Overview

I Artificial Intelligence

II Problem Solving

III Knowledge, Reasoning, Planning

IV Uncertain Knowledge and Reasoning

- 12. Quantifying Uncertainty
- 13. Probabilistic Reasoning
- 14. Probabilistic Reasoning over Time
- 15. Probabilistic Programming
- 16. Making Simple Decisions
- 17. Making Complex Decisions
- **18. Multiagent Decision Making**

V Machine Learning

VI Communicating, Perceiving, and Acting

VII Conclusions

Multiagent Decision Making

- Properties of Multiagent Environments
- Non-Cooperative Game Theory
- Cooperative Game Theory
- Making Collective Decisions

Variants of Multiagent Systems

- One decision maker with multiple actors:
  - **Benevolent agent assumption**: Other agents will simply do what they are told
  - Synchronization of actions necessary
  - Communication constraints require decentralized planning (execution partially decoupled)
- Multiple decision makers:
    - With common goal
  - Without common goal (extreme case: zero-sum games like chess), but not necessary opposed, but with own preferences
    - Agent must consider preferences of other agents (recursive)
    - **Game theory**: Theory of strategic decision making with two goals
      - **Agent design**: Determine best strategy for an agent
      - **Mechanism design:** Define rules for environment of agents so that the collective good of all agents is maximized when each agent acts according to its preferences
        - Examples: Design of internet protocols, laws, economy rules, …

Models of Game Theory

- Game theory provides a range of different models with underlying assumptions
- Most important distinction:
  - **Cooperative games**
    - It is possible to have binding agreements between agents (with verification option)
  - **Non-cooperative games**
    - No binding agreement is possible (but not necessarily competitive)

Multiagent Planning

- Find a plan, that – if executed by the actors – achieves the goal
- Issues:
  - **Concurrency** (plans of different agents may be executed simultaneously)
  - Interactions of actions of different agents (resources, exclusion, precondition)
- Modes of concurrency:
  - **Interleaved excecution:** The sequence of actions of two plans can be arbitrary mixed if the sequence of the actions of each plan remains stable
  - **True concurrency:** Instead of full serialized orderings of the actions only a partial ordering is specified
  - **Synchronization:** A global clock guarantees that action occur truely simultaneously
    - In real world difficult to achieve, but with a simple semantics

Multiagent Planning with Synchronization

- The transition model for a single-agent deterministic case is the function Result (s,a)
  - There might be b different choices for the action
- In multiactor setting with n actors, the single action is replaced by a joint action $\langle a_1, \ldots a_n \rangle$
  - Transition model for b^n different joint actions with a branching factor of b^n
- Solution idea: **Loosely coupled** actors
  - Pretend the problems are decoupled (like for single actors) and then fix up the interactions

Double Tennis Example for Loosely Coupled Actors

$Action(Hit(actor, Ball),$
$\quad \text{CONCURRENT}: \forall b \;\; b \neq actor \Rightarrow \neg Hit(b, Ball)$
$\quad \text{PRECOND}: Approaching(Ball, loc) \land At(actor, loc)$
$\quad \text{EFFECT}: Returned(Ball)).$

$Actors(A, B)$
$Init(At(A, LeftBaseline) \land At(B, RightNet) \land$
$\quad Approaching(Ball, RightBaseline) \land Partner(A, B) \land Partner(B, A)$
$Goal(Returned(Ball) \land (At(x, RightNet) \lor At(x, LeftNet))$
$Action(Hit(actor, Ball),$
$\quad \text{PRECOND}: Approaching(Ball, loc) \land At(actor, loc)$
$\quad \text{EFFECT}: Returned(Ball))$
$Action(Go(actor, to),$
$\quad \text{PRECOND}: At(actor, loc) \land to \neq loc,$
$\quad \text{EFFECT}: At(actor, to) \land \neg At(actor, loc))$

Plan 1: A: [Go (A, RightBaseline, Hit (A, Ball)]

B: [NoOP(B), NoOp (B)]

Extensions: Add concurrent action constraints to action schemas,

- e.g. „Hit“ is not successful, it two actors hit the ball concurrently
- Other actions may successful only, if actors perform the same action (e.g. carry)

Any planning algorithms can be used with these extended kinds of actions schemas

Cooperation and Coordination

- True multiagent setting assumption: Each agent makes its own plan
- Difference to multiagent planning with synchronization: There might be several working plans
- How to agree on the same plan?

