Summaries

Source

MULTIAGENT SYSTEMS, Algorithmic, Game-Theoretic, and Logical Foundations (Shoham)

Summary

Chapter 3

Normal-form game

Game theory is the mathematical study of interaction among independent, self-interested agents. The dominant branch of game theory is called **noncooperative game theory**. In this branch, the basic modelling unit is the individual (including his beliefs, preferences and possible actions). This in contrast to **cooperative game theory**, where the basic modelling unit is the group.

An agent's interests can be modelled using **utility theory**, which aims to quantify an agent's degree of preference across a set of available alternatives. An agent's **utility function** is a mapping from every state of the world to a real number which represents an agent's payoff for being in that state. When an agent needs to perform an action which may lead to different states with different probabilities, an expected utility value can be calculated for each action.

In game theory, the most familiar representation of game's strategic interactions is known as the **normal form**:

Definition 3.2.1 (Normal-form game) *A (finite, n-person)* normal-form game *is a tuple (N, A, u), where:*

- *N is a finite set of n players, indexed by i;*
- *$A = A_1 \times \cdots \times A_n$, where A_i is a finite set of* actions *available to player i. Each vector $a = (a_1, \ldots, a_n) \in A$ is called an* action profile*;*
- *$u = (u_1, \ldots, u_n)$ where $u_i : A \mapsto \mathbb{R}$ is a real-valued* utility *(or* payoff*)* function *for player i.*

As a utility function's domain is the set of states and not the set of possible actions, this definition assumes A = O, with O the n-tuple of outcomes.

An agent's **strategy** is a mapping from the state space to the action space. A **strategy profile** is an n-tuple of strategies for n players.

Analysing games: from optimality to equilibrium

To analyse a game in a single-agent setting, the **optimal strategy** (the strategy that maximizes the agent's expected payoff) is determined. However, the analysis of the optimal strategy is not meaningful in a multiagent setting as the best strategy depends on the choices of other agents who hope to maximise their payoff as well. Two solution concepts to analysing games in a multiagent setting are **Pareto optimality** and **Nash equilibrium**.

Definition 3.3.1 (Pareto domination) *Strategy profile s* Pareto dominates *strategy profile s' if for all $i \in N$, $u_i(s) \geq u_i(s')$, and there exists some $j \in N$ for which $u_j(s) > u_j(s')$.*

Definition 3.3.2 (Pareto optimality) *Strategy profile s is* Pareto optimal, *or* strictly Pareto efficient, *if there does not exist another strategy profile $s' \in S$ that Pareto dominates s.*

Every game must have at least one such optimum. In zero-sum games, all strategy profiles are Pareto optimal.

Intuitively, a **Nash equilibrium** is a **stable strategy profile**: no agent would want to change his strategy if he knew what strategies the other agents were following. (examples: see 3.3.3 page 63)

Chapter 7

7.4 Reinforcement Learning

Q-learning algorithm (+ appendix C)

7.7 Evolutionary Learning

In Evolutionary Learning, the focus is shifted towards **learning of populations of agents**, instead of learning of individual agents. The **replicator dynamic** models a population undergoing frequent interactions.

The Prisoner's Dilemma game is an example of a **symmetric game**, which implies that the agents do not have distinct roles in the game, and the payoff for agents does not depend on their identities. The replicator dynamic describes a population of agents playing such a game in an ongoing fashion. At each point in time, each agent only plays a pure strategy.

What follows is a mathematical model of the replicator dynamic:

The formal model is as follows. Given a normal-form game $G = (\{1,2\}, A, u)$, let $\varphi_t(a)$ denote the number of players playing action a at time t. Also, let

$$\theta_t(a) = \frac{\varphi_t(a)}{\sum_{a' \in A} \varphi_t(a')}$$

be the proportion of players playing action a at time t. We denote with φ_t the vector of measures of players playing each action, and with θ_t the vector of population shares for each action.

The expected payoff to any individual player for playing action a at time t is

$$u_t(a) = \sum_{a'} \theta_t(a') u(a, a').$$

The change in the number of agents playing action a at time t is defined to be proportional to his fitness, that is, his average payoff at the current time,

$$\dot{\varphi}_t(a) = \varphi_t(a) u_t(a).$$

The absolute numbers of agents of each type are not important; only the relative ratios are. Defining the average expected payoff of the whole population as

$$u_t^* = \sum_{a} \theta_t(a) u_t(a),$$

we have that the change in the fraction of agents playing action a at time t is

$$\dot{\theta}_t(a) = \frac{\left[\dot{\varphi}_t(a) \sum_{a' \in A} \varphi_t(a')\right] - \left[\varphi_t(a) \sum_{a' \in A} \dot{\varphi}_t(a')\right]}{\left[\sum_{a' \in A} \varphi_t(a')\right]^2} = \theta_t(a)[u_t(a) - u_t^*].$$

The following definitions and theorems can be used to examine the equilibrium points in the system:

> **Definition 7.7.2 (Steady state)** *A* steady state *of a population using the replicator dynamic is a population state θ such that for all $a \in A$, $\dot{\theta}(a) = 0$.*

> **Definition 7.7.3 (Stable steady state)** *A steady state θ of a replicator dynamic is* stable *if there exists an $\epsilon > 0$ such that for every ϵ-neighborhood U of θ there exists another neighborhood U' of θ such that if $\theta_0 \in U'$ then $\theta_t \in U$ for all $t > 0$.*

> **Definition 7.7.4 (Asymptotically stable state)** *A steady state θ of a replicator dynamic is* asymptotically stable *if it is stable, and in addition there exists an $\epsilon > 0$ such that for every ϵ-neighborhood U of θ it is the case that if $\theta_0 \in U$ then $\lim_{t\to\infty} \theta_t = \theta$.*

> **Theorem 7.7.5** *Given a normal-form game $G = (\{1,2\}, A = \{a_1, \ldots, a_k\}, u)$, if the strategy profile (s, s) is a (symmetric) mixed strategy Nash equilibrium of G then the population share vector $\theta = (s(a_1), \ldots, s(a_k))$ is a steady state of the replicator dynamic of G.*

The inverse of the last theorem does not hold: a **steady state** of the replicator dynamic is not necessarily a Nash equilibrium. However, every **stable steady state** is in fact a Nash equilibrium.