A Tutorial on the Proof of the Existence of Nash Equilibria

Albert Xin Jiang   Kevin Leyton-Brown
Department of Computer Science,
University of British Columbia

1 Game-theoretic preliminaries

In this tutorial we detail a proof of Nash's famous theorem on the existence of Nash equilibria in finite games, first proving Sperner's lemma and Brouwer's fixed-point theorem. We begin with the definition of a finite game.

Definition 1 (Normal-form game) *A (finite,* n*-person)* normal-form game *is a tuple* (N, A, O, μ, u)*, where*

- N *is a finite set of* n *players, indexed by* i*;*
- $A = (A_1, \ldots, A_n)$*, where* A_i *is a finite set of* actions *(or* pure strategies*; we will use the terms interchangeably) available to player* i*. Each vector* $a = (a_1, \ldots, a_n) \in A$ *is called an* action profile *(or* pure strategy profile*);*
- O *is a set of outcomes;*
- $\mu : A \to O$ *determines the outcome as a function of the action profile; and*
- $u = (u_1, \ldots, u_n)$ *where* $u_i : O \to \mathbb{R}$ *is a real-valued* utility *(or* payoff*)* function *for player* i*.*

action, *or* pure strategy

action profile

pure strategy profile

utility function

payoff function

Often we do not need the notion of an outcome as distinct from a strategy profile. In such cases the game has the simpler form N, A, u, and in the remainder of the article we will adopt this form.

We have so far defined the actions available to each player in a game, but not yet his set of *strategies*, or his available choices. Certainly one kind of strategy is to select a single action and play it. We call such a strategy a *pure strategy*. We call a choice of pure strategy for each agent a *pure strategy profile*.

pure strategy

Players could also follow another, less obvious type of strategy: randomizing over the set of available actions according to some probability distribution. Such a strategy is called a mixed strategy. Although it may not be immediately obvious why a player should introduce randomness into his choice of action, in fact in a multiagent setting the role of mixed strategies is critical.

We define a mixed strategy for a normal form game as follows.

Definition 2 (Mixed strategy) *Let $(N, (A_1, \ldots, A_n), O, \mu, u)$ be a normal form game, and for any set X let $\Pi(X)$ be the set of all probability distributions over X. Then the set of* mixed strategies *for player i is $S_i = \Pi(A_i)$.* mixed strategy

Definition 3 (Mixed strategy profile) *The set of* mixed strategy profiles *is simply the Cartesian product of the individual mixed strategy sets, $S_1 \times \cdots \times S_n$.* mixed strategy profile

By $s_i(a_i)$ we denote the probability that an action a_i will be played under mixed strategy s_i. The subset of actions that are assigned positive probability by the mixed strategy s_i is called the *support* of s_i.

Definition 4 (Support) *The* support *of a mixed strategy s_i for a player i is the set of pure strategies $\{a_i | s_i(a_i) > 0\}$.* support of a mixed strategy

Note that a pure strategy is a special case of a mixed strategy, in which the support is a single action.

We have not yet defined the payoffs of players given a particular strategy profile, since the payoff matrix defines those directly only for the special case of pure strategy profiles. But the generalization to mixed strategies is straightforward, and relies on the basic notion of decision theory—*expected utility*. Intuitively, we first calculate the probability of reaching each outcome given the strategy profile, and then we calculate the average of the payoffs of the outcomes, weighted by the probabilities of each outcome. Formally, we define the expected utility as follows (overloading notation, we use u_i for both utility and expected utility). expected utility

Definition 5 (Expected utility of a mixed strategy) *Given a normal form game (N, A, u), the expected utility u_i for player i of the mixed strategy profile $s = (s_1, \ldots, s_n)$ is defined as*

$$u_i(s) = \sum_{a \in A} u_i(a) \prod_{j=1}^{n} s_j(a_j).$$

Now we will look at games from an individual agent's point of view, rather than from the vantage point of an outside observer. This will lead us to the most influential solution concept in game theory, the *Nash equilibrium*.

Our first observation is that if an agent knew how the others were going to play, his strategic problem would become simple. Specifically, he would be left with the single-agent problem of choosing a utility-maximizing action. Formally, define $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_n)$, a strategy profile s without agent i's strategy. Thus we can write $s = (s_i, s_{-i})$. If the agents other than i were to commit to play s_{-i}, a utility-maximizing agent i would face the problem of determining his best response.

