A More Perfect Algorithm

Computing citizens' assemblies more fairly empowers democracy

By Ariel Procaccia

Illustration by Montse Galbany

Ariel Procaccia is Gordon McKay Professor of Computer Science at Harvard University. He is an expert in algorithms and artificial intelligence and is especially interested in questions of societal importance.

N 1983 THE EIGHTH AMENDMENT TO THE IRISH CONSTITUTION ENSHRINED AN ABORTION BAN that had prevailed in the nation for more than a century. Public opinion on the issue shifted in the new millennium, however, and by 2016 it was clear that a real debate could no longer be avoided. But even relatively progressive politicians had long steered clear of the controversy rather than risk alienating voters. Who would be trustworthy and persuasive enough to break the deadlock?

The answer was a bunch of ordinary people. Seriously. The Irish Parliament convened a citizens' assembly, whose 99 members were chosen at random. The selection process ensured that the group's composition represented the Irish population along dimensions such as age, gender and geography. Over several months in 2016 and 2017, the assembly heard expert opinions and held extensive discussions regarding the legalization of abortion. Its recommendation, supported by a significant majority of members, was to allow abortions in all circumstances, subject to limits on the length of pregnancy. These conclusions set the stage for a 2018 referendum in which 66 percent of Ireland's voters chose to repeal the Eighth Amendment, enabling abortion to be legalized. Such an outcome had been almost inconceivable a few years earlier.

The Irish citizens' assembly is just one example of a widespread phenomenon. In recent years hundreds of such groups have convened around the world, their members randomly selected from the concerned population and given time and information to aid their deliberations. Citizens' assemblies in France, Germany, the U.K., Washington State and elsewhere have charted pathways for reducing carbon emissions. An assembly in Canada sought methods of mitigating hate speech and fake news; another in Australia recommended ethical approaches to human genome editing; and yet another in Oregon identified policies for COVID pandemic recovery. Taken together, these assemblies have demonstrated an impressive capacity to uncover the will of the people and build consensus.

The effectiveness of citizens' assemblies isn't surprising. Have you ever noticed how politicians grow a spine the moment they decide not to run for reelection? Well, a citizens' assembly is a bit like a legislature whose members make a pact barring them from seeking another term in office. The randomly selected members are not beholden to party machinations or outside interests; they are free to speak their mind and vote their conscience.

What's more, unlike elected bodies, these assemblies are chosen to mirror the population, a property that political theorists refer to as descriptive representation. For example, a typical citizens' assembly has a roughly equal number of men and women (some also ensure nonbinary participation), whereas the average proportion of seats held by women in national parliaments worldwide was 26 percent in 2021—a marked increase from 12 percent in 1997 but still far from gender balance. Descriptive representation, in turn, lends legitimacy to the assembly: citizens seem to find decisions more acceptable when they are made by people like themselves.

As attractive as descriptive representation is, there are practical obstacles to realizing it while adhering to the principle of random selection. Overcoming these hurdles has been a passion of mine for the past few years. Using tools from mathematics and computer science, my collaborators and I developed an algorithm for the selection of citizens' assemblies that many practitioners around the world are using. Its story provides a glimpse into the future of democracy—and it begins a long time ago.

THE GODDESS OF CHANCE

CITIZENS' ASSEMBLIES are the latest incarnation of an idea called sortition, the random selection of representatives, that dates back to classical Athens. In the fifth century B.C.E. the city-state,

The Greek Democracy Lottery

Ancient Athenians used a kleroterion, a stone slab with a grid of slots, to select jurors from among volunteers in such a way that all of the population's 10 tribes were equally represented. A lottery system enabled the jurors to be randomly chosen on the morning of the trial, minimizing chances of bribery.

1 Each volunteer is issued a token with unique identifying markings. The tokens are slotted into the device, with all tokens for a given tribe going into the column for that tribe, in random order. Some tribes may have more volunteers than others.

2 The total number of marbles equals the minimum number of tokens in any column—in this case, seven. Suppose a jury of 30 is desired, with equal representation of 10 tribes. In that case, three of the marbles are gold ($3 \times 10 = 30$). The marbles are mixed and poured into a funnel.

