Brams 1994b). *Hamlet* is another Shakespearean tragedy that I analyze as a game of incomplete information in section 9.2.

outcomes. The enormity of Desdemona's 16,384 strategies in the game tree (to be discussed in chapter 2) leads Rapoport to remark that "perhaps enough has been said about the practical difficulties of applying game theory in human affairs" (Rapoport 1960, 240). But he argues that "game theory stimulates us to think *about* conflict in a novel way" (Ibid., 242; italics in original) and also shows how interdependent decision situations can be "precisely characterized and rigorously analyzed" (Rapoport 1990).

After analyzing *Othello*, Rapoport turned to *Tosca* (Rapoport 1962), which he portrayed as a 2 × 2 Prisoners' Dilemma. Whereas jealousy fuels the plot in *Othello*, in *Tosca* it is Tosca and Scarpia's mutual betrayal that leads to its tragic end.

In another Shakespearean play, *Much Ado about Nothing*, Chwe (2009) compares the game Beatrice and Benedick play in in their comedic relationship with a game that has exactly the same structure in Richard Wright's dark autobiographical novel, *Black Boy* (Wright 1945). But the outcomes, love in *Much Ado* and hate in *Black Boy*, could not be more different, which can be explained, Chwe argues, because the game—Stag Hunt (sometimes called the Assurance Game or Coordination Game and described in note 9 of chapter 5)—has two Nash equilibria (to be discussed in chapter 2). At the better equilibrium, the players choose a risky strategy, which happened in *Much Ado*, whereas at the worse equilibrium, which happened in *Black Boy*, there is no risk. Chwe also analyzes trickster folk tales, which involve surprisingly sophisticated calculations that game theory helps to elucidate.[10]

In Teodorescu-Brinzcu's (1977) analysis of *Othello*, she assumes that Othello and Iago are involved in a zero-sum game, which, especially from Othello's perspective (who is sympathetic to Iago until the end), seems to me a misinterpretation. Second, she assigns payoffs so that Iago has a dominant strategy, and Othello a best response, but then argues that this "wise [minimax] solution" was not chosen because "it lacks dramatic consistency as it is very commonplace." Instead, "the psychological reality requires that in this clash of passions the Moor's jealousy and Iago's hatred should overcome any lucid calculations and drive them

---

10. In an earlier work, Chwe (2001) analyzed game-theoretic calculations that underlie rituals, especially those that depend on coordination and common knowledge, in both factual and fictional situations.

both to destruction"; indeed, they "die devoured by their own passion" (Teodorescu-Brinzcu 1977, 373).

Coupled with Desdemona's murder, this tragedy suggests to me that there were no winners, making the game decidedly nonzero-sum. Thus, I think the interpretation of this tragedy as zero-sum is untenable.

Teodorescu-Brinzcu's (1977) contention that great drama may require that the characters reach beyond themselves (irrationally?) to seize the moment—sealing their fate and, quite often, their destruction—deserves further comment. This view seems to be a tenet of Marcus's (1977) so-called Romanian school of mathematical linguistics and poetics, because it is also reflected in Lalu's (1977) game-theoretic analysis of Shakespeare's *Richard III*.[11]

Lalu (1977) analyzes this play as an extensive-form nonzero-sum game and concludes that

what the playwright considers as the optimal strategies are in fact optimal for the tension and the rhythm of the performance, seldom for the "actual life" of the character. A cautious hero would be uninteresting. Paradoxically, the optimal strategy of the character is, more often than not, that of "the mad risk." Therefore, the main characters may seldom be considered as perfectly rational players; as far as we view the play in terms of "a slice of life," the characters make mistakes. The optimal strategies for their destinies of actual human beings will seldom be followed; on the contrary, the characters will act following those strategies which the author (perhaps the only rational player) thinks optimal according to an aesthetic criterion. (Lalu 1977, 343)

Lalu (1977, 343) then asks what the point of applying game theory is and answers that she is interested in exploring *deviations* from rationality that are "optimal within the frame of the whole play, regarded as a work of art."