Definition 6 (Best response) *Player i's* best response *to the strategy profile s_{-i} is a mixed strategy $s_i^* \in S_i$ such that $u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i})$ for all strategies $s_i \in S_i$.* best response

The best response is not necessarily unique. Indeed, except in the extreme case in which there is a unique best response that is a pure strategy, the number of best responses is always infinite. When the support of a best response s^* includes two or more actions, the agent must be indifferent between them—otherwise the agent would prefer to reduce the probability of playing at least one of the actions to zero. But thus *any* mixture of these actions must also be a best response, not only the particular mixture in s^*. Similarly, if there are two pure strategies that are individually best responses, any mixture of the two is necessarily also a best response.

Of course, in general an agent won't know what strategies the other players will adopt. Thus, the notion of best response is not a solution concept—it does not identify an interesting set of outcomes in this general case. However, we can leverage the idea of best response to define what is arguably the most central notion in non-cooperative game theory, the Nash equilibrium.

Definition 7 (Nash equilibrium) *A strategy profile $s = (s_1, \ldots, s_n)$ is a* Nash equilibrium *if, for all agents i, s_i is a best response to s_{-i}.* Nash equilibrium

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. This is because in a Nash equilibrium all of the agents simultaneously play best responses to each other's strategies.

2 Proving the existence of Nash equilibria

In this section we prove that every game has at least one Nash equilibrium.

Definition 8 (Convexity) *A set $C \subset \mathbb{R}^m$ is* convex *if for every $x, y \in C$ and $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in C$. For vectors $x^0, \ldots, x^n$ and nonnegative scalars $\lambda_0, \ldots, \lambda_n$ satisfying $\sum_{i=0}^n \lambda_i = 1$, the vector $\sum_{i=0}^n \lambda_i x^i$ is called a* convex combination *of $x^0, \ldots, x^n$.* convexity convex combination

For example, a cube is a convex set in $\mathbb{R}^3$; a bowl is not.

Definition 9 (Affine independence) *A finite set of vectors $\{x^0, \ldots, x^n\}$ in an Euclidean space is* affinely independent *if $\sum_{i=0}^n \lambda_i x^i = 0$ and $\sum_{i=0}^n \lambda_i = 0$ imply that $\lambda_0 = \ldots = \lambda_n = 0$.* affine independence

An equivalent condition is that $\{x^1 - x^0, x^2 - x^0, \ldots, x^n - x^0\}$ are linearly independent. Intuitively, a set of points is affinely independent if no three points from the set lie on the same line, no four points from the set lie on the same plane, and so on. For example, the set consisting of the origin 0 and the unit vectors $e^1, \ldots, e^n$ is affinely independent.

Next we define a simplex, which is an n-dimensional generalization of a triangle.

Definition 10 (n-simplex) *An* n-simplex*, denoted $x^0 \cdots x^n$, is the set of all convex combinations of the affinely independent set of vectors $\{x^0, \ldots, x^n\}$, i.e.* n-simplex

$$x^0 \cdots x^n = \left\{ \sum_{i=0}^n \lambda_i x^i : \forall i \in \{0, \ldots, n\},\ \lambda_i \geq 0; \text{ and} \sum_{i=0}^n \lambda_i = 1 \right\}.$$

Each x^i is called a *vertex* of the simplex $x^0 \cdots x^n$ and each k-simplex $x^{i_0} \cdots x^{i_k}$ is called a *k-face* of $x^0 \cdots x^n$, where $i_0, \ldots, i_k \in \{0, \ldots, n\}$. vertex k-face

Definition 11 (Standard n-simplex) *The* standard n-simplex $\triangle_n$ *is* $\{y \in \mathbb{R}^{n+1} : \sum_{i=0}^{n} y_i = 1, \forall i = 0, \ldots, n,\ y_i \geq 0\}$. standard n-simplex

In other words, the standard n-simplex is the set of all convex combinations of the $n+1$ unit vectors $e^0, \ldots, e^n$.

Definition 12 (Simplicial subdivision) *A* simplicial subdivision *of an n-simplex T is a finite set of simplexes $\{T_i\}$ for which $\bigcup_{T_i \in T} T_i = T$, and for any $T_i, T_j \in T$, $T_i \cap T_j$ is either empty or equal to a common face.* simplicial subdivision

Intuitively, this means that a simplex is divided up into a set of smaller simplexes that together occupy exactly the same region of space and that overlap only on their boundaries. Furthermore, when two of them overlap, the intersection must be an entire face of both subsimplexes. Figure 1 (left) shows a 2-simplex subdivided into 16 subsimplexes.