3 Marbles are revealed one by one, determining the fate of each row, starting from the top. When a gold marble appears, the 10 volunteers whose tokens are in the corresponding row are chosen for the jury. When a white marble is drawn, the row is dismissed.

Here, the first marble is gold, so the first row of 10 volunteers is added to the jury.

In this example, having three gold marbles in the mix ensures that 30 jury members are selected, with three from each tribe. The probability of any given volunteer being selected is three divided by the total number of volunteers from that tribe.

Graphics by Jen Christiansen
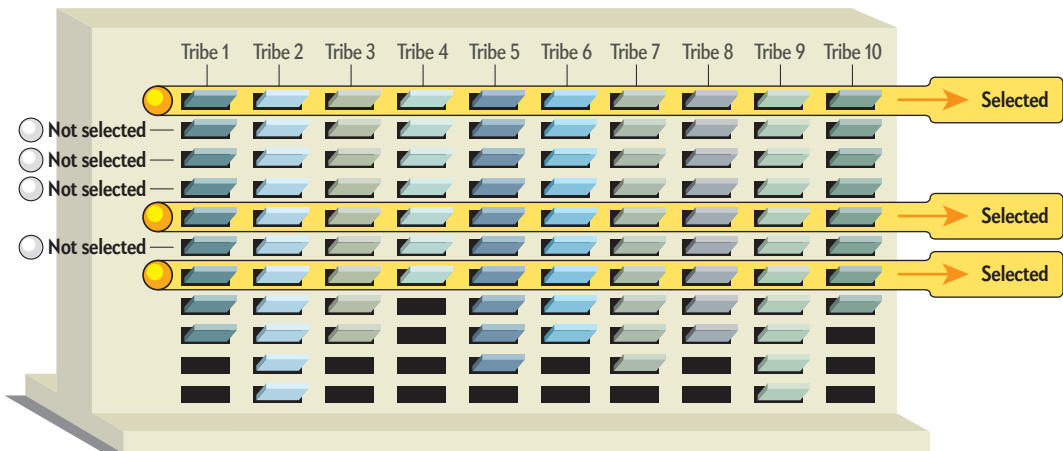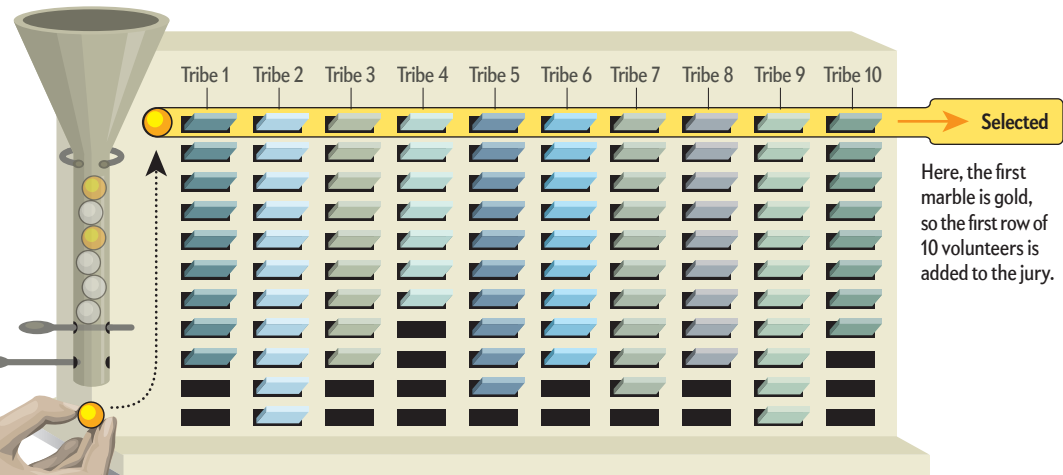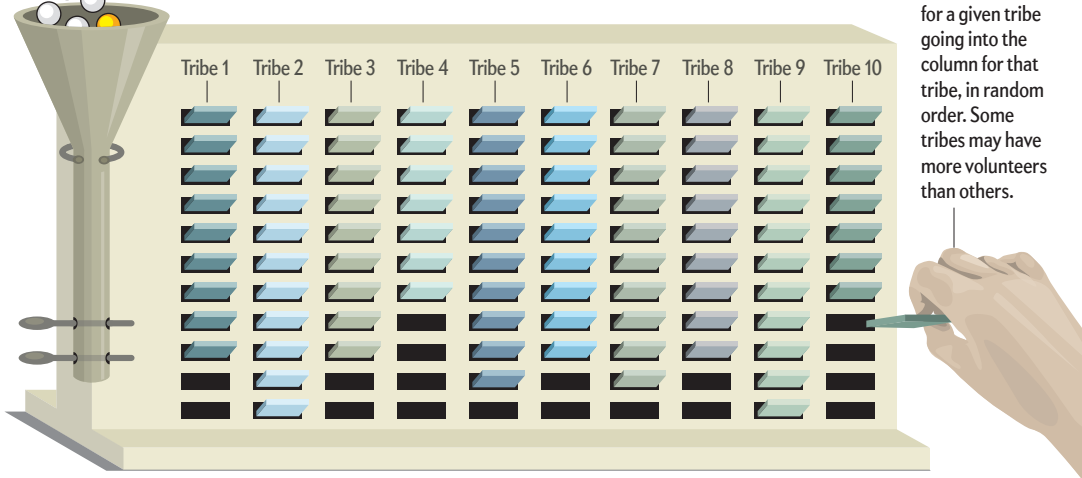
A Skewed Pool of Volunteers