In my opinion, there is considerable arbitrariness in Lalu's assignment of specific numerical values to outcomes and specific probabilities to

---

11. On the other hand, Steriadi-Bogdan (1977), also a disciple of this school, argues in a game-theoretic analysis of Agatha Christie's play *The Mousetrap* that the characters made rational choices. But *The Mousetrap* is a detective story, or whodunit, which is not generally considered to be a great tragedy, whereas "in studying Shakespeare's *Othello*, namely Iago's strategy, you have to observe that Iago does not look for what in the Mathematical Game Theory is called the *best* strategy, but rather for the worst strategy" (Marcus 1990; italics in original). I remain unconvinced that Iago chose, say, a dominated strategy—at least in the beginning, when his plan seemed to be working quite nicely—but I agree that combining "strategic and psychological aspects . . . is a rather delicate task" (Ibid.).

chance events in *Richard III*. These assignments vitiate her claim that Richard chose his worst strategy, though I would not dispute her claim that Shakespeare sought "the ruin of the character [Richard] . . . for the sake of the tension of the performance" (Lalu 1977, 349).

The issue is whether this tension was achieved by making Richard's choices irrational. I think it was not, and an alternative and more defensible game-theoretic analysis—not to be developed here—would demonstrate that Richard was eminently rational. Briefly, the argument underlying this alternative interpretation is that Richard, brilliant and diabolical, knew that he could act boldly with a high likelihood of success; in fact, he rapidly dispatches several of his opponents at the beginning of the play. Although lacking the contemplative character of a Richard II or Hamlet, who seem to weigh options more carefully, as Lalu points out, Richard III, nevertheless, seems no less rational (and tragic) a hero.

I agree with Lalu that Richard III is not prudent, but prudence, which Lalu equates with the minimax principle and estimating the odds in lotteries, is not synonymous with rationality. And neither is a hero's "tragic fall" synonymous with irrationality. In fact, contrary to Lalu, I believe the tragic fall is made more, not less, poignant when characters are driven by an inexorable rationality toward some terrible end.

Interestingly, the hero (or anti-hero) of Joseph Heller's *Catch-22* (1961), John Yossarian, escapes tragedy, even though *catch-22* has come to signify a frustrating situation from which there is no escape. Dixit and Nalebuff (2008, 45) describe the situation in which Yossarian finds himself embroiled in a Prisoners' Dilemma. Brams and Jones (1999) argue that Yossarian played a different game, which I analyze in chapter 10. There I argue that Yossarian faced a catch-22, which can be modeled as a "generic game" that subsumes several specific games. One of these games models Yossarian's dilemma, whereas another models the difficult choices of players in medieval witch trials.

1.6   Coordination Problems, Signaling, and Commitment

Unlike *Othello* and *Tosca*, in which the characters displayed a stunning lack of trust in each other—for good reason in *Tosca* but less so in *Othello*—the theorists who have analyzed O. Henry's short story "The

Gift of the Magi" see conflict arising for almost the opposite reasons. The husband, who sells his watch to buy his wife combs, and the wife, who sells her hair to buy her husband a watch fob, are blinded by their love and perhaps too trusting.

Their blindness leads to a failure to coordinate their gift giving, and great sadness in the end (at least for the reader—more on the game between the author and the reader later) when the consequences of each trying to surprise the other are discovered. Rapoport (1960, 171) speaks of the couple's "misplaced altruism"; Vorob'ev (1968, 370–372) views the game as a battle of the sexes (the usual story illustrating this classic game is given in Luce and Raiffa (1957, 90–94); and Rasmusen (1989, 40) argues that the couple's failure to communicate may, ironically, have been rational, because communication would have ruined the surprise. Indeed, their sacrifices affirmed their great love for each other, despite their misfortune.