Let $y \in x^0 \cdots x^n$ denote an arbitrary point in a simplex. This point can be written as a convex combination of the vertices: $y = \sum_i \lambda_i x^i$. Now define a function that gives the set of vertices "involved" in this point: $\chi(y) = \{i : \lambda_i > 0\}$. We use this function to define a proper labeling.

Definition 13 (Proper labeling) *Let $T = x^0 \cdots x^n$ be simplicially subdivided, and let V denote the set of all distinct vertices of all the subsimplexes. A function $\mathcal{L} : V \to \{0, \ldots, n\}$ is a* proper labeling *of a subdivision if $\mathcal{L}(v) \in \chi(v)$.* proper labeling

One consequence of this definition is that the vertices of a simplex must all receive different labels. (Do you see why?) As an example, the subdivided simplex in Figure 1 (left) is properly labeled.

Definition 14 (complete labeling) *A subsimplex is* completely labeled *if $\mathcal{L}$ assumes all the values $0, \ldots, n$ on its set of vertices.* completely labeled subsimplex

For example in the subdivided triangle in Figure 1 (left), the sub-triangle at the very top is completely labeled.

Lemma 15 (Sperner's lemma) *Let $T_n = x^0 \cdots x^n$ be simplicially subdivided and let $\mathcal{L}$ be a proper labeling of the subdivision. Then there are an odd number of completely labeled subsimplexes in the subdivision.* Sperner's lemma

Proof. We prove this by induction on n. The case $n = 0$ is trivial. The simplex consists of a single point x^0. The only possible simplicial subdivision is $\{x^0\}$. There is only one possible labeling function, $\mathcal{L}(x^0) = 0$. Note that this is a proper labeling. So there is one completely labeled subsimplex, x^0 itself.

We now assume the statement to be true for $n-1$ and prove it for n. The simplicial subdivision of T_n induces a simplicial subdivision on its face $x^0 \cdots x^{n-1}$. This face is an $(n-1)$-simplex; denote it as T_{n-1}. The labeling function $\mathcal{L}$ restricted to T_{n-1} is a proper labeling of T_{n-1}. Therefore by the induction hypothesis there exist an odd number of $(n-1)$-subsimplexes in T_{n-1} that bear the labels

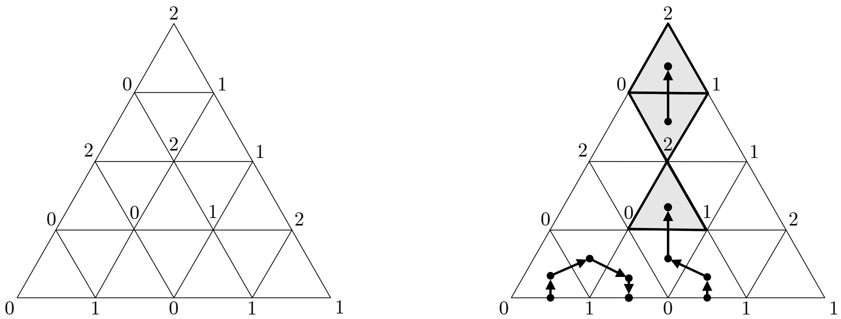

Figure 1: A properly labeled simplex (left), and the same simplex with completely-labeled subsimplexes shaded and three walks indicated (right).

$(0, \ldots, n-1)$. (To provide graphical intuition, we will illustrate the induction argument on a subdivided 2-simplex. In Figure 1 (left), observe that the bottom face x^0x^1 is a subdivided 1-simplex—a line segment—containing four subsimplexes, three of which are completely labeled.)