Plan 1: A: [Go (A, RightBaseline), Hit (A, Ball)]
B: [NoOP(B), NoOp (B)]

Plan 2: A: [Go (A, LeftNet), NoOP(A)],
B: [Go (A, RightBaseline) Hit (B, Ball)]

Two approaches:

- **Convention**: Any constraint on the selection of joint plans
  - Example: Stick to your side of the court
  - Widespread in society, e.g. Right-hand traffic, same language in a country, social laws, …
- **Communication:**
  - Explicit: Verbal communication, e.g. a player shout „Mine“
  - Implicit: One agent starts executing the first part of the plan and the other recognizes it

Non-Cooperative Game Theory

- **Normal form games**: Games with a single move
  - **Players** (agents) make a decision (2 or more)
  - **Actions** that players can choose
  - **Payoff function** that gives utility to each player for all action combinations
    - For two players often represented in a single **payoff matrix**
- Example: Two-finger Morra
  - Two players, O and E, simultaneously display one or two fingers. If the sum of fingers (f) is even, E gets f € from O, if it is odd, O gets f € from E
  - Resulting payoff matrix:

	O: one	*O: two*
E: one	$E=+2, O=-2$	$E=-3, O=+3$
E: two	$E=-3, O=+3$	$E=+4, O=-4$

- To perform rationally, all players must take into account, how the other players act

Prisoners Dilemma

- Famous example:
  - Two alleged burglars, Ali and Bo are caught and interrogated separately. They are offered a deal: If you testify against your partner as the leader of a burglar ring, you‘ll go free, while the partner will serve 10 years in prison. However, if both testify against each other, you‘ll both get 5 years.
  - Resulting payoff matrix:

	Ali:testify	*Ali:refuse*
Bo:testify	$A = -5, B = -5$	$A = -10, B = 0$
Bo:refuse	$A = 0, B = -10$	$A = -1, B = -1$

  - Variants as repeated game or with other payoff matrix including moral reputation

Important Concepts (1)

- **Strategy** in game theory (policy in MDPs)
  - **Pure strategy** (deterministic policy): For single-move games just a single action
  - **Mixed strategy** (randomized policy): Selects actions according to a probability distribution
- **Strategy profile**: Assignment of a strategy to each player
  - Implies an **outcome** for each player (an expected utility in case of mixed strategies)
- **Dominant strategy:**
  - Strategy for one player with a better or at least equal outcome, regardless what the other player chooses (**strongly dominant** or **weakly dominant**)
  - Examples:
    - Two-finger Morra: no dominant strategy
    - Prisoners dilemma: Testify is a dominant strategy (although not optimal for both)
  - A rational player will always choose a dominant strategy
- **Dominant strategy equilibrium**: All players have dominant strategy

Important Concepts (2)

- **Nash equilibrium**:
  - Weaker concept than dominant strategy: If each player has choosen a strategy, no player profit from changing its strategy
  - Local stable point in a game, but a game may contain several Nash equilibria
  - Example with two Nash equilibria (t,l) and (b,r):

	$Ali{:}l$	$Ali{:}r$
$Bo{:}t$	$A = 10, B = 10$	$A = 0, B = 0$
$Bo{:}b$	$A = 0, B = 0$	$A = 1, B = 1$

  - **Coordination problem**: How can both players reach the better Nash equilibrium (t,l)?
    - Possible solution: **Focal points**: Preferring the best Nash equilibrium

Nash Equilibrium with Mixed Strategy

- Not all games have at least one Nash equilibrium as pure strategy, but always as mixed strategy
- Example: coin flipping (winner gets 1 € from other player):

	Ali:heads	*Ali:tails*
Bo:heads	$A = 1, B = -1$	$A = -1, B = 1$
Bo:tails	$A = -1, B = 1$	$A = 1, B = -1$

  - Mixed Nash equilibrium for both players: Choose head and tail with 50% probability
    - If one player chooses another probability, the other can exploit this

