A Course in Game Theory

Reading Notes

Martin J. Osborne and Ariel Rubinstein

Contents

1 Introduction **1**
1.1 Game theory . . . 1
1.2 Games and solutions . . . 1
1.3 Game theory and competitive equilibrium . . . 2
1.4 Rational behavior . . . 2
1.5 The steady state and deductive interpretations . . . 3
1.6 Bounded rationality . . . 3
1.7 Terminology and notation . . . 3

2 Nash Equilibrium **4**
2.1 Strategic games . . . 4
2.2 Nash equilibrium . . . 4
2.3 Examples . . . 5
2.4 Existence of a Nash equilibrium . . . 7
2.5 Strictly competitive games . . . 7
2.6 Bayesian games: strategic games with imperfect information . . . 8

3 Mixed, Correlated, and Evolutionary Equilibrium **10**
3.1 Mixed strategy Nash equilibrium . . . 10
3.2 Interpretations of mixed strategy Nash equilibrium . . . 11
3.3 Correlated equilibrium . . . 13
3.4 Evolutionary equilibrium . . . 14

4 Rationalizability and Iterated Elimination of Dominated Actions **16**
4.1 Rationalizability . . . 16
4.2 Iterated elimination of strictly dominated actions . . . 17
4.3 Iterated elimination of weakly dominated actions . . . 19

5 Knowledge and Equilibrium **21**
5.1 A model of knowledge . . . 21
5.2 Common knowledge . . . 22
5.3 Can people agree to disagree? . . . 23
5.4 Knowledge and solution concepts . . . 24
5.5 The electronic mail game . . . 25

6 Extensive Games with Perfect Information **26**
6.1 Extensive games with perfect information . . . 26
6.2 Subgame perfect equilibrium . . . 27
6.3 Two extensions of the definition of a game . . . 29
6.4 The interpretation of a strategy . . . 29
6.5 Two notable finite horizon games . . . 30
6.6 Iterated elimination of weakly dominated strategies . . . 30

7 Bargaining Games **33**
7.1 Bargaining and game theory . . . 33
7.2 A bargaining game of alternating offers . . . 33
7.3 Subgame perfect equilibrium . . . 34
7.4 Variations and extensions . . . 36

8 Repeated Games **39**
8.1 Repeated interaction and the horizon . . . 39
8.2 Histories, strategies, and evaluations of payoff streams . . . 39
8.3 Strategies represented by machines . . . 41
8.4 Nash folk theorems and permanent punishment . . . 41
8.5 Finite punishments under the limit of means . . . 42
8.6 Punishment under overtaking . . . 42
8.7 Discounting and rewards for punishers . . . 43
8.8 One deviation and optimal punishments . . . 43
8.9 Finite repetition . . . 44

9 Complexity Considerations in Repeated Games **45**
9.1 Complexity as a feature of preferences . . . 45
9.2 The machine game . . . 45
9.3 Every state must be used . . . 46
9.4 Best responses and equal machine sizes . . . 47
9.5 Actions, states, and the structure of the path . . . 47
9.6 Lexicographic preferences in the Prisoner's Dilemma . . . 48

10 Implementation Theory **50**
10.1 Outcomes, preferences, and game forms . . . 50
10.2 Dominant strategy implementation . . . 50
10.3 Public projects and Groves payments . . . 51
10.4 Nash implementation and monotonicity . . . 52
10.5 No veto power and a sufficient condition . . . 53
10.6 Subgame perfect implementation . . . 54
10.7 Virtual implementation . . . 54

11 Extensive Games with Imperfect Information **55**
11.1 Histories and information sets . . . 55
11.2 Perfect recall . . . 56
11.3 Four principles for equivalent extensive forms . . . 56
11.4 Framing and conditional choice . . . 57
11.5 Mixed and behavioral strategies . . . 57
11.6 Outcome equivalence under perfect recall . . . 58
11.7 Failures with imperfect recall . . . 59
11.8 Nash equilibrium in an extensive game . . . 59

12 Sequential Equilibrium **60**
12.1 Strategies and beliefs at every information set . . . 60
12.2 Consistency and sequential equilibrium . . . 60
12.3 Restrictions linking beliefs across information sets . . . 61
12.4 The role of the extensive form . . . 62
12.5 Observable actions and perfect Bayesian equilibrium . . . 62
12.6 Signaling and education . . . 63
12.7 Reputation in the chain store game . . . 64
12.8 Refinements based on incentives to deviate . . . 64
12.9 Trembling hand perfection in strategic games . . . 65
12.10 Perfection in an extensive game . . . 66

13 The Core **67**
13.1 Coalitional games with transferable payoff . . . 67
13.2 Blocking and the core . . . 67
13.3 Veto players and convex games . . . 68
13.4 Balancedness and nonemptiness . . . 68
13.5 Markets with transferable payoff . . . 69
13.6 Competitive payoffs and the core . . . 70
13.7 Games without transferable payoff . . . 70
13.8 Exchange economies . . . 70
13.9 Replication and the competitive limit . . . 71

14 Stable Sets, the Bargaining Set, and the Shapley Value **72**
14.1 Imputations and credible objections . . . 72
14.2 Stable sets . . . 72
14.3 The bargaining set . . . 73
14.4 Excess and the kernel . . . 74
14.5 The nucleolus . . . 74
14.6 The Shapley value and marginal contributions . . . 76
14.7 Balanced contributions . . . 76
14.8 The axiomatic characterization . . . 77
14.9 Cost sharing . . . 78

15 The Nash Solution **79**
15.1 Bargaining problems and risk preferences . . . 79
15.2 Objections and the product characterization . . . 79
15.3 The efficient frontier and worked solutions . . . 80
15.4 Comparative statics of risk aversion . . . 81
15.5 The axioms on agreements and preferences . . . 81
15.6 The standard utility set axioms . . . 82
15.7 Independence of the axioms and alternative solutions . . . 82
15.8 Alternating offers with a risk of breakdown . . . 83
15.9 Exact implementation of the Nash solution . . . 83

1 Introduction

1.1 Game theory

Game theory studies decisions whose consequences depend on the decisions of several people. A player has preferences over the possible consequences and chooses an action in light of beliefs about the other players. Rationality concerns the relation between these preferences, beliefs, and choices. Strategic reasoning concerns the dependence of a choice on what other players may do.

The model separates the rules of an interaction from a proposed account of behavior. A description of an auction, for example, specifies who can bid, which bids are permitted, how the object is allocated, and what each bidder pays. These rules do not themselves determine the bids. A solution concept adds conditions that relate the bids to the bidders' interests and information. The same distinction applies when the players are firms, voters, or groups that make decisions collectively.

The abstraction is useful because different applications may have the same strategic structure. It also determines what the analysis can explain. A model that specifies only preferences and feasible actions cannot explain differences that arise solely from unmodeled differences in computation or perception. The mathematical argument identifies consequences of the stated assumptions; the interpretation identifies the economic situation those assumptions describe.

1.2 Games and solutions

> Definition 1.1 (Game and solution). A game specifies the players, the actions or agreements available to them, their information, and their preferences over the resulting outcomes. A solution concept assigns to each game in a specified class a set of outcomes, strategies, or distributions that satisfy its behavioral conditions.

In a strategic game each player chooses a complete action without observing the other players' choices. Simultaneity has this informational meaning: choices need not occur at exactly the same time. In an extensive game the order of decisions and the information available at each decision are explicit. A strategy then specifies an action at each decision the player could face.

An extensive game has perfect information when the player who moves knows the entire preceding history. With imperfect information, some histories are indistinguishable to a player at the time of choice. Information about previous actions is distinct from information about preferences: players may observe all previous actions while remaining uncertain about another player's valuation.

Noncooperative models take individual actions as primitives and derive their joint consequences. Coalitional models instead specify what groups of players can achieve together. The distinction concerns the mathematical description of feasibility. Cooperation among individuals can be an outcome of a noncooperative game when their incentives support it.

1.3 Game theory and competitive equilibrium

Competitive analysis treats some variables, such as prices, as given to an individual agent. Equilibrium requires those variables to be consistent with the choices of all agents. Strategic analysis lets a player account for how the player's own action changes the relevant consequence, while holding the other players' actions fixed.

Example 1.2 (An activity with an external effect). Let agent i choose activity a_i, and let total pollution be $P = \sum_j a_j$. Suppose utility is $b_i(a_i) - d_i(P)$, with differentiable functions. An agent who treats pollution as a fixed parameter chooses an interior action satisfying $b_i'(a_i) = 0$. A strategic agent instead solves

$$\max_{a_i} \left\{ b_i(a_i) - d_i\left(a_i + \sum_{j \neq i} a_j \right) \right\},$$

and an interior best response satisfies $b_i'(a_i) = d_i'(P)$. A planner who maximizes total utility would include the marginal damage to every agent, giving $b_i'(a_i) = \sum_j d_j'(P)$. These conditions correspond to different decision problems, even though the physical relation defining pollution is the same.

1.4 Rational behavior

Definition 1.3 (Choice under certainty). A decision problem consists of an action set A, a consequence set C, a function $g : A \to C$, and a complete and transitive preference relation $\succeq$ on C. Given a feasible subset $B \subseteq A$, a rational choice is an $a^* \in B$ such that $g(a^*) \succeq g(a)$ for every $a \in B$.

If $U : C \to \mathbb{R}$ represents the preference relation, the choice solves $\max_{a \in B} U(g(a))$. The representation is ordinal: a strictly increasing transformation of U leaves choices among certain consequences unchanged. The behavioral restriction is that the same preferences govern choices from different feasible sets.

Uncertainty may concern an objective parameter, a previous action that was not observed, another player's random choice, or another player's reasoning. Let Ω be a state space, p a probability measure, and $g : A \times \Omega \to C$ the consequence function. Expected utility choice solves

$$\max_{a \in B} \int_{\Omega} u(g(a, \omega))\, dp(\omega).$$

Here p may describe an objective random mechanism or the decision maker's subjective belief. The function u represents preferences over lotteries, so its admissible transformations are positive affine transformations. Replacing it by an arbitrary increasing function can change choices under uncertainty.

Example 1.4 (Ordinal and expected utility comparisons). Suppose three certain consequences have utilities $0, 1, 3$. The decision maker prefers the lottery giving the first and third with equal probability to the middle consequence, because $3/2 > 1$. An increasing transformation that assigns utilities $0, 2, 3$ preserves the ordering of certain consequences but reverses this lottery comparison. Payoff numbers in a mixed strategy analysis must therefore describe preferences over lotteries, not just a ranking of certain outcomes.

1.5 The steady state and deductive interpretations

Under the steady state interpretation, a solution describes a regularity in a population or in a collection of similar interactions. Players have acquired correct expectations about behavior and choose optimally given those expectations. Nash equilibrium expresses this consistency condition without specifying the adjustment process by which it arose.

Under the deductive interpretation, players reason from their knowledge of the game and from assumptions about one another's rationality. The question is which choices this reasoning excludes. Rationality by itself does not require correct beliefs about other players. Consequently, the restrictions obtained from common knowledge of rationality may be weaker than Nash equilibrium.

Repeated observations can support the steady state interpretation of a strategic game when there are no strategic connections between successive plays. If an action today is chosen to affect an opponent's behavior tomorrow, the continuation consequences belong in the game. A repeated game makes those consequences explicit.

1.6 Bounded rationality

The usual model permits every strategy in the stated strategy set and assumes that a player can identify an optimal choice. These assumptions separate existence from computation. A finite game of perfect information may admit a backward induction solution even when its tree is too large for a person to analyze. Knowing that an optimal strategy exists gives no bound on the effort needed to find or implement it.

Bounded rationality models can restrict the strategies a player can implement, make complexity costly, or distinguish players by their information and reasoning. Such changes alter the decision problem itself. In particular, a strategy that is optimal when implementation is free may cease to be optimal when the player also values simplicity.

1.7 Terminology and notation

The finite player set is N. A profile $a = (a_i)_{i \in N}$ specifies one component for each player, and a_{-i} collects every component except i's. Write $A = \prod_{i \in N} A_i$ and $A_{-i} = \prod_{j \neq i} A_j$. The expression (a_i', a_{-i}) means that only player i's component has changed. For a finite set X, $\Delta(X)$ denotes the probability distributions on X.

A preference relation $\succeq$ is complete and transitive; $x \succ y$ means $x \succeq y$ but not $y \succeq x$, and $x \sim y$ means both weak comparisons hold. A payoff function is quasiconcave if each upper contour set is convex. For choice over a convex set, quasiconcavity implies that mixtures of two maximizers are also maximizers.

The book calls $x \in X \subseteq \mathbb{R}^N$ Pareto efficient if no $y \in X$ satisfies $y_i > x_i$ for every i. It calls x strongly Pareto efficient if no feasible y makes every player weakly better off and some player strictly better off. Keeping these two comparisons distinct matters when players may be indifferent.

2 Nash Equilibrium

2.1 Strategic games

Definition 2.1 (Strategic game; the book's Section 2.1). A strategic game is $G = \langle N, (A_i), (\succeq_i) \rangle$, where N is a finite player set, each A_i is a nonempty action set, and $\succeq_i$ is a preference relation on $A = \prod_i A_i$. The game is finite when every A_i is finite. When $u_i : A \to \mathbb{R}$ represents $\succeq_i$, write $G = \langle N, (A_i), (u_i) \rangle$.

Preferences are defined on action profiles because the consequence of one player's decision depends on the decisions of others. If players care directly about consequences in a set C, a consequence function $g : A \to C$ and preferences on C induce preferences on A. An exogenous random variable can similarly make each action profile induce a lottery over consequences. The payoff $u_i(a)$ is then the expected utility of that lottery.

The strategic description records the information available when choosing. Players may submit actions at different dates and still face a strategic game if no player observes another's choice before acting. Conversely, an observed earlier action must be included in a later player's decision problem.

2.2 Nash equilibrium

Definition 2.2 (Nash equilibrium; Definition 14.1). An action profile $a^* \in A$ is a Nash equilibrium if, for every $i \in N$,

$$u_i(a_i^*, a_{-i}^*) \geq u_i(a_i, a_{-i}^*) \qquad \text{for all } a_i \in A_i.$$

The best response correspondence is

$$B_i(a_{-i}) = \arg\max_{a_i \in A_i} u_i(a_i, a_{-i}).$$

Thus equilibrium requires $a_i^* \in B_i(a_{-i}^*)$ for every player.

The comparison holds the opponents' actions fixed. An equilibrium need not maximize a player's payoff over all profiles, and it need not be efficient. It requires the player's action to be optimal against the particular actions the opponents choose. Correct expectations and optimal choice are imposed together.

In a finite payoff matrix, mark each player's maximizing actions against each opponent action. An equilibrium is a cell marked for both players. With continuous actions, one can instead calculate the correspondences B_i and solve their simultaneous conditions. Boundary choices must be included when the first order conditions give an infeasible action.

2.3 Examples

2.3.1 Coordination, conflict, and dominance

In Bach or Stravinsky, the payoff matrix is

	B	S
B	$(2,1)$	$(0,0)$
S	$(0,0)$	$(1,2)$

.

Both (B,B) and (S,S) are equilibria. Each player prefers coordination, but their rankings of the two coordinated outcomes differ. If the diagonal payoffs are instead $(2,2)$ and $(1,1)$, both diagonal profiles remain equilibria even though everyone prefers the first. Equilibrium does not select an outcome solely because it Pareto dominates another equilibrium.

The Prisoner's Dilemma has the matrix

	C	D
C	$(3,3)$	$(0,4)$
D	$(4,0)$	$(1,1)$

.

Action D is strictly better against either opponent action, so (D,D) is the unique equilibrium. The gain from switching from D to C is negative when the opponent's action is held fixed, although both gain from a joint switch from (D,D) to (C,C).

In Hawk and Dove, the payoffs are

	D	H
D	$(3,3)$	$(1,4)$
H	$(4,1)$	$(0,0)$

.

Each player prefers the opposite action to the opponent's, giving the two equilibria (D,H) and (H,D). Matching Pennies has payoffs $(1,-1)$ when actions agree and $(-1,1)$ when they differ. At every pure profile one player wants to change, so no pure equilibrium exists.

2.3.2 Auctions

Consider valuations $v_1 > v_2 > \cdots > v_n > 0$ and nonnegative bids b_i. The highest bidder receives the object; a tie is resolved in favor of the lowest index. In a first price auction the winner pays the winning bid and has payoff $v_i - b_i$; every loser receives zero.

Proposition 2.3 (Complete information first price auction). In every pure Nash equilibrium player 1 wins. The winning bid lies in $[v_2, v_1]$ and is matched by at least one other bid. Conversely, any profile with $b_1 = \max_{j \neq 1} b_j = b \in [v_2, v_1]$ is an equilibrium.

Proof. A winner never bids above its valuation, since losing guarantees zero. If a player $k \neq 1$ wins at a bid at most $v_k < v_1$, player 1 can win profitably by matching that bid, using the tie rule. Thus player 1 wins. If $b_1 < v_2$, player 2 can outbid player 1 slightly and gain. If the highest rival bid were strictly below b_1, player 1 could lower its bid while retaining the object. These facts give necessity. Under the stated sufficient conditions player 1 cannot lower its bid and still win, and no other player can profit by bidding more than $b \geq v_2$.

In a second price auction the winner pays the highest rival bid. Fix that bid at m. A bidder with value v_i wants to win if $v_i > m$, lose if $v_i < m$, and is indifferent if equality holds. Bidding v_i produces one of these optimal outcomes for every rival bid profile, including ties. Truthful bidding is therefore weakly dominant.

Weak dominance does not imply that every equilibrium is truthful or efficient. With two bidders and $v_1 > v_2 > 0$, the profile $b_1 = 0$, $b_2 > v_1$ is an equilibrium: bidder 2 wins and pays zero, whereas bidder 1 would have to pay more than its value to win. The second bidder's high bid has no cost at the equilibrium profile because payment depends on the rival bid.

2.3.3 The war of attrition

Player i chooses a concession time $t_i \geq 0$ and values the object at $v_i > 0$. Waiting costs one per unit of time. The payoff is

$$u_i(t_i, t_j) = \begin{cases} -t_i, & t_i < t_j, \\ v_i/2 - t_i, & t_i = t_j, \\ v_i - t_j, & t_i > t_j. \end{cases}$$

If the concession times differ and the smaller time is positive, the player who concedes can gain by conceding immediately. Equal times cannot be an equilibrium, since a slightly later concession gives the whole object at the same waiting cost. Hence every equilibrium has one player conceding at zero and the other choosing a positive time. More precisely, $(0, t_2)$ is an equilibrium exactly when $t_2 \geq v_1$, and $(t_1, 0)$ is an equilibrium exactly when $t_1 \geq v_2$. The proposed waiting time prevents the conceding player from winning profitably by waiting longer.

2.3.4 Political location

Each potential candidate either stays out or chooses a position in $[0, 1]$. Voters have a strictly positive density of ideal positions and vote for the nearest candidate, sharing equally among candidates at the same position. Candidates prefer a sole victory to a tie for victory, a tie to staying out, and staying out to losing.

With two candidates, both enter and choose the median voter's position. A candidate at the median secures at least half the votes against any other position; moving away when the opponent remains there gives fewer than half. If both candidates choose another common position, one can gain a majority by moving toward the median. If their positions differ, a losing candidate prefers to stay out, while in a tie a candidate can move slightly toward the other and gain votes. With three potential candidates there is no pure equilibrium. A profile with only two at the median invites the third to join the winning tie; three at the median allow one to move slightly to one side and win almost half the electorate against two candidates sharing the rest. More generally, any equilibrium entrant must tie for first, and the extreme positions and possible local deviations preclude such a tie among three.

2.4 Existence of a Nash equilibrium

Theorem 2.4 (Kakutani fixed point theorem). Let X be a nonempty compact convex subset of a finite dimensional Euclidean space. If $F : X \rightrightarrows X$ has nonempty convex values and a closed graph, then some $x^* \in X$ satisfies $x^* \in F(x^*)$.

Proposition 2.5 (Existence; Proposition 20.3). A strategic game has a Nash equilibrium if each A_i is a nonempty compact convex subset of a finite dimensional Euclidean space, each u_i is continuous on A, and $u_i(\cdot, a_{-i})$ is quasiconcave for every a_{-i}.

Proof. Set $B(a) = \prod_i B_i(a_{-i})$. Continuity and compactness imply that each $B_i(a_{-i})$ is nonempty. Quasiconcavity implies that it is convex. To check the closed graph property, suppose $a^k \to a$, $b^k \to b$, and $b^k \in B(a^k)$. For every i and every $c_i \in A_i$,

$$u_i(b_i^k, a_{-i}^k) \geq u_i(c_i, a_{-i}^k).$$

Continuity gives the same inequality at the limit, so $b \in B(a)$. Kakutani supplies a fixed point, which is precisely an equilibrium.

The theorem concerns existence, rather than uniqueness or convergence of an adjustment process. Finite pure action sets usually fail convexity. Their mixed extensions satisfy the theorem because probability simplexes are convex and expected utility is linear in the player's own distribution.

Corollary 2.6 (Symmetric equilibrium). In a symmetric two player game satisfying the preceding assumptions, there is an equilibrium (a^*, a^*).

Proof. The common best response correspondence maps the common action set into itself and satisfies Kakutani's conditions. Its fixed point is a best response to itself. Symmetry makes the same statement true for both players.

2.5 Strictly competitive games

Definition 2.7 (Strict competition and security levels). A two player game is strictly competitive if $a \succeq_1 b$ exactly when $b \succeq_2 a$. Using a payoff representation $(u, -u)$, define

$$\underline{v} = \max_x \min_y u(x, y), \qquad \overline{v} = \min_y \max_x u(x, y).$$

A maximizing choice in the first expression is a maxminimizer for player 1; a minimizing choice in the second is a maxminimizer for player 2.

The formulas presume that the stated extrema exist; infima and suprema give the corresponding bounds in general. Always $\underline{v} \leq \overline{v}$: for every x, y, $\min_{y'} u(x, y') \leq u(x, y) \leq \max_{x'} u(x', y)$. The first quantity is what player 1 can guarantee by choosing first without observing player 2's action. The second is the lowest upper bound player 2 can impose.

Proposition 2.8 (Equilibrium and maxminimization; Proposition 22.2). If (x^*, y^*) is a Nash equilibrium, both components are maxminimizers and

$$\underline{v} = u(x^*, y^*) = \overline{v}.$$

Conversely, if the two values coincide, any pair of maxminimizers is a Nash equilibrium. All equilibria have the same payoffs, and equilibrium strategies are interchangeable.

Proof. At equilibrium the saddle point inequalities are

$$u(x, y^*) \leq u(x^*, y^*) \leq u(x^*, y) \qquad \text{for every } x, y.$$

Taking appropriate extrema gives $\overline{v} \leq u(x^*, y^*) \leq \underline{v}$. The general opposite inequality forces equality and proves the optimality of both components. Conversely, if the common value is v, maxminimizers satisfy $u(x^*, y) \geq v$ and $u(x, y^*) \leq v$. At their intersection the payoff is v, so the saddle point inequalities hold. Pairing components from two equilibria preserves these inequalities.

In pure Matching Pennies, the lower value is -1 and the upper value is 1. The gap is consistent with the absence of pure equilibrium. Strict competition alone does not close it; allowing randomization will do so in finite games. For certain outcomes an ordinal representation can be chosen with payoffs $(u, -u)$. When lotteries are introduced, their preference representation must also be specified before applying expected utility arguments.

2.6 Bayesian games: strategic games with imperfect information

2.6.1 States, signals, and equilibrium

Definition 2.9 (Bayesian game; the book's Section 2.6). A Bayesian game specifies a finite player set N, a finite state space Ω, action sets A_i, finite signal sets T_i, signal functions $\tau_i : \Omega \to T_i$, and priors p_i on Ω. Each possible signal has positive prior probability. Player i has preferences over lotteries on $A \times \Omega$, represented here by expected utility from $u_i(a, \omega)$.

Upon observing t_i, player i conditions on $\Omega_i(t_i) = \{\omega : \tau_i(\omega) = t_i\}$. A pure strategy $s_i : T_i \to A_i$ specifies the action at every signal. Equilibrium requires that, for every i, t_i and $a_i \in A_i$,

$$\sum_{\omega \in \Omega_i(t_i)} p_i(\omega \mid t_i) u_i(s_i(t_i), s_{-i}(\tau_{-i}(\omega)), \omega) \geq \sum_{\omega \in \Omega_i(t_i)} p_i(\omega \mid t_i) u_i(a_i, s_{-i}(\tau_{-i}(\omega)), \omega).$$

The book formulates this as a strategic game with one player for each pair (i, t_i). Each type chooses an action and evaluates it using its conditional belief. The action at an unrealized type still matters to another player's beliefs, which may put positive probability on that type.