We now define rules for “walking” across our subdivided, labeled simplex T_n. The walk begins at an $(n-1)$-subsimplex with labels $(0, \ldots, n-1)$ on the face T_{n-1}; call this subsimplex b. There exists a unique n-subsimplex d that has b as a face; d's vertices consist of the vertices of b and another vertex z. If z is labeled n, then we have a completely labeled subsimplex and the walk ends. Otherwise, d has the labels $(0, \ldots, n-1)$, where one of the labels (say j) is repeated, and the label n is missing. In this case there exists exactly one other $(n-1)$-subsimplex that is a face of d and bears the labels $(0, \ldots, n-1)$. This is because each $(n-1)$-face of d is defined by all but one of d's vertices; since only the label j is repeated, an $(n-1)$-face of d has labels $(0, \ldots, n-1)$ if and only if one of the two vertexes with label j is left out. We know b is one such face, so there is exactly one other, which we call e. (For example, you can confirm in Figure 1 (left) that if a subtriangle has an edge with labels $(0, 1)$, then it is either completely labeled, or it has exactly one other edge with labels $(0, 1)$.) We continue the walk from e. We make use of the following property: an $(n-1)$-face of an n-subsimplex in a simplicially subdivided simplex T_n is either on an $(n-1)$-face of T_n, or the intersection of two n-subsimplexes. If e is on an $(n-1)$-face of T_n we stop the walk. Otherwise we walk into the unique other n-subsimplex having e as a face. This subsimplex is either completely labeled or has one repeated label, and we continue the walk in the same way we did with subsimplex d above.

Note that the walk is completely determined by the starting $(n-1)$-subsimplex. The walk ends either at a completely labeled n-subsimplex, or at a $(n-1)$-subsimplex with labels $(0, \ldots, n-1)$ on the face T_{n-1}. (It cannot end on any other face because $\mathcal{L}$ is a proper labeling.) Note also that every walk can be followed backwards: beginning from the end of the walk and following the same rule as above, we end up at the starting point. This implies that if a walk starts at t on

T_{n-1} and ends at t' on T_{n-1}, t and t' must be different, because otherwise we could reverse the walk and get a different path with the same starting point, contradicting the uniqueness of the walk. (Figure 1 (right) illustrates one walk of each of the kinds we have discussed so far: one that starts and ends at different subsimplexes on the face x^0x^1, and one that starts on the face x^0x^1 and ends at a completely labeled sub-triangle.) Since by the induction hypothesis there are an odd number of $(n-1)$-subsimplexes with labels $(0, \ldots, n-1)$ at the face T_{n-1}, there must be at least one walk that does not end on this face. Since walks that start and end on the face "pair up", there are thus an odd number of walks starting from the face that end at completely labeled subsimplexes. All such walks end at *different* completely labeled subsimplexes, because there is exactly one $(n-1)$-simplex face labeled $(0, \ldots, n-1)$ for a walk to enter from in a completely labeled subsimplex.

Not all completely labeled subsimplexes are led to by such walks. To see why, consider reverse walks starting from completely labeled subsimplexes. Some of these reverse walks end at $(n-1)$-simplexes on T_{n-1}, but some end at other completely labeled n-subsimplexes. (Figure 1 (right) illustrates one walk of this kind.) However, these walks just pair up completely labeled subsimplexes. There are thus an even number of completely labeled subsimplexes that pair up with each other, and an odd number of completely labeled subsimplexes that are led to by walks from the face T_{n-1}. Therefore the total number of completely labeled subsimplexes is odd. ■

Definition 16 (Compactness) *A subset of* $\mathbb{R}^n$ *is* compact *if the set is closed and bounded.* compactness

It is straightforward to verify that $\triangle_m$ is compact. A compact set has the property that every sequence in the set has a convergent subsequence.

Definition 17 (Centroid) *The* centroid *of a simplex* $x^0 \cdots x^m$ *is the "average" of its vertices,* $\frac{1}{m+1}\sum_{i=0}^{m} x^i$. centroid

We are now ready to use Sperner's lemma to prove Brouwer's fixed point theorem.

Theorem 18 (Brouwer's fixed point theorem) *Let* $f : \triangle_m \to \triangle_m$ *be continuous. Then* f *has a fixed point—that is, there exists some* $z \in \triangle_m$ *such that* $f(z) = z$. Brouwer's fixed point theorem

Proof. We prove this by first constructing a proper labeling of $\triangle_m$, then showing that as we make finer and finer subdivisions, there exist a subsequence of completely labeled subsimplexes that converges to a fixed point of f.