Pareto Optimality

- Change of perspective: Design of a game and in particular the pay-off matrix to optimize the outcome for the society as a whole (i.e. for all players)
- **Pareto optimality**: Avoid outcomes that waste utility
  - Def.: There is no other outcome, that would make one player better off without someone else worse off
- Application to Prisoners dilemma: Dominant strategy is not Pareto optimal

Utilitarian social welfare

- **Utilitarian social welfare**: Measure how good an outcome is in the aggregate
  - **Sum of utilities** of all players has two problems:
    - Ignoring the distribution
    - Assuming a common scale for utilities
  - Alternatives for distribution:
    - **Maximin approach**: Maximize utility of worst-off member
    - **Gini Coefficient**: Summarizes, how evenly utility is spread among players
    - Problem: May sacrifice a great deal of social welfare for small distributional gains
  - Alternatives for common scale:
    - Subjective quantity
    - Problem: Utility monster: „I love cookies a thousand times more than anyone else"

Computing Equilibria

- For pure strategies (without randomization)
  - **Exhaustive search** (simplest approach) to find dominant strategies and Nash equilibria
  - **Myopic best response** (**iterated best response**):
    - Start with random strategy profile and check iteratively for each player, whether it can improve by changing its strategy
      - Terminates for important class of games (not all) with a Nash equilibrium
- For mixed strategies and zero-sum games
  - **Maximin** technique (von Neuman, 1928):
    - Exploit the fact, that the gain of one player is the loss of the other
    - Generate n linear equations for n actions (linear programming problem)

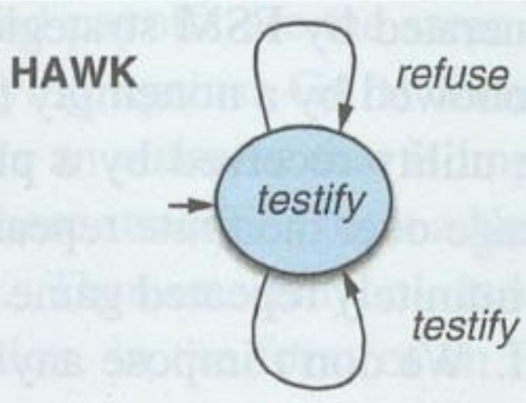
Example of Maximin for Two-finger Morra

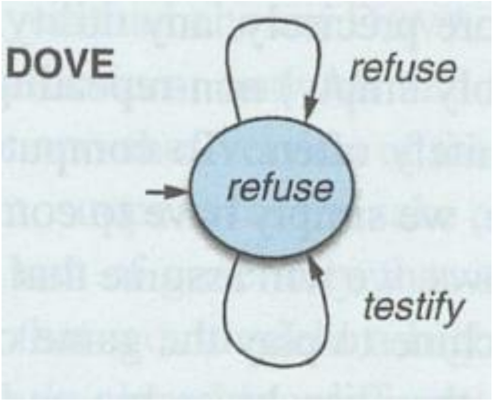
	O: one	*O: two*
E: one	$E=+2, O=-2$	$E=-3, O=+3$
E: two	$E=-3, O=+3$	$E=+4, O=-4$

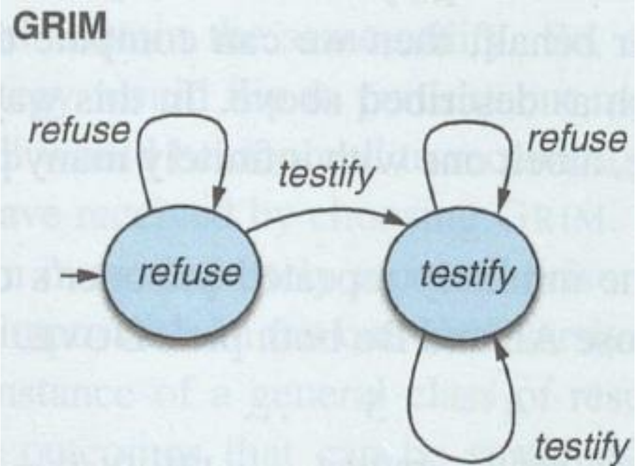
- Define the mixed strategy for E „one“ with probability p and „two“ with probability 1-p
- The outcome for E is:
  - If O chooses always „one“: 2p -3(1-p)
  - If O chooses always „two“: -3p + 4(1-p)
  - In the intersection holds: 5p – 3 = 4 – 7p, i.e. p=7/12 („one“) for E
- Similar for O with the same result (7/12: „one“, 5/12: „two“)
- Result: for O: +1/12; for E: -1/12
- Justification: Mixed strategies for O are not better, if E reveals its strategy
  - Mixed strategy for O (p U_{one} + (1-p) U_{two}) cannot be better than the better of U_{one} or U_{two}

