Stochastic Choice and Dynamics based on Pairwise Comparisons

Felix Brandt
Technical University of Munich

Introduction

A recurring problem in various disciplines is of the following form: there is a set of alternatives and any pair of alternatives can be compared with each other. A pairwise comparison reveals which of the two alternatives is better than the other. Moreover, comparisons may be cyclic (a is better than b, which is better than c, which is better than a) and probabilistic (a beats b with probability $2/3$). Since none of the alternatives needs to be superior to all the other ones, central normative questions in this context are which concepts are suitable to take over the role of the conventional notion of maximality and which axiomatic properties these concepts should satisfy. At the same time, one can consider dynamic processes in which a population of entities, each associated with an alternative, engages in pairwise interactions such that the loser of a pairwise comparison will be associated with the alternative of the winner. Descriptive questions in this context concern the resulting sequence of distributions and its convergence behavior.

Below is a list of six diverse areas where such questions have recently appeared.

Game playing (Balduzzi et al., 2018; Vinyals et al., 2019). When using machine learning to master difficult two-player games via self-play, large populations of bots compete with each other in pairwise matchups. Given the recorded wins and losses, how can a “universal sparring partner” be randomly selected such that, against any other bot, the expected number of wins is at least as large as the expected number of losses?

Voting (Brandl et al., 2016). There is a group of voters, each of which rank-orders the set of available candidates. Which lotteries over the candidates satisfy desirable properties such as Condorcet-consistency (a candidate that beats all other candidates in pairwise comparisons is selected with probability 1), population-consistency (whenever two disjoint electorates agree on a lottery, this lottery should also be chosen for the union of both electorates), and independence of clones (the probability with which a candidate is selected is unaffected by introducing clone of other candidates)?

Machine learning (Dudík et al., 2015; Balsubramani et al., 2016). A company has several online ads targeted at different user groups. Each time a user visits a website he is confronted with two ads and the website tracks which of the ads, if any, the user clicks on. Initially, no information on user behavior is available. Which randomly generated sequence of ads minimizes expected regret? Here, the regret of a sequence of ads is the maximal probability that a fixed sequence of identical ads is more successful in attracting clicks by a random user.

Opinion formation (Brandl and Brandt, 2022). There is a group of agents, each of which entertains one of many possible opinions. Agents come together in random pairwise interactions, in which they try to convince each other of their opinion. There are fixed probabilities with which one opinion beats another and, with some small probability, an agent randomly changes his opinion. What can be said about the distribution of opinions when running this process for a large number of rounds?

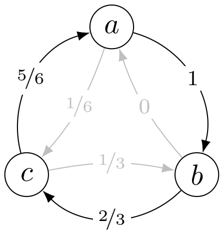
Population biology (Laslier and Laslier, 2017; Grilli et al., 2017). There is a continuum of individuals, which are assigned to some species, and for each pair of species there is a fixed probability with which one supplants the other. In each round, one individual dies and is replaced with an individual whose species is determined as the result of a competition based on pairwise random interactions: three individuals are randomly selected and the winner of the comparison between the first two faces off against the third one. Against which distribution of species does this process converge?

Quantum physics (Knebel et al., 2015). There is a continuum of quantum particles, so-called bosons, which are assigned one of several possible quantum states. Bosons engage in pairwise random interactions and there are fixed probabilities specifying with which probability a boson transitions from one state to the other when encountering another boson. What is the temporal average of quantum state distributions generated by this process in the limit?

As it turns out, the answers to *all* these questions are intimately connected to mixed equilibria of symmetric zero-sum games. That is precisely the class of games that Émile Borel's initial contribution to the theory of games, which is celebrated with this volume, was exclusively concerned with (Borel, 1921, 1953).