A citizens' assembly should resemble the populace, but its members are typically chosen from a pool of volunteers that does not. The 1,727 volunteers for Climate Assembly U.K. were particularly mismatched with the general population in their education levels and views on climate.

	Population	Pool of Volunteers
Climate Views		
Very concerned	52%	67%
Fairly concerned	33%	28%
Not very concerned		
Not at all concerned		
Age		
60+	29%	31%
45–59	25%	26%
30–44	24%	25%
16–29	22%	18%
Geography 1		
Other	55%	55%
South East		
London		
North West		
Scotland		
Geography 2		
Urban	82%	78%
Rural	18%	22%
Education		
Level 2/3	37%	63%
Level 0/1	36%	
Level 4+	27%	27%
Race/Ethnicity		
White	87%	88%
Racial and/or ethnic minorities		
Gender	*Nonbinary data were not included in the 2018 population estimate source*	
Male		53%
Female		46%
Other		

whose patron deity was Athena, embraced sortition to such a degree that one might say it was de facto governed by Tyche, the goddess of chance. A large majority of its public officials were selected by lot from among citizens who volunteered to serve. These included most of the magistrates, who formed the executive branch, thousands of jurors, and the entire Council of 500, a deliberative body with a wide range of responsibilities.

The Athenians' respect for sortition is apparent in the ingenious design of their lottery machine, the *kleroterion,* which was used to select jurors. It's a stone slab with a grid of slots, arranged in 10 vertical columns, corresponding to the 10 Athenian tribes. Citizens who wished to serve as jurors presented their lottery ticket—bronze tokens with identifying information—to a magistrate, who inserted each tribe's tokens into the slots in the appropriate column. The magistrate also poured marbles of two contrasting colors—say, gold and white—through a funnel into a cylinder, where they lined up in random order.