2.6.2 Private values and selection in trade

For a second price auction let $\Omega = V^n$, where V is a finite set of valuations, and let $\tau_i(v) = v_i$. Independent draws from a common distribution are one possible prior. Truthful bidding $s_i(v_i) = v_i$ remains an equilibrium because its optimality holds separately against every realized rival bid profile. No inference about rival values is needed for that dominance argument.

For a different implication of private information, suppose two players draw independent prizes from the same finite support and exchange prizes only if both agree. Suppose both players agree with positive probability. Let x be the largest prize offered for exchange by either player with positive probability. A player holding x can receive at most x conditional on the opponent agreeing, and receives less with positive probability if the opponent offers any lower prize. Agreement would then reduce expected utility. Repeating the argument rules out mutually beneficial exchange of unequal prizes in equilibrium. Thus both players can agree with positive probability only at a common prize x. If a lower prize were in the support, its holder would profit by agreeing whenever the opponent agrees. Hence x must be the smallest possible prize. The willingness to trade contains information about the object offered, so the unconditional expected prize is not the relevant comparison.

2.6.3 Beliefs about information

States may describe knowledge as well as payoff parameters. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$, let player 1's partition be $\{\{\omega_1, \omega_2\}, \{\omega_3\}\}$, and player 2's partition be $\{\{\omega_1\}, \{\omega_2, \omega_3\}\}$. Suppose player 1 prefers b to c in the first two states and reverses that preference in the third. At ω_1, player 2 knows that player 1 prefers b; player 1 does not know whether player 2 has that knowledge, since player 1 also considers ω_2 possible.

The definition permits different priors. A common prior adds a substantive restriction: differences in posterior beliefs must then arise from differences in information. A type can encode a belief about payoff parameters, a belief about others' beliefs, and further orders of beliefs. The state space provides a joint description in which these conditional beliefs can be compared.

3 Mixed, Correlated, and Evolutionary Equilibrium

3.1 Mixed strategy Nash equilibrium

3.1.1 The mixed extension

Let $G = \langle N, (A_i), (u_i) \rangle$ be a finite strategic game. Each u_i is now a von Neumann–Morgenstern utility function: its expectation represents preferences over lotteries on A. A mixed strategy $\alpha_i \in \Delta(A_i)$ assigns probabilities to player i's actions. Its support consists of the actions receiving positive probability.

Definition 3.1 (Mixed extension; Definition 32.1). The mixed extension of G is the strategic game with action sets $\Delta(A_i)$ and payoffs

$$U_i(\alpha) = \sum_{a \in A} \left(\prod_{j \in N} \alpha_j(a_j) \right) u_i(a).$$

A mixed strategy Nash equilibrium is a Nash equilibrium of this extension.

Independence is built into the product distribution. Holding α_{-i} fixed, expected utility is linear in α_i:

$$U_i(\alpha_i, \alpha_{-i}) = \sum_{a_i \in A_i} \alpha_i(a_i) U_i(a_i, \alpha_{-i}).$$

A pure strategy is the degenerate distribution assigning probability one to that action. Every pure equilibrium remains an equilibrium of the mixed extension, since a weighted average of pure deviation payoffs cannot exceed their maximum.

Proposition 3.2 (Existence; Proposition 33.1). Every finite strategic game has a mixed strategy Nash equilibrium.

Proof. For finite A_i, $\Delta(A_i)$ is a nonempty compact convex simplex. Expected payoffs are continuous in the whole profile and linear, hence quasiconcave, in the player's own distribution. Proposition 20.3 therefore applies to the mixed extension.

Lemma 3.3 (The support condition; Lemma 33.2). A mixed profile α^* is an equilibrium if and only if every action in $\mathrm{supp}(\alpha_i^*)$ maximizes $U_i(a_i, \alpha_{-i}^*)$, for every i.

Proof. If a supported action has a lower payoff than another action, shifting its positive probability to the better action raises expected utility. Conversely, if every supported action attains the maximum pure payoff, their weighted average attains it too, and no mixed deviation can do better.

For a proposed support, equilibrium calculation thus has two parts: impose equal payoffs among the supported actions and verify that every excluded action gives no more. Indifference equations alone can give probabilities outside $[0, 1]$ or a profile that admits a profitable excluded action.

With infinite action sets, a useful formulation is that no pure action gives more than the equilibrium payoff and the set of actions giving less has probability zero under the player's strategy. A continuous distribution need not assign positive probability to any individual action.

Mixed existence also extends beyond finite action sets. If each action set is a nonempty compact convex subset of a finite dimensional Euclidean space and payoffs are continuous, a mixed strategy equilibrium exists. Quasiconcavity in a player's own action is not required for this result; it is the additional condition used above for pure equilibrium existence.

3.1.2 Equilibrium calculations

In Bach or Stravinsky, let p be player 1's probability of B and q player 2's probability of B. Player 1 compares $2q$ with $1-q$; player 2 compares p with $2(1-p)$. Both mix only if

$$2q = 1 - q, \qquad p = 2(1-p), \qquad \text{so} \qquad (p,q) = \left(\frac{2}{3}, \frac{1}{3}\right).$$

Each payoff is $2/3$. Together with the two pure equilibria, these are all equilibria. The probability chosen by one player makes the other player indifferent. It is not obtained by maximizing a strictly concave objective in the player's own mixing probability.

Example 3.4 (Guessing the same number). Each player chooses a number in $\{1, \ldots, K\}$. Player 1 receives one from player 2 if their choices match and zero otherwise. Against a distribution q, player 1's best pure payoff is $\max_k q_k$; against p, player 2 can restrict player 1 to $\min_k p_k$. The uniform distribution guarantees $1/K$ for either side, so the value is $1/K$. A nonuniform p has $\min_k p_k < 1/K$, and a nonuniform q has $\max_k q_k > 1/K$. Thus both equilibrium distributions are uniquely uniform.

Example 3.5 (An investment race). Two players choose expenditures $x_i \in [0,1]$. The higher expenditure wins a prize of one, a tie divides the prize, and expenditures are paid regardless of winning. If the opponent uses the uniform distribution on $[0,1]$, an expenditure x wins with probability x and yields expected payoff $x - x = 0$. Ties have probability zero. Consequently, both players using the uniform distribution is a mixed equilibrium, despite the discontinuity of the payoff function at a tie.

For a finite zero sum game, existence and the saddle point result together give equality of the mixed maxmin and minmax values. This conclusion applies when expected utilities are opposites. Opposite rankings of pure outcomes alone do not imply opposite rankings of lotteries: both players may prefer an intermediate certain outcome to a lottery between two extreme outcomes.

3.2 Interpretations of mixed strategy Nash equilibrium

3.2.1 Mixed strategies as objects of choice

A player may deliberately choose a random device and then take the action it selects. Auditing, inspection, and competitive play can involve such choices. At a mixed equilibrium, however, each supported pure action gives the same payoff against the opponents' fixed equilibrium behavior. Optimality therefore does not select a unique randomization at that belief.

In a zero sum game, the equilibrium distribution can nevertheless be the unique way to guarantee the value against every opponent strategy. Matching Pennies illustrates the distinction: either pure action is optimal against a uniform opponent, but only the uniform distribution guarantees zero against all opponents. If an announced randomization affects a later player's behavior, the announcement and the later response belong in an extensive game.

3.2.2 A stochastic steady state

The probabilities can describe stable frequencies in repeated encounters. Each player uses the observed frequencies of opponents' actions to form beliefs, and every action that occurs with positive frequency is optimal against those beliefs. In a population interpretation, one person is drawn from each of several populations for each encounter; α_i describes the distribution of actions in population i.

The independence condition excludes a correlation that a player could use to predict the opponents' actions from additional information. When observed signals contain such information, best responses must be calculated conditionally, leading to correlated equilibrium.

3.2.3 Pure strategies in an extended game

Let player i privately observe an independent signal t_i that does not affect payoffs. A deterministic map $s_i(t_i)$ induces a distribution over actions. If these distributions form a mixed equilibrium, and the map only selects best responses, the maps form a pure Bayesian equilibrium of the extended game.

For example, give each player in Bach or Stravinsky three equally likely independent signals. Player 1 chooses B at two signals and player 2 chooses B at one. Their action frequencies are $2/3$ and $1/3$, respectively, and every type is indifferent between B and S. The same mixed outcome can therefore arise without a player consciously choosing a lottery.

3.2.4 Pure strategies in a perturbed game

Purification makes private information affect preferences. Replace $u_i(a)$ by $u_i(a)+\gamma\epsilon_i(a)$, where each coordinate of the random vector $\epsilon_i = (\epsilon_i(a))_{a\in A}$ lies in $[-1,1]$. Assume the vectors are independent across players and each has an absolutely continuous distribution with a continuously differentiable density. Each player observes its own vector. Against fixed opponent strategies, almost every type then has a unique best response.

The book's Section 3.2.4 states that, for almost every finite game and every fixed perturbation distribution satisfying these conditions, almost every mixed strategy equilibrium is approached by action distributions induced by pure Bayesian equilibria as $\gamma \downarrow 0$. Conversely, every convergent sequence of these equilibrium distributions has a mixed strategy equilibrium of the unperturbed game as its limit. A degenerate game may have equilibria that a particular perturbation does not approach.

Example 3.6 (Purifying the interior equilibrium of BoS). Suppose player i receives an independent δ_i uniform on $[-1, 1]$, and the perturbation adds $\gamma\delta_i$ to the payoff difference between B and S. At interior thresholds the equilibrium probabilities satisfy

$$p = \frac{1}{2} + \frac{3q-1}{2\gamma}, \qquad q = \frac{1}{2} + \frac{3p-2}{2\gamma}.$$

The first equation follows because player 1 chooses B exactly when $3q - 1 + \gamma\delta_1 \geq 0$; the second follows from the analogous comparison for player 2. Solving gives

$$p = \frac{6-\gamma-2\gamma^2}{9-4\gamma^2}, \qquad q = \frac{3+\gamma-2\gamma^2}{9-4\gamma^2}.$$

For sufficiently small positive γ both thresholds are interior, and these probabilities converge to $(2/3, 1/3)$. Almost every type makes a strict choice even though the limiting complete information game has indifferent mixing players.

3.2.5 Mixed strategies as beliefs

Alternatively, α_i can be the other players' common belief about player i's action. The support condition says that all actions to which this belief assigns positive probability are rational given the other equilibrium beliefs. A player may choose one pure action in a particular encounter. Under this interpretation, equilibrium restricts beliefs and best responses; it need not specify an actual random device or individual frequency.

3.3 Correlated equilibrium

Definition 3.7 (Correlated equilibrium; Definition 45.1). A correlated equilibrium consists of a finite probability space (Ω, π), an information partition $\mathcal{P}_i$ for each player, and a strategy $\sigma_i : \Omega \to A_i$ constant on every cell of $\mathcal{P}_i$. For every player i and every alternative strategy τ_i constant on the same cells,

$$\sum_{\omega} \pi(\omega) u_i(\sigma(\omega)) \geq \sum_{\omega} \pi(\omega) u_i(\tau_i(\omega), \sigma_{-i}(\omega)).$$

Since a player can alter behavior at just one information cell, this ex ante inequality is equivalent to conditional optimality at every cell with positive probability. The signal distributions and information partitions are part of the equilibrium description. The common probability measure ensures that all conditional expectations come from the same distribution.

Proposition 3.8 (Direct recommendations; Proposition 47.1). A distribution $\mu \in \Delta(A)$ is the outcome distribution of a correlated equilibrium if and only if, for every i and every $a_i, b_i \in A_i$,

$$\sum_{a_{-i}} \mu(a_i, a_{-i}) \big[u_i(a_i, a_{-i}) - u_i(b_i, a_{-i})\big] \geq 0.$$

Proof. A mediator can draw a from μ and privately recommend a_i to player i. The inequality is the expected gain from obedience on receipt of that recommendation, multiplied by its probability.

Conversely, in any correlated equilibrium collect all states giving the same action profile into its probability mass. A deviation that changes only one recommended action was available at each original information cell producing that action. Summing the original incentive inequalities gives the displayed condition.

Thus correlated equilibrium distributions satisfy finitely many linear inequalities, along with nonnegativity and total mass one. Their set and the associated payoff set are convex. A public lottery over correlated equilibria is another correlated equilibrium. Every mixed equilibrium induces one by taking $\mu(a) = \prod_i \alpha_i(a_i)$.

In Bach or Stravinsky, public randomization equally between (B, B) and (S, S) gives each player $3/2$. Conditional on the public signal, each knows that the other will coordinate. Correlation also allows outcomes beyond public mixtures of Nash equilibria.

Example 3.9 (A correlated payoff outside the Nash payoff hull). Consider

	L	R
T	$(6,6)$	$(2,7)$
B	$(7,2)$	$(0,0)$

.

Put probability $1/3$ on each of $(T, L), (T, R), (B, L)$. A recommendation of B reveals L, against which B is optimal. A recommendation of T implies equal probabilities of L and R, giving 4 from T and $7/2$ from B. The column player's conditions are symmetric. Expected payoffs are $(5, 5)$. The pure equilibrium payoffs are $(2, 7)$ and $(7, 2)$, and the mixed equilibrium gives $(14/3, 14/3)$. Their total payoffs are at most $28/3$, whereas the correlated outcome has total payoff 10, proving that it lies outside their convex hull.

If players assign different probabilities to the signal, their subjective expected payoffs need not be evaluations of one common outcome distribution. Such a model can sustain additional subjective payoff comparisons, but it is distinct from correlated equilibrium with a common prior as defined here.

3.4 Evolutionary equilibrium

Consider a symmetric two player game with common action set B and payoffs $u(a, b)$ to the player using a against b. Pairwise encounters are random. Most members of a population use a resident strategy b^*, while a fraction ϵ use a mutant strategy b. Their expected payoffs are

$$U_R = (1-\epsilon)u(b^*, b^*) + \epsilon u(b^*, b), \qquad U_M = (1-\epsilon)u(b, b^*) + \epsilon u(b, b).$$

Definition 3.10 (Evolutionarily stable strategy; Definition 49.1). An action b^* is evolutionarily stable if, for every $b \neq b^*$, either

$$u(b^*, b^*) > u(b, b^*),$$

or equality holds and $u(b^*, b) > u(b, b)$. Equivalently, (b^*, b^*) is a Nash equilibrium and the second inequality holds for every different best response to b^*.

The definition is equivalent to $U_R > U_M$ for every sufficiently small positive mutant share, with the bound allowed to depend on the mutant. The first comparison concerns encounters with

residents. If that comparison is an equality, encounters involving mutants determine stability. Every strict symmetric equilibrium is evolutionarily stable. A symmetric equilibrium with additional best responses requires the second comparison.

Example 3.11 (Hawk and Dove). Let the prize be one and the cost of a fight be $c > 0$. Payoffs to a player at $(D, D), (D, H), (H, D), (H, H)$ are $1/2, 0, 1, (1 - c)/2$. Against an opponent who chooses H with probability q, the payoff advantage of H over D is $(1 - cq)/2$.

If $c > 1$, the symmetric mixed equilibrium has hawk probability $p = 1/c$. Every mixed strategy is a best response to p. A mutant with hawk probability r satisfies

$$u(p, r) - u(r, r) = \frac{(p - r)(1 - cr)}{2} = \frac{c}{2}(p - r)^2 > 0 \quad (r \neq p).$$

The mixed equilibrium strategy is therefore evolutionarily stable. If $c < 1$, H strictly dominates D, so H is the unique stable strategy. At $c = 1$ the same calculation with $p = 1$ also gives stability of H.

Existence of a symmetric Nash equilibrium does not imply existence of an evolutionarily stable strategy. In the symmetric cyclic game with row payoff matrix

$$\begin{pmatrix} \gamma & 1 & -1 \\ -1 & \gamma & 1 \\ 1 & -1 & \gamma \end{pmatrix}, \qquad 0 < \gamma < 1,$$

the unique symmetric equilibrium is uniform. Each action yields $\gamma/3$ against it. A pure mutant obtains γ against itself, while a resident obtains only $\gamma/3$ against that mutant. The second stability inequality fails, so the symmetric equilibrium is not evolutionarily stable.

4 Rationalizability and Iterated Elimination of Dominated Actions

4.1 Rationalizability

Nash equilibrium requires correct expectations about the opponents' behavior. Rationalizability instead asks whether an action can be justified by beliefs that assign probability only to rational choices, where those choices can themselves be justified in the same way. Beliefs at different stages of the argument need not agree with one another or with actual play.

For a finite strategic game, a belief of player i is a probability distribution $\mu_i \in \Delta(A_{-i})$. The book permits correlation between the actions of different opponents. An action a_i is a best response to μ_i if

$$\sum_{a_{-i}} \mu_i(a_{-i}) u_i(a_i, a_{-i}) \geq \sum_{a_{-i}} \mu_i(a_{-i}) u_i(b_i, a_{-i}) \quad \text{for every } b_i \in A_i.$$

Definition 4.1 (Rationalizable action; Definition 55.1). An action a_i is rationalizable if there are nonempty sets $Z_j \subseteq A_j$, one for each player, such that $a_i \in Z_i$ and every $a_j \in Z_j$ is a best response to some belief supported on $Z_{-j} = \prod_{k \neq j} Z_k$.

The best response is taken among all actions in A_j, including actions outside Z_j. Each proposed action has a supporting belief about opponents' actions, and each of those actions has its own supporting belief. The book's Definition 54.1 writes out an infinite sequence of these justifications. Taking the union of the actions appearing at every stage yields the sets Z_j above. Conversely, one can repeatedly use the supporting beliefs of the Z_j to construct a justification at every depth. These are two descriptions of the same condition.

If two profiles of sets satisfy the condition, their componentwise unions also do: a supporting belief for an action in either original set remains supported on the larger union. Hence there is a largest profile of such sets, the rationalizable action sets R_i.

Proposition 4.2 (Equilibrium actions are rationalizable; Lemma 56.2). Every action used with positive probability in a mixed or correlated equilibrium of a finite game is rationalizable.

Proof. For a correlated equilibrium let Z_i be the actions used with positive probability. Conditional on a recommendation of $a_i \in Z_i$, the distribution of opponents' recommendations is supported on Z_{-i} and makes a_i optimal. These are the required beliefs. Mixed equilibria are a special case.

The implication runs from an equilibrium action to a rationalizable action. Choosing one rationalizable action for each player does not generally produce an equilibrium. In Bach or Stravinsky both actions of both players are rationalizable, because each action is a best response to coordination on it. The mismatched profile (B, S) is therefore a profile of rationalizable actions, although it is not a Nash equilibrium. In the Prisoner's Dilemma, only defection is rationalizable because cooperation is never optimal.

4.1.1 A continuous example with a unique rationalizable action

Consider Cournot quantities $a_i \in [0,1]$ and profits $u_i(a_i, a_j) = a_i(1 - a_i - a_j)$. Against a belief with mean quantity m, expected profit is $a_i(1 - a_i - m)$ and the unique best response is $(1-m)/2$. Thus the first round of rationality restricts quantities to $[0, 1/2]$.

Starting with $I_0 = [0,1]$, define the successive intervals of possible best responses by

$$I_{t+1} = [\ell_{t+1}, h_{t+1}], \qquad \ell_{t+1} = \frac{1 - h_t}{2}, \quad h_{t+1} = \frac{1 - \ell_t}{2}.$$

Every belief on I_t has its mean in I_t, and every mean in that interval is attainable. Thus these intervals describe exactly the remaining best responses. Their lengths satisfy $h_{t+1} - \ell_{t+1} = (h_t - \ell_t)/2$, and both endpoints converge to $1/3$. Any rationalizable action must remain in every interval, so the only such action is $1/3$. It is rationalizable because it is a best response to itself.

4.1.2 Why correlated beliefs matter

In a game with three players, player 1 chooses U, D, player 2 chooses L, R, and player 3 chooses M_1, M_2, M_3, M_4. All players share the payoff specified by the following matrices, one for each choice of player 3:

$$M_1 = \begin{pmatrix} 8 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & 0 \\ 0 & 8 \end{pmatrix}, \quad M_4 = \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix}.$$

Player 3 can justify M_2 by giving equal probability to (U, L) and (D, R): its expected payoff is four, matching M_1 and M_3 and exceeding M_4. Both actions of each of the first two players can be justified by a suitable opponent action and M_2. Therefore M_2 is rationalizable.

With independent beliefs, write $p = \Pr(U)$ and $q = \Pr(L)$. For M_2 to be optimal one would need

$$4\big[pq + (1-p)(1-q)\big] \geq \max\{8pq, 8(1-p)(1-q), 3\}.$$

The first two comparisons force $pq = (1-p)(1-q)$, hence $p+q = 1$. Then the left side is $8p(1-p) \leq 2$, contradicting the comparison with three. Independence therefore eliminates M_2. With two players there is only one opponent, so this distinction disappears.

4.2 Iterated elimination of strictly dominated actions

4.2.1 Actions that are never best responses

> Definition 4.3 (Strict dominance; Definitions 59.1 and 59.2). An action is a never best response if it is not optimal against any belief about opponents' actions. Action a_i is strictly dominated if some mixed strategy α_i satisfies
>
> $$U_i(\alpha_i, a_{-i}) > u_i(a_i, a_{-i}) \qquad \text{for every } a_{-i} \in A_{-i}.$$

Dominance allows a mixed alternative even when the action being eliminated is pure. This is necessary to obtain the full implication of rationality. A pure action can be worse than a mixture against every opponent profile without being worse than any single pure action everywhere.

Lemma 4.4 (Dominance and beliefs; Lemma 60.1). In a finite game, an action is a never best response if and only if it is strictly dominated by a mixed strategy.

Proof. Strict dominance implies a strict gain under every belief by taking expectations. For the converse, fix a_i and form a zero sum game with payoff

$$v(b_i, a_{-i}) = u_i(b_i, a_{-i}) - u_i(a_i, a_{-i}).$$

The maximizing player chooses an alternative b_i and the minimizing player chooses an opponent profile. If a_i is never a best response, then for every distribution μ over opponent profiles, $\max_{b_i} \mathbb{E}_\mu v(b_i, a_{-i}) > 0$. This maximum is continuous in μ; the belief simplex is compact, so its minimum is strictly positive. The finite minimax theorem gives

$$\max_{\alpha_i} \min_{\mu} \mathbb{E}_{\alpha_i,\mu} v = \min_{\mu} \max_{\alpha_i} \mathbb{E}_{\alpha_i,\mu} v > 0.$$

A maximizing α_i consequently gives a positive gain against every pure opponent profile, which is strict dominance.

Finiteness provides compactness and the minimax argument used here. Expected utility provides linearity in beliefs and in mixtures. Both are assumptions of the equivalence stated in the lemma.

4.2.2 The iterative procedure

Start with $X_i^0 = A_i$. At each stage remove any collection of actions strictly dominated in the current restricted game, obtaining $X_i^{t+1} \subseteq X_i^t$. Continue until no action is strictly dominated. The dominating mixture at a stage uses the actions available at that stage.

Proposition 4.5 (Rationalizability and elimination; Proposition 61.2). In a finite game, every completed sequence of iterated elimination of strictly dominated actions leaves exactly the rationalizable actions. The terminal sets do not depend on the order or speed of elimination.

Proof. Take sets Z_i supporting a rationalizable action. Inductively suppose $Z_i \subseteq X_i^t$ for every player. Every action in Z_i is a best response to a belief supported on $Z_{-i} \subseteq X_{-i}^t$, so it cannot be strictly dominated in the current game. Thus no rationalizable action is ever removed.

Conversely, let X_i be the terminal sets. By the lemma, every terminal action is optimal among X_i against some belief on X_{-i}. It is also optimal among the original actions. To see this, maximize expected utility under that belief in any earlier action set. An action removed at that stage cannot be such a maximizer, since its dominating mixture gives strictly more against every current opponent profile, and hence against this belief. The maximum is therefore preserved at every deletion. The terminal action attains the original maximum, and the sets X_i meet the definition of rationalizability.

Example 4.6 (Mixed dominance followed by pure dominance). Consider

	L	R
T	$(3,0)$	$(0,1)$
M	$(0,0)$	$(3,1)$
B	$(1,1)$	$(1,0)$

.

The mixture assigning probability $1/2$ to T and M gives player 1 $3/2$ against either column, strictly dominating B. Once B is removed, player 2 strictly prefers R to L. Against R, player 1 strictly prefers M to T. The unique surviving profile is (M, R). Pure dominance alone would make no deletion in the original game.

The result also gives nonemptiness in finite games: a mixed Nash equilibrium exists, and all its supported actions survive. The outcome is a product of action sets, so the procedure can retain profiles that are not jointly consistent as equilibrium choices.

4.3 Iterated elimination of weakly dominated actions

Definition 4.7 (Weak dominance; Definition 62.1). Action a_i is weakly dominated if some mixed strategy α_i gives at least as much against every opponent profile and strictly more against at least one:

$$U_i(\alpha_i, a_{-i}) \geq u_i(a_i, a_{-i}) \quad \text{for all } a_{-i},$$

with strict inequality for some a_{-i}.