Eric Rasmusen (1990) points out that the couple, in effect, chose a mixed-strategy equilibrium;[12] the pure-strategy equilibria would be the outcome in which either the husband or the wife gives a gift but the other does not. Although game theory tells us that the mixed-strategy equilibrium is inefficient, and may be disastrous when the players choose non-complementary mixed strategies (as occurred in the story), it does not tell us how such a dismal state of affairs may arise. By contrast, the story suggests that

the act of communication would lower utility by eliminating the fun of being surprised. So the example says something about how to apply the theory. The theory also says something about the example: that even if the two people suspected that the ridiculous outcome might occur, they might do it anyway. And it also makes you think about what might have been one of O. Henry's points, that it is the thought that counts in gift giving. (Rasmusen 1990)

Indeed, O. Henry endorses this point of view at the end of the story: "O all who give and receive gifts, such as they are wisest. . . . They are the Magi."

12. Williams (1954/1966, 201–203) discussed such an equilibrium as the solution to a "marriage game" in Shakespeare's *The Merchant of Venice*. But this game, which is between Portia's father, Shylock, and her suitors, is zero-sum, because Shylock wants to frustrate, not coordinate with, Portia's suitors. The father in Giuseppe Verdi's opera *Rigoletto* has a similar goal, but this turns into a tragedy when he mistakenly kills his daughter.

Rasmusen (1990) draws a larger lesson from the story:

In general, examples are good for suggesting wrinkles that might not otherwise occur to the theorist. The easiest way to break out of a paradigm is to have the real world suggest a problem with it, since often the scholars are too used to thinking in one particular way. It is perhaps harder to be surprised by theories than by data.

I concur with these views but do not know of any direct evidence whereby a game-theoretic analysis of fiction has generated significant new theory. However, a large literature on so-called signaling games that has developed in recent years is germane to the strategic exegesis of plots. Avinash Dixit (1990) gives an interesting example:

If you read past all the four-letter words and the graphic violence, the whole theme of *Cogan's Trade* by George V. Higgins [1985] is reputation. For reasons too complicated to explain in brief, the bosses of organized crime in Boston have lost their reputation for protecting the activities they sponsored. How to regain it? This is a signaling game, and as usual there is excessive investment in signaling, in this case quite literally overkill. And the theory of this is almost fully and correctly explained by the enforcer (Cogan) in a conversation with The Man's counsellor.

By contrast, in Shakespeare's *King Lear* there is no enforcer—or any other mechanism—to guarantee that Lear's three daughters will keep their promises to their father, which he learns to his chagrin in the end (Dixit and Nalebuff 2008, 203, 434).[13]

Schelling (1960, 140; 1966, 11, 37) offers examples of the subtle and not-so-subtle signaling of threats in Shakespeare's *Henry V* and *Measure for Measure* and Joseph Conrad's *The Secret Agent*.[14] Citing different passages from *Henry V*, Dixit and Nalebuff (1991, 160–162) show how

13. Lear's willingness to succumb to flattery, abetted by the information problem he faced in learning of the true intentions of his daughters, is analyzed in Chami 1996.

14. Why these literary choices? Schelling (1991) reports:

My use of Henry V in *Arms and Influence* came from just seeing the play in London in 1965; I certainly didn't go to the play looking for illustrative material. I have no recollection of *Measure for Measure*, but I must have seen it in New Haven on the stage because I cannot imagine that I ever would have read it. . . . I do specifically remember how I was led to Conrad's *The Secret Agent*. I heard it from Daniel Ellsberg and when I wanted to use it I called him up and asked whether he was planning to use it in print in the near future and he said no and I asked whether he would release it and he said yes and I read the book and found no other useful examples but did use that one.

Henry inspired his troops, and thereby made his commitment credible, before the battle of Agincourt. His "steel my soldiers' hearts" exhortation echoes Lady Macbeth's plea to the "spirits," as she plans the murder of King Duncan in *Macbeth*, to "Make thick my blood/Stop up access and passage to remorse/That no compunctious visitings of nature shake my fell spirit" (see Brams 1997a, which is also discussed in section 7.5 of this book). Going one step further, in Homer's *Odyssey*, Odysseus (or Ulysses, as he was known in Roman myths) has himself bound to his ship's mast to ensure that he will not give in to the temptation of the sirens (Elster 1979, 36).[15]

Elster (1999, 2009) and Fisher (2002) consider how rationality and emotions mix in other literary works, but they develop no game-theoretic models of this mélange. Neither does Livingston (2001) in his study of rationality in literature.[16]

By contrast, Harmgart, Huck, and Müller (2008, 2009) and Huck (2008) use central ideas from game theory—mixed strategies, counterfactuals, and agreement theorems about beliefs—to render explicable some puzzling choices of the characters in two Wagner operas, *Tannhäuser* and *Lohengrin*. There is not space to describe these, but suffice it to say that the authors use both the music and the words of each opera to offer subtle, detailed interpretations that are persuasive without being overly technical.