The Model

Let A be a finite set of alternatives and $\Delta = \{p \in [0,1]^A \colon \sum_{x \in A} p_x = 1\}$ the set of all lotteries over alternatives. The square matrix $M \in [0,1]^{A \times A}$ describes the pairwise comparisons between alternatives. For two alternatives x and y, $M_{xy} \in [0,1]$ says how much x is preferred to y ($1/2$ represents indifference). By convention, $M_{xx} = 0$ and $M_{xy} + M_{yx} = 1$ for all $x, y \in A$. These numbers could, for example, be interpreted as dominance probabilities, relative strengths of pairwise comparisons, or as fractions of voters who prefer one alternative to another. I will refer to them as dominance



$$M = \underbrace{\begin{array}{c} \\ a \\ b \\ c \end{array}\!\!\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 1 & 1/6 \\ 0 & 0 & 2/3 \\ 5/6 & 1/3 & 0 \end{pmatrix} \end{array}}_{\text{comparison matrix}} \qquad \tilde{M} = \underbrace{\begin{array}{c} \\ a \\ b \\ c \end{array}\!\!\begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{pmatrix} 0 & 1 & -2/3 \\ -1 & 0 & 1/3 \\ 2/3 & -1/3 & 0 \end{pmatrix} \end{array}}_{M - M^\intercal}$$

Figure 1: Example of a comparison matrix M (center), its corresponding directed graph (left), and the induced skew-symmetric matrix $\tilde{M}$ (right).

probabilities here. The important special case where M is a binary matrix, i.e., all entries of M are either 0 or 1, corresponds to complete and antisymmetric relations and thus to tournament graphs. The matrix M induces a skew-symmetric matrix $\tilde{M} = M - M^\intercal$, i.e., $\tilde{M}_{xy} = -\tilde{M}_{yx}$ for all $x, y \in A$.

Maximal Lotteries

Obviously, there need not be a "maximal" alternative $x \in A$ in the sense that it defeats every other alternative with a comparison probability of at least $1/2$, i.e., $\tilde{M}_{xy} \geq 0$ for all $y \in A$. However, it follows from the minimax theorem (von Neumann, 1928) that there has to be a "maximal" lottery $p \in \Delta$ such that $p^\intercal \tilde{M} q \geq 0$ for all lotteries $q \in \Delta$, i.e., p performs at least as well as any other lottery. This amounts to $\max_{p\in\Delta} \min_{q\in\Delta} p^\intercal \tilde{M} q = 0$, which holds because

$$\begin{aligned} \max_{p\in\Delta} \min_{q\in\Delta} p^\intercal \tilde{M} q &= \min_{q\in\Delta} \max_{p\in\Delta} p^\intercal \tilde{M} q = \min_{q\in\Delta} \max_{p\in\Delta} -q^\intercal \tilde{M} p \\ &= -\max_{q\in\Delta} \min_{p\in\Delta} q^\intercal \tilde{M} p = -\max_{p\in\Delta} \min_{q\in\Delta} p^\intercal \tilde{M} q, \end{aligned}$$

where the first equality follows from the minimax theorem and the second one from the skew-symmetry of $\tilde{M}$. In game-theoretic terms, $\tilde{M}$ can be interpreted as a symmetric zero-sum game. The value of any such game is 0 and the maximal lotteries described above are the probability distributions of the maximin or Nash equilibrium strategies of $\tilde{M}$. It does not matter whether one takes the equilibrium strategies of the row or column player since the game is symmetric. In summary,

$$\text{a lottery } p \in \Delta \text{ is } \textit{maximal} \text{ if } p^\intercal \tilde{M} \geq \mathbf{0}$$

and every comparison matrix M admits at least one maximal lottery.