A weakly dominated action can be a best response to a belief that assigns zero probability to every opponent profile where dominance is strict. This is why its elimination uses more than the unrestricted best response requirement underlying rationalizability.

Proposition 4.8 (Admissibility and full support beliefs). In a finite game, a mixed strategy is not weakly dominated if and only if it is a best response to some belief assigning positive probability to every opponent action profile.

Argument. If a strategy is optimal at a full support belief, a weakly dominating strategy would give a strictly higher expectation, a contradiction. For the converse, record each mixed strategy as its vector of payoffs against the finitely many opponent profiles. The feasible payoff vectors form the convex hull of finitely many pure payoff vectors. A vector that is not weakly dominated is an efficient point of this polytope. A supporting linear functional with strictly positive coefficients can be chosen at such a point; normalizing its coefficients to sum to one produces the required full support belief. The finite polytope property is the additional ingredient used in Exercise 64.2.

Unlike strict dominance, repeated weak dominance can depend on the deletion order. The book's example is

	L	R
T	$(1,1)$	$(0,0)$
M	$(1,1)$	$(2,1)$
B	$(0,0)$	$(2,1)$

.

Action M weakly dominates both T and B. Delete T first. In the remaining game R weakly dominates L, and after deleting L every surviving outcome pays $(2,1)$. Alternatively, delete B first. Then L weakly dominates R, and after deleting R all surviving outcomes pay $(1,1)$. Deleting one action removes the profile at which another dominance comparison was strict.

A game is dominance solvable in the book's terminology when the procedure that deletes all weakly dominated actions at each stage leaves outcomes between which every player is indifferent. Specifying simultaneous deletion is part of this definition. It does not imply that every other admissible order of weak deletions produces the same outcome set or payoff profile.

5 Knowledge and Equilibrium

5.1 A model of knowledge

5.1.1 Information functions and partitions

Let Ω be a state space. A state describes the features of the situation relevant to the analysis. When the question concerns interactive knowledge, these features include what each player knows about the state and about the information of other players.

Definition 5.1 (Information function; Definition 68.1). An information function assigns to each $\omega \in \Omega$ a nonempty set $P(\omega) \subseteq \Omega$. At state ω, the decision maker considers precisely the states in $P(\omega)$ possible. The function is partitional if $P(\omega)$ is the cell containing ω in a partition of Ω.

Lemma 5.2 (Partition conditions; Lemma 68.3). An information function is partitional if and only if

$$\omega \in P(\omega), \qquad \omega' \in P(\omega) \implies P(\omega') = P(\omega).$$

Proof. A partition satisfies both conditions. Conversely, the first makes the information sets cover Ω. If two such sets intersect at ω'', the second makes both equal to $P(\omega'')$. Distinct information sets are therefore disjoint and form a partition.

The first condition excludes certainty of a false event. The second requires the decision maker's account of possible states to be the same at all states the decision maker cannot distinguish. For example, observing the first four decimal digits of a number in $[0, 1)$ produces a partition into intervals of length 10^{-4}. By contrast, the rule $P(n) = \{n-1, n, n+1\}$ for integers is not partitional: neighboring information sets overlap without coinciding.

5.1.2 Knowledge operators

Definition 5.3 (Knowledge of an event). An event is a subset $E \subseteq \Omega$. Its knowledge event is

$$K(E) = \{\omega \in \Omega : P(\omega) \subseteq E\}.$$

Thus $\omega \in K(E)$ means that at ω the information rules out every state outside E.

For any information function, $K(\Omega) = \Omega$, knowledge is monotone with respect to set inclusion, and $K(E \cap F) = K(E) \cap K(F)$. With partitional information one also has

$$K(E) \subseteq E, \qquad K(K(E)) = K(E), \qquad K(\Omega \setminus K(E)) = \Omega \setminus K(E).$$

The first property is truth of knowledge. The second says that knowing an event includes knowing that one knows it. The third says that failure to know an event is itself known. These last two properties follow because $K(E)$ is a union of information cells, as is its complement.

The information set can be recovered from the knowledge operator by

$$P(\omega) = \bigcap \{E \subseteq \Omega : \omega \in K(E)\}.$$

Every event known at ω contains $P(\omega)$, and $P(\omega)$ itself is known there. For a finite state space, the reverse construction also recovers a knowledge operator satisfying preservation of the whole space, monotonicity, and finite intersections, provided the resulting information sets are nonempty.

Proposition 5.4 (More information in an individual decision problem). Fix a prior π on a finite state space, a nonempty action set A, and a bounded payoff $u(a, \omega)$. If partition $\mathcal{P}$ refines partition $\mathcal{Q}$, the supremum expected payoff with $\mathcal{P}$ is at least the supremum with $\mathcal{Q}$. When optimal acts exist, this compares their attained maximum payoffs.

Proof. A feasible act is constant on each information cell. Every act constant on the larger cells of $\mathcal{Q}$ is also constant on the smaller cells of $\mathcal{P}$. Refining information expands the set of feasible acts.

The argument holds the payoff environment fixed. In a game, changing one player's information can change other players' equilibrium actions, so an equilibrium payoff need not obey the same comparison.

5.1.3 The puzzle of the hats

There are n individuals, each wearing a white or black hat and observing everyone else's hat. A public announcement says that at least one hat is white. After each count, anyone who knows their own color raises a hand. Let

$$\Omega = \{c \in \{W, B\}^n : \#\{i : c_i = W\} \geq 1\}.$$

Individual i initially considers the states agreeing with the observed c_{-i} possible. Usually there are two such states. If all observed hats are black, only the state with $c_i = W$ remains.

If exactly one hat is white, its wearer knows this at the first count. If nobody raises a hand, everyone removes all states with exactly one white hat. A person seeing exactly one white hat can then infer that their own hat is white and raises a hand at the second count. Inductively, after $k-1$ counts with no raised hand, the remaining states have at least k white hats. If the actual number is k, all k wearers of white hats first raise their hands at count k.

The announcement and the publicly observed absence of a response change what is known about others' knowledge. Even when everyone already sees a white hat, the public announcement changes the higher orders of knowledge needed for this induction.

5.2 Common knowledge

For player information functions P_i, write K_i for their knowledge operators. Mutual knowledge of E means membership in $\bigcap_i K_i(E)$. Common knowledge requires every finite iteration of the operators to hold, including knowledge of others' knowledge at arbitrarily many levels.

Definition 5.5 (Self evident events; Definition 73.2). An event F is self evident to the players if $P_i(\omega) \subseteq F$ for every $\omega \in F$ and every i. An event E is common knowledge at ω if some self evident event F satisfies $\omega \in F \subseteq E$.

For partitions, F is self evident precisely when it is a union of cells of every player's partition. Equivalently, $K_i(F) = F$ for all i. The common knowledge cell $C(\omega)$ consists of all states reachable from ω by a finite sequence of steps, each step moving between two states in some player's information cell. It is the smallest self evident event containing ω.

Proposition 5.6 (Equivalent common knowledge conditions). For finite Ω with partitional information, E is common knowledge at ω if and only if every finite iteration of the knowledge operators holds there, and if and only if $C(\omega) \subseteq E$.

Proof. A self evident subset F of E is preserved by every knowledge operator, so all their finite iterations contain F. Conversely, failure of $C(\omega) \subseteq E$ gives a finite sequence of information steps from ω to a state outside E. The corresponding ordered iteration of knowledge operators then fails at ω. The reachable set is itself closed under every information step, establishing the remaining equivalence.

Example 5.7 (Mutual knowledge without common knowledge). Let the partitions on $\{1,2,3,4,5,6\}$ be

$$\mathcal{P}_1 = \{\{1,2\},\{3,4,5\},\{6\}\}, \qquad \mathcal{P}_2 = \{\{1\},\{2,3,4\},\{5\},\{6\}\}.$$

For $E = \{1,2,3,4\}$, both players know E at states 1 and 2. However, $C(1) = C(2) = \{1,2,3,4,5\}$, so E is not common knowledge there. The connection through player 2's cell $\{2,3,4\}$ and player 1's cell $\{3,4,5\}$ eventually reaches state 5, outside E.

5.3 Can people agree to disagree?

Theorem 5.8 (Agreement; Proposition 75.1). Suppose a finite state space has partitional information and a common prior ρ. At a state in a positive probability common knowledge cell, suppose it is common knowledge that player 1's posterior probability of E is η_1 and player 2's is η_2. Then $\eta_1 = \eta_2$.

Proof. Let F be a positive probability self evident event on which both specified posterior equalities hold. Since F is a union of player 1's information cells, denoted D, the law of total probability gives

$$\rho(E \cap F) = \sum_{D \subseteq F} \rho(D)\rho(E \mid D) = \eta_1 \rho(F).$$

Summing over player 2's cells gives $\rho(E \cap F) = \eta_2 \rho(F)$. Dividing by $\rho(F) > 0$ proves the result.

Different private information can produce different posteriors. What the theorem rules out is common knowledge of two specified unequal posterior values under a common prior. Even common knowledge that the values differ is weaker: their ordering may vary across states in the common knowledge cell. By contrast, it cannot be common knowledge that player 1's posterior is strictly above player 2's everywhere in a positive probability cell, because their conditional averages over that cell both equal $\rho(E \mid F)$.

The same reasoning applies to expectations. Two players with a common prior cannot have it be common knowledge that one evaluates a common random variable strictly above a number η while the other evaluates it strictly below η. Averaging each player's conditional expectation over the common knowledge event would give contradictory inequalities for the same expectation.

5.4 Knowledge and solution concepts

Fix a strategic game. A state now specifies the action $a_i(\omega)$, information set $P_i(\omega)$, and belief $\mu_i(\omega) \in \Delta(A_{-i})$ of every player. A belief is consistent with knowledge if

$$\operatorname{supp} \mu_i(\omega) \subseteq \{a_{-i}(\omega') : \omega' \in P_i(\omega)\}.$$

Player i is rational at ω when $a_i(\omega)$ is a best response to that belief. Consistency here is a restriction on support; it does not require a common prior or Bayesian conditioning.

> Proposition 5.9 (Knowledge of actions; Proposition 77.1). If at a state every player knows the opponents' actions, has a belief consistent with knowledge, and is rational, the action profile is a Nash equilibrium.

Proof. Knowing opponents' actions makes the action profile constant over the information set. Consistency forces the belief to assign probability one to that profile, and rationality gives the Nash best response inequalities.

> Proposition 5.10 (Knowledge of beliefs with two players; Proposition 78.1). In a two player finite game, suppose each player knows the other's belief, has a belief consistent with knowledge, and knows that the other is rational. Then
>
> $$(\alpha_1, \alpha_2) = (\mu_2(\omega), \mu_1(\omega))$$
>
> is a mixed strategy Nash equilibrium.

Proof. If a_i is in the support of $\mu_j(\omega)$, consistency supplies a state $\omega' \in P_j(\omega)$ where player i takes that action. Player j knows that i is rational, so a_i is optimal against $\mu_i(\omega')$. Player j also knows i's belief, making this belief equal to $\mu_i(\omega)$. Every supported action is therefore optimal against the other component of α.

With more than two players, different players may hold different beliefs about the same third player's action. Mutual knowledge of beliefs does not by itself make these marginal beliefs identical. The book gives a three player construction in which player 1 assigns probability $1/3$ to player 3's action A and player 2 assigns probability $2/3$, although at the state under study everyone knows the other players' beliefs and rationality. A common prior together with stronger common knowledge of beliefs removes this particular disagreement through the agreement argument.

> Proposition 5.11 (Common knowledge of rationality; Proposition 80.1). If rationality and consistency of beliefs with knowledge are common knowledge at a state, the actions chosen there are rationalizable.

Proof. Let F be a self evident event containing the state on which these properties hold, and set $Z_i = \{a_i(\omega') : \omega' \in F\}$. An action in Z_i is optimal against its belief at some $\omega' \in F$. Consistency restricts that belief to actions arising in $P_i(\omega')$, and self evidence makes $P_i(\omega') \subseteq F$. Thus the belief is supported on Z_{-i}. The sets Z_i meet the rationalizability condition.

Finally, if actions are constant on information cells, beliefs everywhere are conditional beliefs from one common prior, and every player is rational at every state, the information structure and action functions themselves form a correlated equilibrium. Sum conditional optimality over the cells of each player's partition to obtain its ex ante incentive inequality.

5.5 The electronic mail game

Two players choose A or B. With probability $1-p$ the payoff matrix is G_a, and with probability $0<p<1/2$ it is G_b, where $L>M>1$:

$$G_a = \begin{array}{c|cc} & A & B \\ \hline A & (M,M) & (1,-L) \\ B & (-L,1) & (0,0) \end{array}, \qquad G_b = \begin{array}{c|cc} & A & B \\ \hline A & (0,0) & (1,-L) \\ B & (-L,1) & (M,M) \end{array}.$$

Player 1 initially knows the game. In state b, its computer sends a message, and each received message generates a confirmation. Each message independently fails to arrive with probability $\epsilon \in (0,1)$, ending the exchange. A player observes only the number of messages its own computer has sent.

Write a state as (Q_1, Q_2), where $Q_1 = Q_2$ or $Q_1 = Q_2 + 1$. The state $(0,0)$ means G_a; all other states mean G_b. The common prior is

$$\rho(0,0) = 1-p, \quad \rho(q+1,q) = p\epsilon(1-\epsilon)^{2q}, \quad \rho(q+1,q+1) = p\epsilon(1-\epsilon)^{2q+1}.$$

For $q \geq 1$, player 1 observing q considers $(q,q-1)$ and (q,q) possible. Player 2 observing $q \geq 0$ considers (q,q) and $(q+1,q)$ possible. Every finite state is linked through these information cells to $(0,0)$. Hence G_b is never common knowledge, however many messages have arrived. More messages do establish progressively higher finite orders of knowledge of G_b.

> Proposition 5.12 (Electronic mail equilibrium; Proposition 83.1). The electronic mail game has a unique Bayesian Nash equilibrium: every type of each player chooses A.

Proof. Player 1's type zero knows G_a, where A is strictly dominant. Player 2's type zero considers states $(0,0)$ and $(1,0)$. Choosing A gives expected payoff at least $(1-p)M/(1-p+p\epsilon)$, whereas choosing B gives at most $[-(1-p)L + p\epsilon M]/(1-p+p\epsilon)$. The first is larger because $p<1/2$ and $L>M$.

Suppose player 2's type $q-1$ must choose A. Player 1's type q assigns probability

$$z = \frac{\epsilon}{\epsilon + (1-\epsilon)\epsilon} = \frac{1}{2-\epsilon} > \frac{1}{2}$$

to player 2 being that lower type. Thus B gives at most $-zL + (1-z)M < 0$, whereas A guarantees a nonnegative payoff. Player 1's type q must choose A. The same comparison then forces player 2's type q to choose A, since it assigns probability z to player 1's type q rather than $q+1$. Induction covers all finite types. The resulting profile is indeed an equilibrium, since A is strictly optimal against an opponent who always chooses A in either game.

For every $\epsilon > 0$, $z > 1/2$, so the restriction at the least informed type extends to all types. Even as ϵ tends to zero, every type chooses A in equilibrium. A public signal identifying G_b instead makes (B,B) a Nash equilibrium by making the payoff matrix common knowledge.

6 Extensive Games with Perfect Information

6.1 Extensive games with perfect information

6.1.1 Histories and the player function

An extensive game specifies the order of decisions and what has happened when each decision is made. In a game with perfect information, the moving player observes the full preceding history. Initially consider games without chance moves and with one player moving at each history.

> Definition 6.1 (Extensive game; Definition 89.1). An extensive game with perfect information is $\Gamma = \langle N, H, P, (\succeq_i)\rangle$. The history set H contains the empty history, every prefix of every history in H, and every infinite sequence whose finite prefixes belong to H. A history is terminal if it is infinite or has no continuation in H; write Z for the terminal histories. The function $P : H \setminus Z \to N$ identifies the player who moves. Each $\succeq_i$ is a preference relation on Z.

The available actions after h are $A(h) = \{a : (h, a) \in H\}$. An extensive game form specifies N, H, P without preferences. A finite game has finitely many histories. A game has a finite horizon if there is a finite bound on the length of all histories. Finite horizon does not imply finiteness when a player can choose from an infinite set of actions.

An infinite sequence is already a terminal history: the definition permits no further action after it. A preference ranking over Z must therefore say how players evaluate infinite play as well as finite termination.

6.1.2 Strategies and outcomes

> Definition 6.2 (Strategy; Definition 92.1). A strategy s_i assigns an action $s_i(h) \in A(h)$ to every nonterminal history at which $P(h) = i$. A strategy profile s induces the terminal history $O(s)$ obtained by following these prescriptions from the empty history onward.

A strategy specifies choices even after histories inconsistent with the player's own earlier choices. Suppose player 1 first chooses A or B, where B ends the game. Following A, player 2 chooses C or D; following (A, C), player 1 chooses E or F. Player 1 has four strategies AE, AF, BE, BF. The last two induce the same outcome against every strategy of player 2, but differ in their prescription after (A, C).

For Nash equilibrium one can use reduced strategies that omit choices made irrelevant by the player's own earlier prescriptions. In the example these are AE, AF, B. A further reduction identifies strategies that give every player the same payoff against every opponent strategy, even if they generate different terminal histories. This produces the reduced strategic form of Definition 95.1. Subgame perfection requires the full contingent strategy, since behavior after every history is evaluated.

6.1.3 Nash equilibrium in the strategic form

The strategic form has one action for each extensive form strategy and preferences induced by O. Its Nash equilibrium condition is

$$u_i(O(s^*)) \geq u_i(O(s_i, s^*_{-i})) \qquad \text{for every } i \text{ and } s_i.$$

This condition evaluates deviations from the beginning of the game. An action at a history that never occurs may have no effect on the player's equilibrium payoff, even when that action would be strictly inferior if the history occurred.

Example 6.3 (A Nash equilibrium supported by an unprofitable threat). Player 1 chooses B, ending at $(1, 2)$, or A, after which player 2 chooses L for $(0, 0)$ or R for $(2, 1)$. There are two pure Nash equilibria, (B, L) and (A, R). In (B, L), player 1 prefers B given the threatened L, and player 2's prescribed choice does not affect the outcome. After A actually occurs, however, player 2 strictly prefers R.

6.2 Subgame perfect equilibrium

Definition 6.4 (Subgames and subgame perfection; the book's Section 6.2). The subgame $\Gamma(h)$ following history h contains all continuations h' for which $(h, h') \in H$, with the inherited player function and preferences. The strategy restriction is $s_i|h(h') = s_i(h, h')$. A profile s^* is subgame perfect if $s^*|h$ is a Nash equilibrium of $\Gamma(h)$ for every history h.

Thus the strategy must be optimal after histories generated by equilibrium play and after histories generated by deviations. In the preceding example, subgame perfection forces player 2 to choose R after A and player 1 to choose A initially. The other Nash equilibrium fails in the subgame following A.

Example 6.5 (Dividing two indivisible objects). Player 1 proposes $(2, 0), (1, 1)$, or $(0, 2)$; player 2 accepts or rejects, and rejection gives both zero. In every subgame perfect equilibrium player 2 accepts offers giving it one or two objects. It is indifferent at an offer of zero. If it accepts zero, player 1 proposes $(2, 0)$; if it rejects zero, player 1 proposes $(1, 1)$. Hence there are two subgame perfect equilibria, with different allocations. Indifference at a later decision can affect an earlier player's optimal choice.

Lemma 6.6 (The one deviation property; Lemma 98.2). In a finite horizon game with perfect information, a strategy profile is subgame perfect if and only if no player can gain, in any subgame where that player moves first, by changing only its initial action and then returning to the prescribed strategy.

Proof. Necessity follows from the definition. If a profitable deviation exists, only finitely many of its actions are relevant on its induced path. Among profitable deviations choose one differing from the prescribed strategy at as few relevant histories as possible. At the last difference, either the deviation is profitable in that subgame, yielding a profitable single deviation, or returning to

the prescribed action is weakly better. In the second case the earlier profitable deviation remains profitable with fewer differences, a contradiction.

Example 6.7 (Failure at an infinite horizon). One player can stop or continue at each date. Stopping at any finite date gives zero; continuing forever gives one. The strategy that always stops has no profitable single deviation: one continuation followed by stopping still gives zero. Continuing at every subsequent date is profitable. The finite horizon assumption in the lemma cannot be removed without another condition on preferences at infinity.

Theorem 6.8 (Backward induction; Proposition 99.2). Every finite extensive game with perfect information has a pure subgame perfect equilibrium.

Proof. At each history immediately preceding termination, choose a maximizing action for the moving player. Replace that decision by the terminal outcome it induces. Repeat at the histories of the resulting shorter game until reaching the initial history. Finiteness guarantees that a maximizing action exists at each step and that the procedure ends. The resulting strategy profile has no profitable single deviation after any history, hence is subgame perfect.

If no player is indifferent between distinct terminal histories, each maximization has a unique solution, giving a unique subgame perfect equilibrium. A weaker condition is that whenever one player is indifferent between two histories, every player is indifferent. Under that condition all subgame perfect outcomes give the same payoff profile, and equilibrium strategies are interchangeable. Induction proves this because different maximizers at a decision then give identical continuation payoffs to everyone.

Finiteness of the history set is substantive. A one decision game in which the player chooses $x \in [0, 1)$ and receives x has finite horizon but no optimal choice. Finite action sets at each history are also insufficient without a bound on the horizon: a player who can stop at date t for $1 - 1/(t+2)$, or continue forever for zero, has no maximizing strategy.

6.2.1 Sequential quantity choice

In a Stackelberg game player 1 chooses a_1, and player 2 observes it before choosing a_2. Subgame perfection requires a best response selection $s_2(a_1) \in \arg\max_{a_2} u_2(a_1, a_2)$ at every a_1, followed by $a_1^* \in \arg\max_{a_1} u_1(a_1, s_2(a_1))$.

For Cournot profits $a_i(1 - a_1 - a_2)$ on $[0, 1]^2$, the follower chooses $(1 - a_1)/2$. The leader maximizes $a_1(1 - a_1)/2$, giving quantities $(1/2, 1/4)$ and profits $(1/8, 1/16)$. The simultaneous game instead gives quantities $(1/3, 1/3)$ and profits $(1/9, 1/9)$.

When the follower has several best responses, maximizing the leader's payoff over all pairs satisfying the follower's best response condition selects the most favorable available response for the leader. Other subgame perfect equilibria can use a less favorable selection. For example, let the follower be indifferent between all outcomes, let the leader obtain one after choosing O, and let choices (I, L) and (I, R) give it two and zero. The follower strategy choosing R after I supports O in a subgame perfect equilibrium, even though maximizing over favorable best responses would choose (I, L).

6.3 Two extensions of the definition of a game

6.3.1 Exogenous uncertainty

Chance moves allow $P(h) = c$ at designated histories, with a probability distribution $f_c(\cdot \mid h)$ on $A(h)$. Chance is specified by the model and has no preferences or strategy. A player strategy profile induces a lottery over terminal histories, so preferences must rank these lotteries. With expected utility, backward induction replaces a chance decision by its expected continuation payoff vector. The one deviation property and existence of a pure subgame perfect equilibrium continue to hold in finite games with perfect information and chance moves.

6.3.2 Simultaneous moves

One can let $P(h)$ be a set of players and set $A(h) = \prod_{i \in P(h)} A_i(h)$. Each active player observes the history but not the other current actions. Histories are then sequences of action profiles. The subgame perfection condition is unchanged, and the finite horizon one deviation argument still applies.

Pure existence can fail because a simultaneous continuation game may have no pure Nash equilibrium. A single simultaneous Matching Pennies stage is already a counterexample. In finite games, allowing mixed choices at each simultaneous stage permits backward induction using a mixed Nash equilibrium of each continuation stage.

> Example 6.9 (Unanimous acceptance by two responders). Player 1 proposes a division of a pie among three players, after which players 2 and 3 simultaneously accept or reject. The division is implemented only if both accept. Even when both receive positive shares, both rejecting is a Nash equilibrium of the response game: a unilateral acceptance does not change the zero outcome. Both accepting is another equilibrium. The simultaneous response structure therefore permits continuation equilibria that would be excluded if each responder made a final effective decision alone.

6.4 The interpretation of a strategy

A full strategy can describe planned behavior together with other players' expectations about behavior after departures from that plan. In the earlier example, a player 1 strategy that chooses B still specifies E or F after (A, C). Player 2 uses this prescription to evaluate C after A, and player 1 in turn uses that anticipated response to evaluate its initial choice.

Under this interpretation, a strategy restriction also restricts beliefs about behavior following deviations. With several opponents, a single strategy for player i represents a common expectation about i at every history. Subgame perfection retains rational continuation play after any observed history, even if that history could only result from actions that the proposed equilibrium would not prescribe. Whether an observer instead revises beliefs about preferences or rationality requires an expanded model of the uncertainty involved.