Part 1: $\mathcal{L}$ is a proper labeling. Let $\epsilon > 0$. We simplicially subdivide $\triangle_m$ such that the Euclidean distance between any two points in the same m-subsimplex is at most ϵ. We define a labeling function $\mathcal{L} : V \to \{0, \ldots, m\}$ as follows. For each v we choose a label satisfying

$$\mathcal{L}(v) \in \chi(v) \cap \{i : f_i(v) \leq v_i\}, \tag{1}$$

where v_i is the i^{th} component of v and $f_i(v)$ is the i^{th} component of $f(v)$. In other words, $\mathcal{L}(v)$ can be any label i such that $v_i > 0$ and f weakly decreases the i^{th}

component of v. To ensure that $\mathcal{L}$ is well-defined, we must show that the intersection on the right side of Equation (1) is always nonempty. (Intuitively, since v and $f(v)$ are both on the standard simplex $\triangle_m$, and on $\triangle_m$ each point's components sum to 1, there must exist a component of v that is weakly decreased by f. This intuition holds even though we restrict to the components in $\chi(v)$ because these are exactly all the positive components of v.) We now show this formally. For contradiction, assume otherwise. This assumption implies that $f_i(v) > v_i$ for all $i \in \chi(v)$. Recall from the definition of a standard simplex that $\sum_{i=0}^{m} v_i = 1$. Since by the definition of χ, $v_j > 0$ if and only if $j \in \chi(v)$, we have

$$\sum_{j \in \chi(v)} v_j = \sum_{i=0}^{m} v_i = 1. \tag{2}$$

Since $f_j(v) > v_j$ for all $j \in \chi(v)$,

$$\sum_{j \in \chi(v)} f_i(v) > \sum_{j \in \chi(v)} v_j = 1. \tag{3}$$

But since $f(v)$ is also on the standard simplex $\triangle_m$,

$$\sum_{j \in \chi(v)} f_i(v) \leq \sum_{i=0}^{m} f_i(v) = 1. \tag{4}$$

Equations (3) and (4) lead to a contradiction. Therefore $\mathcal{L}$ is well-defined; it is a proper labeling by construction.

Part 2: As $\epsilon \to 0$, completely labeled subsimplexes converge to fixed points of f. Since $\mathcal{L}$ is a proper labeling, by Sperner's lemma (15) there is at least one completely labeled subsimplex $p^0 \cdots p^m$ such that $f_i(p^i) \leq p^i$ for each i. Let $\epsilon \to 0$ and consider the sequence of centroids of completely labeled subsimplexes. Since $\triangle_m$ is compact, there is a convergent subsequence. Let z be its limit; then for all $i = 0, \ldots, m$, $p^i \to z$ as $\epsilon \to 0$. Since f is continuous we must have $f_i(z) \leq z_i$ for all i. This implies $f(z) = z$, because otherwise (by an argument similar to the one in Part 1) we would have $1 = \sum_i f_i(z) < \sum_i z_i = 1$, a contradiction. ■

Theorem 18 cannot be used directly to prove the existence of Nash equilibria. This is because a Nash equilibrium is a point in the set of mixed strategy profiles S. This set is not a simplex but rather a *simplotope*: a Cartesian product of simplexes. (Observe that each individual agent's mixed strategy *can* be understood as a point in a simplex.) However, it turns out that Brouwer's theorem can be extended beyond simplexes to simplotopes.[1] In essence, this is because every simplotope is topologically the same as a simplex (formally, they are *homeomorphic*).

simplotope

Definition 19 (Bijective function) *A function f is* injective *(or one-to-one) if $f(a) =$*

[1] An argument similar to our proof below can be used to prove a generalization of Theorem 18 to arbitrary convex and compact sets.

$f(b)$ *implies* $a = b$. *A function* $f : X \to Y$ *is* onto *if for every* $y \in Y$ *there exists* $x \in X$ *such that* $f(x) = y$. *A function is* bijective *if it is both injective and onto.* bijective

Definition 20 (Homeomorphism) *A set* A *is* homeomorphic *to a set* B *if there exists a continuous, bijective function* $h : A \to B$ *such that* h^{-1} *is also continuous. Such a function* h *is called a* homeomorphism. homeomorphism

Definition 21 (Interior) *A point* x *is an* interior point *of a set* $A \subset \mathbb{R}^m$ *if there is an open* m*-dimensional ball* $B \subset \mathbb{R}^m$ *centered at* x *such that* $B \subset A$. *The* interior *of a set* A *is the set of all its interior points.* interior

Corollary 22 (Brouwer's fixed point theorem, simplotopes) *Let* $K = \prod_{j=1}^{k} \triangle_{m_j}$ *be a simplotope and let* $f : K \to K$ *be continuous. Then* f *has a fixed point.*

Proof. Let $m = \sum_{j=1}^{k} m_j$. First we show that if K is homeomorphic to $\triangle_m$, then a continuous function $f : K \to K$ has a fixed point. Let $h : \triangle_m \to K$ be a homeomorphism. Then $h^{-1} \circ f \circ h : \triangle_m \to \triangle_m$ is continuous, where $\circ$ denotes function composition. By Theorem 18 there exists a z' such that $h^{-1} \circ f \circ h(z') = z'$. Let $z = h(z')$, then $h^{-1} \circ f(z) = z' = h^{-1}(z)$. Since h^{-1} is injective, $f(z) = z$.