Consider, for example, pairwise comparisons among three alternatives $A = \{a, b, c\}$ given in Figure 1. The matrix $\tilde{M}$ admits a unique maximal lottery $p = (1/6, 1/3, 1/2)^\intercal$ since

$$p^\intercal \tilde{M} = \begin{pmatrix} 1/6 \\ 1/3 \\ 1/2 \end{pmatrix}^\intercal \begin{pmatrix} 0 & 1 & -2/3 \\ -1 & 0 & 1/3 \\ 2/3 & -1/3 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \geq \mathbf{0}.$$

The term *maximal lottery* goes to back to Fishburn (1984) who introduced this concept in the context of social choice theory. The idea to circumvent the Condorcet paradox—which lies at the heart of Arrow's impossibility theorem—by introducing the notion of a randomized Condorcet winner can be traced back to Kreweras (1965). Maximal lotteries have been rediscovered several times under various names. Depending on whether the comparison matrix is binary or not, the support of maximal lotteries has been analyzed under the name *bipartisan set* (a term proposed by Roger Myerson) or *essential set* (Laffond et al., 1993; Dutta and Laslier, 1999). Felsenthal and Machover (1992) refer to maximal lotteries as the *game theory procedure* and Rivest and Shen (2010) as the *game theory method*. More recently, maximal lotteries have been rediscovered again in machine learning, where they are called *von Neumann winners* in the context of dueling multi-armed bandits (Dudík et al., 2015; Balsubramani et al., 2016) or *Nash averaging* when evaluating the performance of players in round-robin tournaments (Balduzzi et al., 2018; Vinyals et al., 2019).

Uniqueness and Support Size

Fisher and Ryan (1992) and Laffond et al. (1993) have shown independently that there is a *unique* maximal lottery when M is binary. This statement was strengthened by Laffond et al. (1997) who extended it to matrices whose entries are fractions with odd denominators. An even weaker sufficient condition based on congruencies was provided by Le Breton (2005). In a more abstract sense, maximal lotteries are almost always unique since the set of all matrices M that admit multiple maximal lotteries has measure zero and is nowhere dense in $[0,1]^{A\times A}$. This follows from considering the dimension of the kernel of sub-matrices of $\tilde{M}$. Fisher and Reeves (1995) and Brandl (2017) analyzed the size of the support of maximal lotteries for random matrices. Under fairly general symmetric distributions of matrices, maximal lotteries randomize over half of the alternatives on average.

Axiomatic Properties

Maximal lotteries satisfy a number of axiomatic properties that are deemed desirable in various contexts. In the following, the ML probability of an alternative is the probability that a maximal lottery assigns to the alternative based on a given comparison matrix.

- Any alternative that defeats every other alternative with probability greater than $1/2$ has ML probability 1. Such an alternative corresponds to a strict pure Nash equilibrium of $\tilde{M}$. Similarly, any alternative that is defeated by every other alternative with probability greater than $1/2$ has ML probability 0 (Laffond et al., 1993; Laslier, 2000).

- A lottery that is maximal for two comparison matrices is also maximal for any convex combination of both matrices (Brandl et al., 2016; Brandl and Brandt, 2019).

- A lottery remains maximal when removing alternatives with ML probability 0 (Laffond et al., 1993; Laslier, 2000).
- A maximal lottery is unaffected by reducing the dominance probabilities of alternatives with ML probability 0. In particular, the dominance probabilities between two such alternatives are irrelevant (Laslier, 1997).
- When increasing the dominance probability of x against y, then the ML probability of x relative to the ML probability of y does not decrease (Brandl et al., 2022).
- The ML probability of an alternative is unaffected by cloning other alternatives. More generally, maximal lotteries of comparison matrices that admit a substitution decomposition can be decomposed into the maximal lotteries of each component (Laffond et al., 1996; Laslier, 2000; Brandl et al., 2016).

In probabilistic social choice, some of these properties have been used to completely characterize maximal lotteries. There are characterizations using Arrow's independence of irrelevant alternatives and Pareto efficiency (Brandl and Brandt, 2020) as well as population-consistency and composition-consistency (Brandl et al., 2016), and Condorcet-consistency and a strong version of Moulin's participation condition (Brandl et al., 2019).[1] The support of maximal lotteries has been characterized using axioms that prescribe consistency with respect to variable sets of alternatives (Laslier, 2000; Brandt et al., 2018). Moreover, versions of these axioms have been used to characterize Nash equilibrium in zero-sum games (Brandl and Brandt, 2019).