6.5 Two notable finite horizon games

6.5.1 The chain store game

A chain store faces one possible entrant in each of K cities in sequence. The entrant chooses entry or staying out. Following entry, the chain store fights or accommodates. In a city, the payoffs to the chain store and entrant are $(5, 1)$ from staying out, $(0, 0)$ from entry and fighting, and $(2, 2)$ from entry and accommodation. The chain store sums its payoffs across cities, and everyone observes preceding actions.

In the last city, the chain store strictly prefers accommodation after entry, so the entrant enters. Future play in that city is therefore fixed following every earlier history. The same comparison applies in the preceding city, and induction gives entry and accommodation in every city as the unique subgame perfect equilibrium.

Nash equilibrium admits histories containing either no entry or accommodated entry in each city. A prescription to fight a particular entrant can deter entry even though it would be unprofitable if that entrant appeared. The chain store's equilibrium payoff does not test an unused response. The backward induction prediction follows from known preferences and the finite horizon. A reputation argument in which early fights change beliefs about the chain store's preferences requires those preferences to be uncertain in the model.

6.5.2 The centipede game

Two players alternate between stopping and continuing for T dates, where T is even. Each prefers stopping at its present opportunity to letting the opponent stop at the next opportunity, while continued play beyond both opportunities can benefit both. In the six date example, the successive stopping payoffs are

$$(1, 0),\ (0, 2),\ (3, 1),\ (2, 4),\ (5, 3),\ (4, 6),$$

and continuing to the end gives $(6, 5)$.

The last player prefers to stop rather than continue to the terminal payoff. The preceding player then prefers to stop one date earlier, and so on. The unique subgame perfect strategy profile prescribes stopping after every history and ends at the first date.

Every Nash equilibrium has the same initial stopping outcome. If the first stop were at date $t \geq 2$, the previous mover would gain by stopping at $t - 1$; continuing to the final outcome is also defeated by a last date deviation. To support stopping at date 1, player 2 must prescribe stopping at date 2. Nash equilibrium places no further restriction on later choices, whereas subgame perfection does. Thus equality of equilibrium outcomes does not imply equality of equilibrium strategy sets.

6.6 Iterated elimination of weakly dominated strategies

6.6.1 Relation with subgame perfection

In a finite perfect information game with no indifference between distinct terminal histories, one can reproduce the backward induction outcome by an appropriate sequence of weak dominance deletions in the strategic form. At a final decision, any strategy consistent with reaching that decision and choosing an inferior action is weakly dominated by the otherwise identical strategy choosing the best action. The improvement is strict when the opponents make the decision reachable, and irrelevant when they do not. After these deletions, repeat at earlier decisions.

Every remaining strategy profile then generates the unique backward induction outcome. Some surviving strategies may still prescribe inferior actions at histories that their own earlier choices

make unreachable. Thus reproducing the outcome is weaker than retaining only subgame perfect strategy profiles.

Other deletion orders can remove the subgame perfect equilibrium itself. Consider the strategic form

	C	D
AE	$(2,0)$	$(1,1)$
AF	$(0,2)$	$(1,1)$
BE	$(3,3)$	$(3,3)$
BF	$(3,3)$	$(3,3)$

.

Its extensive form is the three decision example above, with B ending the game. Backward induction gives (BE, D). If AE is deleted first, column C weakly dominates D. Removing D and then AF leaves (BE, C) and (BF, C), neither of which is subgame perfect. Both still produce the backward induction outcome B.

With indifference, even the equivalence of surviving and subgame perfect outcomes can fail. Weak dominance may remove an equilibrium outcome whose support relies on a player's indifference, or preserve an outcome with an inferior continuation that cannot be eliminated strictly anywhere relevant.

6.6.2 Forward induction

Suppose player 1 can take an outside option paying $(2,2)$ or enter Bach or Stravinsky with payoffs $(3,1)$ at (B,B), $(1,3)$ at (S,S), and zero at mismatches. In the reduced strategic form, entering and choosing S is strictly dominated by the outside option. After deleting it, B weakly dominates S for player 2. Finally, entering and choosing B strictly dominates the outside option. The surviving outcome is (B,B).

This reasoning uses the observed refusal of an available payoff to infer what a player intends to do next. If player 1 has voluntarily given up two, player 2 attributes to it a continuation that can earn more than two. That inference selects B and makes coordination on B optimal for player 2.

Example 6.10 (Burning money; Example 111.1). Before the same concert game, player 1 may discard one unit of payoff, observed by player 2, or discard nothing. Write $0B, 0S, DB, DS$ for player 1's reduced strategies and BB, BS, SB, SS for player 2's responses after $0, D$. The payoff matrix is

	BB	BS	SB	SS
$0B$	$(3,1)$	$(3,1)$	$(0,0)$	$(0,0)$
$0S$	$(0,0)$	$(0,0)$	$(1,3)$	$(1,3)$
DB	$(2,1)$	$(-1,0)$	$(2,1)$	$(-1,0)$
DS	$(-1,0)$	$(0,3)$	$(-1,0)$	$(0,3)$

.

One valid sequence deletes DS, then SS, then BS, then $0S$, then SB, then DB. The dominating strategies at those steps are respectively $0B, SB, BB, DB, BB, 0B$. The surviving pair is $(0B, BB)$: no money is burned, and player 1 obtains its preferred coordination outcome.

The inference begins from player 1's guaranteed payoff of 3/4 in the concert game without disposal. Disposal followed by S can give at most zero, so a rational reason to dispose must involve choosing B. Anticipating player 2's response, disposal can then yield two. Choosing no disposal

is subsequently interpreted as an intention to choose B, since S yields at most one. The relevant reasoning involves alternatives that are never taken in the surviving outcome.

A costless message has a different formal role. If player 1 may send any message before the concert game and payoffs do not depend on its content, there are subgame perfect equilibria in which player 2 ignores every message and the players coordinate on either pure equilibrium. The ordinary game form does not assign a binding interpretation to the wording of a message.

7 Bargaining Games

7.1 Bargaining and game theory

Bargaining arises when several agreements are better for both parties than disagreement, but the parties disagree about which agreement to choose. An extensive game describes the offers, responses, and opportunities to end negotiation. Its equilibrium division depends on those rules and on the parties' preferences over agreement and delay.

The alternating offers model gives each party an opportunity to propose and to reject. Rejection postpones agreement and changes who makes the next proposal. A responder therefore compares an available agreement with the continuation payoff from becoming the proposer. That comparison determines how much the current proposer must offer.

7.2 A bargaining game of alternating offers

Definition 7.1 (Alternating offers; the book's Section 7.2). Let X be a compact connected subset of a finite dimensional Euclidean space whose elements are agreements. In period 0, player 1 proposes $x \in X$, and player 2 accepts or rejects. Acceptance implements x immediately. Rejection leads to period 1, when player 2 proposes and player 1 responds. The players alternate in this way without a final date. If no proposal is accepted, the outcome is disagreement D.

A finite agreement outcome is denoted (x, t), where t is the acceptance date. Preferences depend on the accepted agreement and its date, rather than the preceding sequence of rejected proposals. Every agreement is weakly preferred to disagreement. Delay weakly worsens an agreement and strictly worsens one that is strictly better than disagreement. Preferences are continuous in agreements and stationary: shifting both dates by the same amount leaves their comparison unchanged.

Under the representation discussed in the book's Section 7.2, preferences over dated agreements can be written as $\delta_i^t w_i(x)$ for a chosen $\delta_i \in (0,1)$ and a continuous nonnegative function w_i. For a division of a unit pie,

$$X = \{(x_1, x_2) \in \mathbb{R}_+^2 : x_1 + x_2 = 1\},$$

one can take $w_i(x) = w_i(x_i)$, where w_i is increasing and continuous and $w_i(0) = 0$. Disagreement then gives zero.

A discount factor is meaningful relative to the utility index used with it. If $\delta^t u(x)$ represents preferences and $u \geq 0$, the same preferences are represented by $\epsilon^t u(x)^{\log \epsilon / \log \delta}$ for another $\epsilon \in (0,1)$. Comparing discount factors alone establishes a patience comparison when the instantaneous utility function is held fixed.

Nash equilibrium allows every immediately accepted division. For any x^*, let both players always offer x^* and accept only x^*. Each obtains x^* by following the strategy, and an alternative agreement cannot be obtained unilaterally. In a pie division game, delayed agreement can likewise be supported by prescribed early rejections. These strategies can require a responder to reject a current share that exceeds the discounted share it will receive after rejection. Subgame perfection tests that comparison.

7.3 Subgame perfect equilibrium

7.3.1 Characterization for general agreements

The uniqueness result uses four additional assumptions. First, distinct agreements do not give the same utility to both players. Second, the best agreement b_i for player i leaves the other player indifferent between receiving it now, receiving it one period later, and disagreement. Third, the Pareto frontier is strictly monotone: if x is efficient, no different agreement makes both players weakly better off. Fourth, there is a unique pair of efficient agreements (x^*, y^*) satisfying

$$(x^*, 1) \sim_1 (y^*, 0), \qquad (y^*, 1) \sim_2 (x^*, 0).$$

These are assumptions A1–A4 of the book's Section 7.3. The last condition is a uniqueness condition on the two responders' indifference comparisons.

> Theorem 7.2 (Alternating offers equilibrium; Proposition 122.1). Under A1–A4 and the preference assumptions above, the game has an essentially unique subgame perfect equilibrium. Player 1 proposes x^* whenever it proposes, and player 2 proposes y^*. Player 1 accepts y^* and every proposal it strictly prefers to y^*, and rejects every proposal it strictly dislikes relative to y^*. Player 2 uses the corresponding rule with threshold x^*. Agreement occurs immediately at every proposal subgame. If every agreement is efficient, the equilibrium is unique.

The possible indeterminacy concerns an inefficient proposal that a responder values exactly as much as its equilibrium threshold, while the proposer values it less. Responses to such proposals can differ without changing equilibrium offers or outcomes.

Existence argument. After player 2 rejects an offer, the prescribed continuation gives y^* one period later. The indifference condition and stationarity make its acceptance threshold exactly its payoff from x^* now. The same argument applies to player 1. Among accepted offers, efficiency and strict monotonicity make x^* optimal for player 1 and y^* optimal for player 2. An offer that is rejected delays the proposer until the other player's offer, which is worse than its own immediate equilibrium offer. Thus no single deviation is profitable after any history. Discounted continuous preferences on compact X make arbitrarily remote continuation payoffs uniformly small, so a profitable deviation would have a profitable finite approximation. The finite deviation argument then yields a profitable single deviation, proving subgame perfection.

Why the result covers nonstationary strategies. Stationarity of preferences makes all subgames beginning with the same proposer strategically equivalent. Let G_i denote such a subgame, and let $M_i(G_i)$ and $m_i(G_i)$ be the supremum and infimum of player i's payoffs across all its subgame perfect equilibria. Use a common discount factor δ in the preference representations and write the efficient utility frontier as $u_2 = \phi(u_1)$, with ϕ continuous and decreasing. The responder's ability to reject and the proposer's ability to offer a slightly larger current payoff imply

$$m_2(G_2) \geq \phi(\delta M_1(G_1)), \qquad M_1(G_1) \leq \phi^{-1}(\delta m_2(G_2)).$$

The first inequality follows by offering player 1 just more than its largest possible discounted continuation payoff. The second follows because player 2 can secure at least its smallest discounted continuation payoff by rejection.

Combining gives

$$M_1(G_1) \le \phi^{-1}\big(\delta\phi(\delta M_1(G_1))\big).$$

The equality version is exactly the pair of indifference equations. At player 1's best agreement the inequality is reversed strictly. If the supremum exceeded $u_1(x^*)$, continuity would produce a second solution of the indifference equations between that supremum and the best agreement, contradicting A4. The analogous comparisons for the infima and the other proposer give

$$m_1(G_1) = M_1(G_1) = u_1(x^*), \qquad m_2(G_2) = M_2(G_2) = u_2(y^*).$$

Efficiency and strict impatience then exclude delay, and the acceptance thresholds determine the proposals. Stationary equilibrium offers are a conclusion of the argument, not a restriction imposed on strategies.

7.3.2 Linear utility for a pie

Suppose agreement (x, t) gives player i payoff $\delta_i^t x_i$, where $0 < \delta_i < 1$. Let player 1 demand a when proposing and player 2 demand b when proposing. The responder indifference conditions are

$$1 - b = \delta_1 a, \qquad 1 - a = \delta_2 b.$$

Their unique solution is

$$a = \frac{1-\delta_2}{1-\delta_1\delta_2}, \qquad b = \frac{1-\delta_1}{1-\delta_1\delta_2}.$$

The initial division is therefore

$$x^* = \left(\frac{1-\delta_2}{1-\delta_1\delta_2}, \frac{\delta_2(1-\delta_1)}{1-\delta_1\delta_2}\right),$$

and the division following a rejection and a player 2 proposal is $y^* = (\delta_1 a, b)$.

A direct uniqueness proof. Let m_1, M_1 bound player 1's equilibrium payoff when it proposes, and m_2, M_2 bound player 2's when it proposes. A responder's rejection guarantee and a proposer's ability to offer slightly more than the largest continuation payoff imply

$$1 - \delta_2 M_2 \le m_1 \le M_1 \le 1 - \delta_2 m_2,$$

with the symmetric inequalities for player 2. Consequently,

$$M_1 - m_1 \le \delta_2(M_2 - m_2) \le \delta_1\delta_2(M_1 - m_1).$$

Since $\delta_1\delta_2 < 1$, both payoff intervals have zero length. The equalities above then determine a, b. At the initial subgame player 2 can obtain at least $\delta_2 b = 1 - a$, so the sum of equilibrium payoffs is at least one. Any delayed agreement has discounted total payoff below one. Agreement must therefore be immediate, and the unique feasible division is $(a, 1-a)$.

> Example 7.3 (Numerical division and acceptance). If $\delta_1 = 0.9$ and $\delta_2 = 0.8$, then $a = 5/7$ and $b = 5/14$. Player 1 offers $(5/7, 2/7)$; player 2 is indifferent between accepting $2/7$ now and proposing for $5/14$ next period, since $0.8(5/14) = 2/7$. Player 2 offers $(9/14, 5/14)$ in its proposal subgames, and player 1's threshold is $0.9(5/7) = 9/14$.

An acceptance rule must include the equilibrium offer. If the responder rejected exactly its threshold but accepted every strictly higher amount, the proposer could approach its best payoff with ever smaller concessions without attaining it. In the divisible pie game this cannot be an equilibrium.

7.3.3 Properties and comparative statics

With equal discount factors δ, the first proposer gets $1/(1+\delta) > 1/2$ and the responder gets $\delta/(1+\delta)$. The first proposer avoids the delay required to exercise the other proposal opportunity. As $\delta \uparrow 1$, this first mover advantage vanishes.

Holding the linear utility indices fixed,

$$\frac{\partial a}{\partial \delta_1} = \frac{\delta_2(1-\delta_2)}{(1-\delta_1\delta_2)^2} > 0, \qquad \frac{\partial a}{\partial \delta_2} = -\frac{1-\delta_1}{(1-\delta_1\delta_2)^2} < 0.$$

Greater patience improves a player's share. More generally, a preference relation is at least as impatient as another if they rank immediate agreements identically and every delayed agreement formerly indifferent to a current agreement is now weakly worse. Proposition 126.1 shows that making one player more impatient in this sense weakly worsens that player's equilibrium agreement, provided both games satisfy A1–A4.

As a useful time comparison, let offers be separated by Δ units of time and put $\delta_i = e^{-r_i\Delta}$. Then

$$\lim_{\Delta \downarrow 0} a = \frac{r_2}{r_1 + r_2}.$$

Frequent offers remove the advantage from the initial proposal but retain the division determined by relative rates of impatience.

Fixed costs of delay describe a different preference specification. If utility is $x_i - c_i t$, even an agreement giving zero pie becomes worse when delayed. This violates the endpoint indifference assumption A2. The two interior indifference equations would require the transfers in both directions to equal the respective costs, which has no solution when $c_1 \neq c_2$ and can have many solutions when $c_1 = c_2$. The discounted uniqueness argument therefore does not apply to this specification.

7.4 Variations and extensions

7.4.1 The bargaining procedure

If player 1 makes every offer, a rejection no longer gives player 2 a turn as proposer. Under A1–A3, the equilibrium implements player 1's best agreement b_1 immediately. In the linear pie case, if M_2 is the largest equilibrium payoff player 2 can obtain at the start of such a subgame, player 1 can offer just above $\delta_2 M_2$ and retain the rest. This gives $M_2 \leq \delta_2 M_2$, so $M_2 = 0$.

If players instead make simultaneous proposals each period and agreement requires compatibility, every efficient division can be sustained in a subgame perfect equilibrium. Prescribe compatible proposals for that division after every history. A unilateral change produces disagreement and delay, so it cannot improve on the prescribed current agreement. The power to offer the other player an immediately acceptable alternative is central to the alternating offers result.

7.4.2 Indivisible units

If the pie is divided only in multiples of $\epsilon > 0$, the proposer cannot always improve an offer by an arbitrarily small amount. With sufficiently patient players, every feasible division can be supported as an immediate subgame perfect outcome, and suitable continuation equilibria can also support delayed agreement or perpetual disagreement. For fixed discount factors, however, making the unit sufficiently small brings equilibrium payoffs close to those of the divisible pie game and restores immediate agreement, as examined in Exercise 128.1. The order of the limits in patience and divisibility matters.

7.4.3 Opting out

Suppose only player 2 can opt out, and only when responding to an offer. For common discount factor δ, let its outside option pay $b < 1$ at that date, and let player 1 receive zero if it is exercised. The usual first period share of player 2 is $\delta/(1+\delta)$.

If $b \leq \delta/(1+\delta)$, the outside option does not improve player 2's acceptance threshold: rejecting and bargaining next period is at least as good. If $b > \delta/(1+\delta)$, equilibrium proposals are

$$\widehat{x} = (1-b, b), \qquad \widehat{y} = (\delta(1-b), 1-\delta(1-b)).$$

Player 2 accepts b because its outside option pays b, and its discounted payoff from rejection is no greater:

$$\delta[1-\delta(1-b)] \leq b \quad \Longleftrightarrow \quad b \geq \frac{\delta}{1+\delta}.$$

Player 1's response threshold is $\delta(1-b)$. Thus a binding outside option determines the responder's share; a lower one has no effect. The timing of the opt out opportunity is part of this result. Allowing it at other points creates different continuation choices and can change the equilibrium set.

7.4.4 Risk of breakdown

Suppose players do not discount time, but after each rejection bargaining ends with probability $\alpha \in (0,1)$. Normalize the von Neumann–Morgenstern utility of breakdown to zero. The responder now compares a certain agreement with a lottery giving next period's agreement with probability $1-\alpha$ and breakdown otherwise. The indifference equations become

$$u_1(y^*) = (1-\alpha)u_1(x^*), \qquad u_2(x^*) = (1-\alpha)u_2(y^*).$$

When they have a unique efficient solution, the same offer and acceptance logic gives immediate agreement. Linear utility for pie makes the effective discount factor $1-\alpha$, so player 1 receives $1/(2-\alpha)$. With nonlinear utility for shares, the equations use utility rather than pie quantities, and risk attitudes affect the resulting division.

7.4.5 More than two players

Consider three players dividing a pie, with common discount factor δ. Player 1 proposes first, followed by player 2 and then player 3 in subsequent periods after rejection. Each proposal requires the acceptance of both other players in sequence. For $\delta \geq 1/2$, every division can be an immediately accepted subgame perfect outcome.

One construction uses a continuation division y, initially the desired division x. Players propose y and accept offers giving at least δy_i. If proposer i demands more than y_i, choose another player j whose offered share is below $1/2$ and change the continuation division to the unit allocation e_j. That responder can obtain the whole pie after rejection, worth $\delta \geq 1/2$ now, so rejects the deviating offer. The continuation rewards a responder for rejecting a proposal that would benefit the deviating proposer. The prescribed division is consequently accepted immediately, and no proposer gains by changing it.

Restricting attention to stationary strategies gives the particular initial division

$$\left(\frac{1}{1+\delta+\delta^2}, \frac{\delta}{1+\delta+\delta^2}, \frac{\delta^2}{1+\delta+\delta^2} \right).$$

The shares reflect the discounted proposal opportunities. In the unrestricted game, continuation divisions can depend on earlier proposals, producing the much larger equilibrium set. Unlike the two player theorem, stationarity here restricts the class of strategies being considered.

8 Repeated Games

8.1 Repeated interaction and the horizon

A repeated game allows current behavior to affect future play. A player may therefore prefer an action that gives a lower current payoff when it preserves a valuable continuation. The essential incentive constraint compares the gain from deviation with the resulting change in continuation payoffs. Subgame perfection adds a second requirement: the players who impose a punishment must themselves prefer to carry it out.

Example 8.1 (The Prisoner's Dilemma in the book's Section 8.1). The stage payoffs are

	C	D
C	$(3,3)$	$(0,4)$
D	$(4,0)$	$(1,1)$

.

Suppose each player chooses C until a deviation occurs and then chooses D forever. Cooperation gives $3/(1-\delta)$, while a deviation gives $4+\delta/(1-\delta)$. Cooperation is optimal precisely when

$$3 \geq 4(1-\delta)+\delta, \qquad \text{or} \qquad \delta \geq \frac{1}{3}.$$

The punishment is credible because (D,D) is a stage Nash equilibrium. These strategies consequently form a subgame perfect equilibrium for these discount factors.

A known final period changes the argument. In a finite repetition of this game, defection is optimal in the last period. Since the final payoff is then independent of earlier behavior, the same reasoning applies to the preceding period and continues backward. An infinite horizon instead describes a situation in which, after every finite history, further interaction remains possible. The book's Section 8.2 emphasizes that the relevant horizon is part of the players' strategic situation. The mere physical finiteness of a relationship does not specify how its end enters their decisions.

8.2 Histories, strategies, and evaluations of payoff streams

Let $G = \langle N, (A_i), (u_i) \rangle$, with compact action sets and continuous payoffs, and write $A = \prod_i A_i$. The public history before period t is $h^t = (a^1, \ldots, a^{t-1}) \in A^{t-1}$. Every player observes all past actions and chooses simultaneously in each period. A pure strategy is a function

$$s_i : \bigcup_{t \geq 0} A^t \longrightarrow A_i.$$

The function specifies behavior after histories inconsistent with the player's own earlier strategy as well as after histories generated by that strategy.

Definition 8.2 (Payoff criteria; the book's Section 8.3). For bounded payoff streams $v = (v^t)$ and $w = (w^t)$, the three criteria are:

$$\text{discounting:} \quad U_\delta(v) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} v^t, \quad 0 < \delta < 1;$$

$$\text{limit of means:} \quad v \succ w \iff \liminf_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T}(v^t - w^t) > 0;$$

$$\text{overtaking:} \quad v \succ w \iff \liminf_{T\to\infty} \sum_{t=1}^{T}(v^t - w^t) > 0.$$

When the average payoff converges, its limit is the payoff under the limit of means criterion. The definitions using differences remain meaningful when the individual averages do not converge.

A finite change in payoffs has no effect under the limit of means criterion. Overtaking retains such changes: $(-1, 2, 0, \ldots)$ is better than $(0, 0, \ldots)$ under overtaking but has the same limiting average. Discounting also values the timing of a change; $(1, -1, 0, \ldots)$ has positive discounted value and zero eventual cumulative gain. These distinctions explain why the punishment constructions below differ.

Definition 8.3 (Feasibility and enforceability). Following the book, a feasible payoff w is a rational convex combination of stage payoff vectors:

$$w = \sum_{a\in A} \alpha_a u(a), \qquad \alpha_a \in \mathbb{Q}_+, \quad \sum_a \alpha_a = 1,$$

with finitely many nonzero coefficients. It can be generated by a finite cycle. Player i's pure minmax payoff is

$$v_i = \min_{a_{-i}\in A_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i}).$$

The payoff w is enforceable if $w_i \geq v_i$ for every i, and strictly enforceable if all these inequalities are strict. Let p_{-i} attain the minimum and let $b_i(a_{-i})$ denote a best response.

The restriction to pure actions applies to punishments as well as equilibrium play. Allowing mixed punishments can change the minmax values. Under a common discount factor, all players attach the same weights to the sequence of stage payoff vectors. With different discount factors their normalized payoffs need not lie in the same convex hull.

8.3 Strategies represented by machines

A machine for player i is $(Q_i, q_i^0, f_i, \tau_i)$, where Q_i is a state set, q_i^0 is the initial state, $f_i : Q_i \to A_i$ is the output function, and $\tau_i : Q_i \times A \to Q_i$ is the transition function. It chooses $a_i^t = f_i(q_i^t)$ and updates to $q_i^{t+1} = \tau_i(q_i^t, a^t)$. Dependence on the full action profile makes the associated strategy well defined after a deviation by the player himself.

