Mathematical Language and Game Theory

Pierpaolo Battigalli
Bocconi University

September 2023

Abstract

This note complements *Game Theory: Analysis of Strategic Thinking*, explaining in a concise and systematic way its mathematical language. The reader should try to prove all the claims, remarks etc. that are reported here without proof.

1 Sets, functions, and correspondences

The building blocks of the language used to describe games are sets, functions, correspondences, and mathematical structures, that is, lists of sets, functions, and (possibly) correspondences. Here we take for granted that the meaning of "set," "function," "correspondence," and "Cartesian product of sets" is understood. We also take for granted that standard set-theoretic notation is understood. Despite this, it is worth clarifying how we use arrows to denote functions, or correspondences.

Functions A **function** (or **map**) f from domain X to codomain Y is denoted $f : X \to Y$. When we want to specify that an element x of X is associated with an element of Y according to a particular *formula* expressed by function f we use the $\mapsto$ arrow[1] and write

$$x \mapsto f(x).$$

The **image**, or **range** of f is the set[2]

$$f(X) := \{y \in Y : \exists x \in X, f(x) = y\}.$$

[1] Written "\mapsto" in Latex; hence, the "mapsto arrow."

[2] The symbol ":=" means the the left-hand side is defined by (equal by definition to) the right-hand side. Similarly, the symbol "=:" means that the right-hand side is defined by the left-heand side.

Example 1 *For example, if* $f : \mathbb{R} \to \mathbb{R}$ *is the "square of" function that associates each real number with its square, we write*

$$f : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R}_+ \\ x & \mapsto & x^2 \end{array},$$

where x^2 *is the formula expressing the "square of" function. The range of the "square of" function* f *is* $f(\mathbb{R}) = \mathbb{R}_+$. ▲

It is sometimes useful to stack the two representations one below the other to clarify every feature of the given function. In Example 1, $f : \mathbb{R} \to \mathbb{R}$ gives us the domain ($\mathbb{R}$), codomain ($\mathbb{R}$), and the label (f) that we use to denote the "square of" function, whereas $x \mapsto x^2$ tells us that we associate each real number x with its square. When we are concerned with an abstract analysis of all functions in a given class, we use the general notation $f : X \to Y$. For example, for any real-valued function $f : X \to \mathbb{R}$, we can define $\sup_{x\in X} f(x) := \sup f(X)$ (least upper bound of $f(X)$) and $\inf_{x\in X} f(x) := \inf f(X)$ (greatest lower bound of $f(X)$).[3] For every real-valued and *continuous* function defined on a *compact* subset X of a topological space,[4] both $\max_{x\in X} f(x) := \max f(X)$ and $\min_{x\in X} f(x) := \min f(X)$ are well defined.

Correspondences A **correspondence** F that associates each element of domain X with a subset of codomain Y is written $F : X \rightrightarrows Y$, where the multiple arrow $\rightrightarrows$ is meant to remind the reader that the image set $F(x)$ of x may contain more than one element. However, when we specify with a formula the set associated with a particular element of the domain, we still use the $\mapsto$ arrow.

Example 2 *If* $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}$ *is the correspondence that associates each pair of real numbers* (a, b) *with the closed interval between* a *and* b*, we write*

$$F : \begin{array}{ccc} \mathbb{R}^2 & \rightrightarrows & \mathbb{R} \\ (a,b) & \mapsto & \{x \in \mathbb{R} : a \le x \le b\} \end{array}.$$

▲

[3] If $f(X)$ is not bounded above (respectively, below), then $\sup f(X) = +\infty$ (respectively, $\inf f(X) = -\infty$).

[4] Such as a closed and bounded subsets of a $\mathbb{R}^n$, or—more generally—a compact subset of a metric space.

It is convenient to allow $F(x)$ to be the *empty set*. For example, the "argmax" correspondence may be empty-valued.

Example 3 *Fix a real-valued function* $u : X \to \mathbb{R}$ *and a subset* $C \subseteq X$ *of its domain. With this,*

$$\arg\max_{x \in C} u(x) := \{x^* \in C : \forall x \in C, u(x^*) \geq u(x)\}$$

is the—possibly empty—set of maximizers of u *on the "constraint set"* C*. Let correspondence*

$$B: \quad P \quad \rightrightarrows \quad X$$

denote a parameterization of the constraint.[5] *Then,*

$$\begin{array}{rccl} F: & P & \rightrightarrows & X \\ & p & \mapsto & \arg\max_{x \in B(p)} u(x) \end{array}$$

may be empty-valued if, for example, u *is not continuous, or* $B(p)$ *is not compact for some* $p \in P$*.* ▲

Index sets as domains Finally, many examples of functions and correspondences will be defined on a domain $X = I$ that we interpret as an **index set**, such as the set of all **individuals** in a given group. Note that *we do not make any general assumption that* I *be an ordered set*, although it may be an ordered set in some special cases. Prominent examples of such functions in game theory are the **profiles** of actions (or strategies) that associate each player $i \in I$ with an action (or strategy) of i.

1.1 Set of functions (Cartesian products)

Cartesian products are important examples of sets, which can be interpreted as *sets of functions*.

Sets of profiles Specifically, consider an index set I, a "universal" set[6] $\bar{A}$ of "alternatives," and a correspondence $i \mapsto A_i$ that associates each element i of I with a nonempty subset $A_i \subseteq \bar{A}$. We interpret the **Cartesian product**

[5] Correspondence B could be the budget map, which associates each price vector p with the set of affordable consumption bundles given some endowment ω.

[6] We are using the term "universal" informally to mean that the set contains all the objects that we may want to consider in a given problem.

$\times_{i \in I} A_i$ as the set of **selections** from the correspondence $i \mapsto A_i$, that is, the *set of functions*

$$\begin{array}{rccc} a: & I & \rightarrow & \bar{A} \\ & i & \mapsto & a_i \end{array}$$

such that

$$\forall i \in I, a_i \in A_i.$$

We often write such functions as follows:

$$a = (a_i)_{i \in I}\,.$$

When I is interpreted as a set of individuals, functions with domain I are called "**profiles**."