We must still show that $K = \prod_{j=1}^{k} \triangle_{m_j}$ is homeomorphic to $\triangle_m$. K is convex and compact because each $\triangle_{m_j}$ is convex and compact, and a product of convex and compact sets is also convex and compact. Let the *dimension* of a subset of an Euclidean space be the number of independent parameters required to describe each point in the set. For example, an n-simplex has dimension n. Since each $\triangle_{m_j}$ has dimension m_j, K has dimension m. Since $K \subset \mathbb{R}^{m+k}$ and $\triangle_m \subset \mathbb{R}^{m+1}$ both have dimension m, they can be embedded in $\mathbb{R}^m$ as K' and $\triangle'_m$ respectively. Furthermore, whereas $K \subset \mathbb{R}^{m+k}$ and $\triangle_m \subset \mathbb{R}^{m+1}$ have no interior points, both K' and $\triangle'_m$ have non-empty interior. For example, a standard 2-simplex is defined in $\mathbb{R}^3$, but we can embed the triangle in $\mathbb{R}^2$. As illustrated in Figure 2 (left), the product of two standard 1-simplexes is a square, which can also be embedded in $\mathbb{R}$. We scale and translate K' into K'' such that K'' is strictly inside $\triangle'_m$. Since scaling and translation are homeomorphisms, and a chain of homeomorphisms is still a homeomorphism, we just need to find a homeomorphism $h : K'' \to \triangle'_m$. Fix a point a in the interior of K''. Define h to be the "radial projection" with respect to a, where $h(a) = a$ and for $x \in K'' \setminus \{a\}$,

$$h(x) = a + \frac{||x' - a||}{||x'' - a||}(x - a),$$

where x' is the intersection point of the boundary of $\triangle'_m$ with the ray that starts at a and passes through x, and x'' is the intersection point of the boundary of K'' with the same ray. Because K'' and $\triangle'_m$ are convex and compact, x'' and x' exist and are unique. Since a is an interior point of K'' and $\triangle_m$, $||x' - a||$ and $||x'' - a||$ are both positive. Intuitively, h scales x along the ray by a factor of $\frac{||x'-a||}{||x''-a||}$. Figure 2 (right) illustrates an example of this radial projection from a square simplotope to a triangle.

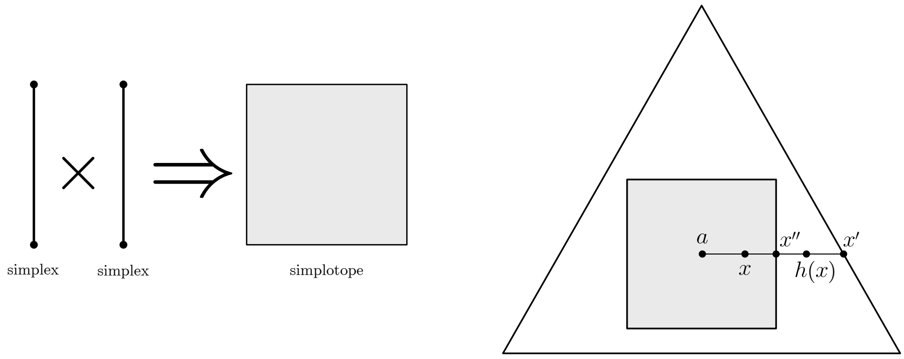

Figure 2: A product of two standard 1-simplexes is a square (a simplotope; left). The square is scaled and put inside a triangle (a 2-simplex), and an example of radial projection h is shown (right).