The grim strategy uses two states: a cooperative state with output C and an absorbing punishment state with output D. A punishment lasting exactly L periods generally requires states recording how many punishment periods remain. If all machines are finite and deterministic, the joint state eventually repeats. Subsequent play is periodic. A finite machine can therefore generate an initial sequence followed by a cycle, but cannot represent every history dependent strategy in an infinite game.

8.4 Nash folk theorems and permanent punishment

Proposition 8.4 (Propositions 144.1 and 144.3). Every Nash equilibrium payoff under discounting, or under the limit of means when the payoff exists, is enforceable. Conversely, every feasible enforceable payoff is a Nash equilibrium payoff under the limit of means criterion.

Proof. Against a fixed pure strategy profile s_{-i}, player i can choose $b_i(s_{-i}(h))$ after each history and obtain at least v_i each period. Thus an equilibrium cannot give him less. For the converse, choose a finite cycle with average w. Follow this cycle unless a single player deviates; after the first such deviation by i, every other player uses p_{-i} forever. A deviator gets at most v_i in all subsequent periods. His gain before punishment has zero limiting average, and $v_i \le w_i$.

Proposition 8.5 (Discounted Nash folk theorem, Proposition 145.2). For every strictly enforceable feasible payoff w and every $\varepsilon > 0$, there is $\bar{\delta} < 1$ such that each $\delta > \bar{\delta}$ admits a Nash equilibrium with payoff within ε of w.

For a constant target a^*, permanent minmax punishment deters a deviation when

$$(1-\delta)\big[\max_{a_i} u_i(a_i, a_{-i}^*) - u_i(a^*)\big] \le \delta\big[u_i(a^*) - v_i\big].$$

The right side stays positive as δ approaches one, whereas the left side vanishes. For a cycle, the discounted payoff from every position approaches its cycle average; there are finitely many positions, so one discount factor bound works for all of them. Permanent minmax punishment need not be optimal for its participants. The Nash construction therefore does not by itself establish subgame perfection.

8.5 Finite punishments under the limit of means

Theorem 8.6 (Proposition 146.2). Every feasible strictly enforceable payoff is a subgame perfect equilibrium payoff under the limit of means criterion.

Let the target cycle have length γ and average w, and let g bound the gain from changing one action at any stage profile. Choose a positive integer m, divisible by γ, such that

$$\gamma g \leq m(w_i - v_i) \qquad \text{for every } i.$$

When a player first deviates during an unpunished cycle, record his identity, complete that cycle, and then minmax him for m periods. Afterward restart the target cycle. During punishment, the punished player best responds to p_{-i}; the punishment ends at its appointed time even if an action differs from its prescription.

Why the construction works. Any fixed initial punishment disappears from limiting averages. If a player deviates only finitely often, his continuation has the same average as the target. If he deviates repeatedly, each affected cycle gives at most γg extra payoff, followed by a loss of at least $m(w_i - v_i)$ relative to the target average. Incomplete initial and final blocks have uniformly bounded lengths and do not create a strictly positive limiting average gain. A punisher can change his actions during a finite punishment without improving his limiting average. The same argument applies after every history.

8.6 Punishment under overtaking

Theorem 8.7 (Proposition 149.1). Every strictly enforceable stage outcome a^* can be repeated forever in a subgame perfect equilibrium under the overtaking criterion.

Here a punisher's current sacrifice matters. If the prescribed punishment ends regardless of his behavior, he may gain by refusing to participate. Introduce a normal state and states $P(j,t)$ specifying that player j has t punishment periods remaining. Normal play is a^*; in $P(j,t)$ the others choose p_{-j} and player j best responds.

Let M bound all stage payoffs. A deviation by j in the normal state starts a punishment long enough that

$$M + tv_j < (t+1)u_j(a^*).$$

A deviation by a single punisher $i \neq j$ replaces the existing punishment with a punishment of i. Its length is chosen to make the cumulative payoff from deviation strictly lower than the cumulative payoff from completing the prescribed punishment and returning to a^*. Such a length exists because $u_i(a^*) > v_i$. Deviations by the currently punished player cannot improve his stage payoff and do not postpone the return to normal play. Punishment lengths can increase after further refusals to punish. The construction consequently allows an unbounded set of states.

8.7 Discounting and rewards for punishers

Unbounded punishment lengths create a difficulty under discounting: rewards arbitrarily far in the future may have too little value. Proposition 151.1 uses a fixed punishment length and continuation outcomes that distinguish the players.

> Theorem 8.8 (Proposition 151.1). Let a^* be strictly enforceable. Suppose there are strictly enforceable outcomes $a(i)$, one for each player, satisfying
>
> $$u_i(a^*) > u_i(a(i)), \qquad u_i(a(j)) > u_i(a(i)) \quad (j \neq i).$$
>
> For every sufficiently large common discount factor, repeating a^* forever is a subgame perfect equilibrium outcome.

Use continuation states $C(0), C(1), \ldots, C(n)$ with outputs $a^*, a(1), \ldots, a(n)$, and punishment states $P(i,t)$ for $1 \leq t \leq L$. A deviation by i in a continuation state starts L periods of minmax punishment, followed by $C(i)$. A deviation by a punisher starts a punishment of that punisher. The punished player best responds and cannot affect the punishment's remaining length.

First choose L so that, writing $a(0) = a^*$,

$$M - u_i(a(j)) < L\big[u_i(a(j)) - v_i\big] \quad \text{for all } i, j.$$

For δ close enough to one, the discounted counterpart also holds, and deviation from a continuation state is unprofitable. A punisher $i \neq j$ who complies reaches $a(j)$ after at most L periods. Refusal leads instead to $a(i)$ after punishment. The strictly positive difference $u_i(a(j)) - u_i(a(i))$ generates a continuation gain of order $1/(1-\delta)$, which outweighs the bounded cost of carrying out L punishment periods. This establishes incentives in every state.

8.8 One deviation and optimal punishments

> Lemma 8.9 (Lemma 153.1). In a discounted repeated game with bounded stage payoffs, a strategy profile is subgame perfect if and only if no player can gain by changing his action for one period after any history and then following the profile.

Proof. Necessity is immediate. If a profitable unrestricted deviation exists, truncate it sufficiently far in the future that its discounted tail has smaller value than the gain. The truncated deviation is still profitable and changes only finitely many actions. At its last changed action, a profitable improvement must be possible unless that action can be replaced by the prescribed action without loss. Repeating this replacement backward contradicts the absence of a profitable single deviation.

Under the compactness and continuity assumptions, the set of subgame perfect equilibrium payoffs is closed (Lemma 153.3). Thus, when this set is nonempty, each player's lowest equilibrium payoff is attained. Proposition 154.1 uses these worst equilibria to support any given equilibrium path: follow that path until a single deviation by i, then start an equilibrium giving i his lowest possible equilibrium payoff. Any later deviation by j starts the corresponding worst equilibrium for j. A player's punishment continuation depends on his identity; the earlier history need not affect it. Since these continuations are no better for a deviator than the original equilibrium's continuations, the one deviation inequalities remain satisfied.

8.9 Finite repetition

The T period game assigns payoff $T^{-1}\sum_{t=1}^{T} u_i(a^t)$. Its final equilibrium action profile must be a stage Nash equilibrium. Two distinct restrictions follow.

> Proposition 8.10 (Propositions 155.1 and 157.2). If every stage Nash equilibrium gives every player his minmax payoff, every Nash equilibrium path of every finite repetition consists of stage Nash equilibria. If the stage game has a unique Nash equilibrium payoff vector, the action profile after every history in every subgame perfect equilibrium is a stage Nash equilibrium.

For the first claim, consider the last period on an equilibrium path in which play is not a stage equilibrium. A player can gain in that period and thereafter guarantee his minmax payoff, losing nothing relative to the remaining path. For the second, all final subgames give the same payoff vector. Backward induction then makes continuation payoffs independent of the current action at each preceding date.

> Proposition 8.11 (Finite Nash folk theorem, Proposition 156.1). Suppose the stage game has a Nash equilibrium $\hat{a}$ with $u_i(\hat{a}) > v_i$ for every player. For any strictly enforceable stage outcome a^* and any $\varepsilon > 0$, every sufficiently long finite repetition has a Nash equilibrium whose average payoff lies within ε of $u(a^*)$.

Play a^* for $T - L$ periods and $\hat{a}$ for the last L, with permanent minmax punishment following an earlier deviation. Choose L independently of T so that

$$\max_{a_i} u_i(a_i, a^*_{-i}) - u_i(a^*) \le L\big[u_i(\hat{a}) - v_i\big].$$

The final block deters deviation; its contribution to the average vanishes as T grows. For subgame perfection, the terminal rewards and punishments must themselves be equilibria. In the modified Prisoner's Dilemma of the book's Section 8.10, an additional action E gives $(1/2, 1/2)$ when both use it and zero otherwise. Both (D, D) and (E, E) are equilibria. Cooperation through $T - 3$, followed by three periods of (D, D), is sustainable by switching permanently to (E, E) after a deviation: the current gain of one is smaller than the terminal loss $3(1 - 1/2)$.

Proposition 160.1 gives a finite perfect folk theorem when each player has two stage Nash equilibria with different payoffs and the distinct continuation outcomes $a(i)$ satisfy the inequalities in Proposition 151.1. The construction has a long target phase, a fixed period for completing punishments and rewarding punishers, and a terminal sequence of stage equilibria. The latter two lengths do not grow with T, so the average converges to the target payoff.

9 Complexity Considerations in Repeated Games

9.1 Complexity as a feature of preferences

The equilibrium strategies of a repeated game may require states used only to carry out threats. Chapter 9 asks what happens when players also prefer simpler strategies. Complexity is an object of choice: each player selects a finite machine and compares both its payoff and its number of states. This differs from placing a fixed bound on the memory available to a player. The player can choose a larger machine, but must find its additional complexity worthwhile.

The model concerns two players in a discounted repeated game with a finite stage game. The two player restriction matters because a player can reproduce the state of his single opponent's machine. With several opponents, reproducing their joint state may require more states than any one opponent uses.

9.2 The machine game

Definition 9.1 (Definition 165.1). A finite machine of player i is $M_i = (Q_i, q_i^0, f_i, \tau_i)$, where $f_i : Q_i \to A_i$ and $\tau_i : Q_i \times A_j \to Q_i$ for $j \neq i$. Its complexity is $c(M_i) = |Q_i|$. The pair (M_1, M_2) generates

$$q_i^1 = q_i^0, \qquad a_i^t = f_i(q_i^t), \qquad q_i^{t+1} = \tau_i(q_i^t, a_j^t)$$

and discounted payoff

$$U_i(M_1, M_2) = (1-\delta)\sum_{t\geq 1} \delta^{t-1} u_i(a^t).$$

In the machine game, each player's action is a finite machine. Preferences are strictly increasing in U_i holding complexity fixed and strictly decreasing in complexity holding U_i fixed.

The transition function observes only the other player's action. It describes a plan executed by the machine, whose own output is already determined by its state. Chapter 8 instead used transitions depending on the full action profile to represent an extensive form strategy after all histories.

Two useful specifications of preferences are

$$U_i(M_1, M_2) - \gamma_i c(M_i), \qquad \gamma_i > 0,$$

and lexicographic preferences that compare U_i first and use complexity only to break payoff ties. Under lexicographic preferences, even an arbitrarily small strict payoff improvement justifies any finite increase in complexity. Yet a reduction in complexity is decisive when it preserves the payoff exactly.

Example 9.2 (Why grim punishment fails, Example 166.1). In this chapter the Prisoner's Dilemma has payoffs $(3,3)$ at (C,C), $(1,1)$ at (D,D), and temptation payoff 5 against C, with sucker payoff zero. Two grim machines form an equilibrium of the repeated game when $\delta \geq 1/2$. They do not form an equilibrium of the machine game. Against a grim opponent, a one state machine that always chooses C gives the same payoff of three as the two state grim machine. It saves a state. Each player therefore has a profitable change of machine.

The argument compares unilateral machine choices while the other machine stays fixed. After the change, the opponent might have a profitable deviation of his own, but this does not remove the first player's profitable deviation from the proposed equilibrium.

Example 9.3 (An initial punishment phase, Example 166.2). Use states P_1, P_2, P_3, R, initially P_1, with output D in the first three and C in R. A player observing D advances through the three punishment states and then enters R. Observing C during punishment restarts P_1; observing D in R also restarts P_1, whereas observing C leaves R unchanged. Two such machines play three periods of (D,D), then cooperate.

Defecting in R gains two immediately and causes at least three periods of punishment. The inequality

$$5 + \delta + \delta^2 + \delta^3 < 3 + 3\delta + 3\delta^2 + 3\delta^3$$

holds for δ sufficiently close to one. Deviating during the initial phase cannot accelerate its end or increase its stage payoff. Moreover, obtaining the same payoff requires playing D through the three punishment dates and C at the first cooperative date. These four different positions require four states, since the opponent has supplied the same input D at the first three transitions. Thus the pair is an equilibrium under lexicographic preferences for sufficiently patient players.

9.3 Every state must be used

Lemma 9.4 (Lemma 168.1). In a Nash equilibrium of the machine game, every state of each player's machine is reached at some date on the induced path.

Proof. Delete any state never reached on the path and redirect transitions into it arbitrarily to retained states. The induced path against the fixed opponent does not change, so payoff stays the same and complexity falls.

Consequently, a punishment state retained only for a deviation cannot be justified by its deterrent effect. In an equilibrium where punishment requires separate states, those states also have to occur during equilibrium play. The four state example satisfies this condition because the punishment states occur at the beginning.

9.4 Best responses and equal machine sizes

Lemma 9.5 (Lemma 168.2). In an equilibrium (M_1^*, M_2^*) of the two player machine game,

$$c(M_1^*) = c(M_2^*).$$

The corresponding strategies also form a Nash equilibrium of the discounted repeated game.

Proof. Fix $M_i = (Q_i, q_i^0, f_i, \tau_i)$. Player j faces a discounted control problem with state $q \in Q_i$. Using unnormalized continuation payoffs, its Bellman equation is

$$V_j(q) = \max_{a_j \in A_j} \{u_j(f_i(q), a_j) + \delta V_j(\tau_i(q, a_j))\}.$$

Choose an optimal stationary action $a_j^*(q)$ at every state. A machine for player j can record the opponent's state, choose $a_j^*(q)$, and update to $\tau_i(q, a_j^*(q))$. It uses at most $|Q_i|$ states and attains the best repeated game payoff against M_i.

If M_j^* had more states than M_i^*, this response would weakly raise payoff and strictly lower complexity. If the payoff improvement were strict, add unused states to the response until it has $c(M_j^*)$ states. The enlarged response is strictly preferred to M_j^* at the same complexity, and removing the added states is strictly preferred at the same payoff. If the payoffs are equal, the smaller response is directly preferred. Hence $c(M_j^*) \leq c(M_i^*)$; interchanging the players gives equality. A failure to attain the best repeated game payoff would then permit a strict payoff increase at the same complexity.

This argument places a strong restriction on an equilibrium even when the tradeoff between payoff and complexity has not been specified numerically. It uses only the two separate monotonicity requirements.

9.5 Actions, states, and the structure of the path

Lemma 9.6 (Lemma 170.1). On an equilibrium path of a two player machine game, actions are paired uniquely: for any dates s, t,

$$a_i^s = a_i^t \quad \Longrightarrow \quad a_j^s = a_j^t.$$

The proof uses an economical representation of an optimal response. If the opponent takes different actions in two of his states while the player takes the same action, that player can merge the corresponding two states of his response machine. The opponent's observed action then identifies which transition is required. The merged machine still implements an optimal response but uses one less state, contradicting Lemma 168.2.

This restriction concerns actions, not only payoffs. In a game with two actions per player, an equilibrium path that uses more than one outcome must pair the two actions bijectively. In the Prisoner's Dilemma its outcomes are therefore contained either in

$$\{(C, C), (D, D)\} \qquad \text{or in} \qquad \{(C, D), (D, C)\}.$$

A cycle containing $(C, C), (C, D), (D, C)$ cannot occur in an equilibrium of this machine game even if a profile of repeated game strategies can support it.

Proposition 9.7 (Proposition 171.1). For an equilibrium pair of machines there are integers t^* and $\ell < t^*$ such that, for each player, the states visited before date t^* are distinct and

$$q_i^t = q_i^{t-\ell} \qquad \text{for all } t \geq t^*.$$

The initial states and the states used in the recurring cycle are disjoint; each cycle state occurs exactly once in a cycle.

Proof idea. Take the first date on which either machine repeats a state. Its action repeats, so Lemma 170.1 implies that the other player's action repeats as well. The transition rules then reproduce the first machine's future state sequence. Every state must be used, and both machines have the same size, so the other machine must also repeat a state at this date. If the two repetitions referred back to different dates, one player's intervening state could be removed while preserving the path. Thus both machines repeat in the same period and remain synchronized thereafter.

The machine game is a strategic game in which machines are selected once. If equilibrium were additionally required after every repeated game history when players could replace their machines, an initial phase would be impossible: once the cycle begins, deleting the initial states reduces complexity without changing any future outcome. The stronger requirement therefore removes many of the equilibria described above.

9.6 Lexicographic preferences in the Prisoner's Dilemma

With lexicographic preferences, the action pairing restriction still leaves two distinct types of equilibrium. For sufficiently patient players, any prescribed positive rational frequency of cooperation can occur in a cycle containing only (C, C) and (D, D). A path containing (C, C) must also have an initial phase; the states that make cooperation enforceable cannot all be confined to an otherwise unused punishment.

A second type alternates the asymmetric outcomes. The book's Section 9.4 shows that for positive integers n_1, n_2 satisfying

$$\frac{5n_i}{n_1 + n_2} > 1 \qquad (i = 1, 2),$$

a sufficiently large discount factor supports a cycle with n_1 occurrences of (D, C) followed by n_2 occurrences of (C, D), without an initial phase. The inequalities make each player's average payoff exceed his minmax payoff.

Example 9.8 (Alternation with two states). Give each player states D and C, with corresponding outputs. In state D, remain there if the opponent chooses D and move to C if the opponent chooses C. In state C, move to D after either action. Player 1 starts in D and player 2 in C. The outcome alternates between (D, C) and (C, D).

When the opponent is in D, complying with the prescribed action C gives the unnormalized payoff

$$V_C = \frac{5\delta}{1-\delta^2}.$$

Defecting gives one immediately and leaves the opponent in D, so the relevant inequality is $V_C \geq 1 + \delta V_C$, or

$$\frac{5\delta}{1+\delta} \geq 1 \quad \Longleftrightarrow \quad \delta \geq \frac{1}{4}.$$

When the opponent is in C, choosing D is optimal and the opponent moves to D irrespective of that action. For $\delta > 1/4$, each player strictly prefers the alternating continuation to a one state response: always D eventually gives one per period, while always C gives only the cooperation payoff against C and zero against D. The two state response is optimal and no one state machine attains its payoff. The pair is therefore a machine game equilibrium under lexicographic preferences.

The observability of actions matters to these conclusions. When a stage game is represented in extensive form, some contingent actions may never be observed. Machines then cannot use those actions to distinguish states. Exercise 174.1 shows that replacing a strategic representation by its extensive form can change machine equilibria even when the standard repeated game Nash analysis is unchanged. Complexity is attached to the information needed to execute the plan, so the information supplied at the end of each stage must be specified.

10 Implementation Theory

10.1 Outcomes, preferences, and game forms

An implementation problem specifies desired outcomes as a function of the players' preferences and asks for rules of interaction that produce precisely those outcomes. The players know the preference profile; the planner does not. The planner chooses the game form and its consequence function, but cannot choose the players' preferences or their actions.

Let N be a finite set of players, C a consequence set, and $\mathcal{P}$ a domain of preference profiles. A choice rule is a nonempty correspondence $f : \mathcal{P} \rightrightarrows C$; a singleton valued rule is a choice function. A strategic game form is $G = \langle N, (A_i), g \rangle$, with $g : \prod_i A_i \to C$. A preference profile $R = (\succeq_i)_i$ induces preferences over action profiles by comparing their consequences. An extensive game form instead assigns a consequence to every terminal history.

Definition 10.1 (Implementation, Definition 179.1). Given a solution concept $\mathcal{S}$, the game form G implements f if

$$g(\mathcal{S}(G, R)) = f(R) \qquad \text{for every } R \in \mathcal{P}.$$

Every prescribed consequence must be the consequence of a solution, and every solution must produce a prescribed consequence.

Definition 10.2 (Truthful implementation, Definition 179.2). In a direct game form each player reports a preference profile. It truthfully implements f if, for every true profile R, unanimous truthful reporting is a solution and its consequence belongs to $f(R)$.

Truthful implementation permits other solutions with different consequences. It can also select only one member of a choice correspondence. This distinction is central when using revelation principles: a truthful equilibrium may reproduce one desired outcome while the direct mechanism admits additional equilibria.

10.2 Dominant strategy implementation

An action a_i^* is dominant if

$$g(a_i^*, a_{-i}) \succeq_i g(a_i, a_{-i}) \qquad \text{for every } a_i, a_{-i}.$$

A dominant strategy equilibrium consists of dominant actions for all players. Dominance here is weak optimality against every opponent action profile; it does not require a strict improvement against any particular profile.

Proposition 10.3 (Revelation principle, Lemma 181.4). If a choice rule is implementable in dominant strategies, it is truthfully implementable in dominant strategies. Moreover, a direct mechanism can be chosen in which each player reports only his own preference relation and truthful reports form a dominant strategy equilibrium.

Proof. In an implementing game form, select a dominant action $a_i(R_i)$ for each possible preference relation of player i. Dominance depends on R_i and the game form, not on the other players' true preferences. Define

$$g^*(R'_1, \dots, R'_n) = g(a_1(R'_1), \dots, a_n(R'_n)).$$

Any profitable false report would induce a profitable replacement of the selected dominant action against the corresponding actions of the other players. Thus truthful reporting is dominant, and the truthful outcome belongs to $f(R)$.

Theorem 10.4 (Gibbard–Satterthwaite, Proposition 181.2). Suppose C has at least three elements, the preference domain is unrestricted, and for every $c \in C$ there is a profile R such that $f(R) = \{c\}$. If f is implementable in dominant strategies, it is dictatorial: there is a player j such that every consequence in $f(R)$ is a most preferred consequence of j at every profile R.

The unrestricted domain and the range condition carry substantial content. The result does not apply to every restricted economic environment. In a quasilinear public project problem, the dependence of utility on money permits nontrivial truthful mechanisms.

10.3 Public projects and Groves payments

Let $x \in \{0, 1\}$ indicate whether a project of cost $\gamma \geq 0$ is undertaken. Player i has utility $\theta_i x - m_i$, where m_i is his payment and $\theta_i \in \mathbb{R}$. Efficiency requires

$$x(\theta) = 1 \quad \Longleftrightarrow \quad \sum_i \theta_i \geq \gamma.$$

The planner can use unrestricted monetary payments; balancing those payments against the project cost is a separate condition.

Proposition 10.5 (Groves mechanism, Proposition 184.2). Let each player report $a_i \in \mathbb{R}$ and set

$$x(a) = \mathbf{1}\{\sum_i a_i \geq \gamma\},$$

$$m_i(a) = x(a) \left(\gamma - \sum_{j \neq i} a_j \right) + h_i(a_{-i}),$$

where h_i is arbitrary. Reporting $a_i = \theta_i$ is a dominant action.

Proof. Fix a_{-i}. Player i's utility is

$$x(a)\left(\theta_i + \sum_{j \neq i} a_j - \gamma\right) - h_i(a_{-i}).$$

The final term is independent of his report. His report affects the first term only by determining whether x equals zero or one. Truthful reporting selects one exactly when the coefficient is nonnegative and hence chooses a maximizing alternative.

The payment form also follows from incentive compatibility. Fix opponents' reports and write $k = \gamma - \sum_{j \neq i} a_j$. Reports below k all select $x = 0$, and those above k all select $x = 1$. Payments must be constant within each group, since otherwise a player could preserve his allocation and reduce his payment. Let these payments be m^0, m^1. Truthfulness for types on either side of k implies

$$\theta_i \geq m^1 - m^0 \quad (\theta_i \geq k), \qquad \theta_i \leq m^1 - m^0 \quad (\theta_i < k).$$

Taking types arbitrarily close to k gives $m^1 - m^0 = k$. Thus $m_i = xk + h_i(a_{-i})$.

Example 10.6 (Truthful dominance and additional Nash equilibria). With two players, $\gamma = 2$, and $\theta_1 = \theta_2 = 1$, truthful reports undertake the project. The report profile $(0, 0)$ is also a Nash equilibrium: either player who alone induces the project must pay an additional two while valuing it at one. It selects $x = 0$. Truthful implementation in dominant strategies therefore does not imply full implementation in Nash equilibrium.

