Computational Game Theory
The Basic Definitions

Oliver Schulte
Simon Fraser University
School of Computing Science

February 3, 2010

1 Relations and Functions

Let A be a set. An **ordered pair** from A is a pair (a, b) such that both a and b are from A. To say that a pair (a, b) is ordered is to say that we distinguish it from the pair (b, a) (whereas with sets, $\{a, b\} = \{b, a\}$). A (binary) **relation** R **on** A is a set of ordered pairs from A. If (a, b) is an ordered pair in a relation R, we write aRb (read "R holds between a and b"). A relation R on a set A is **reflexive** iff for all $x \in A$, xRx holds. A relation R on a set A is **transitive** iff for all $x, y, z \in A$, *if* it is the case that xRy and yRz both hold, *then* xRz holds. A relation R on a set A is **total** iff for all $x, y \in A$, either xRy or yRx (or both) holds.

We are especially interested in relations that rank a given set of options A. For a given agent and a given set of options A, we assume that the agent has a (**weak**) **preference** relation $\succeq$ on the options A. The agent's **strict preference relation** $\succ$ on A is then defined for all $x, y \in A$ by: $x \succ y$ iff $x \succeq y$ and *not* $y \succeq x$. The agent's **indifference relation** $\sim$ is defined for all $x, y \in A$ by: $x \sim y$ iff $x \succeq y$ and $y \succeq x$. A weak preference $\succeq$ relation is **rational** iff it is total, reflexive and transitive.

A **score function**, or **utility function**, u assigns a number to each option in a given set A. A utility function u **represents** a preference relation $\succeq$ on a set of options A just in case for all options $x, y \in A : x \succeq y$ holds iff $u(x) \geq u(y)$.

Theorem 1 *Let A be a finite set of options and $\succeq$ a preference relation on A. Then there is a utility function that represents $\succeq$ just in case $\succeq$ is rational.*

2 Decision Problems

In all decision problems, we start with a set of options A. We distinguish between decision-making under certainty and decision-making under uncertainty.

2.1 Decision-Making With Certainty

In a decision problem with certainty, we have a finite set $D = \{d_1, ..., d_n\}$ of **dimensions** or **attributes** of interest. We assume that for each dimension $d_i \in D$, the agent has a rational preference ordering $\succeq_i$over the options A. An option $x \in A$ **strongly Pareto-dominates** another option $y \in A$ iff[1] for all dimensions $d_i \in D$ it is the case that $x \succ_i y$. An option $x \in A$ **weakly Pareto-dominates** another option $y \in A$ iff for *all* dimensions $d_i \in D$ it is the case that $x \succeq_i y$, and for *at least one* dimension $d_k \in D$, we have that $x \succ_k y$.

We obtain another interpretation of decisions under certainty if we take the set of "dimensions" to be a set of *members of society*, such that the preference orderings $\succeq_i$represent the preferences of person i. Strong Pareto-dominance then means that *all* members of the society prefer the dominant option over the dominated one, and weak Pareto-dominance means that *some* members of the society prefer the dominant option over the dominated one, while others prefer the dominant option at least as much as the dominated one (but not necessarily more).

2.2 Decision-Making Under Uncertainty

We represent uncertainty by a finite set $S = \{s_1, s_2, ..., s_j\}$ of **states of the world.** We assume as well that a set of possible **outcomes** O is given, and that the decision-maker has a rational preference relation $\succeq$ represented by a utility function u over the outcomes O. An **act** a is a function that assigns to each state of the world $s \in S$ an outcome $a(s) \in O$.

An act a **strictly dominates** an act b iff for all states of the world $s \in S$, it is the case that $u(a(s)) > u(b(s))$. An act a **weakly dominates** an act b iff for *all* states of the world $s \in S$, it is the case that $u(a(s)) \geq u(b(s))$, and for *at least one* state of the world $s \in S$, it is the case that $u(a(s)) > u(b(s))$.

The worst-case outcome for an act a is the smallest number r such that for some state of the world s, r is the payoff that a yields in s (i.e., $r = u(a(s))$). Let $Worst(a)$ denote this number, for any act a. An act a is a **maximin** option iff there is no other act b such that $Worst(b) > Worst(a)$.

In certain circumstances, we may assume that an agent chooses in accordance with the principles of *expected utility.* In these circumstances, we assume that the agent has a *subjective probability function* p that assigns a number (a probability) to each state of the world. This function p is a **probability function** iff the following conditions obtain.