Sequences The set of *all* functions with domain I and codomain $\bar{A}$ is denoted $\bar{A}^I$. When I is finite, there are at least $|I|!$ bijections between the set of functions $\bar{A}^I$ and the set of ordered $|I|$-tuples $\bar{A}^{|I|}$, one for each strict order (or permutation) on I.[7] Indeed, when $I = \{1, ..., n\}$, then $\bar{A}^I$ and $\bar{A}^{|I|} = \bar{A}^n$ are essentially the same mathematical object under the canonical bijection

$$(a_i)_{i \in I} \leftrightarrow (a_1, ..., a_n)\,,$$

which preserves the order of natural numbers, that is, the set of ordered n-tuples (finite sequences of length n) of elements of $\bar{A}$. For example, vectors of real numbers are a special case where $I = \{1, ..., n\}$ and $\bar{A} = \mathbb{R}$. Then, the relevant Cartesian product is $\mathbb{R}^{\{1,...,n\}}$, or equivalently $\mathbb{R}^n$. (Sometimes the set of the first n natural numbers is written $[n] := \{1, ..., n\}$; thus, $\mathbb{R}^{\{1,...,n\}} = \mathbb{R}^{[n]}$.) Another important example is the set of all countable **sequences** of elements of $\bar{A}$, such as the set $\mathbb{R}^{\mathbb{N}}$ of all sequences of real numbers when $\bar{A} = \mathbb{R}$. Such sequences are equivalently denoted by $(a_n)_{n \in \mathbb{N}}$ or $(a_n)_{n=1}^{\infty}$.

A comment on sequences It is common in the mathematical literature to write sequences as follows: $\{a_n\}_{n=1}^{\infty}$. Such use of brackets in *inconsistent with the standard notation about sets* and may induce the reader to confuse a sequence with its range. For example, it is not clear whether $\{(-1)^n\}_{n=1}^{\infty}$ denotes the sequence $(1, -1, 1, -1, ...)$, or its range, which is the doubleton $\{-1, 1\}$. Such confusing notation is just a traditional bad habit. Like all habits, also the bad ones die hard. But this is not a reason to knowingly indulge in bad habits.

[7] $|\cdot|$ denotes the cardinality of a set. Thus, $|I|$ is the number of elements of the finite set I.

Cardinalities The following observation further explains why the "power notation" for sets of functions is explicative and convenient.

Remark 1 *If both the index set (domain) I and the set of alternatives (codomain) $\bar{A}$ are finite, then the number of functions from I to $\bar{A}$ (i.e., the cardinality of $\bar{A}^I$) is*

$$\left|\bar{A}^I\right| = \left|\bar{A}\right|^{|I|}.$$

Cartesian products and products of numbers are related in a similar way.

Remark 2 *The number of selections from the correspondence $i \mapsto A_i$, that is, the number of elements of the Cartesian product $\times_{i \in I} A_i$, is given by the product formula*

$$|\times_{i \in I} A_i| = \prod_{i \in I} |A_i|.$$

In words, the *cardinality of a Cartesian product of sets is the product of the cardinalities of such sets.*

Power sets Another convenient "power notation" is used to denote the collection of all subsets of I, including the empty set $\emptyset$ and I itself, that is, the **power set** 2^I.

Remark 3 *If I is finite,*

$$\left|2^I\right| = 2^{|I|}.$$

Indicator functions The natural number $2^{|I|}$ (with I finite) is also the cardinality of the set of functions from I to a binary codomain. In particular, let $\bar{A} = \{0, 1\}$, and recall that the **indicator function** $\mathbf{1}_J : I \to \{0, 1\}$ of a subset $J \subseteq I$ is

$$\mathbf{1}_J(i) = \begin{cases} 1, & \text{if } i \in J, \\ 0, & \text{if } i \in I \backslash J. \end{cases}$$

With this, $\bar{A}^I = \{0, 1\}^I$ can be interpreted as the set of all indicator functions of subsets of I. Specifically, for each $J \subseteq I$, $a = \mathbf{1}_J$ if and only if $a^{-1}(1) = J$.

1.2 Sections and projections

Fix a (nonempty-valued) correspondence $i \mapsto A_i \subseteq \bar{A}$, the corresponding Cartesian product $A = \times_{i \in I} A_i$ and a strict subset $J \subset I$ of the index set. To ease notation, let $A_J = \times_{j \in J} A_j$ denote the "sub-product" over the index subset J, with typical element a_J. When $J = I \backslash \{i\}$ for some $i \in I$, we further simplify and write A_{-i}, with typical element a_{-i}. For all sub-functions (profiles) $a_J = (a_j)_{j \in J} \in A_J$ and $a_{I \backslash J} = (a_k)_{k \in I \backslash J} \in A_{I \backslash J}$, we obtain a function (profile) $i \mapsto a_i$ denoted by $\left(a_J, a_{I \backslash J}\right)$, where a_i is determined by a_J if $i \in J$, and by $a_{I \backslash J}$ if $i \in I \backslash J$. Since $A = \times_{i \in I} A_i$ is the set of functions from I to $\bar{A}$ such that $a_i \in A_i$ for each i, it is clear that $\left(a_J, a_{I \backslash J}\right) \in A$, because $\left(a_J, a_{I \backslash J}\right)$ is one such function.[8]

Sections of sets Fix any subset $B \subseteq A = \times_{i \in I} A_i$; hence, B may be a Cartesian product or not. For each $i \in I$ and $a_i \in A_i$, the **section at a_i of set** B is the subset $B_{a_i} \subseteq A_{-i}$ of all the profiles a_{-i} that form with a_i an element of B, that is,

$$B_{a_i} := \{a_{-i} \in A_{-i} : (a_i, a_{-i}) \in B\}.$$

Example 4 *Suppose that $A = \mathbb{R}^3$, B is the unit ball centered at $(0,0,0)$, and $a_3 = 0$; then, B_{a_3} is the unit disk in $\mathbb{R}^2$ centered at $(0,0)$.* ▲

More generally, for each index subset $J \subset I$ and $a_J \in A_J$, the section of B at a_J is

$$B_{a_J} := \left\{a_{I \backslash J} \in A_{I \backslash J} : \left(a_J, a_{I \backslash J}\right) \in B\right\}.$$