10.4 Nash implementation and monotonicity

The Nash revelation principle (Lemma 185.2) reproduces a selected equilibrium by asking each player to report the entire preference profile. Select an equilibrium $a(R)$ of an implementing mechanism and map reports $(R^1, \ldots, R^n)$ into $g(a_1(R^1), \ldots, a_n(R^n))$. At unanimous truthful reports, a deviation only selects an alternative action in the original game and cannot improve the reporting player's payoff. Reporting the entire profile is natural here because the equilibrium action can depend on all preferences.

Definition 10.7 (Monotonicity, Definition 186.1). A choice rule is monotonic if whenever $c \in f(R)$ but $c \notin f(R')$, there are a player i and consequence b such that

$$c \succeq_i b \qquad \text{and} \qquad b \succ'_i c.$$

Equivalently, if every player's lower contour set at a selected c expands, then c remains selected.

Proposition 10.8 (Proposition 186.2). Every Nash implementable choice rule is monotonic.

Proof. Let a be an equilibrium at R with consequence c. If c is excluded at R', full implementation requires that a cease to be an equilibrium. Some player i must then have a deviation a'_i with consequence b such that $b \succ'_i c$. Since a was an equilibrium at R, $c \succeq_i b$.

Example 10.9 (Solomon's predicament, Example 186.3). Let a assign an object to player 1, b assign it to player 2, and d destroy it. The possible preferences are

	1	2
R	$a \succ b \succ d$	$b \succ d \succ a$
R'	$a \succ d \succ b$	$b \succ a \succ d$.

The desired rule selects a at R and b at R'. No consequence that was below a for either player moves above a when preferences change to R'. The rule fails monotonicity and cannot be Nash implemented with this consequence set. Adding monetary consequences can create the comparisons needed to distinguish the two profiles.

10.5 No veto power and a sufficient condition

Definition 10.10 (Definition 187.1). A choice rule has no veto power if a consequence belongs to $f(R)$ whenever it is most preferred by at least $|N| - 1$ players.

Theorem 10.11 (Proposition 187.2). With at least three players, a monotonic choice rule with no veto power is Nash implementable when arbitrary strategic game forms are available.

Each player reports a preference profile, a consequence, and a nonnegative integer. Agreement on (R, c, m) with $c \in f(R)$ yields c. If all but one player agree on such a report, the exceptional player j can obtain his announced consequence b only when $c \succeq_j b$ under the agreed profile; otherwise c remains the outcome. In the remaining cases, an agent announcing a highest integer selects the consequence, with a fixed rule for ties.

Unanimous truthful reports of any $c \in f(R)$ are an equilibrium: a unilateral change can only replace c by something the deviator weakly dislikes under his true preferences. An equilibrium based on unanimous false reports must still select $f(R)$, because monotonicity supplies a profitable challenge otherwise. At other putative equilibria, all but possibly one player can obtain any desired consequence by increasing their integer. Equilibrium therefore requires the current consequence to be most preferred by at least $|N| - 1$ players. No veto power then places it in $f(R)$.

The integer component eliminates undesired disagreement equilibria; its unboundedness ensures that a player can always exceed the other reports. The construction establishes possibility for a broad class of rules. For particular domains, simpler mechanisms suffice. For example, if exactly one player wants an object and all others prefer not to receive it, asking who wants it and assigning it to the first affirmative respondent implements the desired assignment. If two players want it and both assignments must be possible, that priority rule excludes one desired outcome. Full implementation requires a rule that preserves both possibilities.

10.6 Subgame perfect implementation

Sequential choices can reveal preferences by presenting a player with a direct comparison whose answer determines an earlier player's incentives.

Example 10.12 (Example 191.2). An object's legitimate owner values it at v_H, the other player at v_L, with $v_H > v_L > 0$. Choose $v_L < M < v_H$ and $\varepsilon > 0$. First ask player 1 whether the object is hers. A negative answer assigns it to player 2 without payment. Following an affirmative answer, player 2 chooses whether to claim it. If he does not, player 1 receives it without payment. If he does, player 2 receives it and pays M, while player 1 pays ε.

If player 2 is not the owner, claiming gives $v_L - M < 0$, so he yields. Player 1 then claims and obtains the object. If player 2 is the owner, claiming gives $v_H - M > 0$, so he claims whenever reached. Player 1 consequently yields at the start, avoiding the fine. Backward induction selects the desired owner without an equilibrium payment in either state.

10.7 Virtual implementation

The book's Section 10.5 allows lotteries over deterministic consequences and monetary fines, with expected utility $u_i(L) - m_i$. Virtual implementation requires that for each $\varepsilon > 0$ the desired deterministic consequence receive probability at least $1-\varepsilon$ in the unique subgame perfect equilibrium. A small probability of auxiliary consequences supplies incentives to identify preferences.

The finite domain construction in Proposition 193.1 uses a finite set of distinct preferences over consequences, separated by lottery comparisons. For each pair v, v' choose alternatives L, L' such that v strictly prefers L and v' strictly prefers L'. Let $L^*(u, v, v')$ be the alternative preferred by reported type u. When all pairs receive equal probability, truthful reporting weakly improves every comparison and strictly improves the comparison between the true and reported types. Finiteness gives a common positive gain bound B.

The mechanism has K stages in which players sequentially report the full profile, followed by a stage in which they report their individual types. Each of the first K stages contributes probability $(1-\varepsilon)/K$ to the final lottery. If all but at most one player report the same profile, this component selects f of that profile; otherwise it selects a fixed consequence c^*. The final reports determine the separating lotteries, whose total probability is ε. A fine η is charged to the last player in the first K stages whose report differs from the final reported profile. In each of these stages, an additional fine η/K is charged to a player if all other players report the same profile and his report differs from theirs.

If D bounds the utility difference between deterministic consequences, choose

$$0 < \eta < \frac{\varepsilon B}{|N|}, \qquad \frac{(1-\varepsilon)D + \eta}{K} < \eta.$$

The first inequality makes truthful reporting in the final stage strictly optimal after every history, even allowing for the effect on fines. Given that truthful final report, a last false report in an earlier stage either changes no consequence and incurs an unnecessary disagreement fine, or saves the larger fine η when corrected at a cost bounded by the left side of the second inequality. Backward induction removes every such false report. With at least three players, the exceptional dissenter is well defined, and the desired consequence obtains probability at least $1 - \varepsilon$.

11 Extensive Games with Imperfect Information

11.1 Histories and information sets

In an extensive game with imperfect information, a player may not observe actions that have already occurred. The game specifies which histories he can distinguish when choosing an action. Beliefs may permit further inferences, but the information partition describes the observations supplied by the situation.

Definition 11.1 (Definition 200.1). An extensive game consists of a finite player set N, a set H of histories, a player function $P : H \setminus Z \to N \cup \{c\}$, chance probabilities $f_c(\cdot \mid h)$, information partitions $(\mathcal{I}_i)_i$, and expected utility preferences over terminal histories Z. Histories include the empty history, contain every prefix of each member, and include an infinite sequence whenever all its finite prefixes belong to H.

For every nonterminal history h, the feasible actions are $A(h) = \{a : (h, a) \in H\}$. Each $\mathcal{I}_i$ partitions the histories at which i acts. If h, h' belong to the same information set I, then $A(h) = A(h') =: A(I)$.

Equal action sets at indistinguishable histories are required for the interpretation: facing different available actions would itself reveal which history occurred. Chance moves have specified probabilities and terminal payoffs $u_i(z)$ evaluate the induced lotteries. Perfect information is the special case in which every information set is a singleton.

Example 11.2 (An unobserved action). Player 1 first chooses L or R. Choosing R ends the game. After L, player 2 chooses A or B, and player 1 then chooses ℓ or r without observing player 2's choice. Her information partition is

$$\mathcal{I}_1 = \big\{\{\varnothing\}, \{(L, A), (L, B)\}\big\},$$

while player 2 has the singleton information set $\{L\}$. Player 1 observes that she previously chose L, but not the intervening choice. The continuation after L can also represent simultaneous choices by players 1 and 2: one move is written first in the tree without being observed by the other player.

Definition 11.3 (Pure strategy, Definition 203.1). A pure strategy of player i chooses an action $s_i(I) \in A(I)$ at every information set $I \in \mathcal{I}_i$.

Strategies cannot condition on which member of an information set has occurred. In the example, player 1 has four pure strategies, $(L, \ell), (L, r), (R, \ell), (R, r)$. The two strategies beginning with R differ at an information set they themselves prevent from being reached. In the reduced strategic form they are identified when they induce the same terminal lottery against every strategy of the opponents.

11.2 Perfect recall

Let $X_i(h)$ record player i's own earlier information sets and actions along history h, in their order of occurrence. The record includes everything he was informed of at his decision dates and the actions he selected.

Definition 11.4 (Definition 203.3). The game has perfect recall if

$$X_i(h) = X_i(h') \qquad \text{whenever } h, h' \in I \in \mathcal{I}_i.$$

At an information set, a player with perfect recall has not forgotten his earlier information or his own actions. He can still be uncertain about opponents' actions and chance moves that he never observed. Imperfect recall can take several forms: forgetting a private signal, forgetting one's own earlier action, or not knowing whether the current decision is a first or second visit to the same information set. The last possibility is allowed by the general definition of an extensive game but is incompatible with perfect recall.

Perfect recall is a restriction on the information structure, independent of whether a particular equilibrium path exposes any forgetting. It is also the condition that permits a player's ex ante randomization over complete plans to be replaced by suitable randomization at each information set.

11.3 Four principles for equivalent extensive forms

The book's Section 11.2 studies transformations that preserve the reduced strategic form. They express judgments about which details of a tree matter for strategic choice. Their applicability depends on how information and decisions are interpreted.

Inflation and deflation of information sets. Suppose two parts of an information set can already be distinguished from the player's own earlier actions at an earlier information set. Splitting or merging those parts preserves the reduced strategic form. This principle treats the player as able to infer his past choice from his selected plan. If the merged information set instead represents actual forgetting, the behavioral interpretation can change.

Adding or removing a superfluous move. A move is superfluous when every available action leaves the same future possibilities, information, and terminal consequences. Its branches can be collapsed, or duplicate branches can be introduced. Consequences must remain the same for every player; mere indifference of the player making the move is not enough. The transformation also preserves the information relations in the continuation.

Coalescing moves. If a player's action leads directly to another information set of the same player without new information, the two decisions can be combined. Choosing ℓ and then A or B, with an alternative initial choice r, becomes a single choice among $\ell A, \ell B, r$. A complete plan already specifies the second action, so this transformation leaves the available contingent choices unchanged.

Interchanging moves. When one player chooses after another but does not observe the earlier action, the order can be reversed with the information sets adjusted accordingly. The resulting

terminal histories encode the same action combinations and consequences. This is the relation between a simultaneous move game and its sequential drawing with an information set joining the second player's nodes.

For finite games in which an information set is never visited twice on a path, the book reports that these transformations connect extensive forms having the same reduced strategic form. A solution concept based only on that form is invariant under all four. A solution concept that distinguishes some of the trees attaches significance to information or timing that at least one of the transformations suppresses.

11.4 Framing and conditional choice

Preserving a reduced strategic form does not settle every question about behavior at an information set. The book's Section 11.3 compares a decision made before a chance outcome with a decision made after learning that a particular chance outcome occurred. Adding a superfluous move can preserve ex ante outcome possibilities while changing the conditional decision that a player perceives.

In the strictly competitive strategic form displayed in Figure 212.1, player 1's payoffs are

	ℓ	r
ℓ	4	0
r	3	1

.

If she chooses ℓ with probability p, her worst expected payoff is

$$\min\{4p + 3(1-p),\, 1-p\} = 1-p.$$

Her ex ante maxmin choice is therefore pure r. In one extensive representation, however, she chooses only after observing the relevant chance branch, where the conditional payoff matrix has payoffs two for matching actions and zero for mismatching them. Conditional maxmin choice then uses each action with probability one half. The difference concerns the point at which the criterion is applied and the conditioning information. A claim of equivalence must specify which of these decisions its solution concept evaluates.

11.5 Mixed and behavioral strategies

> Definition 11.5 (Definition 212.1). A mixed strategy σ_i is a probability distribution on player i's pure strategies. A behavioral strategy β_i assigns an independent probability distribution $\beta_i(\cdot \mid I)$ to the actions at each information set. Two strategies are outcome equivalent if, against every pure strategy profile of the other players, they induce the same distribution over terminal histories.

Mixed strategies randomize over complete contingent plans. Their actions at different information sets may be correlated. Behavioral strategies specify local randomizations. In a finite game, a behavioral profile gives terminal history $z = (a^1, \ldots, a^K)$ probability

$$\Pr_{\beta}(z) = \prod_{k=0}^{K-1} \beta_{P(h^k)}(a^{k+1} \mid h^k),$$

where $h^k = (a^1, \ldots, a^k)$ and the factor for a chance move is its specified chance probability.

For a mixed strategy, let $\pi_i(h)$ be the total probability of pure plans consistent with the actions of i occurring along h. Then

$$\Pr_{\sigma}(z) = \pi_c(z) \prod_i \pi_i(z).$$

The two factorizations look different because the mixed strategy already contains joint probabilities across a player's several decisions.

11.6 Outcome equivalence under perfect recall

Proposition 11.6 (Kuhn's theorem, Proposition 214.1). In a finite extensive game with perfect recall, every mixed strategy has an outcome equivalent behavioral strategy. Every behavioral strategy also has an outcome equivalent mixed strategy.

From behavior to a mixture. Choose independently, at the start, the action to be used at every information set. The probability assigned to a pure strategy s_i is

$$\sigma_i(s_i) = \prod_{I \in \mathcal{I}_i} \beta_i(s_i(I) \mid I).$$

Along any terminal history each information set is visited at most once, so the products give the same terminal probabilities as local randomization. This direction needs only the absence of repeated visits to an information set along a path.

From a mixture to behavior. For $h \in I$, define

$$\beta_i(a \mid I) = \frac{\pi_i(h, a)}{\pi_i(h)} \quad \text{if } \pi_i(h) > 0.$$

Perfect recall makes the numerator and denominator independent of the choice of $h \in I$: every such history requires the same previous actions at the same information sets of player i. Choose arbitrary probabilities when the denominator is zero. Along a history consistent with the mixed strategy, successive conditional probabilities telescope to $\pi_i(h)$. Against any fixed opponents' strategies, chance and opponent contributions remain unchanged, establishing outcome equivalence.

Example 11.7 (Conditional randomization). Player 1 chooses A or B at the start and has a later information set reached only after B. Her mixture puts probabilities $0.4, 0.1, 0.5$ on $(B, r), (B, \ell), (A, \ell)$, respectively. Its equivalent behavioral strategy chooses B with probability 0.5 and, conditional on reaching the later information set, chooses r with probability $0.4/0.5 = 0.8$. Using the unconditional probability 0.4 at that information set would give the wrong terminal distribution.

11.7 Failures with imperfect recall

Consider a player who chooses L or R, then chooses L or R again while forgetting the first choice. A mixed strategy selecting LL and RR equally induces probabilities $(1/2, 0, 0, 1/2)$ on LL, LR, RL, RR. A behavioral strategy with probabilities p and q of L at the two information sets gives

$$(pq,\ p(1-q),\ (1-p)q,\ (1-p)(1-q)).$$

Making both middle probabilities zero forces at least one of the desired diagonal probabilities to be zero. Local independent randomizations cannot reproduce the correlation contained in the mixed strategy.

Conversely, suppose the same information set may be visited twice. At the first visit, action b terminates play with zero, while a leads to another visit. At that visit, a yields zero and b yields one. Every pure plan chooses the same action at both visits and earns zero. A behavioral rule choosing a with probability p at each visit earns $p(1-p)$, maximized at $p = 1/2$ with value $1/4$. No mixture of the pure plans has that outcome.

11.8 Nash equilibrium in an extensive game

A mixed strategy profile σ is a Nash equilibrium if

$$\mathbb{E}_{O(\sigma)} u_i \geq \mathbb{E}_{O(\sigma_i', \sigma_{-i})} u_i \qquad \text{for all } i, \sigma_i'.$$

The analogous definition allows behavioral deviations. In a finite game with perfect recall, Kuhn's theorem makes the two equilibrium outcome sets coincide: each deviation of either type has an equivalent deviation of the other type. The finite strategic form supplies existence of a mixed equilibrium and hence of an equivalent behavioral equilibrium.

With imperfect recall, these deviation sets need not coincide, as the repeated information set example demonstrates. Even with perfect recall, Nash equilibrium does not require optimal action at an unreached information set. Extending sequential rationality therefore requires a probability distribution over the histories in that set. Equilibrium strategies determine that distribution by Bayes' rule when the set is reached with positive probability; further restrictions are needed when its probability is zero.

12 Sequential Equilibrium

12.1 Strategies and beliefs at every information set

At an information set containing several histories, optimal behavior depends on which history the player believes has occurred. A strategy profile determines this belief through Bayes' rule whenever the information set has positive probability. If the information set is not reached, the profile alone gives no conditional probabilities. Sequential equilibrium includes these beliefs in the solution and restricts how they relate to the strategies.

Throughout the finite game results, assume perfect recall and remove chance branches of probability zero. Every remaining information set is then reached with positive probability under completely mixed behavioral strategies.

Definition 12.1 (Assessment and sequential rationality). An assessment is a pair (β, μ), where β is a behavioral strategy profile and $\mu(\cdot \mid I)$ is a probability distribution on the histories in each information set I. It is sequentially rational if, at every information set I of every player i,

$$\mathbb{E}_{O(\beta,\mu|I)} u_i \geq \mathbb{E}_{O((\beta_i',\beta_{-i}),\mu|I)} u_i \quad \text{for every continuation strategy } \beta_i'.$$

Here $O(\beta, \mu \mid I)$ first draws the history in I according to μ and then continues according to the behavioral profile and chance probabilities.

Sequential rationality applies after every information set, including those assigned zero probability by equilibrium play. The comparison holds beliefs at the current information set fixed while considering a continuation deviation. It concerns the player's entire continuation, not merely the action chosen now.

12.2 Consistency and sequential equilibrium

Definition 12.2 (Definitions 224.2 and 225.1). An assessment (β, μ) is consistent if there are completely mixed behavioral profiles $\beta^k \to \beta$ whose Bayesian belief systems $\mu^k \to \mu$. A sequential equilibrium is a consistent, sequentially rational assessment.

For every information set I reached with positive probability under β, consistency implies

$$\mu(h \mid I) = \frac{\Pr_\beta(h)}{\sum_{h' \in I} \Pr_\beta(h')}.$$

At an unreached set, the relative rates at which the probabilities of different histories vanish determine the limiting belief. One common sequence must justify all information sets at once. The perturbations need not themselves be equilibria or be sequentially rational.

Example 12.3 (An unreached information set). Player 1 chooses L, M, R. Action L terminates play with payoff $(2, 2)$. After M or R, player 2 acts without observing which was chosen. His payoffs from his own L, R are $(1, 2)$ following M and $(2, 1)$ following R. If α is his conditional probability of M, then

$$U_2(L) = 2 - \alpha, \qquad U_2(R) = 1 + \alpha.$$

He prefers R if and only if $\alpha \geq 1/2$. In the game of Figure 220.1, player 1 choosing L and player 2 planning R is sequentially rational whenever this inequality holds. Every $\alpha \in [0, 1]$ can be obtained as a limiting belief by assigning suitably small probabilities to M and R. For an interior α, use probabilities $\alpha\varepsilon$ and $(1 - \alpha)\varepsilon$; for endpoint beliefs, use probabilities of orders ε and ε^2.

Proposition 12.4 (Implications of sequential equilibrium). The behavioral strategy profile of a sequential equilibrium is a Nash equilibrium. In a game with perfect information, sequential equilibrium and subgame perfect equilibrium coincide. For a consistent assessment in a finite game with perfect recall, sequential rationality holds if and only if no player can improve at any information set by changing only his action distribution there and leaving his other choices fixed.

The last statement is the one deviation property in Exercise 227.1. Consistency ensures that conditional continuation probabilities agree as a player moves from one of his information sets to a later one. A profitable continuation deviation can therefore be shortened from its last changed information set backward, as in backward induction. This makes local incentive comparisons sufficient once consistency has been established.

12.3 Restrictions linking beliefs across information sets

Structural consistency, as defined in Definition 228.1, requires that each information set's belief arise from some strategy profile that reaches it with positive probability. The profile may differ across information sets. It is different from sequential consistency, which takes limits along one common sequence.

Example 12.5 (A consistent limit that no profile generates). Two players independently choose L with probability ε and R with probability $1 - \varepsilon$. A third player observes that (R, R) did not occur but does not observe which of $(L, L), (L, R), (R, L)$ occurred. His conditional belief is

$$\frac{(\varepsilon^2,\ \varepsilon(1-\varepsilon),\ \varepsilon(1-\varepsilon))}{2\varepsilon - \varepsilon^2} \longrightarrow (0, 1/2, 1/2).$$

No independent strategy profile generates this limiting conditional belief exactly: making (L, L) impossible forces at least one player never to choose L, which also makes one of $(L, R), (R, L)$ impossible. Thus a consistent limiting belief need not be structurally consistent in the exact sense.

Conversely, separate rationalizations need not admit a common perturbation. In Example 230.1, two players choose among L, M, R, and a third learns the unordered pair of actions. Suppose he assigns probability $2/3$ to (M, L) rather than (L, M), to (L, R) rather than (R, L), and to (R, M)

rather than (M, R). Independent positive action probabilities would require each of

$$\frac{p_1(M)p_2(L)}{p_1(L)p_2(M)}, \qquad \frac{p_1(L)p_2(R)}{p_1(R)p_2(L)}, \qquad \frac{p_1(R)p_2(M)}{p_1(M)p_2(R)}$$

to converge to two. Their product is identically one, so the three beliefs cannot be jointly consistent. This restriction uses only the information structure and independence of players' randomizations.

Another implication is that observing an unexpected action by one player does not generally justify revising an unrelated earlier belief about a different player. If the unexpected player's action has the same probability at all histories in his information set, that probability cancels from the relevant Bayesian likelihood ratio. Example 229.2 uses this cancellation to rule out a sequentially rational assessment with separately rationalizable beliefs.

12.4 The role of the extensive form

Sequential equilibrium can change under coalescing of moves. In Example 226.2, an initial choice among three actions is replaced by a choice to stop or continue, followed by a further choice of the same player. In the first game, the relative probabilities of two unchosen initial actions can converge at different rates. In the second, the corresponding continuation choice is itself subject to sequential rationality, which can force one action with probability one. That restriction determines the later opponent's belief and can eliminate an equilibrium of the coalesced game.

The definition also permits information sets of different players whose order depends on chance. At a current information set, the belief may attribute a past unexpected action to a player who could instead act later at another history in the same set. Continuation play still uses that player's prescribed strategy. The book's Section 12.2 examines this interaction between explanations of past events and expectations about future actions; the extensive form specifies both.

12.5 Observable actions and perfect Bayesian equilibrium

In a Bayesian extensive game with observable actions, chance independently draws each player's type θ_i from a finite set Θ_i using a positive prior p_i. Each player observes his own type and all subsequent action histories. Let $\sigma_i(\theta_i, h)$ describe the strategy of type θ_i after public history h, and let $\mu_i(h)$ be the common posterior belief about player i's type.

Definition 12.6 (Perfect Bayesian equilibrium, Definition 232.1). The strategies and type beliefs form a perfect Bayesian equilibrium if:

1. Every type's continuation strategy is optimal after every public history, given the other types' strategies and the product of posterior type beliefs.
2. Initial beliefs are $\mu_i(\varnothing) = p_i$.
3. Beliefs about player i change only in response to his own action. If he does not move, his posterior stays fixed; at simultaneous moves, other players' current actions do not affect that posterior.
4. Whenever the observed action has positive probability under current beliefs, updating follows
$$\mu_i(h, a)(\theta_i) = \frac{\mu_i(h)(\theta_i)\sigma_i(\theta_i, h)(a_i)}{\sum_{\theta_i'} \mu_i(h)(\theta_i')\sigma_i(\theta_i', h)(a_i)}.$$

This is the particular definition used in the book's Section 12.3, including its restrictions on how beliefs can respond to actions. Proposition 234.1 shows that a sequential equilibrium of the associated finite extensive game induces such a perfect Bayesian equilibrium. The converse can fail: beliefs after two different unexpected histories may separately satisfy these requirements but require contradictory relative probabilities of the same earlier deviations.

For instance, in Example 234.3, types x and y both choose Out initially, but would choose L and M, respectively, if they continued. At the unexpected history (C, L), certainty of type y requires the probability of type x continuing to be negligible relative to that of type y. Certainty of type x at (C, M) requires the reverse. No common sequence of perturbations has both properties. Also, a posterior probability of zero need not remain zero after a later action that has zero probability under the current assessment. The book's definition permits a new inference at such a history.