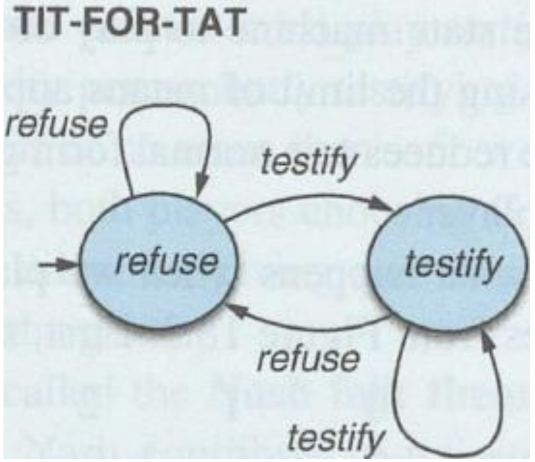
Repeated (Iterated) Games

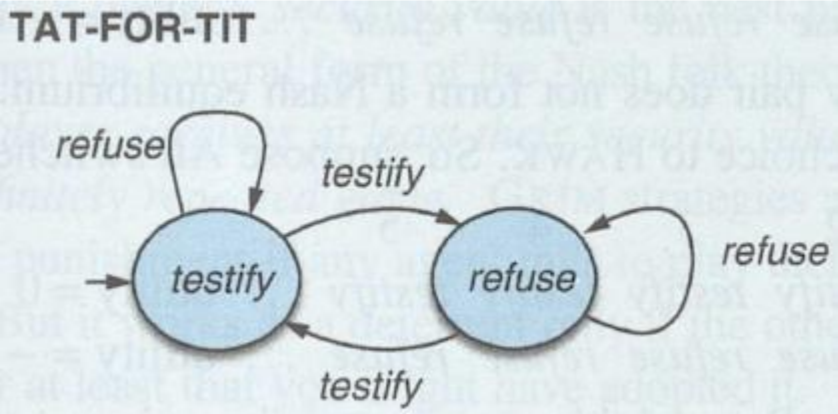
- Simplest form of multiple-move games: Players repeatedly play rounds of a single-move game
  - If both players know the number of rounds, their strategies don‘t change
    - Reason: The last round is a single-move game. Therefore, the last but one round is also a single-move game etc.
  - With infinetely many rounds (or an unknown number of many rounds) strategies change:
  - Sequence utility: $\lim_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T} U_t$
  - Strategy evolution: Dove/Dove → Hawk/Dove → Hawk/Hawk → Hawk/Grim → Grim/Grim











Nash Folk Theorem

- **Nash Folk Theorem:** Every outcome in which every player receives at least their security value can be sustained as a Nash equilibrium in an infinitely repeated game
  - Security value: The payoff that the player could guarantee to obtain
  - Grim strategies are the key to the folk theorems (i.e. the mutual threat of punishment if any agent fails to play its part in the desired outcome), if the other player believes you have adopted this strategy

Further Issues

- Arbitrary sequential games
- Sequential games with chance are similar to stochastic games
- Simultaneous moves with additional constraint
- Capturing imperfect information: Extends stochastic partially observable games (like Kriegsspiel) by representing the belief states (information sets) of all players at once
- Uncertain payoffs
- Assistance games (a robot assistant tries to find out the preferences of a human)

Cooperative Game Theory

- If groups of agents cooperate, the group as a whole should get a higher utility than the sum of utilities each agent gets individually
- This gain can be distributed back to the agents
- Problems:
  - Coalition structures
    - A grand coalition (with all players) should be better for each player than any other coalition
  - How should the gain be distributed?