Projections Let $A = \times_{i \in I} A$ be a given and understood Cartesian product. The **projection function** from $A = \times_{i \in I} A$ onto A_J is the map that associates each profile $a = \left(a_J, a_{I \backslash J}\right)$ with its component $a_J \in A_J$:

$$\begin{array}{llll} \mathrm{proj}_{A_J} : & A & \to & A_J \\ & \left(a_J, a_{I \backslash J}\right) & \mapsto & a_J \end{array}.$$

Recall that, for every function $f : X \to Y$ and subset $B \subseteq X$, the image of B through f is denoted

$$f(B) := \{y \in Y : \exists x \in B, f(x) = y\} = \bigcup_{x \in B} \{f(x)\}.$$

With this, the **projection of set** $B \subseteq A$ **onto** A_J is the image of B through the projection function $\mathrm{proj}_{A_J} : A \to A_J$, that is, the set of profiles a_J such that $\left(a_J, a_{I \backslash J}\right) \in B$ for at least one $a_{I \backslash J} \in A_{I \backslash J}$:

$$\mathrm{proj}_{A_J}(B) = \left\{a_J \in A_J : \exists a_{I \backslash J} \in A_{I \backslash J}, \left(a_J, a_{I \backslash J}\right) \in B\right\}.$$

[8] Do not be distracted by the fact that a_J is written before $a_{I \backslash J}$. This is irrelevant, because we are not assuming any order on I; hence, *here the order in which symbols are written does not matter.* What matters and has to be clear is how each $i \in I$ is associated with some element of the universal set of alternatives $\bar{A}$.

Often we suppress the parentheses and write $\mathrm{proj}_{A_J} B$ instead of $\mathrm{proj}_{A_J}(B)$.

Example 5 *The projection of the unit ball in* $\mathbb{R}^3$ *centered at* $(2,2,2)$ *onto* $\mathbb{R}^2$ *is the unit disk centered at* $(2,2)$. ▲

One can check by inspection of the definitions that the projection of a set is a union of sections:

Remark 4 $\mathrm{proj}_{A_J} B = \bigcup_{a_{I\setminus J} \in A_{I\setminus J}} B_{a_{I\setminus J}}$.

Sections of functions Fix a function $f : A \to Y$, where $A = \times_{i\in I} A_i$ and Y is an arbitrary codomain, e.g., a set of "outcomes" or "outputs." For each i and $a_i \in A_i$, we obtain from f a function f_{a_i} of a_{-i} called the **section at** a_i **of function** f:

$$f_{a_i} : \begin{array}{ccc} A_{-i} & \to & Y \\ a_{-i} & \mapsto & f_{a_i}(a_{-i}) = f(a_i, a_{-i}) \end{array}.$$

Example 6 *Consider a Cobb-Douglas production function*

$$f : \begin{array}{ccc} \mathbb{R}_+^{\{K,L\}} & \to & \mathbb{R}_+ \\ (k,\ell) & \mapsto & (k\ell)^{\frac{1}{2}} \end{array},$$

where K *and* L *are respectively the indexes for capital and labor. If capital is fixed at* $k=1$ *in the short run, then the short-run production function depends only on labor, and it is the section of* f *at* $k=1$*:*

$$f_{k=1} : \begin{array}{ccc} \mathbb{R}_+ & \to & \mathbb{R}_+ \\ \ell & \mapsto & \sqrt{\ell} \end{array}.$$

▲

Similarly, for each strict subset $J \subset I$, we can consider the section at $a_J = (a_j)_{j\in J}$ of f:

$$f_{a_J} : \begin{array}{ccc} A_{I\setminus J} & \to & Y \\ a_{I\setminus J} & \mapsto & f_{a_J}\left(a_{I\setminus J}\right) = f\left(a_{I\setminus J}, a_J\right) \end{array}.$$

We obtain an important special case when $J = I\setminus\{i\}$, that is, when we fix all the values of profile a except a_i.

The section notation is convenient in several cases. One of them is when we have a function $f : A \to Y$, we *fix* a_i and y, and we want to consider all the a_{-i} such that $f(a_i, a_{-i}) = y$, that is, the set

$$\{a_{-i} \in A_{-i} : f(a_i, a_{-i}) = y\}.$$

Using the section notation, this set can be compactly written as the inverse image of y through function f_{a_i}:

$$f_{a_i}^{-1}(y) := \{a_{-i} \in A_{-i} : f_{a_i}(a_{-i}) = y\} = \{a_{-i} \in A_{-i} : f(a_i, a_{-i}) = y\}.$$

Sections of sets and of functions are related. First recall that the **graph** of a function $f : A \to Y$ is the set of pairs $(a, y) \in A \times Y$ such that $f(a) = y$, that is,

$$\mathrm{Graph}(f) := \{(a, y) \in A \times Y : f(a) = y\}.$$

Remark 5 *The graph of the section of a function is the section of the graph of such function:* $\mathrm{Graph}(f_{a_J}) = (\mathrm{Graph}(f))_{a_J}$.

Example 7 *In Example 6 (Cobb-Douglas production function),* $\mathrm{Graph}(f)$ *is*

$$\left\{(k, \ell, y) \in \mathbb{R}_+^{\{K,L,O\}} : y = \sqrt{k\ell}\right\},$$

where O is the index of output. Then $\mathrm{Graph}(f_{k=1}) = (\mathrm{Graph}(f))_{k=1}$ *is the graph of function $\ell \mapsto \sqrt{\ell}$, that is,*

$$\left\{(\ell, y) \in \mathbb{R}_+^{\{L,O\}} : y = \sqrt{\ell}\right\}.$$

▲

2 Games, sets, and functions

As an illustration of the aforementioned mathematical language and notation, consider the mathematical description of the essential aspects of the rules of a game with simultaneous-moves.

Game forms A **game form** is a mathematical structure

$$\left\langle I, Y, (A_i)_{i \in I}, g \right\rangle$$

where

- I is the set of **players**,
- Y is a set **outcomes**, or consequences,
- A_i is the set of feasible **actions** of player i,
- $g : \times_{i \in I} A_i \to Y$ is the **outcome** (or consequence) **function**.

Structure $\langle I, Y, (A_i)_{i \in I}, g \rangle$ represents the rules of the game under the implicit assumption that players move simultaneously (if moves were not simultaneous, then the order of moves should appear in the mathematical representation). To avoid some measure-theoretic technicalities, here we assume that the game form is *finite*, but most of the analysis below does not depend on this assumption.