12.6 Signaling and education

A signaling game has an informed sender, a message m, and a receiver who observes m before choosing a. Payoffs depend on the type, message, and response. Pooling means that different types choose the same message; separation means that their messages reveal their types.

Example 12.7 (Education, Example 237.2). A worker has ability $\theta \in \{\theta_L, \theta_H\}$, where $0 < \theta_L < \theta_H$, and prior high type probability p_H. Education is $e \geq 0$, and worker utility is $w - e/\theta$. The employer's utility is $-(w - \theta)^2$, so its optimal wage is

$$w(e) = \mathbb{E}[\theta \mid e].$$

In a pooling equilibrium with education e^*, the wage is $\bar{\theta} = p_H\theta_H + (1 - p_H)\theta_L$. Assigning low type beliefs to every other education level makes the best deviation choose zero education. Both types are deterred exactly when the more demanding low type inequality holds:

$$\theta_L \leq \bar{\theta} - \frac{e^*}{\theta_L}, \quad \text{or} \quad 0 \leq e^* \leq \theta_L p_H(\theta_H - \theta_L).$$

In a separating equilibrium the low type chooses zero: she cannot earn less than the lowest ability wage, so positive education is wasteful for her when her type is already revealed. With wages θ_L, θ_H, the two imitation constraints are

$$\theta_L \geq \theta_H - \frac{e_H}{\theta_L}, \qquad \theta_H - \frac{e_H}{\theta_H} \geq \theta_L.$$

Hence

$$\theta_L(\theta_H - \theta_L) \leq e_H \leq \theta_H(\theta_H - \theta_L).$$

Low type beliefs following any education other than e_H support every level in this interval.

12.7 Reputation in the chain store game

The finite chain store game changes substantially if there is initially a small probability $\varepsilon > 0$ of a type that prefers fighting an entrant to accommodation. The regular incumbent earns -1 from fighting, zero from accommodation, and $a > 1$ when an entrant stays out. The tough type reverses its ranking of fighting and accommodation. Entrant k earns $b \in (0, 1)$ from accommodated entry, $b - 1$ if fought, and zero from staying out.

Let q be the current posterior probability of toughness and K the number of entrants. In the equilibrium of Example 239.1, the tough type always fights, and the regular type accommodates the last entrant. Before the last period, after entrant k enters, the regular type fights for sure if $q \geq b^{K-k}$; below that threshold it fights with probability

$$r_k(q) = \frac{(1 - b^{K-k})q}{(1 - q)b^{K-k}}.$$

An entrant stays out for $q > b^{K-k+1}$, enters for $q < b^{K-k+1}$, and at equality stays out with probability $1/a$.

The mixing formula follows directly from Bayesian updating. Fighting must raise the posterior to the next entrant's threshold:

$$\frac{q}{q + (1 - q)r_k(q)} = b^{K-k}.$$

At $q = b^{K-k+1}$, the probability of fighting is $q + (1 - q)r_k(q) = b$, so entry gives expected payoff $b - b = 0$. The entrant's probability $1/a$ of staying out provides the incumbent's continuation indifference. Accommodation reveals the regular type and subsequent entrants enter. A decision to stay out supplies no new information, so it leaves q unchanged.

Before the first entry, the posterior remains ε. Entry becomes possible when $b^{K-k+1} \geq \varepsilon$: at equality the entrant enters with probability $1 - 1/a$, and under a strict inequality entry is certain. The number of these final periods is bounded independently of K. A longer game therefore has a longer initial interval without entry. During that interval the posterior stays at the prior: the regular incumbent's willingness to fight sustains deterrence even though no fighting occurs on the path.

12.8 Refinements based on incentives to deviate

Sequential consistency restricts beliefs through probabilities and information. Refinements can also use a type's possible gain from a deviation. If a message cannot benefit a type even under its most favorable receiver response, but can benefit another type under a response optimal for that type, the latter has a stronger reason to send it. The inference is made relative to the proposed equilibrium payoffs.

Example 12.8 (Beer or Quiche, Example 244.2). The sender is strong with probability 0.9 and weak with probability 0.1. She gets two from avoiding a fight and one from consuming her preferred breakfast: beer for the strong type and quiche for the weak type. The receiver prefers fighting a weak sender and not fighting a strong one.

Pooling on quiche with no fight gives the weak type three and the strong type two. If beer is unexpectedly observed, the weak type cannot gain: even without a fight her payoff would be only two. The strong type would earn three if recognized and not fought. Excluding the weak type after this deviation makes beer profitable for the strong type and rules out the pooling quiche outcome. Pooling on beer does not have the same defect: a weak sender can want to switch to quiche if doing so avoids a fight, so a belief that an unexpected quiche sender is weak is compatible with the incentive comparison.

In education signaling, a pooling equilibrium can be defeated by an education level that would benefit the high type if paid θ_H but would hurt the low type even at that wage. In a separating equilibrium with $e_H > \theta_L(\theta_H - \theta_L)$, a small reduction in the high type's education has the same property. The incentive based refinement in Example 246.1 thus selects the least education that deters imitation:

$$e_L = 0, \qquad e_H = \theta_L(\theta_H - \theta_L).$$

12.9 Trembling hand perfection in strategic games

Definition 12.9 (Definition 248.1). A mixed profile σ of a finite strategic game is trembling hand perfect if there are completely mixed profiles $\sigma^k \to \sigma$ such that each σ_i is a best response to σ^k_{-i} for every k.

Continuity implies that every perfect equilibrium is Nash. A weakly dominated strategy is not a best response to any completely mixed opponents' profile, since an outcome at which dominance is strict then has positive probability. Thus perfection excludes weakly dominated strategies.

Proposition 12.10 (Proposition 248.2). In a finite two player strategic game, a mixed profile is trembling hand perfect if and only if it is a Nash equilibrium and neither player's mixed strategy is weakly dominated.

Proof of sufficiency. An undominated mixed strategy is a best response to some completely mixed opponent strategy τ_j. If σ is also Nash, then σ_i is a best response to

$$\sigma_j^\varepsilon = (1-\varepsilon)\sigma_j + \varepsilon\tau_j$$

by linearity. These profiles are completely mixed and converge to σ. The result is specific to two players: with several opponents, an arbitrary fully supported belief over their joint actions need not be a product of their individual mixed strategies.

Proposition 12.11 (Existence, Proposition 249.1). Every finite strategic game has a trembling hand perfect equilibrium.

Require every action to receive a small positive minimum probability. The resulting restricted simplices are compact and convex, and expected payoffs retain the conditions for Nash existence. Let the lower bounds tend to zero and select a convergent sequence of equilibria. Actions assigned positive probability in the limit eventually exceed their lower bounds and must therefore be best responses. The limiting profile satisfies perfection.

12.10 Perfection in an extensive game

Perfection of the ordinary strategic form does not necessarily impose optimal behavior at every information set. A pure strategy may prescribe an inferior later action at a node that its own earlier action prevents from being reached. Perturbing only opponents' complete plans leaves that later choice irrelevant.

The agent strategic form assigns a separate agent to each information set. Every agent chooses an action at that information set, and agents belonging to the same original player share that player's terminal payoff. A mixed profile of agents corresponds to a behavioral profile in the extensive game.

Definition 12.12 (Definition 251.1). A trembling hand perfect equilibrium of a finite extensive game is a behavioral profile whose associated profile in the agent strategic form is trembling hand perfect.

Perturbations now include the actions of a player's other agents. Every information set receives positive probability, including ones ruled out by the player's intended earlier behavior. This makes optimality at that set substantive.

Proposition 12.13 (Proposition 251.2 and Corollary 253.2). Every trembling hand perfect behavioral equilibrium of a finite extensive game with perfect recall is the strategy component of a sequential equilibrium. Every such game has both a trembling hand perfect equilibrium and a sequential equilibrium.

Proof. Take the completely mixed sequence supporting perfection in the agent form. Its Bayesian beliefs have a convergent subsequence because there are finitely many finite information sets. The limit assessment is consistent. Each agent's optimality against the perturbed profiles implies local optimality conditional on its information set in the limit. The one deviation property gives sequential rationality. Existence follows by applying strategic perfection existence to the finite agent form.

The converse fails: in an extensive representation of simultaneous moves, every Nash equilibrium has sequentially consistent beliefs, while weakly dominated strategies may prevent perfection. Extensive game perfection and perfection of the ordinary strategic form are also not nested in general. They perturb different objects: local decisions in one case and complete plans in the other. Figures 250.1 and 253.1 illustrate why the game form must remain part of the statement of the refinement.

13 The Core

13.1 Coalitional games with transferable payoff

A coalitional game describes what each group of players can achieve together. It leaves the internal procedure for choosing and enforcing a joint action unspecified. In a game with transferable payoff, the possibilities of a coalition are summarized by a single total worth, which can be divided freely among its members. The worth is taken to be independent of outsiders' actions.

Definition 13.1 (Definitions 257.1 and 258.1). A transferable payoff game is (N, v), where N is finite and $v(S) \in \mathbb{R}$ is the worth of every nonempty coalition $S \subseteq N$. Set $v(\varnothing) = 0$. The game is cohesive if

$$v(N) \geq \sum_{k=1}^{K} v(S_k)$$

for every partition $\{S_1, \ldots, S_K\}$ of N. It is superadditive if $v(S \cup T) \geq v(S) + v(T)$ for every disjoint S, T.

Superadditivity implies cohesiveness. Cohesiveness makes formation of the grand coalition efficient relative to forming disjoint coalitions. It does not by itself ensure that the grand coalition's worth can be allocated so that no smaller coalition wants to leave.

13.2 Blocking and the core

Write $x(S) = \sum_{i \in S} x_i$. A feasible grand coalition payoff satisfies $x(N) = v(N)$.

Definition 13.2 (Definition 258.2). A coalition S blocks x if it can distribute $v(S)$ in a vector y with $y_i > x_i$ for every $i \in S$. The core is

$$\mathcal{C}(v) = \{x \in \mathbb{R}^N : x(N) = v(N),\ x(S) \geq v(S) \text{ for every } S \subseteq N\}.$$

If $v(S) > x(S)$, its excess can be divided into strictly positive gains for every member, so S blocks. Conversely, a strict gain for every member sums to $v(S) > x(S)$. The linear inequalities therefore express exactly the absence of blocking. They make the core closed and convex. Singleton inequalities give $x_i \geq v(\{i\})$, and efficiency supplies upper bounds, so a nonempty core is compact.

Example 13.3 (Three player majority game, Example 259.1). Let $v(N) = 1$, $v(\{i\}) = 0$, and $v(\{i, j\}) = \alpha \in [0, 1]$. The core conditions reduce to

$$x_1 + x_2 + x_3 = 1, \qquad 0 \leq x_i \leq 1 - \alpha.$$

Summing the three upper bounds gives $1 \leq 3(1 - \alpha)$, or $\alpha \leq 2/3$. This is sufficient because $(1/3, 1/3, 1/3)$ then satisfies every inequality. At $\alpha = 2/3$ the core is a singleton; above it the core is empty.

Example 13.4 (An indivisible good market, Example 260.1). Each seller owns one good worth zero to him; each buyer values one unit at one. For buyer set B and seller set L,

$$v(S) = \min\{|S \cap B|, |S \cap L|\}.$$

If $|B| > |L|$, the core gives each seller one and each buyer zero. To see this, let x_b and x_ℓ be the smallest buyer and seller payoffs. Pairwise blocking requires $x_b + x_\ell \geq 1$, while efficiency gives

$$|L| \geq |B|x_b + |L|x_\ell \geq |L| + (|B| - |L|)x_b.$$

Individual rationality implies $x_b = 0$ and then every seller must receive one. If $|B| = |L|$, the core instead consists of equal buyer payoffs α and equal seller payoffs $1 - \alpha$, for $0 \leq \alpha \leq 1$. Equality of payoffs within each side follows by summing all buyer–seller inequalities and using efficiency.

13.3 Veto players and convex games

A simple game has $v(S) \in \{0, 1\}$ and $v(N) = 1$. A coalition of worth one is winning, and a veto player belongs to every winning coalition. At a core payoff, every player excluded from some winning coalition must receive zero: the winning coalition already requires the entire unit, while singleton inequalities make all payoffs nonnegative. Thus a simple game without a veto player has an empty core. With veto players, its core consists of all nonnegative divisions of the unit among those players.

Proposition 13.5 (Convex games, Exercise 260.4). Suppose

$$v(S) + v(T) \leq v(S \cup T) + v(S \cap T) \quad \text{for every } S, T.$$

For any ordering of the players, give each player his marginal contribution to the predecessors in that ordering. The resulting vector is in the core.

Proof. Convexity implies increasing marginal contributions: if $S \subseteq T$ and $i \notin T$, then $v(S \cup \{i\}) - v(S) \leq v(T \cup \{i\}) - v(T)$. Marginal contributions along the full ordering telescope to $v(N)$. For a coalition R, each member's full predecessor set contains his predecessors within R. Summing the increasing marginal contribution inequalities over members of R gives $x(R) \geq v(R)$.

13.4 Balancedness and nonemptiness

Definition 13.6 (Balanced weights). Nonnegative weights $(\lambda_S)_{\varnothing \neq S \subseteq N}$ are balanced if

$$\sum_{S \ni i} \lambda_S = 1 \qquad \text{for every } i.$$

The game is balanced if every such collection satisfies $\sum_S \lambda_S v(S) \leq v(N)$.

A partition corresponds to balanced weights equal to zero or one. Balancedness also compares the grand coalition with fractional combinations of overlapping coalitions. It is consequently stronger than cohesiveness.

Theorem 13.7 (Bondareva–Shapley, Proposition 262.1). A finite transferable payoff game has a nonempty core if and only if it is balanced.

Proof. If x is in the core, then for any balanced weights,

$$\sum_S \lambda_S v(S) \le \sum_S \lambda_S x(S) = \sum_i x_i \sum_{S \ni i} \lambda_S = x(N) = v(N).$$

For sufficiency, consider the linear program

$$\min_x \sum_i x_i \quad \text{subject to} \quad x(S) \ge v(S) \quad (S \neq \varnothing).$$

It is feasible by choosing all coordinates sufficiently large, and its objective is bounded below by the singleton constraints. Its dual is

$$\max_{\lambda \ge 0} \sum_S \lambda_S v(S) \quad \text{subject to} \quad \sum_{S \ni i} \lambda_S = 1 \quad (i \in N).$$

Balancedness bounds the dual value by $v(N)$, while $\lambda_N = 1$ attains $v(N)$. Strong duality supplies a primal optimum with total $v(N)$, which is a core payoff. This is the linear programming form of the separation argument in the book.

For the three player majority game, put weight $1/2$ on each pair. Every player is counted once, and balancedness requires $3\alpha/2 \le 1$. This recovers the exact nonemptiness condition without solving for individual payoffs.

13.5 Markets with transferable payoff

Agent i owns input bundle $\omega_i \in \mathbb{R}^\ell_+$ and produces a transferable output using a continuous, nondecreasing, concave function $f_i : \mathbb{R}^\ell_+ \to \mathbb{R}_+$. The associated worth is

$$v(S) = \max \left\{ \sum_{i \in S} f_i(z_i) : z_i \ge 0,\ \sum_{i \in S} z_i = \sum_{i \in S} \omega_i \right\}.$$

Only coalition members' endowments and technologies enter this problem.

Proposition 13.8 (Proposition 264.2). Every market with transferable payoff satisfying these assumptions has a nonempty core.

Proof. For balanced weights, let z_i^S be an optimal input allocation for coalition S, and set $z_i^* = \sum_{S \ni i} \lambda_S z_i^S$. Balancedness implies

$$\sum_i z_i^* = \sum_S \lambda_S \sum_{i \in S} \omega_i = \sum_i \omega_i.$$

Thus z^* is feasible for the grand coalition. Since the weights of coalitions containing i sum to one, concavity gives

$$v(N) \ge \sum_i f_i(z_i^*) \ge \sum_i \sum_{S \ni i} \lambda_S f_i(z_i^S) = \sum_S \lambda_S v(S).$$

Apply the balancedness theorem.

13.6 Competitive payoffs and the core

A competitive equilibrium of the production market is an input price vector $p \geq 0$ and a feasible allocation z^* such that every z_i^* solves

$$\max_{z_i \geq 0} \{f_i(z_i) - p \cdot (z_i - \omega_i)\}.$$

The associated payoff is $x_i = f_i(z_i^*) - p \cdot (z_i^* - \omega_i)$.

> Proposition 13.9 (Proposition 267.1). Every competitive payoff vector belongs to the market's core.

Proof. For every feasible coalition allocation $(z_i)_{i \in S}$, individual optimality and cancellation of the coalition's input trades imply

$$\sum_{i \in S} x_i \geq \sum_{i \in S} [f_i(z_i) - p \cdot (z_i - \omega_i)] = \sum_{i \in S} f_i(z_i).$$

Maximizing the right side gives $x(S) \geq v(S)$. Applying the same comparison to N and using feasibility of z^* gives $x(N) = v(N)$.

If there is one input and all agents share a concave production function f, equal input holdings $\bar{\omega} = \sum_i \omega_i / |N|$ maximize production. A supporting slope p at $\bar{\omega}$ gives the affine function $g(w) = f(\bar{\omega}) + p(w - \bar{\omega})$. Since $f(w) \leq g(w)$, the payoff $x_i = g(\omega_i)$ satisfies all coalition constraints. This exhibits directly how competitive input values support a core division.

13.7 Games without transferable payoff

A general coalitional game specifies a consequence set X, a feasible consequence set $V(S) \subseteq X$ for each coalition, and preferences $(\succeq_i)_i$ on X. Its core consists of $x \in V(N)$ for which no coalition S has a $y \in V(S)$ with $y \succ_i x$ for every $i \in S$ (Definitions 268.2 and 268.3). The scalar coalition inequalities of a transferable payoff game are replaced by comparisons of attainable consequences. The criterion remains unanimous strict improvement within the blocking coalition.

13.8 Exchange economies

An exchange economy has endowments $\omega_i \in \mathbb{R}^\ell_+$ and continuous, nondecreasing, quasiconcave preferences over consumption bundles. Aggregate endowment has a positive amount of each good. An allocation x satisfies $\sum_i x_i = \sum_i \omega_i$. A coalition can redistribute only its own endowment, and every agent cares only about his own consumption.

A competitive equilibrium consists of nonzero prices $p \geq 0$ and an allocation x with

$$p \cdot x_i \leq p \cdot \omega_i, \qquad x_i \succeq_i y_i \quad \text{whenever } p \cdot y_i \leq p \cdot \omega_i.$$

In a two person economy, the core is the set of individually rational allocations that cannot be strictly improved for both players by a feasible reallocation. In an Edgeworth box with smooth strictly convex preferences, it is the portion of the contract curve lying above both agents' endowment indifference curves.

Proposition 13.10 (Proposition 272.1). Every competitive allocation of an exchange economy is in its core.

Proof. If a coalition could give every member $y_i \succ_i x_i$, individual optimality would imply $p \cdot y_i > p \cdot \omega_i$ for every member. Summing contradicts the coalition resource equality $\sum_{i \in S} y_i = \sum_{i \in S} \omega_i$.

13.9 Replication and the competitive limit

The replica kE contains k copies of every original agent, with the same preferences and endowment. Copies of agent i have type i.

Lemma 13.11 (Equal treatment, Lemma 272.2). If preferences are increasing and strictly quasiconcave, every allocation in the core of a replica gives all agents of the same type the same bundle.

Proof. Take the least satisfied copy of each type. Give each selected copy the average bundle assigned to its type in the proposed allocation. This is feasible for the selected coalition because type averages sum to one copy of aggregate endowment. Quasiconcavity makes every selected copy weakly better off. If copies of some type received different bundles, strict quasiconcavity makes that type's selected copy strictly better off. Continuity permits a small reduction of its bundle while retaining that gain; increasing preferences allow the released goods to make all other selected agents strictly better off. The resulting coalition blocks.

Theorem 13.12 (Replica core theorem, Proposition 273.1). Suppose preferences are increasing and strictly quasiconcave and every agent has a positive endowment of every good. If the equal treatment replication of an allocation x belongs to the core of kE for every positive integer k, then x is a competitive allocation of E.

Proof idea. Consider the convex hull Q of net trades z_i such that $z_i + \omega_i \succ_i x_i$. If zero belonged to Q, a convex combination of strictly improving net trades would sum to zero. Rational coefficients would give a blocking coalition with corresponding integer numbers of each type in a sufficiently large replica. Continuity and positive endowments permit the same conclusion by approximation for irrational coefficients. Thus $0 \notin Q$.

A separating hyperplane gives nonzero $p \geq 0$ with $p \cdot z \geq 0$ for $z \in Q$. Applying this to small improvements of each x_i and using aggregate feasibility gives $p \cdot x_i = p \cdot \omega_i$. If $y_i \succ_i x_i$, continuity allows a slightly smaller bundle θy_i still preferred to x_i, where $\theta < 1$. Separation then implies $\theta p \cdot y_i \geq p \cdot \omega_i > 0$, and hence $p \cdot y_i > p \cdot \omega_i$. No strictly preferred bundle is affordable, so x is competitive.

Competitive allocations survive every replication by the preceding core inclusion result. The theorem therefore identifies the allocations surviving all replica coalition tests with the competitive allocations, under its stated preference and endowment assumptions.

14 Stable Sets, the Bargaining Set, and the Shapley Value

14.1 Imputations and credible objections

The core allows every feasible coalition improvement to block an allocation. Chapter 14 restricts objections by considering their possible answers. Stable sets require the proposed alternative to belong to the same set of stable allocations. The bargaining set, kernel, and nucleolus instead ask whether an objection has a suitable counterobjection. The Shapley value relates a player's payoff to his marginal contributions across coalitions.

Let (N, v) be a cohesive transferable payoff game, with $v(\varnothing) = 0$. Its imputation set is

$$X = \{x \in \mathbb{R}^N : x(N) = v(N),\ x_i \geq v(\{i\}) \text{ for all } i\}.$$

Cohesiveness makes X nonempty. It is compact and convex. All solutions based on imputations in this chapter use this domain.

14.2 Stable sets

An imputation x dominates y through coalition S, written $x \succ_S y$, if $x_i > y_i$ for every $i \in S$ and $x(S) \leq v(S)$. Coalition S can support the proposed payoff for its members using its own worth, although the full proposal x distributes the grand coalition's worth.

Definition 14.1 (Definition 279.1). A set $Y \subseteq X$ is stable if it satisfies:

1. Internal stability: no member of Y dominates another member through any coalition.
2. External stability: every imputation outside Y is dominated through some coalition by a member of Y.

Let $D(Y)$ be the set of imputations dominated by a member of Y. Both conditions together are $Y = X \setminus D(Y)$. Unlike the core, which is one well defined set, a game may have several stable sets or no stable set.

Proposition 14.2 (Proposition 279.2). Every stable set contains the core. No stable set is a proper subset of another stable set. If the core itself is stable, it is the unique stable set.

Proof. A core imputation has no dominating imputation, so external stability prevents it from being outside Y. If $Y \subsetneq Y'$ were both stable, any $x \in Y' \setminus Y$ would be dominated by a member of Y, contradicting internal stability of Y'. The final assertion follows from these two facts.

Example 14.3 (Three player majority game, Example 279.3). Let every pair and the grand coalition have worth one, and every singleton have worth zero. Its core is empty. The set

$$Y = \{(1/2, 1/2, 0), (1/2, 0, 1/2), (0, 1/2, 1/2)\}$$

is stable. Between two members, only one player gains, so no winning coalition strictly gains. Outside this set, some two players each receive less than $1/2$ and can dominate through their equal division.

There are also infinitely many stable sets. Fix a player i and $c \in [0, 1/2)$, and set $Y_{i,c} = \{x \in X : x_i = c\}$. Internal stability follows because a strict gain for one of the other players requires a loss for the other. If $z_i > c$, the two others can both gain by reducing i's payoff to c. If $z_i < c$, choose the less rewarded of the other two players, say j, and give (i, j) the shares $(c, 1 - c)$, with zero to the remaining player. Both i and j strictly gain. Thus external stability also holds.

14.3 The bargaining set

Definition 14.4 (Definition 282.1). At imputation x, an objection of i against j is (y, S) with $i \in S$, $j \notin S$, $y(S) = v(S)$, and $y_k > x_k$ for every $k \in S$. A counterobjection is (z, T) with $j \in T$, $i \notin T$, $z(T) = v(T)$, and

$$z_k \geq y_k \quad (k \in T \cap S), \qquad z_k \geq x_k \quad (k \in T \setminus S).$$

The bargaining set consists of imputations at which every objection has a counterobjection.