1.7   The Devil and God

In Goethe's *Faust*, Faust gambles not just his wealth or reputation but also his life in making a compact with the devil. By selling his soul to Mephisto in exchange for knowledge and power for twenty-four years (in other versions of the Faust legend, sex or youth is the lure), Faust appears to commit himself irrevocably to eternal damnation when the

---

15. Mehlmann (2000, 132–142) develops an elaborate signaling game to model different choices of Ulysses. He offers other examples from literature (as well as movies), one of which I briefly discuss in section 1.7.

16. But other literary scholars have proposed simple ordinal games to model the choices of characters in several French literary works by Pierre Corneille, Guy de Maupassant, and Alain Robbe-Grillet (de Ley 1988) and the work of Polish writer Stanislaw Lem (Swirski 1996, 2007). I do not include these writings in the list in table 1.1 so as not to extend this list unduly.

"supreme moment" arrives. Fortunately for Faust, his final repentance saves his immortal soul from Mephisto, though not all versions of this legend have such a felicitous ending.

Mehlmann (2000, 72–78) uses differential game theory to analyze Goethe's great drama, making certain assumptions about the linkage between the players' beliefs about the timing of the supreme moment and also about how the players' payoffs are affected by each other's activities (repentance by Faust, temptation by Mephisto). He demonstrates consequences of these assumptions for the equilibrium path, arguing that Faust's "will to strive" (i.e., to repent), as the supreme moment approaches, explains his salvation.

Alexander Mehlmann (1990) cites other purported explanations (literary, legal) for Faust's salvation but claims that his mathematical model has "all the ingredients needed." Although I am not convinced that he has captured the essence of the drama in the parameters and functions he assumes, his application illustrates how advanced tools of game theory can be employed in literary exegesis.

Mehlmann (1990) reports that he has "always been interested in unusual applications of mathematics" and believes that "mathematics should play the role of an art rather than that of a science." Searching for a dynamic conflict situation to which he could apply differential game theory, "by chance . . . *Faust* came into my mind." He says that this application makes contributions both to the mathematical theory and to the modeling of player beliefs.

My motivation for applying game theory to the Hebrew Bible came from teaching a humanities seminar at New York University for freshmen and sophomores, which required that primary sources be used. I hoped to show, through a careful reading of certain narratives in the Hebrew Bible, how elementary game theory could lend coherence to the strategic interpretation of these stories. I also hoped that the analysis of several individual stories would allow me to draw general conclusions about the games that the biblical characters (God included) played. The seminar, which included orthodox Jews, devout Catholics, fundamentalist Protestants, and others turned out to be very stimulating and led to *Biblical Games* (Brams 1980/2003), which I draw examples from in chapters 2 and 5 (section 5.5). Dixit and Nalebuff (2008) also cite several passages from the Bible, including the New Testament, in discussing the "art" of strategy.

In *Biblical Games*, I show that biblical characters are, by and large, rational in the twenty or so stories of conflict and intrigue that I analyze. God is a "superlative strategist," but having granted free will to His human subjects (which justifies a game-theoretic treatment), He is besieged by problems that their freedom engenders. These cause Him great anguish, leading to very human-like displays of anger, frustration, and jealousy.[17]

God's wrath is especially great when his chosen people, the Israelites, cross Him (see section 5.5). It is sometimes expressed in petty, sometimes vindictive, behavior, but He is also merciful, always stopping short of wiping the slate clean, at least for His chosen people.