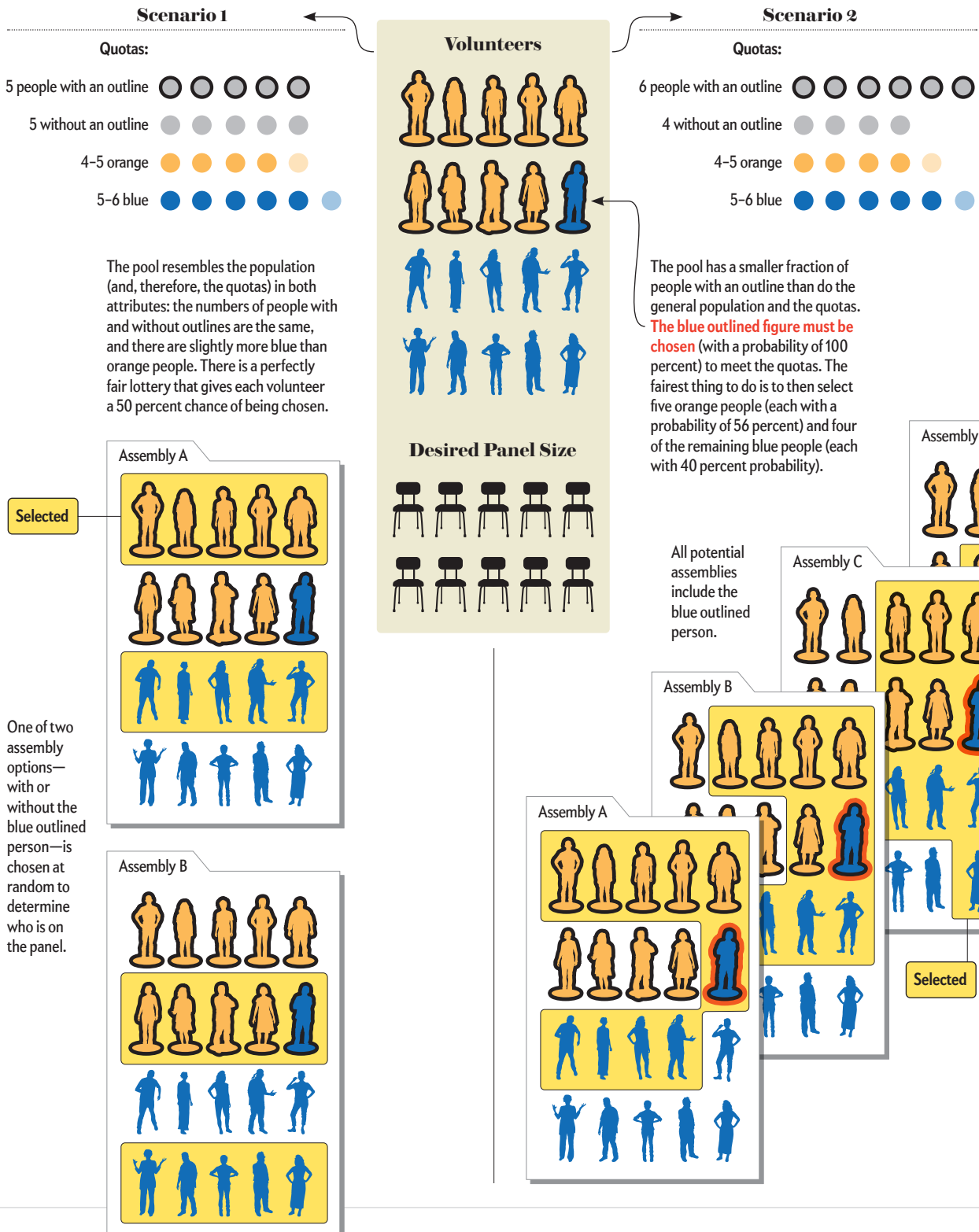
Then, the magistrate used a mechanism to reveal the marbles one by one. If the first marble was gold, the 10 citizens whose tokens appeared in the top row were added to the jury; if it was white, they were all dismissed. And so on, down the column of marbles and the rows of citizens: gold meant in; white meant out. To select a jury of 30 citizens, for example, the magistrate would include three gold marbles in the mix. Because each gold marble picks precisely one citizen from each tribe, any jury selected in this way would necessarily have an equal number of members from each tribe. This passed for descriptive representation in a society that practiced slavery and excluded women from the political process.

As clever as a kleroterion is, the present-day selection process for citizens' assemblies is more complicated because our concept of descriptive representation is much more nuanced. A citizens' assembly is expected to reflect many demographic attributes of the population, not just one. Take Climate Assembly U.K., which the House of Commons commissioned in 2019 to discuss how the nation should reach its target of zero greenhouse gas emissions by the year 2050. Organizers selected the 110 members randomly while seeking to represent the populace according to seven criteria: gender, age, geographic region, education, ethnicity, rural or urban residence, and climate views. Consider the rural-or-urban criterion: in the U.K., about 80 percent of the population lives in urban areas, so out of the 110 seats, 88 seats (or 80 percent) were reserved for urbanites, and 22 seats (or 20 percent) were allocated to country dwellers. Quotas were calculated similarly for each of the other criteria.