Computation and Dynamics

Maximal lotteries correspond to equilibrium strategies of a zero-sum game and can thus be computed in polynomial time via linear programming. This entails a number of convenient technical properties, e.g., the set of maximal lotteries for a given comparison matrix is convex and the function that maps a comparison matrix to the corresponding set of maximal lotteries is upper hemi-continuous. It follows from the well-known reduction from linear programs to symmetric two-player zero-sum games due to Dantzig (1951) that deciding whether an action is played with positive probability in a symmetric zero-sum game is at least as hard as deciding feasibility of a linear program. The latter problem is known to be P-complete, which implies P-completeness of the former (see, also, Brandt and Fischer, 2008). As a consequence,

computing maximal lotteries belongs to the "hardest" problems that can be solved in polynomial time.

As alluded to at the beginning of this essay, maximal lotteries emerge through several simple dynamic processes in which entities change their type based on the comparison matrix. These types of entities could be agents that entertain the same opinion, creatures

[1] Note, however, that maximal lotteries violate strategyproofness. In fact, only random dictatorships are strategyproof and efficient (Gibbard, 1977; Hylland, 1980).

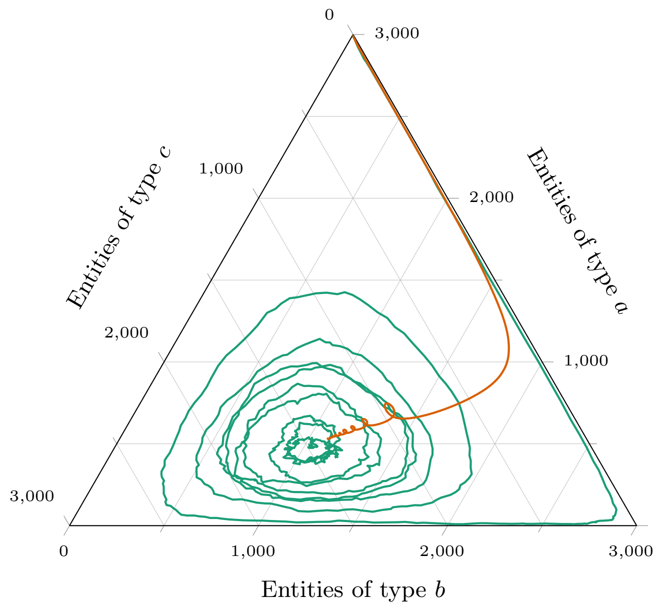


Figure 2: Simulation of a simple replacement process that approaches the unique maximal lottery of the comparison matrix M given in Figure 1.
There are 3,000 entities, who are initially all of type a, and the process runs for 300,000 rounds. In each round, two random entities engage in a pairwise comparison with probability 0.98. With the remaining probability, the type of a random entity "mutates" to a random type. The green line depicts the distribution of types among the entities over time while the red line depicts the temporal average of type distributions until the given round. The type distribution corresponding to the maximal lottery is $(500, 1000, 2500)$.

that belong to the same species, or particles that are in the same quantum state. Let me here consider the discrete and probabilistic model due to Brandl and Brandt (2022), where there is a fixed finite number of entities, each of which belongs to some type. The initial distribution of types is irrelevant. Types correspond to alternatives and are thus elements of A. In each round of the process two random entities come together and are randomly compared with respect to the dominance probabilities of the comparison matrix M. The entity of the inferior type is replaced with one of the superior type. Additionally, every now and then, the type of a random entity "mutates" to a random type.[2] It can be shown that,

when the number of entities is sufficiently large and the process is run for sufficiently many rounds, the distribution of types among the entities is close to a maximal lottery of the comparison matrix with high probability.