Action profiles Let $\bar{A} = \bigcup_{i \in I} A_i$ denote the set of *all* actions. The Cartesian product $\times_{i \in I} A_i$ is the set of **action profiles**, that is, the set of all functions

$$\begin{array}{rccc} a: & I & \to & \bar{A} \\ & i & \mapsto & a_i \end{array}$$

such that $a_i \in A_i$ for each $i \in I$. With this, the set of action profiles A is also the set of all *selections* from the feasible actions correspondence $i \mapsto A_i$.

Co-players From now on, we will take the perspective of a particular player $i \in I$. Thus, A_{-i} is the set of action profiles of the opponents of i (also called "**co-players**"). This is the set of functions $a_{-i} : I \backslash \{i\} \to \bar{A}$ such that $a_j \in A_j$ for each $j \in I \backslash \{i\}$. Then, it makes sense to write $a = (a_i, a_{-i})$, which means that action profile a is the function that associates i with a_i and j with a_j for each j different from i. Since we think of A as a set of functions rather than a set of ordered $|I|$-tuples, it also makes sense to write $A = A_i \times A_{-i}$. Thus, if we want to emphasize the perspective of player i, we can write the outcome function as follows:

$$g : A_i \times A_{-i} \to Y.$$

2.1 Sections, beliefs, and preferences

Here we use the mathematical language introduced so far to express some ideas from the theory of choice in connection with games.

Choice under certainty The section of the outcome function g at any particular a_{-i}, $g_{a_{-i}}$, shows how i affects the outcome when the actions of the other players are kept fixed at a_{-i}. For example, if i is *certain* that the opponents (or co-players) play a_{-i}, then his conjectured action-outcome map is $g_{a_{-i}} : A_i \to Y$. If i's preferences are represented by a utility function $v_i : Y \to \mathbb{R}$, then a rational i with such (deterministic) conjecture solves

$$\max_{a_i \in A_i} v_i \left(g_{a_{-i}} \left(a_i \right) \right),$$

a problem of choice under (subjective) certainty. The solution of a problem of choice under certainty is invariant to strictly increasing transformations of the utility function.

Remark 6 *Let $f : v_i (\mathbb{R}) \to \mathbb{R}$ be strictly increasing. Then, for every outcome function $g : A \to Y$ and co-players' action profile $a_{-i} \in A_{-i}$,*

$$\arg\max_{a_i \in A_i} v_i \left(g_{a_{-i}} \left(a_i \right) \right) = \arg\max_{a_i \in A_i} \left(f \circ v_i \right) \left(g_{a_{-i}} \left(a_i \right) \right).$$

Before we consider choice under risk and/or uncertainty, we recall some concepts from elementary probability theory.

Probability measures In probability theory, a probability space is a pair (Ω, μ), where μ is a probability measure on Ω. When Ω is *finite*, μ can be equivalently represented either as a function $\mu : 2^{\Omega} \to [0,1]$ such that (i) $\mu(\Omega) = 1$, and (ii) $\mu(E' \cup E'') = \mu(E'') + \mu(E'')$ if $E' \cap E'' = \emptyset$, or a (probability mass) function $\mu : \Omega \to \mathbb{R}_+$ such that $\sum_{\omega \in \Omega} \mu(\omega) = 1$, letting

$$\mu(E) = \sum_{\omega \in E} \mu(\omega).$$

(We ignore the complications related to infinite state spaces.) With this, we let[9]

$$\begin{aligned} \Delta(\Omega) \quad &:= \\ & \left\{ \mu \in [0,1]^{2^{\Omega}} : \mu(\Omega) = 1, \forall E', E'' \in 2^{\Omega}, E' \cap E'' = \emptyset \Rightarrow \mu(E' \cup E'') = \mu\left(E''\right) + \mu\left(E''\right) \right\} \\ \cong \quad & \left\{ \mu \in \mathbb{R}_+^{\Omega} : \sum_{\omega \in \Omega} \mu(\omega) = 1 \right\} \end{aligned}$$

denote the set of all probability measures on Ω.

[9] Symbol $\cong$ means "isomorphic to". Intuitively, two sets are isomorphic if there is a bijection between them that preserves all the structures we care about. Roughly speaking, two isomorphic sets represent the same thing.

Dirac measures and embedding of Ω in $\Delta(\Omega)$ Note that set Ω can be identified with the set of degenerate probability measures that assign probability 1 to one element of Ω, also called **Dirac measures**. Formally, for each $\omega \in \Omega$, the Dirac measure $\delta_\omega \in \Delta(\Omega)$ is defined as follows: for every event E,

$$\delta_\omega(E) = \begin{cases} 1, & \text{if } \omega \in E, \\ 0, & \text{if } \omega \notin E. \end{cases}$$

Obviously, there is a bijection between Ω and the set of Dirac measures. Therefore, it makes sense to write, with a slight *abuse of notation*, $\Omega \subseteq \Delta(\Omega)$.[10]

Expectation functionals A **random variable** (or **random number**) on probability space (Ω, μ) is a real-valued function $v : \Omega \to \mathbb{R}$. The **expected value** of v (given μ) is the weighted average

$$\mathbb{E}_\mu(v) = \sum_{\omega \in \Omega} v(\omega)\,\mu(\omega).$$

We can interpret the $\mathbb{E}_\mu(\cdot)$ notation as a special case of our notation on functions and sections. For a fixed state space Ω, let us consider all the probability measures $\mu \in \Delta(\Omega)$ and variables/functions $v \in \mathbb{R}^\Omega$. We call "functional" a real-valued function whose domain is a set of functions, or—more generally—a set of n-tuple of functions. The **expectation functional**

$$\begin{array}{rccl} \mathbb{E}: & \mathbb{R}^\Omega \times \Delta(\Omega) & \to & \mathbb{R} \\ & (v, \mu) & \mapsto & \sum_{\omega \in \Omega} v(\omega)\,\mu(\omega) \end{array}$$

associates each pair (v, μ) with the expected value of v under probability measure μ. When we fix $\mu \in \Delta(\Omega)$, we obtain the section

$$\begin{array}{rccl} \mathbb{E}_\mu: & \mathbb{R}^\Omega & \to & \mathbb{R} \\ & v & \mapsto & \sum_{\omega \in \Omega} v(\omega)\,\mu(\omega) \end{array},$$

that is, the expectation functional defined on the set of random variables over probability space (Ω, μ). (Of course, it also makes sense to define the "dual" functional $\mu \mapsto \sum_{\omega \in \Omega} v(\omega)\,\mu(\omega)$ for any fixed variable/function $v : \Omega \to \mathbb{R}$; but this is less important here.)