I believe that the Bible, as well as other religious works regarded as sacred, can be viewed at two levels. One level is as a literary work, with the stories it tells being susceptible to the same kind of game-theoretic analysis that helps to make the strategic aspects of secular stories perspicuous. The other level takes account of religious questions, such as the rationality of belief in a superior being or the problem of evil, which I have addressed in another work (Brams 1983/2007) and draw examples from in chapter 3.

The profound and profane may not be so different, at least in terms of the kinds of game-theoretic models needed to explicate their strategic structures. For example, if a superior being is immortal, it must be concerned with its reputation, which in fact obsesses the biblical God, especially in the Torah (the first five books of the Hebrew Bible). Thus, it makes sense to consider a concern with reputation as a correlate of immortality, to which the substantial literature on reputation in repeated games is pertinent.

1.8 Reputation and Intrapsychic Games

In Verdi's opera, *Rigoletto*, an assassin, Sparafucile, reports that Rigoletto "pays me and he buys my loyalty." It is Rigoletto's "strong reputation," Dixit and Nalebuff (2008, 210) argue, that prevents Sparafucile from killing him (though it does not prevent tragedy from

---

17. The anger and frustration of human characters are modeled in Brams 1997a and Brams and Jones 1999, which will be discussed in chapters 7 and 10.

befalling Rigoletto's daughter, Gilda). But what if one loses one's reputation? In Shakespeare's *Othello*, Cassio experienced all too well the calamity that ensues: "Reputation, reputation, reputation! O, I have lost my reputation! I have lost the immortal part of myself, and what remains is bestial."

Reputations are based on beliefs, and the modeling of such beliefs is a central feature of O'Neill's (1991) application of game theory to *Sir Gawain and the Green Knight*, a Middle English poem of the late fourteenth century that was only rediscovered in the nineteenth century. This poem describes the sudden appearance of a Green Knight of immense size, who challenges the hero, Sir Gawain, to behead him in exchange for a return blow.[18] After accepting the dare, which results in the beheading but not the death of the Green Knight, the poem recounts Gawain's search for the Green Knight, including tests of chivalry he must endure, before the Green Knight is allowed his turn to behead Sir Gawain a year and a day later. Feigning a beheading, the Green Knight inflicts a minor wound on Sir Gawain, presumably to symbolize Gawain's imperfection.

Now considered a great literary work that is rivaled only by Chaucer's poetry of the same period, *Sir Gawain* "engages modern readers by addressing modern problems," in particular "the predicament of how to follow one's ideals when the world maneuvers them into opposition to each other" (O'Neill 1990). Although the story might seem fantastic, the Green Knight is not described in just supernatural terms but is given a distinct human dimension, suggesting him to be vulnerable emotionally if not physically.

O'Neill (1991) analyzes two games, the first having to do with Gawain's reputation, which the Green Knight throws into doubt by his bold challenge to Arthur and the Round Table (Gawain persuades Arthur to let him stand in for him). O'Neill (Ibid.), postulating different beliefs that the players might have in different versions of a game of incomplete information, analyzes why the Green Knight throws down the gauntlet, and why Gawain accepts.

In one version, for example, he argues that Gawain seeks to enhance his reputation by placing a high value on his reputation, defined recur-

---

18. O'Neill 1991 also analyzes *The Feast of Bricriu*, an Old Irish medieval tale that describes another beheading, but its analysis is similar to that of *Sir Gawain*, so I do not discuss it here.

sively. In other words, Gawain wants to be seen as someone to be reckoned with generally, independent of the specific challenge he faces. The Green Knight makes a similar calculation in uttering his dare, and the players compete in a contest to bolster their *relative* reputations.

The second game O'Neill (1991) analyzes is that between Gawain's two natures—one chivalrous and the other self-preserving—that echoes the conflict between the id and the superego in Freud's theory. (The third component in Freud's theory, the ego, might come into play if there were a mediator or arbitrator involved.) In effect, Gawain must play against himself, not knowing whether the Green Knight is (1) chivalrous and vulnerable or (2) malevolent and invulnerable, which would make the game fair or unfair, respectively. If (1), then Gawain owes the Green Knight fair play, which will be reciprocated; if (2), then Gawain is absolved of his duty to rise to the challenge and should instead avoid being killed.