As if this isn't complicated enough, organizers of citizens' assemblies often face the challenge that they can select members only from among volunteers, and the pool of willing candidates may look nothing like the population. Typically the organizers issue invitations by mail or phone to a large number of people, but only a fraction of invitees opt in. For example, the organizers of Climate Assembly U.K. sent invitation letters to 30,000 households and mustered 1,727 volunteers. Of the latter, 63 percent had attained the highest level of education (in the British system), whereas a mere 27 percent of Britons fell into that category. It should also come as no surprise that the distribution of climate views among volunteers was skewed, with those concerned about the issue being overrepresented, compared with the general population: it is a rare climate skeptic who relishes

The Challenge of Fair Representation

With a skewed volunteer pool, choosing an assembly that resembles the general population in multiple attributes is usually accomplished by filling quotas. If half the population is rural and half urban, so must the members be. In this example, 10 members reflecting the populace in two characteristics (*color* and *outline*) must be chosen from 20 volunteers.

Scenario 1

Quotas:

5 people with an outline

5 without an outline

4–5 orange

5–6 blue

The pool resembles the population (and, therefore, the quotas) in both attributes: the numbers of people with and without outlines are the same, and there are slightly more blue than orange people. There is a perfectly fair lottery that gives each volunteer a 50 percent chance of being chosen.

Selected

Assembly A

One of two assembly options—with or without the blue outlined person—is chosen at random to determine who is on the panel.

Assembly B

Volunteers

Desired Panel Size

Scenario 2

Quotas:

6 people with an outline

4 without an outline

4–5 orange

5–6 blue

The pool has a smaller fraction of people with an outline than do the general population and the quotas. **The blue outlined figure must be chosen** (with a probability of 100 percent) to meet the quotas. The fairest thing to do is to then select five orange people (each with a probability of 56 percent) and four of the remaining blue people (each with 40 percent probability).

All potential assemblies include the blue outlined person.

Assembly E

Assembly D

Assembly C

Assembly B

Assembly A

Selected

Visualizing Random Selection

As with the kleroterion, visualizing the process of random selection helps to increase trust. In this example, each of the 256 squares is a potential assembly meeting all the necessary criteria. The lottery selects one of these candidate assemblies as the actual one, with perfect randomness (but not necessarily fairness). This visualization takes the view of a volunteer, with the red squares being the potential assemblies that the volunteer is a member of. At each step, we spin the "wheel of fortune," and it tells us which corner of the grid to focus on. For example, the first spin is northeast, so that corner of the grid is selected. Three more spins, and we end up with a single assembly—the selected one.
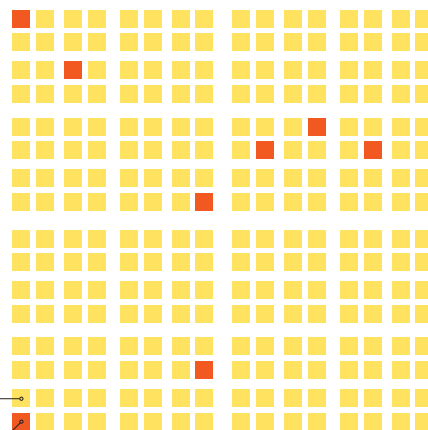
Each square represents an assembly

Red assemblies include volunteer "A"

Spin 1	Spin 2	Spin 3	Spin 4

Randomly selected assembly includes volunteer "A"