[10] The inclusion is strict if Ω has at least 2 elements.

Probability functionals The simplest random variables on probability spaces (Ω, μ) $(\mu \in \Delta(\Omega))$ are the indicator functions of events

$$\mathbf{1}_E : \Omega \to \{0,1\} \quad (E \subseteq \Omega).$$

The expected value of $\mathbf{1}_E$ given μ is the probability of E given μ. This gives the **probability functional**

$$\begin{array}{llll} \mathbb{P}: & 2^{\Omega} \times \Delta(\Omega) & \to & \mathbb{R} \\ & (E, \mu) & \mapsto & \mathbb{E}(\mathbf{1}_E, \mu) = \sum_{\omega \in E} \mu(\omega) \end{array}.$$

When we fix $\mu \in \Delta(\Omega)$, we obtain the section

$$\begin{array}{llll} \mathbb{P}_\mu: & 2^{\Omega} & \to & \mathbb{R} \\ & E & \mapsto & \sum_{\omega \in E} \mu(\omega) \end{array}.$$

Savage acts, choice under uncertainty If we consider the section of outcome function g at any action $a_i \in A_i$, we obtain functions $g_{a_i} : A_{-i} \to Y$ $(a_i \in A_i)$. This shows that *an action can always be interpreted as a function* from what the decision maker i does not control (in this case, the actions of other players a_{-i}) to the outcomes, or consequences. In decision theory, such functions are called **acts** in the sense of Savage, or **Savage acts.** According to the **subjective expected utility (SEU)** theory of choice under uncertainty of Savage (1954), i has a preference relation $\succsim_i \subseteq Y^{A_{-i}} \times Y^{A_{-i}}$ over the set $Y^{A_{-i}}$ of *all conceivable* acts,[11] which contains the set

$$\left\{ f \in Y^{A_{-i}} : \exists a_i \in A_i, f = g_{a_i} \right\}$$

of all *feasible acts.*[12] If $\succsim_i$ satisfies the axioms of SEU theory, there exists a utility function $v_i : Y \to \mathbb{R}$ (unique, up to positive affine transformations) and a probability measure $\mu^i \in \Delta(A_{-i})$ such that, for all acts $f', f'' \in Y^{A_{-i}}$,[13]

$$f' \succsim_i f'' \Leftrightarrow \mathbb{E}_{\mu^i}\left(v_i \circ f'\right) \geq \mathbb{E}_{\mu^i}\left(v_i \circ f''\right).$$

Thus, a *rational* player i with (probabilistic) conjecture μ^i chooses one of the actions in the set

$$\arg\max_{a_i \in A_i} \mathbb{E}_{\mu^i}(v_i \circ g_{a_i}) := \left\{ a_i^* \in A_i : \forall a_i \in A_i, \mathbb{E}_{\mu^i}\left(v_i \circ g_{a_i^*}\right) \geq \mathbb{E}_{\mu^i}(v_i \circ g_{a_i}) \right\}.$$

[11] Recall, a **binary relation** over a set X is a subset of $R \subseteq X \times X$. We write $x' R x''$ if $(x', x'') \in R$.

[12] What we say about acts and subjective expected utility is "in the same spirit" as Savage (1954). But he considers an uncountable A_{-i}; therefore, his theory does not exactly apply here.

[13] Note that, for every $f : A_{-i} \to Y$ and $\mu^i \in \Delta(A_{-i})$, function $v_i \circ f : A_{-i} \to \mathbb{R}$ is a random variable on the probability space $\left(A_{-i}, \mu^i\right)$.

For each pair of constants k and m, we denote by $(mv_i + k)$ the following affine transformation

$$\begin{array}{rccl} (mv_i + k): & Y & \to & \mathbb{R} \\ & y & \mapsto & mv_i(y) + k \end{array}.$$

Uniqueness up to positive affine transformations implies the following.

Remark 7 *For all $(k, m) \in \mathbb{R} \times \mathbb{R}_{++}$, $g \in Y^A, a_i \in A_i$, and $\mu^i \in \Delta(A_{-i})$,*

$$\mathbb{E}_{\mu^i}((mv_i + k) \circ g_{a_i}) = m\mathbb{E}_{\mu^i}(v_i \circ g_{a_i}) + k;$$

therefore,

$$\arg\max_{a_i \in A_i} \mathbb{E}_{\mu^i}(v_i \circ g_{a_i}) = \arg\max_{a_i \in A_i} \mathbb{E}_{\mu^i}((mv_i + k) \circ g_{a_i}).$$

Stochastic outcomes Some games have rules that make the outcome depend on actions only stochastically, that is, with some noise. This can be represented in two ways:

- There is a "chance player" that chooses a deterministic outcome function $g \in Y^A$ at random according to some objective probability measure $\gamma \in \Delta(Y^A)$. Then, in the finite case, for all $a \in A$ and $y \in Y$, the probability of outcome y given action profile a is the sum of the probabilities of all the outcome functions g such that $g(a) = y$:

$$\hat{g}(y|a) = \sum_{g \in Y^A : g(a) = y} \gamma(g).$$

- There is a "chance player" that, for each $a \in A$, selects the outcome at random according to an objective probability measure $\bar{g}(a) \in \Delta(Y)$. This yields a **stochastic outcome function**

$$\begin{array}{rccl} \bar{g}: & A & \to & \Delta(Y) \\ & a & \mapsto & \bar{g}(a) \end{array}.$$