There is psychic harmony in this game if Gawain's two natures agree on the character of the Green Knight, but each of the natures prefers a different interpretation: The chivalrous nature prefers (1), and the self-preserving nature prefers (2). If the two natures disagree, there is tension, which is worse for both players (i.e., Gawain's two natures) than harmony. The resulting game, in which the two natures are locked in battle, is the classic battle of the sexes, which has two Nash equilibria in pure strategies and one in mixed strategies.

The lack of an obvious solution, O'Neill (1991) argues, renders the outcome equivocal, which "makes for a good literary plot." Unlike the Romanian school, however, O'Neill (Ibid.) does not contend that a character must act irrationally in order to dramatize the conflict. Instead, the players' harrowing choices, due largely to a coordination problem caused by the lack of information on how to regard the Green Knight, sustain our keen interest in the story.[19]

Which, if either, persona of Gawain has its preferred outcome chosen (the chivalrous nature prefers a chivalrous Green Knight, the self-preserving nature a malevolent Green Knight) depends on how the intrapsychic battle between Sir Gawain's two natures is resolved. The

19. A reader, in my view, is much more likely to identify with a rational protagonist than an irrational one, especially one, like the Green Knight, who seems so unbelievable from the start.

actual resolution in favor of chivalry validates Sir Gawain's acceptance of the dare, but through most of the narrative the rationality of this course of action is anything but apparent.

Barry O'Neill's motivation for analyzing a literary work is very different from Mehlmann's:

I wasn't looking for a place to apply game theory. Instead I was reading the work . . . and then it occurred to me that it was an interesting problem to formulate the hero's situation as a game. . . . Some of the hero's problems were the same as problems in my life at the time . . ., and this led me to think very hard about the poem. I read it and pondered on it. I would walk around thinking about it. It was not just for entertainment. (O'Neill 1990)

Like other theorists, O'Neill (1990) believes that game theory can clarify a literary work. Nonetheless, he points out that some "past applications of game theory . . . did not take the literary work seriously in its details" or take account of "good ideas scattered through the informal literature."

1.9   Wherein Lies the Future?

Besides taking the textual details of a literary work seriously, O'Neill (1990) claims that "it is also necessary to relate our work to the vocabulary already in use" if game theory is to make a contribution to literary analysis. (He is less sanguine that the game-theoretic analysis of literature will lead to mathematical advances.) More practically, O'Neill (Ibid.) is concerned that neither literary nor mathematics journals are generally open to a linkage of these very different interests.

It is difficult to say how much the lack of publishing outlets has retarded interdisciplinary work. My own belief is that linkages between mathematics and literature are not viewed as worth exploring by young scholars in either field if they are interested in advancing their careers. Aggravating this problem is that there is no interdisciplinary training for people who might be interested in the combination, with the possible exception of the Romanian school mentioned in section 1.5.

A related problem is that several of the applications I have discussed are no more than off-the-cuff illustrations. While most of the authors are mathematically sophisticated, they have made little effort to find non-

trivial applications of game theory. Of course, they cannot be faulted if a probing literary analysis was not their objective, but still one might hope for a more serious concern with the literary work. O'Neill (1990) speculates that Vorob'ev, a respected Russian mathematician who offered cursory analyses of several fictional works (see table 1.1) but did not report his own views, "perhaps regarded his study of game theory and literature as an interesting diversion, reading for the masses."

By contrast, in cases where the literary work was primary, the game-theoretic analysis was sometimes flawed (true of some of the Romanian authors). The opposite problem plagues Mehlmann (2000), where the mathematical structure is impressive but is not persuasively related to the narrative.

Howard (1971) and O'Neill (1991) use nontrivial game theory to construct plausible strategic interpretations of the play and poem, respectively, each analyzes. Interestingly enough, both authors, as noted earlier, indicated that they did not set out to "apply" game theory, but the literary works themselves riveted their attention.