As a consequence, the temporal average of type distributions converges to a lottery close to a maximal lottery with probability 1. Moreover, the probability that the distribution is close to a maximal lottery increases exponentially fast. Figure 2 shows an example run of this process for the comparison matrix given in Figure 1. The contin-

[2] Without these mutations, the distribution of types will eventually be degenerate with probability 1.

uous and deterministic version of this process is related to the replicator equation in population dynamics and evolutionary game theory (see, e.g., Taylor and Jonker, 1978; Schuster and Sigmund, 1983; Hofbauer and Sigmund, 1998).

Apart from providing a strong argument for the natural emergence of maximal lotteries, this result is also of interest from a computational perspective. Since computing a maximal lottery is equivalent to solving a linear program, the process can be seen as a probabilistic algorithm that approximates a polynomial-time computable function represented by the comparison matrix. In contrast to traditional computing devices such as Turing machines, the dynamic process is based on unordered elementary entities that randomly interact according to very simple replacement rules.

Outlook

In this essay, I gave several examples where equilibrium strategies of symmetric zero-sum games, whose study was initiated by Émile Borel, appear unexpectedly in diverse fields such as economics, computer science, biology, and physics. Some of the attractive properties of equilibrium distributions such as their robustness with respect to clones have been observed independently and repeatedly with different applications in mind. I believe this area offers a lot of potential for inter-disciplinary research, in particular, with respect to the recent developments in machine learning and dynamic processes.

Acknowledgements

This chapter is based on work supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/11-2 and BR 2312/12-1. I am grateful to Florian Brandl for many captivating discussions on maximal lotteries.

References

D. Balduzzi, K. Tuyls, J. Perolat, and T. Graepel. Re-evaluting evaluation. In *Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS)*, pages 3272–3283, 2018.

A. Balsubramani, Z. Karnin, R. E. Schapire, and M. Zoghi. Instance-dependent regret bounds for dueling bandits. In *Proceedings of the 29th Annual Conference on Learning Theory (COLT)*, pages JMLR 49:1–24, 2016.

E. Borel. La théorie du jeu et les équations intégrales à noyau symétrique. *Comptes Rendus de l'Académie des Sciences*, 173:1304–1308, 1921.

E. Borel. The theory of play and integral equations with skew symmetric kernels. *Econometrica*, 21(1), 1953.

F. Brandl. The distribution of optimal strategies in symmetric zero-sum games. *Games and Economic Behavior*, 104:674–680, 2017.

F. Brandl and F. Brandt. Justifying optimal play via consistency. *Theoretical Economics*, 14(4): 1185–1201, 2019.

F. Brandl and F. Brandt. Arrovian aggregation of convex preferences. *Econometrica*, 88(2): 799–844, 2020.

F. Brandl and F. Brandt. A natural adaptive process for collective decision-making. 2022. Working paper.

F. Brandl, F. Brandt, and H. G. Seedig. Consistent probabilistic social choice. *Econometrica*, 84(5):1839–1880, 2016.

F. Brandl, F. Brandt, and J. Hofbauer. Welfare maximization entices participation. *Games and Economic Behavior*, 14:308–314, 2019.

F. Brandl, F. Brandt, and C. Stricker. An analytical and experimental comparison of maximal lottery schemes. *Social Choice and Welfare*, 58(1):5–38, 2022.

F. Brandt and F. Fischer. Computing the minimal covering set. *Mathematical Social Sciences*, 56(2):254–268, 2008.

F. Brandt, M. Brill, H. G. Seedig, and W. Suksompong. On the structure of stable tournament solutions. *Economic Theory*, 65(2):483–507, 2018.

G. B. Dantzig. A proof of the equivalence of the programming problem and the game problem. In T. C. Koopmans, editor, *Activity Analysis of Production and Allocation*, pages 330–335. John Wiley & Sons Inc., 1951.

M. Dudík, K. Hofmann, R. E. Schapire, A. Slivkins, and M. Zoghi. Contextual dueling bandits. In *Proceedings of the 28th Annual Conference on Learning Theory (COLT)*, pages JMLR 40:563–587, 2015.

B. Dutta and J.-F. Laslier. Comparison functions and choice correspondences. *Social Choice and Welfare*, 16(4):513–532, 1999.