1. The probability function p assigns a number $p(s)$ to each state of the world in S.

2. The probability function p assigns no negative numbers. That is, for all states of the world $s \in S$, $p(s) \geq 0$.

[1] I use the abbreviation "iff" for "if and only if".

3. The probabilities of the various states of the world add up to 1. In symbols: $p(s_1) + p(s_2) + \cdots + p(s_j) = 1$.

Now we can define the **expected utility of an act** a given probability function p and utility function u:

$$EU(a) = p(s_1) \times u(a(s_1)) + p(s_2) \times u(a(s_2)) + \cdots + p(s_j) \times u(a(s_j)).$$

The principle of expected utility says that an agent should strictly prefer an act a to an act b just in case $EU(a) > EU(b)$.

3 Cooperative Games

Cooperative Game Theory deals with the question of what sort of agreements a group of agents would commit themselves to, for example as a result of a negotiation. The definition of a cooperative game is very simple.[2]

Definition 2 *A cooperative game is a pair* (N, v)*, where*

1. N *is a finite set of players, and*
2. v *assigns to each nonempty subset* C *of* N *a number* $v(C)$.

A nonempty subset C of the set of players N is called a **coalition**. The number $v(C)$ is called the **worth of the coalition**. Intuitively, $v(C)$ is the amount that the players forming the coalition C can obtain by joining forces. The result of a cooperative game with n players is an **allocation** $\mathbf{x} = (x_1, x_2, ..., x_n)$ that specifies the payoff to each player, namely x_i for player i. For a given coalition C and an allocation $\mathbf{x} = (x_1, x_2, ..., x_n)$, we obtain the payoff to the coalition C by adding the payoffs of its members; we write $\mathbf{x}(C)$ for the payoff to coalition C from $\mathbf{x}$. For example, suppose that in a cooperative game with three players, the resulting payoff vector is $\mathbf{x} = (4, 2, 8)$. Then the payoff to the coalition of player 1 and 3 is $\mathbf{x}(\{1, 3\}) = x_1 + x_3 = 4 + 8 = 12$.

A coalition C is **effective** for an allocation $\mathbf{x} = (x_1, x_2, ..., x_n)$ iff $v(C) \geq \mathbf{x}(C)$. The central concept for predicting the outcome of a cooperative game is the notion of the **core** of a game. We predict that an allocation will not be the result of a cooperative game if there is another allocation that is preferred to the first by some coalition that has the power to bring about its preferred allocation. To formulate this idea precisely, we say that an allocation $\mathbf{y} = (y_1, y_2, ..., y_n)$ **dominates** another allocation $\mathbf{x} = (x_1, x_2, ..., x_n)$ **via** coalition C iff for all members i of $C, y_i > x_i$. Finally, an allocation is feasible if it does not award more payoff than there is to be distributed; formally, an allocation $\mathbf{x}$ is **feasible** iff $\mathbf{x}(N) \leq v(N)$.

To avoid awkward phrasing, the next definition defines what payoffs vectors are *not* in the core; all others are of course in the core.

[2]Strictly speaking, our definition is that of a cooperative game with transferable payoffs, a special case of the more general definition of a cooperative game.

Definition 3 *Let (N, v) be a cooperative game with n players. An allocation $\mathbf{x} = (x_1, x_2, ..., x_n)$ is* not *in the core of (N, v) iff there is another feasible allocation $\mathbf{y} = (y_1, y_2, ..., y_n)$ and a coalition C such that*

1. *C is effective for $\mathbf{y}$ (i.e., $\mathbf{y}(C) \leq v(C)$), and*

2. *$\mathbf{y}$ dominates $\mathbf{x}$ via C (i.e., $y_i > x_i$ for all members $i \in C$).*

4 Simultaneous Matrix Games

This section introduces the notion of a *simultaneous-move game*, also known as a **normal form game**, or a **matrix game**, or a **strategic form game**. Here and in the next section, I will give the definitions for 2-player games; the definitions generalize to the case of n-player games in more or less obvious ways.