---
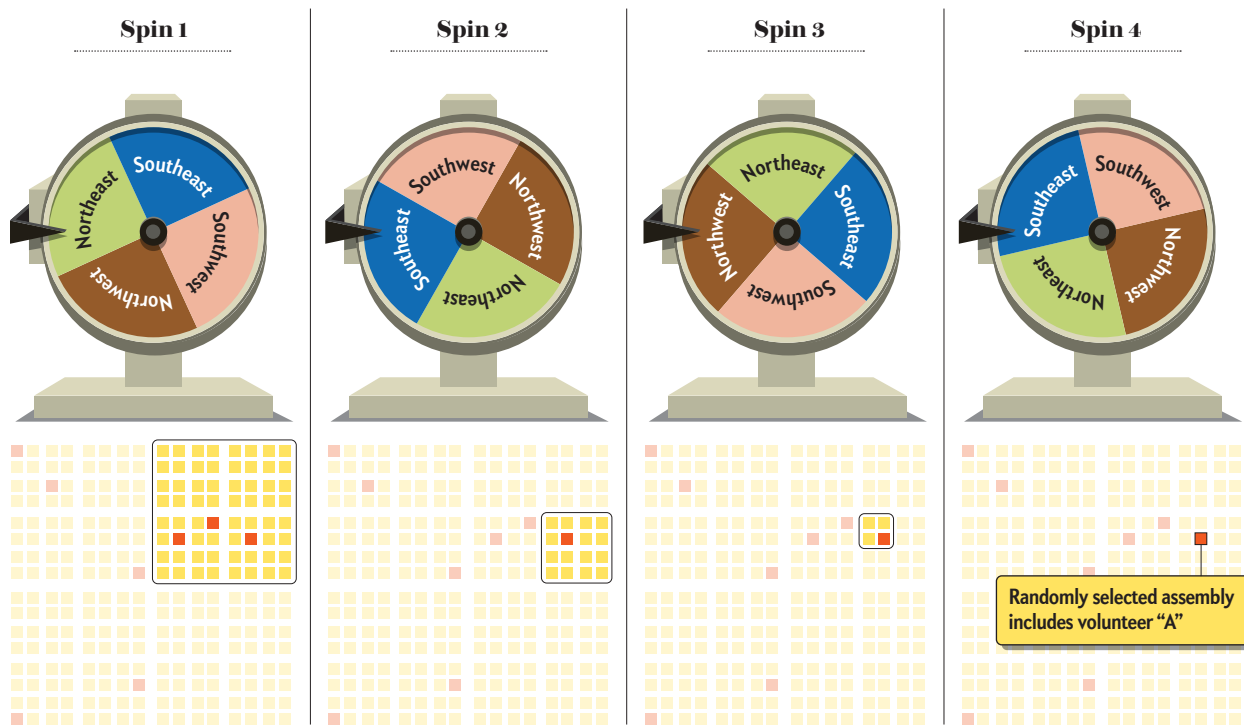
the opportunity to spend long weekends charting a course to zero emissions.

To summarize, we need a modern-day kleroterion that can select a citizens' assembly that is representative in terms of multiple criteria—and can do so starting from an unrepresentative pool of volunteers. Thankfully, we've progressed from stone slabs to computers, so this problem boils down to the design of the right algorithm.

Until recently, the prevalent approach relied on what computer scientists call a "greedy algorithm." This is a bit of misnomer because such an algorithm is really guilty of sloth rather than greed: It takes the action that seems best right now, without making an effort to understand what would work well in the long term. To select an assembly, a greedy algorithm adds volunteers one by one in a way that makes the most immediate progress toward filling the quotas. For example, the algorithm might determine that, right now, the assembly is sorely missing individuals in the 30-to-44 age group, and among all volunteers in this age group, it would choose one at random to join the assembly. Next, it might identify a shortage of Londoners and select someone from that group.

The algorithm may make some bad choices and end up in a situation where it is unable to fill the quotas, but in that case, it can simply restart, and experience shows that it will eventually catch a lucky break. In fact, a particular greedy algorithm developed by a U.K.-based nonprofit, the Sortition Foundation, was used to select that country's climate assembly and many other consequential assemblies.

TO BE FAIR

IT WAS AN EXAMINATION of the greedy algorithm that instigated my own work on the selection of citizens' assemblies, done in collaboration with Bailey Flanigan and Anupam Gupta, both at Carnegie Mellon University, Paul Gölz of Harvard University and Brett Hennig of the Sortition Foundation. We realized that, in the greedy algorithm's short-sighted pursuit of filling quotas, it may sacrifice another important goal: giving all volunteers a fair chance of serving on the assembly. Political theorists view fairness as key to achieving democratic ideals such as equality of opportunity. To be sure, some imbalance is inevitable: Because the objective is descriptive representation of the entire population, volunteers who belong to groups that are underrepresented in the pool are more likely to be selected than those in overrepresented groups. In practice, however, the greedy algorithm excludes some volunteers from the process, even when it is unnecessary.