The connection between $\bar{g}$ and $\hat{g}$ is the following:

$$\forall a \in A, \forall y \in Y,\ \bar{g}(a)(y) = \hat{g}(y|a).$$

Choice under risk Suppose that the game form is stochastic and player i is *certain of* a particular action profile a_{-i}. How is i supposed to rank his actions and choose? In this case each action a_i induces an *objective* probability measure over outcomes, or **lottery**, $\bar{g}_{a_{-i}}(a_i) \in \Delta(Y)$. According to the **objective expected utility** theory of von Neumann and Morgenstern (1947), i's preference relation $\succsim_i \subseteq \Delta(Y) \times \Delta(Y)$ over the whole set of conceivable lotteries is such that there is a utility function $v_i : Y \to \mathbb{R}$ (unique up to positive affine transformations) such that, for all lotteries $\lambda', \lambda'' \in \Delta(Y)$,

$$\lambda' \succsim_i \lambda'' \Leftrightarrow \mathbb{E}_{\lambda'}(v_i) \geq \mathbb{E}_{\lambda''}(v_i).$$

Note that, if i is certain of a_{-i} (i.e., his subjective belief is represented by the Dirac measure $\delta_{a_{-i}}$), then he thinks that the set of *feasible* lotteries is

$$\left\{\lambda \in \Delta(Y) : \exists a_i \in A_i, \bar{g}_{a_{-i}}(a_i) = \lambda\right\},$$

that is, the set of all lotteries induced by some action of i given conjecture a_{-i}. Thus, a rational player i with such conjecture chooses one of the actions in the set

$$\arg\max_{a_i \in A_i} \mathbb{E}_{\bar{g}_{a_{-i}}(a_i)}(v_i).$$

Choice under risk and uncertainty What if i is not certain about a_{-i}? In this case, each action a_i induces a function $\bar{g}_{a_i} : A_{-i} \to \Delta(Y)$ that associates each a_{-i} with an objective probability measure on Y. Such functions (the elements of $\Delta(Y)^{A_{-i}}$) are called "**Anscombe-Aumann acts,**" or AA-acts. According to the **subjective expected utility theory** of Anscombe and Aumann (1963), agent i has a preference relation $\succsim_i \subseteq \Delta(Y)^{A_{-i}} \times \Delta(Y)^{A_{-i}}$ over the set of all *conceivable* AA-acts, where $\succsim_i$ has the property that there exist a utility function $v_i : Y \to \mathbb{R}$ (unique, up to positive affine transformations) and a subjective probability measure $\mu^i \in \Delta(A_{-i})$ such that, for all $f', f'' \in \Delta(Y)^{A_{-i}}$,

$$f' \succsim_i f'' \Leftrightarrow \sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \mathbb{E}_{f'(a_{-i})}(v_i) \geq \sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \mathbb{E}_{f''(a_{-i})}(v_i).$$

Note that the set of *feasible* AA-acts is

$$\left\{f \in \Delta(Y)^{A_{-i}} : \exists a_i \in A_i, f = \bar{g}_{a_i}\right\}.$$

Therefore, a rational player i with (probabilistic) conjecture μ^i chooses one of the actions in the set

$$\begin{aligned} & \arg\max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \mathbb{E}_{\bar{g}_{a_i}(a_{-i})} (v_i) \\ = & \arg\max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \sum_{y \in Y} \bar{g}_{a_i}(a_{-i}) (y) \, v_i (y) \, , \end{aligned}$$

where $\bar{g}_{a_i}(a_{-i}) \in \Delta(Y)$ is the lottery induced by action a_i when the co-players choose a_{-i}. Uniqueness up to positive affine transformation implies the following.

Remark 8 *For all* $(k, m) \in \mathbb{R} \times \mathbb{R}_{++}$, $\bar{g} \in \Delta(Y)^A, a_i \in A_i$ *and* $\mu^i \in \Delta(A_{-i})$,

$$\sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \mathbb{E}_{\bar{g}_{a_i}(a_{-i})} (mv_i + k) = m \left(\sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \mathbb{E}_{\bar{g}_{a_i}(a_{-i})} (v_i) \right) + k;$$

therefore,

$$\begin{aligned} & \arg\max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \mathbb{E}_{\bar{g}_{a_i}(a_{-i})} (v_i) \\ = & \arg\max_{a_i \in A_i} \sum_{a_{-i} \in A_{-i}} \mu^i(a_{-i}) \mathbb{E}_{\bar{g}_{a_i}(a_{-i})} ((mv_i + k)) \, . \end{aligned}$$

Mixed actions There is another way to obtain AA-acts and it works even if the outcome function g is deterministic: Player i can delegate the choice of his action a_i to a randomization device with distribution $\alpha_i \in \Delta(A_i)$. By choosing such device, i induces a **mixed action** α_i. Since the induced mixed action is all that matters, we can think of a set of *conceivable* choices $\Delta(A_i)$ whereas the set of *feasible choices* includes A_i (which we identify with the set of Dirac measures on A_i) and all the non-Dirac measures α_i that are induced by some feasible randomization device.

Each mixed action α_i induces a corresponding AA-act:

$$\alpha_i \mapsto \left(\hat{g}_{a_{-i}} (\alpha_i) \right)_{a_{-i} \in A_{-i}} \in \Delta(Y)^{A_{-i}} \, ,$$

where

$$\begin{aligned} \hat{g}_{a_{-i}} : \quad & \Delta(A_i) && \to && \Delta(Y) \\ & \alpha_i && \mapsto && \alpha_i \circ g_{a_{-i}}^{-1} \end{aligned}$$

is the **pushforward** function such that

$$\hat{g}_{a_{-i}}(\alpha_i)(E) = \alpha_i\left(g_{a_{-i}}^{-1}(E)\right) = \alpha_i\left(\left\{a_i \in A_i : g_{a_{-i}}(a_i) \in E\right\}\right)$$

for each outcome-event $E \subseteq Y$.