Definition 4 *A 2-player simultaneous matrix game is a quadruple $G = (S_1, S_2, \succeq_1, \succeq_2)$ where*

1. *S_i is the set of **strategies** for player i,*

2. *$\succeq_i$ is player i's preference relation over all pairs of strategies (s_1, s_2) where $s_1 \in S_1, s_2 \in S_2$.*

We assume that each player's preference relation $\succeq_i$is rational and hence can be represented by a **utility** or **payoff** function u_i. Thus player i strictly prefers a pair of strategies (s_1, s_2) to another pair (s_1', s_2') just in case $u_i(s_1, s_2) > u_i(s_1', s_2')$.

Consider a strategic form game $G = (S_1, S_2, \succeq_1, \succeq_2)$ and payoff functions u_i that represent $\succeq_i$. A strategy $s_1 \in S_1$ **strictly dominates** another strategy $s_1' \in S_1$ just in case for *all* strategies $s_2 \in S_2$, we have that $u_1(s_1, s_2) > u_1(s_1', s_2)$. A strategy $s_1 \in S_1$ **weakly dominates** another strategy $s_1' \in S_1$ just in case for *all* strategies $s_2 \in S_2$, we have that $u_1(s_1, s_2) \geq u_1(s_1', s_2)$ and for *at least one* strategy $s_2 \in S$, we have that $u_1(s_1, s_2) > u_1(s_1', s_2)$. Dominance is defined similarly for player 2's strategies.

The worst-case outcome for a strategy $s_1 \in S_1$ is the smallest number r such that for some strategy $s_2 \in S_2$ of player 2, r is the payoff that s_1 yields against s_2 (i.e., $u_1(s_1, s_2) = r$). Let $Worst(s_1)$ denote this number, for any strategy $s_1 \in S_1$. A strategy $s_1 \in S_1$ is a **maximin** strategy iff there is no other strategy $s_1' \in S_1$ such that $Worst(s_1) > Worst(s_1')$. We define player 2's maximin strategies in the analogous way.

Dominance and maximin are essentially the same concepts as in individual decision theory, with the other player's strategies playing the role of states of the world. An important new notion in games is that of a *best reply* and a Nash equilibrium.

Definition 5 *Let $G = (S_1, S_2, \succeq_1, \succeq_2)$ be a two-player simultaneous-move game. Assume that utility function u_i represents the preference relation $\succeq_i$.*

1. *A strategy $s_1 \in S_1$ of player 1 is a* ***best reply*** *against a strategy $s_2 \in S_2$ iff there is no other strategy $s'_1 \in S_1$ of player 1 such that $u_1(s'_1, s_2) > u_1(s_1, s_2)$. Similarly, a strategy $s_2 \in S_2$ of player 2 is a* ***best reply*** *against a strategy $s_1 \in S_1$ of player 1 iff there is no other strategy $s'_2 \in S_2$ of player 2 such that $u_2(s_1, s'_2) > u_2(s_1, s_2)$.*

2. *A pair of strategies (s_1, s_2) with $s_1 \in S_1, s_2 \in S_2$ is a* ***Nash equilibrium*** *iff s_1 is a best reply against s_2* and *s_2 is a best reply against s_1.*

5 Sequential Game Trees

A **directed graph** G is a pair $G = (V, E)$ where V is the set of vertices (or nodes), and E is the set of edges connecting the vertices,[3] such that if there is an edge from node u to node v, there is no edge back from node v to node u. A tree T is a directed graph with a distinguished node r, the **root**, such that

1. for every node in the tree, there is exactly one path connecting the node to the root,

2. the root has no predecessors,

3. each node other than the root has exactly one predecessor, and

4. the successors of each node are ordered from left to right.

Nodes without successors in a tree T are called **leaves**, or **terminal** nodes.

We define finite *game trees* for two players as follows. Game trees are also known as **sequential games,** or as **games in extensive form.**

Definition 6 *A finite* ***game tre****e for two players is a triple $(T, \succeq_1, \succeq_2)$ with* ***information sets*** *$I_1^1, I_1^2, ..., I_1^k$ for player 1, and $I_2^1, I_2^2, ..., I_2^m$ for player 2 with the following properties.*

1. *p is a* ***turn function*** *that assigns to each non-terminal node v in T a player i. Intuitively, if $p(v) = i$, it is player i's turn to move at v.*

2. *$\succeq_i$represents player i's preferences over the terminal nodes in T.*

3. *Each non-terminal node v in T is contained in exactly one information set I_i^k such that $i = p(v)$.*

A game tree is a game of **perfect information** if each information set contains exactly one node.

[3]Formally, E is a relation on V.