To see how the greedy algorithm is unfair, we can revisit the selection process of Climate Assembly U.K. by simulating the different assemblies put together by the algorithm, each of which could, in principle, have been the actual one. It turns out that the algorithm selects some of the 1,727 volunteers with a minuscule probability of less than 0.03 percent, whereas it is possible to guarantee that even the least fortunate volunteer is chosen with a probability of at least 2.6 percent—86 times higher—while meeting the same quotas.

To create a fairer algorithm, my collaborators and I adopt a holistic approach. Instead of considering volunteers one at a time, we consider the entire ensemble of potential assemblies, each of which meets all the demographic quotas. Each candidate assembly is given a lottery ticket that specifies its probability of being selected as the actual assembly. The probabilities are determined later, in such a way that they add up to 100 percent, and there's only one winning ticket.

Imagine that each volunteer is given a copy of the lottery ticket of every assembly of which they are a member. The volunteer is selected if any of their lottery tickets wins; in other words, the probability that a volunteer is selected is the sum of probabilities associated with all the potential assemblies that include them. Of all possible lotteries, our algorithm seeks to construct the fairest one, in the sense that the selection probability of the volunteer who is least likely to be chosen is as high as possible.

Now all we need to do is to go over all potential assemblies and … oh wait, the number of potential assemblies is beyond astronomical. A common way to illustrate "astronomical" is to compare the quantity in question with the number of atoms in the observable universe, estimated to be at most 10^{82}. But even that doesn't quite cut it: if you took every atom in the universe and replaced it with an entire universe, each with 10^{82} atoms, the total number of atoms you'd get is still much smaller than the number of ways to select the 110 members of Climate Assembly U.K. from the 1,727 volunteers (without quotas).

Fortunately, computational problems at this mind-boggling scale are routinely solved by machinery from the field of optimization. To apply these techniques, one must construct a mathematical model that includes an objective (in this case, maximizing fairness) and defines a set of possible solutions. The goal is to find the optimal (fairest) solution out of all possible solutions.

In another example, when a navigation app such as Google Maps plans a trip from one location to another, it is solving an optimization problem wherein every feasible route is a possible solution and the objective is to find the shortest possible travel time. In a large city, the number of routes can be enormous, yet we take it for granted that our phones will comb through all these possible trips in seconds. The problem of finding the fairest lottery of the potential assemblies is a much harder problem, but it, too, can be conquered by the right combination of optimization tools.

Our algorithm was released as open source in 2020 and has since become a common method for selecting citizens' assemblies. It was initially adopted by our partners at the Sortition Foundation, who have used it to select, among others, Scotland's climate assembly, convened by the Scottish government; a citizens' jury on assisted dying in Jersey Island, which led to its par-

We love to talk of "democratizing AI" or "democratizing finance," but democracy itself demands our attention.

liament's decision to allow the practice in principle; and a public advisory group created by the U.K.'s National Health Service to discuss how the government should use data in its response to the COVID pandemic. Other organizations have employed our algorithm to select major citizens' assemblies in Germany, France and the U.S., including a panel in Michigan to chart a pathway for pandemic recovery. Last year, thanks to an effort led by Gölz and Gili Rusak, a doctoral student at Harvard, our algorithm became freely accessible through the website Panelot.org (panel selection by lot), making it even easier for practitioners to apply it.

THE DEMOCRACY CODE

AN AMERICAN TIME TRAVELER visiting the present from the late 18th century would find an almost unrecognizable world, but one thing, at least, would look eerily familiar: the way our system of democracy works. Although the endurance of the political system is a tribute to the framers of the constitution, it's abundantly clear that not all is well. In America and in some other democracies around the world, faith in governments has hit rock bottom, and even the most popular legislation often fails to be enacted. There's an urgent need to rethink the practice of democracy using modern tools.

I believe that mathematicians and computer scientists have a significant role to play in this endeavor. We love to talk of "democratizing AI" or "democratizing finance," but democracy itself demands our attention. An algorithmic approach is crucial to the construction of new frameworks to engage citizens and give them a voice. But this apparatus of democracy comes with uniquely challenging instructions: "random assembly required." **SA**