Distribution of Gains in Coalitions to Players

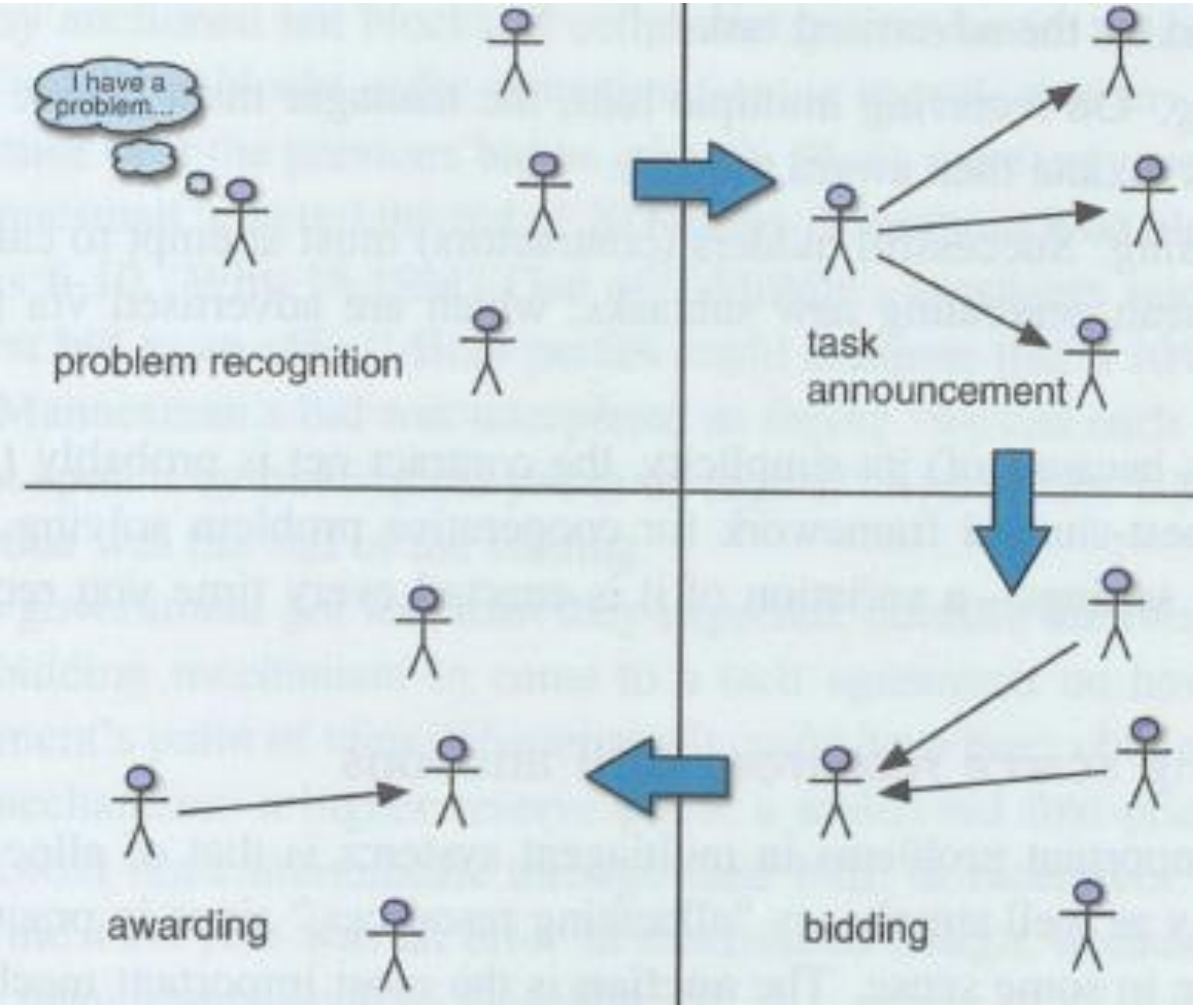
- Not evenly, but proportional to the player‘s contribution to the coalition
  - Simple, but unsatisfactory approach: Proportional to difference between the payoff of the coalition with and without a player
    - Problem: Assumption, that a player i is the last player to enter the coalition
  - Best approach: **Shapley value**: A player should get the average marginal contribution that player i makes, over all possible orderings of the players, to the set of players preceding player i in the ordering