Example 8 *Consider the outcome function*

$g:$	ℓ	r
t	G	B
m	M	M
d	B	G

and let $A_i = \{t, m, d\}$ *(hence,* $A_{-i} = \{\ell, r\}$*). Then, the mixed action* $\alpha_i = \frac{1}{2}\delta_t + \frac{1}{2}\delta_m$*—which randomizes* $\frac{1}{2} : \frac{1}{2}$ *on* t *and* m*—yields the following AA-act* f

f	B	M	G
$\hat{g}_\ell\left(\frac{1}{2}\delta_t + \frac{1}{2}\delta_m\right)$	0	$\frac{1}{2}$	$\frac{1}{2}$
$\hat{g}_r\left(\frac{1}{2}\delta_t + \frac{1}{2}\delta_m\right)$	$\frac{1}{2}$	$\frac{1}{2}$	0

▲

Battigalli *et al.* (2017) show that it is not necessary to consider the whole set $\Delta(Y)^{A_{-i}}$ of conceivable AA-acts to develop an axiomatic theory of choice under uncertainty with a subjective EU representation, because the subset

$$\left\{f \in \Delta(Y)^{A_{-i}} : \exists \alpha_i \in \Delta(A_i) : f = \left(\hat{g}_{a_{-i}}(\alpha_i)\right)_{a_{-i} \in A_{-i}}\right\}$$

of AA-acts induced by mixed actions is "sufficiently rich." That is, one can provide an axiomatic theory of choice under uncertainty looking at preferences over mixed actions, $\succsim_i \subseteq \Delta(A_i) \times \Delta(A_i)$.

3 Predicates, sets, and games

Predicate calculus is a branch of mathematical logic that analyzes sentences about the properties of individual objects and relationships between objects. A predicate is a statement that some property or relationship holds for an object, or a tuple of objects. Here we are only interested in the formal representation of such sentences, how they are used to represent sets, and some simple rules. We illustrate with examples from the theory of games.

Games A **simultaneous-move game** is given by a game form $\langle I, Y, (A_i)_{i\in I}, g\rangle$ and a profile of utility functions $(v_i : Y \to \mathbb{R})_{i\in I}$. For each player $i \in I$, let

$$u_i = v_i \circ g : A \to \mathbb{R}$$

denote the so called "**payoff function**" of i.[14] Then, we can consider a kind of reduced form

$$G = \langle I, (A_i, u_i)_{i\in I}\rangle$$

of the game. The basic set of objects under consideration is the set of mixed actions $\Delta(A_i)$ of a given player i. As explained above, we regard A_i as a subset of $\Delta(A_i)$: $A_i \subseteq \Delta(A_i)$, because choosing action a_i is equivalent to choosing the degenerate probability measure (Dirac measure) δ_{a_i} that assigns probability 1 to a_i. We also consider the set $\Delta(A_{-i})$ of conjectures about the action profiles of the co-players; again, it makes sense to write $A_{-i} \subseteq \Delta(A_{-i})$.

Justifiability and dominance Here are two examples of (informally stated) sentences in predicate calculus:

- "action a_i is justifiable" and
- "action a_i is dominated."

Both examples involve the "for all" quantifier $\forall$ and the "there exists" quantifier $\exists$. Specifically, *by definition*,

- a_i is **justifiable** iff (if and only if) there exists a conjecture $\mu^i \in \Delta(A_{-i})$ such that $\mathbb{E}_{\mu^i}(u_{i,a_i}) \geq \mathbb{E}_{\mu^i}(u_{i,b_i})$ for all $b_i \in A_i$,[15]
- a_i is **dominated** iff there exists a mixed action α_i such that $\mathbb{E}_{\alpha_i}(u_{i,a_{-i}}) > u_{i,a_i}(a_{-i})$ for all $a_{-i} \in A_{-i}$.

[14] It is somewhat confusing to call "payoff function" the map that associates each action profile with a player's utility of the induced outcome. This has a historical origin. The early work on game theory assumed that *outcomes are* profiles $y = (y_i)_{i\in I} \in \mathbb{R}^I$ of *monetary payoffs* and that *players are risk neutral*, hence they just want to maximize their own expected monetary payoff. In this case $v_i(y) = y_i$, and $u_i(a) = g_i(a)$, where $g_i(a)$ is i's "gain" given action profile a.

[15] Here we take for granted without proof that, for every conjecture, the subjective expected utility of a mixed action cannot be higher than the largest subjective expected utility of pure actions. Hence, justifiability can be expressed with reference to pure actions.

Note, when we state such sentences with logical quantifiers in the natural language, we sometimes put the "for all" quantification at the end of the sentence, as in the examples above. Sometimes we put the "there exists" quantification at the end of the sentence, but with different words, such as "for some ..." or "for at least one ...". As long as the meaning is clear, this is fine.

Predicates about justifiability and dominance However, it is often convenient to express such sentences with formulas containing $\forall$ and/or $\exists$. Such *predicate formulas must be written with all the quantifiers at the beginning and in the correct order.* With this, there are simple rules to manipulate such formulas according to logical calculus that help us in our reasoning. Here we focus on the rule to obtain the *negation* of a sentence in predicate calculus.

The formula for "a_i is justifiable" is

$$\exists \mu^i \in \Delta\left(A_{-i}\right), \forall b_i \in A_i, \mathbb{E}_{\mu^i}\left(u_{i,a_i}\right) \geq \mathbb{E}_{\mu^i}\left(u_{i,b_i}\right).$$

The formula for "a_i is dominated" is

$$\exists \alpha_i \in \Delta\left(A_i\right), \forall a_{-i} \in A_{-i}, \mathbb{E}_{\alpha_i}\left(u_{i,a_{-i}}\right) > u_{i,a_i}\left(a_{-i}\right).$$

Note that *the order of quantifiers is crucial*, as illustrated by the following example.

Example 9 (Justifiability and order of quantifiers) *Consider the following payoff function u_i:*

$u_i:$	ℓ	r
t	0	1
m	$\frac{1}{3}$	$\frac{1}{3}$
d	1	0

where $A_i = \{t, m, d\}$ (u_i is obtained by appending utility function $v_i(G) = 1$, $v_i(M) = 1/3$, $v_i(B) = 0$ to the game form of Example 8). It can be checked that action m is not justifiable*: indeed, the special case for this payoff function and action m of the justifiability formula written above can be re-written as follows*

$$\exists \mu^i \in \Delta\left(\{\ell, r\}\right), \left(\frac{1}{3} \geq \mu^i(r)\right) \wedge \left(\frac{1}{3} \geq \mu^i(\ell)\right),$$