D. S. Felsenthal and M. Machover. After two centuries should Condorcet's voting procedure be implemented? *Behavioral Science*, 37(4):250–274, 1992.

P. C. Fishburn. Probabilistic social choice based on simple voting comparisons. *Review of Economic Studies*, 51(4):683–692, 1984.

D. C. Fisher and R. B. Reeves. Optimal strategies for random tournament games. *Linear Algebra and its Applications*, 217:83–85, 1995.

D. C. Fisher and J. Ryan. Optimal strategies for a generalized "scissors, paper, and stone" game. *American Mathematical Monthly*, 99(10):935–942, 1992.

A. Gibbard. Manipulation of schemes that mix voting with chance. *Econometrica*, 45(3):665–681, 1977.

J. Grilli, G. Barabás, M. J. Michalska-Smith, and S. Allesina. Higher-order interactions stabilize dynamics in competitive network models. *Nature*, 548:210–213, 2017.

J. Hofbauer and K. Sigmund. *Evolutionary games and population dynamics*. Cambridge University Press, 1998.

A. Hylland. Strategyproofness of voting procedures with lotteries as outcomes and infinite sets of strategies. Mimeo, 1980.

J. Knebel, M. F. Weber, T. Krüger, and E. Frey. Evolutionary games of condensates in coupled birth–death processes. *Nature Communications*, 6(6977), 2015.

G. Kreweras. Aggregation of preference orderings. In *Mathematics and Social Sciences I: Proceedings of the seminars of Menthon-Saint-Bernard, France (1–27 July 1960) and of Gösing, Austria (3–27 July 1962)*, pages 73–79, 1965.

G. Laffond, J.-F. Laslier, and M. Le Breton. The bipartisan set of a tournament game. *Games and Economic Behavior*, 5(1):182–201, 1993.

G. Laffond, J. Lainé, and J.-F. Laslier. Composition-consistent tournament solutions and social choice functions. *Social Choice and Welfare*, 13(1):75–93, 1996.

G. Laffond, J.-F. Laslier, and M. Le Breton. A theorem on symmetric two-player zero-sum games. *Journal of Economic Theory*, 72(2):426–431, 1997.

B. Laslier and J.-F. Laslier. Reinforcement learning from comparisons: Three alternatives is enough, two is not. *The Annals of Applied Probability*, 27(5):2907–2925, 2017.

J.-F. Laslier. *Tournament Solutions and Majority Voting*. Springer-Verlag, 1997.

J.-F. Laslier. Aggregation of preferences with a variable set of alternatives. *Social Choice and Welfare*, 17(2):269–282, 2000.

M. Le Breton. On the uniqueness of equilibrium in symmetric two-player zero-sum games with integer payoffs. *Économie publique*, 17(2):187–195, 2005.

R. L. Rivest and E. Shen. An optimal single-winner preferential voting system based on game theory. In *Proceedings of the 3rd International Workshop on Computational Social Choice (COMSOC)*, pages 399–410, 2010.

P. Schuster and K. Sigmund. Replicator dynamics. *Journal of Theoretical Biology*, 100(3): 533–538, 1983.

P. D. Taylor and L. B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1–2):145–156, 1978.

O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev, J. Oh, D. Horgan, M. Kroiss, I. Danihelka, A. Huang, L. Sifre, T. Cai, J. P. Agapiou, M. Jaderberg, A. S. Vezhnevets, R. Leblond, T. Pohlen, V. Dalibard, D. Budden, Y. Sulsky, J. Molloy, T. L. Paine, C. Gulcehre, Z. Wang, T. Pfaff, Y. Wu, R. Ring, D. Yogatama, D. Wünsch, K. McKinney, O. Smith, T. Schaul, T. Lillicrap, K. Kavukcuoglu, D. Hassabis, C. Apps, and D. Silver. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 575:350–354, 2019.

J. von Neumann. Zur Theorie der Gesellschaftspiele. *Mathematische Annalen*, 100(1):295–320, 1928.