A countercoalition must match the objector's offer to members recruited from the objecting coalition. It need only preserve the original payoff of members who did not join that coalition. The objector is excluded, and the player against whom the objection was made is included. Since core imputations have no objections, the core is contained in the bargaining set.

Example 14.5 (The majority game's bargaining set). In the three player majority game, an objection by i against j uses the pair $\{i, h\}$, where h is the third player. It can give h any payoff strictly below $1 - x_i$ and above x_h. A counterobjection by j can use $\{j, h\}$ only if $x_j + y_h \leq 1$; the singleton $\{j\}$ suffices when $x_j = 0$. If $x_j > x_i$, then $x_j > 0$ and an offer y_h sufficiently close to $1 - x_i$ has no counterobjection. Thus membership requires $x_i \geq x_j$ for every ordered pair, giving

$$\mathcal{B}(v) = \{(1/3, 1/3, 1/3)\}.$$

At equal division every objection can indeed be answered by the other pair.

14.4 Excess and the kernel

The excess of a coalition at x is $e(S, x) = v(S) - x(S)$. A positive excess is the gain available to the coalition if it leaves; a negative excess means its members jointly receive more than its standalone worth. For $i \neq j$, define

$$s_{ij}(x) = \max_{S:\ i \in S,\ j \notin S} e(S, x).$$

This is the largest excess available to a coalition including i and excluding j.

Definition 14.6 (Kernel, Definition 284.1). An imputation x belongs to the kernel if, for every ordered pair i, j,

$$s_{ij}(x) \leq s_{ji}(x) \qquad \text{or} \qquad x_j = v(\{j\}).$$

If both players receive more than their singleton worth, the two maximal excesses must be equal. If j already receives only his singleton worth, a larger s_{ij} cannot justify reducing j's payoff further within the imputation set. Comparing coalition excesses requires transferable and comparable payoff units.

Lemma 14.7 (Lemma 285.1). The kernel is contained in the bargaining set.

Proof. Take a bargaining set objection (y, S) of i against j. If $x_j = v(\{j\})$, the singleton $\{j\}$ is a countercoalition. Otherwise choose T containing j but not i attaining $s_{ji}(x)$. The kernel gives

$$v(T) - x(T) \geq v(S) - x(S) = \sum_{k \in S} (y_k - x_k).$$

Since $i \in S \setminus T$ and all gains in S are positive, the right side strictly exceeds $\sum_{k \in S \cap T} (y_k - x_k)$. Coalition T can therefore match y on its intersection with S and preserve x for its other members.

14.5 The nucleolus

Let $E(x)$ list all coalition excesses in weakly decreasing order. The first entry is the largest dissatisfaction, the next is the second largest, and so on. Lexicographic comparison first compares the largest entry, then the next entry if the first entries tie.

Definition 14.8 (Definitions 286.1 and Lemma 286.2). The nucleolus is the set of imputations that lexicographically minimize $E(x)$. Equivalently, whenever an alternative imputation y lowers some coalition's excess, $e(S, y) < e(S, x)$, there is a coalition T with

$$e(T, y) > e(T, x), \qquad e(T, y) \geq e(S, x).$$

The equivalent formulation requires a counterobjection whenever a proposed change reduces a coalition's excess: another coalition must then suffer an increase to at least the original excess being reduced. It compares complete imputations, whereas the bargaining set compares coalition offers.

Proposition 14.9 (Propositions 288.1 and 288.4). For a finite game with nonempty imputation set, the nucleolus exists and is unique.

Existence. Each ordered excess $E_k(x)$ is continuous, being an order statistic of finitely many affine functions. Minimize E_1 over compact X, then minimize E_2 over the compact minimizer set, and continue. Every stage leaves a nonempty compact set, and there are finitely many excesses.

Uniqueness. Suppose distinct x, y both minimize the ordered excess vector. At the first excess level for which the coalitions attaining that level differ, the midpoint $(x + y)/2$ leaves every earlier excess unchanged but reduces either the size or the number of excesses at that level. Affinity of each coalition excess gives this strict lexicographic improvement. Thus each coalition must have the same excess at x and y. Applying this to singleton coalitions gives $x = y$.

Proposition 14.10 (Lemma 287.1 and Corollary 288.3). The nucleolus is contained in the kernel, which is contained in the bargaining set. Hence both the kernel and the bargaining set are nonempty for the games under consideration.

If $s_{ij}(x) > s_{ji}(x)$ and $x_j > v(\{j\})$, transfer a sufficiently small amount from j to i. Excesses of coalitions containing i but not j decrease, and those containing j but not i increase. The largest excesses affected are in the first group. Choosing a small transfer makes the ordered excess vector lexicographically smaller, so x cannot be the nucleolus. Also, when the core is nonempty, the nucleolus belongs to it: the grand coalition's excess is always zero, so minimizing the largest excess attains zero exactly on the core.

Example 14.11 (One large participant and three small participants). Let a coalition win one unit if it contains player 1 and any other player, or if it contains all of players 2, 3, and 4. The core is empty. The bargaining set calculated in Example 282.3 is

$$\{(1 - 3\alpha, \alpha, \alpha, \alpha) : 1/7 \leq \alpha \leq 1/5\}.$$

Within this set, $s_{12} = 2\alpha$ and $s_{21} = 1 - 3\alpha$. Both players are above their singleton payoffs, so the kernel requires equality, giving

$$x = (2/5, 1/5, 1/5, 1/5).$$

The kernel is a singleton and contains the nucleolus, so this is also the nucleolus. The example distinguishes a set of defensible coalition offers from the sharper comparison of maximum excesses.

14.6 The Shapley value and marginal contributions

A value assigns a unique efficient payoff vector to every game. For an ordering R of N, let $S_i(R)$ be the players preceding i. Player i's marginal contribution to a coalition S not containing him is $\Delta_i(S) = v(S \cup \{i\}) - v(S)$.

Definition 14.12 (Definition 291.2). The Shapley value is

$$\begin{aligned}\phi_i(v) &= \frac{1}{n!}\sum_R \Delta_i(S_i(R)) \\ &= \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!}\,[v(S \cup \{i\}) - v(S)].\end{aligned}$$

The coefficient is the probability that exactly the players in S precede i in a uniformly random ordering: order S, then place i, then order the remaining players. Along every ordering the marginal contributions telescope to $v(N)$, so $\sum_i \phi_i(v) = v(N)$.

14.7 Balanced contributions

For $S \subseteq N$, let v_S be the game restricted to players in S. A value ψ has balanced contributions if, for every distinct i, j,

$$\psi_i(N, v) - \psi_i(N \setminus \{j\}, v_{N\setminus\{j\}}) = \psi_j(N, v) - \psi_j(N \setminus \{i\}, v_{N\setminus\{i\}}).$$

The payoff change imposed on i by the removal of j equals the change imposed on j by the removal of i. The condition evaluates both smaller games by the same value used for the original game.

Proposition 14.13 (Proposition 291.3). The Shapley value is the unique efficient value with balanced contributions.

Uniqueness. Induct on the number of players. If two such values agree on smaller games, balanced contributions implies that their differences in the current game are the same for every player. Efficiency makes the sum of these differences zero, so each difference is zero. For the Shapley formula, existence follows by grouping coalitions excluding i, j and using

$$\Delta_i(S) - \Delta_j(S) = \Delta_i(S \cup \{j\}) - \Delta_j(S \cup \{i\}).$$

The associated ordering coefficients sum to the coefficients for the game with one player removed.

14.8 The axiomatic characterization

Two players are interchangeable if their marginal contributions to every coalition containing neither are equal. Player i is a dummy if $\Delta_i(S) = v(\{i\})$ for every S excluding him.

Theorem 14.14 (Proposition 293.1). On games with a fixed player set, the Shapley value is the unique efficient value satisfying symmetry between interchangeable players, payment of $v(\{i\})$ to every dummy, and additivity

$$\psi(v + w) = \psi(v) + \psi(w).$$

Proof. The marginal contribution formula satisfies each property directly. To prove uniqueness, define the unanimity game $u_T(S) = 1$ if $T \subseteq S$ and zero otherwise. These games form a basis: coefficients in $v = \sum_{T \neq \varnothing} \alpha_T u_T$ can be determined recursively by

$$\alpha_T = v(T) - \sum_{S \subsetneq T} \alpha_S.$$

In $\alpha_T u_T$, players outside T are zero dummies and players in T are interchangeable. Efficiency and symmetry require equal shares $\alpha_T/|T|$ for members of T. Additivity therefore determines the value of every game.

Each axiom imposes an independent restriction. Averaging marginal contributions only over orderings with a fixed player first preserves additivity and dummy payments but can violate symmetry. Equal division of $v(N)$ preserves symmetry and additivity but can reward a zero dummy. Paying dummies their worths and equally splitting the remainder preserves symmetry and dummy payments but need not be additive.

Example 14.15 (Voting and market applications). In the weighted voting game with weights $(1, 1, 1, 2)$ and quota three, player 4 is pivotal exactly when placed second or third in an ordering. Therefore

$$\phi = (1/6, 1/6, 1/6, 1/2).$$

These voting weights produce the same winning coalitions as the preceding four player example, after relabeling. The Shapley value differs from its nucleolus because it averages marginal contributions instead of minimizing ordered excesses.

With one seller and two identical buyers, the worth is one whenever the seller and at least one buyer are present. The seller is pivotal in four of the six orderings, while each buyer is pivotal in one. Thus $\phi = (2/3, 1/6, 1/6)$, whereas the core is $(1, 0, 0)$. In larger replicas with strictly more buyers than sellers, a seller is increasingly likely to meet excess buyer demand among predecessors. His Shapley payoff converges to one and a buyer's to zero.

In a convex game, every ordering's marginal contribution vector lies in the core by Exercise 260.4. The core is convex, so their average, the Shapley value, also lies in the core. The market example shows that this inclusion is not a general property of arbitrary games.

14.9 Cost sharing

Let $C(S)$ be the cost of supplying a service to coalition S. Applying the Shapley formula to C assigns player i his expected incremental cost when users are added in random order. Efficiency becomes exact cost recovery; symmetry treats users with identical incremental costs equally; the dummy condition charges a user whose incremental cost is constant exactly that cost. Additivity means that charges for separate services add.

Example 14.16 (Additional example: a common facility). Suppose any nonempty set of users requires a fixed cost F, and serving user i additionally costs c_i. Then

$$C(S) = F + \sum_{i \in S} c_i \qquad (S \neq \varnothing).$$

Player i adds c_i in every ordering and adds the fixed cost only when first, an event of probability $1/n$. His charge is therefore

$$\phi_i(C) = c_i + \frac{F}{n}.$$

These charges sum to $C(N)$. For a cost game the coalition stability inequality would require its members' total charge not to exceed their standalone cost, which reverses the payoff core inequality.

15 The Nash Solution

15.1 Bargaining problems and risk preferences

A bargaining problem specifies the agreements two players can jointly obtain, the consequence of disagreement, and their preferences over lotteries on these consequences. Unlike an alternating offers game, it does not specify who makes offers or how long negotiations take. Risk attitudes enter through preferences over lotteries, including the risk of disagreement.

Definition 15.1 (Definition 300.1). A bargaining problem is $(X, D, \succeq_1, \succeq_2)$, where X is a compact set of agreements, $D \in X$ is disagreement, and the preferences on lotteries over X are continuous and satisfy expected utility. Every agreement is weakly preferred to disagreement by both players, and some agreement is strictly preferred by both. The agreement set is rich enough that every lottery between two agreements is equivalent for both players to a deterministic agreement. Agreements giving both players identical utilities are identified. Each player has a unique best agreement B_i, and the other player is indifferent between B_i and disagreement.

Choose expected utility representations u_i with $u_i(D) = 0$. The attainable utility set

$$U = \{(u_1(x), u_2(x)) : x \in X\}$$

is compact and convex, lies in $\mathbb{R}^2_+$, and contains a point with two positive coordinates. Nonredundancy identifies a utility pair with a unique agreement. In the usual utility set formulation, disagreement can instead be any $d \in U$, and the relevant gains are $u_i - d_i$.

Write $p \cdot x$ for the lottery that selects agreement x with probability p and disagreement otherwise. With the normalization above, $u_i(p \cdot x) = p u_i(x)$. The same deterministic ordering can be represented by many utility functions, but preferences over lotteries determine the utility scale up to a positive affine transformation.

15.2 Objections and the product characterization

Definition 15.2 (Definition 301.2). An agreement x^* is the Nash solution if, for either player i and the other player j, every $p \in [0, 1]$ and $x \in X$ satisfy

$$p \cdot x \succ_i x^* \quad \Longrightarrow \quad p \cdot x^* \succeq_j x.$$

A player who demands x while accepting a breakdown probability $1 - p$ must prefer the resulting risky demand to the current agreement. The other player can answer by preferring to retain the current agreement under the same risk rather than concede x. At the Nash solution every objection has such an answer.

Theorem 15.3 (Proposition 302.1). The Nash solution exists, is unique, and maximizes

$$u_1(x)u_2(x) \qquad \text{over } x \in X$$

when utilities are normalized by $u_i(D) = 0$. With disagreement payoff d, it maximizes $(u_1 - d_1)(u_2 - d_2)$ over individually rational attainable payoffs.

From product maximization to objections. A maximizing agreement has positive utilities because some feasible agreement gives both players positive utility. If $pu_i(x) > u_i(x^*)$, then

$$pu_i(x)u_j(x^*) > u_i(x^*)u_j(x^*) \geq u_i(x)u_j(x).$$

Dividing by $u_i(x) > 0$ gives $pu_j(x^*) > u_j(x)$, so the other player can counterobject.

From objections to product maximization. An agreement satisfying the objection condition must give both players positive utility. If one received zero, an agreement with two positive utilities and a sufficiently small positive p would yield an unanswered objection. Now take x with $u_i(x) > u_i(x^*)$ and both utilities positive. Every $p > u_i(x^*)/u_i(x)$ yields an objection, so $pu_j(x^*) \geq u_j(x)$. Letting p decrease to that ratio gives

$$u_i(x^*)u_j(x^*) \geq u_i(x)u_j(x).$$

If neither player's utility is higher at x, the product comparison holds directly.

Existence and uniqueness. Compactness gives a maximizer, which lies where both utilities are positive. There the logarithm of the product is $\log u_1 + \log u_2$, a strictly concave function of the utility pair. Convexity of U therefore gives a unique maximizing pair. Nonredundancy gives a unique agreement.

Positive rescaling of either normalized utility multiplies every product by the same positive constant and does not alter the maximizer. In the formulation with $d \neq 0$, using gains $u_i - d_i$ also makes the solution invariant to shifts of the utility origin.

15.3 The efficient frontier and worked solutions

Suppose the individually rational frontier is $u_2 = g(u_1)$, with positive utilities at the solution and differentiable g. Maximizing $u_1 g(u_1)$ gives

$$g(u_1) + u_1 g'(u_1) = 0, \qquad -g'(u_1) = \frac{u_2}{u_1}.$$

Thus the frontier slope equals the ratio of the two utility gains. Equality of physical shares follows only when the preferences and feasible set make the problem symmetric.

Example 15.4 (A divisible surplus). Two risk neutral players divide a surplus V and have disagreement utilities d_1, d_2, with $V > d_1 + d_2$. On the efficient frontier $x_2 = V - x_1$, maximize

$$(x_1 - d_1)(V - x_1 - d_2).$$

The derivative is $V - d_2 + d_1 - 2x_1$, so

$$x_1 = d_1 + \frac{V - d_1 - d_2}{2}, \qquad x_2 = d_2 + \frac{V - d_1 - d_2}{2}.$$

The solution divides the surplus over disagreement equally.

Example 15.5 (Risk aversion and the physical division). For a unit pie and zero disagreement, let player 1 have utility x^a and player 2 utility $(1-x)^b$, with $a, b > 0$. The Nash solution maximizes $a \log x + b \log(1-x)$ and satisfies

$$\frac{a}{x} = \frac{b}{1-x}, \qquad x = \frac{a}{a+b}.$$

When $0 < a, b \leq 1$, these are concave utilities for money. Reducing a makes player 1 more risk averse and reduces his physical share.

15.4 Comparative statics of risk aversion

Player 1's preferences become more risk averse when her ordering of deterministic agreements stays fixed and a lottery equivalent to a certainty under the old preferences becomes no better than that certainty under the new preferences. For utility representations, this corresponds to an increasing concave transformation $u_1' = f(u_1)$, normalized so $f(0) = 0$.

Proposition 15.6 (Proposition 304.1). If only player 1 becomes more risk averse and both resulting problems satisfy the bargaining assumptions, the new Nash agreement is no better for player 1 than the original agreement, evaluated by her unchanged ordering of deterministic agreements.

The product characterization shows why. For $r > 0$, concavity and $f(0) = 0$ imply that $f(r)/r$ is nonincreasing. The new product is the old product multiplied by $f(u_1)/u_1$. Suppose the new maximizer gave player 1 strictly more utility than the original one. The old product would be strictly lower there, by uniqueness of the old solution, and its multiplier would be weakly lower. It could not maximize the new product. Greater aversion to the risk of disagreement weakens the player's position under this solution concept.

15.5 The axioms on agreements and preferences

The book's axiomatization works directly with agreements and preferences. Pareto efficiency requires that no agreement weakly improve both players and strictly improve one. A symmetric problem has a relabeling $\varphi : X \to X$ that fixes disagreement, satisfies $\varphi(\varphi(x)) = x$, and interchanges the players' preferences over every lottery. Symmetry of a solution requires its selected agreement to be fixed by φ.

The book's independence axiom changes a player's risk preferences while keeping the agreement set and the ranking of deterministic agreements fixed. If the selected agreement is x^*, the allowed change to player i has two features:

1. Whenever $x \succeq_i x^*$ and $p \cdot x \sim_i x^*$, the new preferences make $p \cdot x$ no better than x^*.
2. Whenever $x \preceq_i x^*$ and $x \sim_i p \cdot x^*$, that indifference is preserved.

The first condition makes demands above x^* no more attractive. The second preserves the player's willingness to resist the other player's demands. Independence requires the selected agreement to remain x^* under this change.

Theorem 15.7 (Proposition 307.2). The Nash solution is the unique bargaining solution satisfying Pareto efficiency, symmetry, and the independence axiom of the book's Section 15.3.1.

Product maximization gives efficiency, and relabeling the players preserves the product and hence gives symmetry. The objection characterization establishes independence: the changed preferences create no new effective objection by i and preserve his counterobjections.

For uniqueness, normalize utilities at the Nash agreement to $u_1(x^*) = u_2(x^*) = 1$. Every feasible agreement then satisfies

$$u_1(x) + u_2(x) \leq 2.$$

Otherwise, mixing a small amount of that agreement with x^* would increase the product: for $\lambda > 0$ small,

$$\prod_{i=1}^{2}[1 + \lambda(u_i(x) - 1)] = 1 + \lambda[u_1(x) + u_2(x) - 2] + O(\lambda^2) > 1.$$

The proof changes risk preferences above and below x^* to produce a symmetric problem with linear utility frontier $u_1 + u_2 = 2$. Efficiency and symmetry select $(1, 1)$ there. The independence axiom then returns to the original preferences one player at a time without changing the selected agreement.

15.6 The standard utility set axioms

In the standard formulation (U, d), the corresponding characterization uses Pareto efficiency, symmetry, covariance under separate positive affine transformations of utilities, and contraction consistency:

$$U \subseteq U', \quad F(U', d) \in U \quad \Longrightarrow \quad F(U, d) = F(U', d).$$

Contraction consistency is the usual independence of irrelevant alternatives axiom. It compares different feasible utility sets. This differs in formulation from the book's axiom, which holds the physical agreement set fixed and changes risk preferences.

The geometric proof is particularly simple. Normalize $d = (0, 0)$ and the Nash point to $(1, 1)$. The supporting inequality just proved places U inside the triangle $\{u \geq 0 : u_1 + u_2 \leq 2\}$. Symmetry and efficiency select $(1, 1)$ from the triangle; contraction consistency selects the same point from U. Affine covariance reverses the normalization.

15.7 Independence of the axioms and alternative solutions

The axioms do not duplicate one another. Always selecting disagreement preserves symmetry and independence but violates efficiency. The weighted product

$$u_1(x)^{\alpha} u_2(x)^{1-\alpha}, \qquad 0 < \alpha < 1,$$

gives an efficient solution satisfying independence, but violates symmetry when $\alpha \neq 1/2$. Its uniqueness and invariance to utility scales follow from the same logarithm argument as for the symmetric product.

The Kalai–Smorodinsky solution is efficient and symmetric but generally differs from the Nash solution. With zero disagreement, it selects equal fractions of the ideal utilities:

$$\frac{u_1(x)}{u_1(B_1)} = \frac{u_2(x)}{u_2(B_2)}.$$

Its dependence on each player's best attainable utility explains why it can change when an unchosen feasible alternative changes an ideal point. It therefore does not satisfy the Nash independence axiom in general.

15.8 Alternating offers with a risk of breakdown

Consider alternating offers over X. After each rejected offer, negotiations break down with probability $\alpha \in (0,1)$ and disagreement D occurs. Let $x^*(\alpha)$ and $y^*(\alpha)$ be the efficient stationary proposals of players 1 and 2. Under the regularity assumptions for the alternating offers result, responder indifference gives

$$u_1(y^*(\alpha)) = (1-\alpha)u_1(x^*(\alpha)), \qquad u_2(x^*(\alpha)) = (1-\alpha)u_2(y^*(\alpha)).$$

Both proposal points consequently have the same Nash product.

> Proposition 15.8 (Proposition 310.3). As the probability of breakdown following rejection tends to zero, both equilibrium proposals converge to the Nash solution.

The first player's proposal lies on his preferred side of the Nash solution, while the second player's proposal lies on the other side. On the efficient frontier, the product is strictly quasiconcave, so two distinct points with equal product bracket its unique maximizer. The responder equations make the utility difference between these points vanish as $\alpha \to 0$. Compactness and uniqueness then force their common limit to be the Nash point. The bargaining protocol and the axiomatic problem use different primitives, but small symmetric breakdown risk yields the same limiting agreement.

15.9 Exact implementation of the Nash solution

The book's Section 15.5 gives a finite stage procedure. Player 1 first proposes $y \in X$. Player 2 then chooses $x \in X$ and $p \in [0,1]$. With probability $1-p$ the game ends at disagreement. Conditional on survival, player 1 chooses between x and the lottery $p \cdot y$. Thus choosing the latter produces total unconditional agreement probability p^2.

> Proposition 15.9 (Proposition 312.1). This game form implements the Nash solution in subgame perfect equilibrium.

Continuation incentives. Normalize disagreement utilities to zero. At the final decision player 1 chooses x when $u_1(x) > pu_1(y)$ and chooses $p \cdot y$ when the inequality is reversed. Let the Nash utilities be (n_1, n_2) and let $F = n_1 n_2$. Write $Y_i = u_i(y)$ and $X_i = u_i(x)$.

If player 1 initially proposes the Nash agreement, acceptance of x requires $X_1 \geq pn_1$. Player 2's resulting payoff satisfies

$$pX_2 \leq \frac{X_1 X_2}{n_1} \leq n_2.$$

If player 1 chooses $p \cdot y$, player 2 gets $p^2 n_2 \leq n_2$. The choice $p = 1, x = y$ attains n_2 and gives player 1 n_1. Equality in either bound requires the final agreement to be the Nash agreement with probability one.

Why another initial proposal cannot improve player 1's payoff. If $Y_1 < n_1$, player 2 can choose the Nash agreement with $p = 1$, which player 1 strictly accepts, securing n_2. Any accepted agreement with $X_1 > n_1$ would have $X_2 < n_2$ by the unique product maximum and cannot be a best response for player 2. Any rejected agreement yields player 1 only $p^2 Y_1 < n_1$.

If $Y_1 > n_1$, player 2 can offer the Nash agreement with p just below n_1/Y_1, securing payoffs arbitrarily close to F/Y_1. In any best response that player 1 accepts, therefore,

$$pX_2 \geq \frac{F}{Y_1}, \qquad pY_1 \leq X_1 \leq \frac{F}{X_2} \leq pY_1.$$

All inequalities must be equalities. Thus x is the Nash agreement and $p = n_1/Y_1$, giving player 1 unconditional payoff $n_1^2/Y_1 < n_1$. Rejection cannot attain player 2's bound because $p^2 Y_2 \leq Y_2 < F/Y_1$, using strict product maximality for $Y_1 \neq n_1$.

For $Y_1 = n_1$, the same product bound shows that a best response gives player 1 at most n_1, with equality only through the Nash agreement. Hence player 1 can secure n_1 by proposing the Nash agreement and cannot profit from another proposal. A subgame perfect strategy profile exists by choosing x at final indifferences and maximizing player 2's feasible payoff in each continuation. Every subgame perfect equilibrium consequently has the same outcome: the Nash agreement without disagreement risk.