(where $\wedge$ means "and"), which is false *because either $\mu^i(r) \geq \frac{1}{2}$ or $\mu^i(\ell) \geq \frac{1}{2}$. Yet, inverting the order of quantifiers in the justifiability formula, we obtain*

$$\forall b_i \in \{t, m, d\}, \exists \mu^i \in \Delta(\{\ell, r\}), \mathbb{E}_{\mu^i}(u_{i,m}) \geq \mathbb{E}_{\mu^i}(u_{i,b_i}),$$

which is true, *because $\mathbb{E}_{\mu^i}(u_{i,m}) \geq \mathbb{E}_{\mu^i}(u_{i,t})$ if $\mu^i(r) \leq \frac{1}{3}$, and $\mathbb{E}_{\mu^i}(u_{i,m}) \geq \mathbb{E}_{\mu^i}(u_{i,d})$ if $\mu^i(\ell) \leq \frac{1}{3}$. A similar example applies to dominance.*[16] ▲

Predicate formulas with quantifiers can be used to express sets. For example, the set of justifiable actions is

$$J_i := \left\{a_i \in A_i : \exists \mu^i \in \Delta(A_{-i}), \forall b_i \in A_i, \mathbb{E}_{\mu^i}(u_{i,a_i}) \geq \mathbb{E}_{\mu^i}(u_{i,b_i})\right\},$$

and the set of dominated actions is

$$D_i := \left\{a_i \in A_i : \exists \alpha_i \in \Delta(A_i), \forall a_{-i} \in A_{-i}, \mathbb{E}_{\alpha_i}\left(u_{i,a_{-i}}\right) > u_{i,a_i}(a_{-i})\right\}.$$

Predicates and negation Is there a relatively simple and explicit way to write the sets of unjustifiable and undominated actions? Yes, it is enough to apply the following rule:

- *To obtain the negation of a predicate formula with quantifiers, replace each $\forall$ with $\exists$, replace each $\exists$ with $\forall$, and replace the predicate in the formula with its negation (e.g., replace $<$ with $\geq$).*

For example, the negation of the false formula

$$\forall n \in \mathbb{N}, \exists m \in \mathbb{N}, m < n$$

is the true formula

$$\exists n \in \mathbb{N}, \forall m \in \mathbb{N}, m \geq n.$$

To check that the first formula is false and the second (its negation) is true, just consider the smallest natural number $n = 1$.[17]

[16] The crucial importance of the order of quantifiers explains why predicate formulas must list quantifiers on the left, so that their order reflects the reading order from left to right. Some mathematicians and theorists, casually write the universal quantifier $\forall$ on the right, but only in formulas that do not contain quantifier $\exists$. In particular, this is very common among decision theorists.

[17] We use $\mathbb{N}$ to denote the set of *strictly positive* integers. We denote the set of nonnegative integers by $\mathbb{N}_0$.

We can use this simple rule to express the sets of unjustifiable and undominated actions as follows:

$$A_i \backslash J_i = \left\{a_i \in A_i : \forall \mu^i \in \Delta\left(A_{-i}\right), \exists b_i \in A_i, \mathbb{E}_{\mu^i}\left(u_{i,a_i}\right) < \mathbb{E}_{\mu^i}\left(u_{i,b_i}\right)\right\}$$

and

$$A_i \backslash D_i = \left\{a_i \in A_i : \forall \alpha_i \in \Delta\left(A_i\right), \exists a_{-i} \in A_{-i}, \mathbb{E}_{\alpha_i}\left(u_{i,a_{-i}}\right) \leq u_{i,a_i}\left(a_{-i}\right)\right\}.$$

Predicates and optimization We can also relate predicate formulas involving inequalities with the sup and inf operations, which become max and min operations when we can take for granted that the sup and inf are always attained. For example, given the assumption that the game is finite, we can express justifiability as follows: a_i is justifiable iff, for some conjecture μ^i, there is a non-positive incentive to deviate from a_i to any action b_i, that is,

$$\exists \mu^i \in \Delta\left(A_{-i}\right), \forall b_i \in A_i, \mathbb{E}_{\mu^i}\left(u_{i,b_i}\right) - \mathbb{E}_{\mu^i}\left(u_{i,a_i}\right) \leq 0,$$

which is equivalent to

$$\min_{\mu^i \in \Delta(A_{-i})} \max_{b_i \in A_i} \left(\mathbb{E}_{\mu^i}\left(u_{i,b_i}\right) - \mathbb{E}_{\mu^i}\left(u_{i,a_i}\right)\right) \leq 0.$$

To see this, note that—for every given μ^i—$\mathbb{E}_{\mu^i}\left(u_{i,b_i}\right) - \mathbb{E}_{\mu^i}\left(u_{i,a_i}\right) \leq 0$ for all b_i in the finite set A_i iff

$$\max_{b_i \in A_i} \left(\mathbb{E}_{\mu^i}\left(u_{i,b_i}\right) - \mathbb{E}_{\mu^i}\left(u_{i,a_i}\right)\right) \leq 0.$$

Let $H\left(\mu^i\right)$ denote the above maximum. Then a_i is justifiable iff $H\left(\mu^i\right) \leq 0$ for some μ^i. By standard arguments, H is a continuous function over the compact domain $\Delta\left(A_i\right)$, therefore $\min_{\mu^i \in \Delta(A_{-i})} H\left(\mu^i\right)$ is well defined and $H\left(\mu^i\right) \leq 0$ for some μ^i iff $\min_{\mu^i \in \Delta(A_{-i})} H\left(\mu^i\right) \leq 0$.

Using arguments of this kind and the celebrated Maxmin Theorem of von Neumann, one can prove the following lemma of Wald and Pearce: $J_i = A_i \backslash D_i$ (see the Appendix of Chapter 3 of *Game Theory: Analysis of Strategic Thinking*).

References

[1] Anscombe, F. J. and R. J. Aumann (1963): "A Definition of Subjective Probability," *Annals of Mathematics and Statistics*, 34, 199-205.

[2] Battigalli, P., E. Catonini, and N. De Vito (2023): *Game Theory: Analysis of Strategic Thinking*. Typescript, Bocconi University.

[3] Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2017): "Mixed Extensions of Decision Problems under Uncertainty," *Economic Theory*, 63, 827-866.

[4] Savage, L.J. (1954): *The Foundations of Statistics*. New York: Wiley.

[5] von Neumann, J. and O. Morgenstern (2004 [1944]): *Theory of Games and Economic Behavior,* Princeton NJ: Princeton University Press.