Making Collective Decisions

- The problem of designing the right game for a collection of agents to play (mechanism design)
  - A language for describing the set of allowable strategies that agents may adopt
  - A distinguished agent („center“) that collects reports of strategy choices from the agents (e.g. the auctioneer is the center of an auction)
  - An outcome rule, known to all agents, that the center uses to determine the payoffs to each agent, given their strategy choices

Allocating Tasks with Contract Net

- Contract net protocol
  - Agent recognizes problem
  - Agent announces task to other agents
  - Other agents make offers (bidding)
  - Agent select best offer (awarding)
- Best studied framework for cooperative problem solving with the main tasks:
  - Task announcement processing
  - Bid processing
  - Award processing



Allocating Scarce Resources with Auctions

- Specification of an auction:
  - Single resource
  - Each bider has a private value for the item (not known to others)
- English auction (ascending-bid auction)
  - Center starts with mininum bid and increases it stepwise, until nobody is willing to pay it
  - Last bider wins the item
  - Biders have simple dominant strategy
  - Problem: High communication overhead, collusion possible (illegal agreement of bidders)
- Sealed bid auction
  - Each bidder makes single sealed bid to center.
  - Highest bid wins (little communication overhead)
  - Problem: No dominant strategy, since bidder should bid just more than the second bidder
- **Vickrey auction** *(*Sealed bid second price auction; *avoids both problems)*
  - Each bidder makes a single sealed bid to center
  - Highest bid wins, but pays second highest price

Vickrey-Clarke-Groves (VCG) Auction for Several Items

- Auction for n items (e.g. three free wireless internet tranceivers in a city)
- Each bider (e.g. each neighborhood of the city) makes a sealed offer
- The best n bidders get the item and but only pay the price of the (n+1) bidder

Common Goods

- Tragedy of the commons: If nobody has to pay for using a common resource, then it may be exploited and all suffer from this (e.g. environment pollution)
- Solution approach:
  - Change mechanism that one charges each agent for using the common
  - Difficult part: Setting the prices, e.g. with VCG auctions
  - Example: Carbon taxes

Voting

- What is the best mechanism for voting procedures? → **Social Choice Theory**
- **Concordoret‘s Paradox**:
  - Suppose there are three candidates (C1, C2, C3) and three voters (V1, V2, V3)
  - V1 has preferences (C1 > C2 > C3); V2: (C3 > C1 > C2); V3: (C2 > C3 > C1)
  - Whatever candidate is elected, 2/3 of the voters would prefer another candidate

Voting procedures

- **Simple majority voting** (for two candidates)
- **Plurality voting** (> two candidates): winner with most top choices, even without majority
- **Borda count**: Voters give all candidates a rank and candidate with highest rank sum wins
- **Approval voting**: Voters choose a subset of candidates (without ranking).
- **Instant runoff voting**: Voters rank all candidates. Candidate with majority of first-place votes wins, otherwise candidate with fewest first-place votes is eliminated and process is repeated based on the remaining ranked lists.
- **True majority voting**: Voters rank all candidates. Winner beats all other candidates in pairwise comparisons (however, in Condorcet paradox no candidate wins a majority)

- Multiple voting procedures, e.g. combination of plurality voting and simple majority voting with the two best candidates or leave-one-out voting (the voters who voted for a candidate who is left out, vote for a new candidate)

Bargaining (Negotiation)

- Bargaining is used when agents need to reach agreement on a matter of common interest
- Agents make offers (proposals, deals) to each other under specific protocols and either accept or recect offer.
  - **Alternating offers protocol:** Agents alternately make offers until one is accepted
    - Problem: infinitely many negotiation rounds
  - **Impatient agents:** The value of the item is discounted with a factor γ in each round
    - Fast agreement, but unrealistic protocol
  - **Monotonic concession protocol:** In each round the agents must make a concession, otherwise negotiation terminates.
    - Both agents make proposals simultaneously
      - Alternative: Sequentially?
    - **Zeuthen strategy:** An agent should make the smallest concession so that the other is willing to accept the deal