> Finally, it remains to show that h is a homeomorphism. It is relatively straightforward to verify that h is continuous. Since we know that $h(x)$ lies on the ray that starts at a and passes through x, given $h(x)$ we can reconstruct the same ray by drawing a ray from a which passes through $h(x)$. We can then recover x' and x'', and find x by scaling $h(x)$ along the ray by a factor of $\frac{||x''-a||}{||x'-a||}$. Thus h is injective. h is onto because given any point $y \in \triangle'_m$, we can construct the ray and find x such that $h(x) = y$. So h^{-1} has the same form as h except that the scaling factor is inverted, thus h^{-1} is also continuous. Therefore h is a homeomorphism.■

We are now ready to prove the existence of Nash equilibrium. Indeed, now that we have Corollary 22 and notation for discussing mixed strategies (Section 1), it is surprisingly easy. The proof proceeds by constructing a continuous $f : S \to S$ such that each fixed point of f is a Nash equilibrium. Then we use Corollary 22 to argue that f has at least one fixed point, and thus that Nash equilibria always exist.

Theorem 23 (Nash 1951) *Every game with a finite number of players and action profiles has at least one Nash equilibrium.*

> **Proof.** Given a strategy profile $s \in S$, for all $i \in N$ and $a_i \in A_i$ we define
>
> $$\varphi_{i,a_i}(s) = \max\{0, u_i(a_i, s_{-i}) - u_i(s)\}.$$
>
> We then define the function $f : S \to S$ by $f(s) = s'$, where
>
> $$\begin{aligned} s'_i(a_i) &= \frac{s_i(a_i) + \varphi_{i,a_i}(s)}{\sum_{b_i \in A_i} s_i(b_i) + \varphi_{i,b_i}(s)} \\ &= \frac{s_i(a_i) + \varphi_{i,a_i}(s)}{1 + \sum_{b_i \in A_i} \varphi_{i,b_i}(s)}. \end{aligned} \tag{5}$$

Intuitively, this function maps a strategy profile s to a new strategy profile s' in which each agent's actions that are better responses to s receive increased probability mass.

The function f is continuous since each φ_{i,a_i} is continuous. Since S is convex and compact and $f : S \rightarrow S$, by Corollary 22 f must have at least one fixed point. We must now show that the fixed points of f are the Nash equilibria.

First, if s is a Nash equilibrium then all φ's are 0, making s a fixed point of f.

Conversely, consider an arbitrary fixed point of f, s. By the linearity of expectation there must exist at least one action in the support of s, say a'_i, for which $u_{i,a'_i}(s) \leq u_i(s)$. From the definition of φ, $\varphi_{i,a'_i}(s) = 0$. Since s is a fixed point of f, $s'_i(a'_i) = s_i(a'_i)$. Consider Equation (5), the expression defining $s'_i(a'_i)$. The numerator simplifies to $s_i(a'_i)$, and is positive since a'_i is in i's support. Hence the denominator must be 1. Thus for any i and $b_i \in A_i$, $\varphi_{i,b_i}(s)$ must equal 0. From the definition of φ, this can only occur when no player can improve his expected payoff by moving to a pure strategy. Therefore s is a Nash equilibrium. ■

3 History and references

In 1950 John Nash introduced the concept of what would become known as the "Nash equilibrium" [Nash, 1950; Nash, 1951], without a doubt the most influential concept in game theory to this date. The proof in [Nash, 1950] uses Kakutani's fixed point theorem; our proof of Theorem 23 follows [Nash, 1951]. Lemma 15 is due to Sperner [1928] and Theorem 18 is due to Brouwer [1912]; our proofs follow Border [1985]. We thank Éva Tardos and Tim Roughgarden for making available notes that we drew on for our proofs of Sperner's lemma and the general form of Brouwer's fixed-point theorem, respectively. This report is an excerpt from [Shoham & Leyton-Brown, 2007].

References

Border, K. C. (1985). *Fixed point theorems with applications to economics and game theory*. Cambridge University Press.

Brouwer, L. E. J. (1912). Uber abbildung von mannigfaltikeiten. *Mathematische Annalen*, *71*, 97–115.

Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences USA*, *36*, 48–49. Reprinted in H. Kuhn (ed.), *Classics in Game Theory*, Princeton University Press, Princeton, New Jersey, 1997.

Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, *54*, 286–295.

Shoham, Y., & Leyton-Brown, K. (2007). Multiagent systems: Algorithmic, game-theoretic and logical foundations. Book in preparation.

Sperner, E. (1928). Neuer beweis fur die invarianz der dimensionszahl und des gebietes. *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universitat*, *6*, 265–272.