Other tools of mathematical analysis have been applied to literature, but they generally give short shrift to plot (some citations are given in O'Neill 1991). Game theory makes plot front and center; when there is no strong plot or story line, as is the case in much contemporary fiction, then the theory has little to offer. I share Howard's (1990) view that "plot is essential for the kind of great art which really changes people," so I am not worried that game theory will suffer from lack of literary material to which to apply its methods, some contemporary fiction notwithstanding.

Howard (1990), who reported that he analyzed "every incident and conversation as a set of interlinked games" in Anthony Trollope's *The Warden* and then transposed the novel into a modern setting (Howard's reworked version was not published), indicates that game theory may have other roles to play, such as

to help writers construct plots. In film-making, where many people have to cooperate, it would be exceedingly useful to work with a clear game-theoretically analyzed plot—just as musicians find it useful to have a score.

He added that this kind of analysis can also help game theory, because game theorists

benefit from the great store of intuitive wisdom about human behaviour contained in the world's fiction. They should continually be testing their theories against this. If it doesn't make sense to Shakespeare, perhaps it doesn't make sense!

Game theory, in my view, should be able to do more than suggest that there is a problem in a relationship. The fact that Scarpia and Tosca are enmeshed in a Prisoners' Dilemma, or the husband and wife in "The Gift of the Magi" have a coordination problem à la the battle-of-the sexes game, is not particularly enlightening. Why are these stories compelling and not just humdrum illustrations of these games?

O'Neill (1990) suggests that the tragic or surprising aspects of these stories require that we look more deeply into the information available to the players, and how it was used, in order to understand their human dramas. Indeed, the lack of information may itself be a central strategic feature of a story, as I showed in the players' choices of mixed strategies in *Light in August*.[20]

The game played between the author and the reader, as the reader progressively acquires more information (not necessarily accurate, such as the false clues in a mystery), is one that does not seem to have been analyzed for any literary work.[21] An appropriate framework for such an analysis might be the theory of psychological games (Geanakoplos, Pearce, and Stacchetti 1989) or "information-dependent games" (Gilboa and Schmeidler 1988), in which the players' payoffs depend on whether certain postulated beliefs are fulfilled.

Thus, a reader may be either thrilled or disappointed not only by the way a story evolves but also whether tension builds or he or she is surprised by the ending. If a horrific ending turns out only to have been a

20. Holler and Klose-Ullmann (2008) also suggest the use of mixed strategies in Friedrich von Schiller's play trilogy, *Wallenstein* (1800); they are not interpreted as "probabilities or chance" but as "a level of expected action." In my opinion, this interpretation is a sensible one that game theorists might well incorporate into their models, notwithstanding the authors' disclaimer that "the intention of this article is to convince theatergoers and people who work in the theatrical arts that it is worthwhile to study some game theory. . . . It is not this article's purpose to teach game theorists."

21. However, this subject is the main theme of a horror novel, *Misery* (1987), by Stephen King, which was made into a movie in 1990. In the novel, a reader takes revenge on an author for killing off her favorite character in the last of a series of novels, forcing the author to burn the manuscript of his next novel and resurrect this character in a new novel.

dream, a reader may feel either manipulated or relieved—depending on the reader's prior expectations—by the author's choice.[22]

Situated, as they are, in different worlds, game theory and literature have their own coordination problem, with game theorists and humanists not often benefiting from each others' insights. What makes a literary creation succeed is not just its overall structure but also its details, including the emotional lives of its characters. Game theorists need to ponder these and adapt their theory accordingly, just as literary scholars need to appreciate that game theory has its own richness that goes beyond mathematical symbols and abstract forms.

---

22. We sometimes use the term *cop-out* when we feel betrayed by the author. To avoid this feeling, authors might try to take account of the expectations of their readers in constructing plots and portraying characters. Thus, mystery writers might aim to surprise their readers, whereas other writers might prefer no surprises in order to stress the unrelieved boredom of the human condition (as in Samuel Beckett's play *Waiting for Godot*). Still others may search for an appropriate resolution to some conflict. If game theory were used to help authors in this manner, it would radically change the theory's purpose: Instead of using it to show that characters in a text act rationally, one would start with the characters' (or the author's) motives and write the text to show the rational working out of these motives, reversing the order in which game theory is applied from before to after the text is written.