Randomization and Simplification

By

Ehud Kalai[1] and Eilon Solan[2,3]

Abstract

Randomization may add beneficial flexibility to the construction of optimal simple decision rules in dynamic environments. A decision maker, restricted to the use of simple rules, may find a stochastic rule that strictly outperforms all deterministic ones. This is true even in highly separable Markovian environments where the set of feasible choices is stationary and the decision maker's choices have no influence on future payoff functions. In separable environments, however, the period selection of an action can still be deterministic; only the transitions in the evolution of his behavior may require randomization.

---

[1] MEDS Dept., Kellogg Graduate School of Management and Dept. of Mathematics, College of Arts and Sciences, Northwestern University, Evanston IL 60208-2001. e-mail: kalai@nwu.edu. Kalai wishes to acknowledge financial support from the National Science foundation, grant number SES 9904931
[2] MEDS Dept., Kellogg Graduate School of Management, Northwestern University, Evanston IL 60208-2001. e-mail: e-solan@nwu.edu.
[3] The authors wish to thank Adam Kalai and Ehud Lehrer for helpful conversations.

<u>Randomization and Simplification</u>

By Ehud Kalai and Eilon Solan

Randomization serves several useful purposes in multi-person decision making. In a play against an antagonist in a static environment, von Neumann [1928] showed that a player can increase his maxmin payoff by randomizing, making his choice of an action unpredictable. In a static non-antagonistic environment, Aumann [1974] showed that all players may be made better off by the use of a correlation device that allows randomization. In general, such uses of randomization are not needed in single-person decision making. Even in Rabin's [1980] design of a randomizing computational algorithm for a single optimizer, the objective is to maximize expected payoff under a worst case scenario, and thus against an "imaginary antagonist."

This note highlights another role of randomization, useful even in one-person decision problems. It shows by way of an example, that in a dynamically changing environment, a simple decision rule that involves randomization may strictly outperform all simple deterministic rules. Randomization may generate beneficial flexibility that is not possible under rigid deterministic rules.

The environment is a finite state Markov chain, with one characteristic of interest associated to every one of its states. The decision maker selects one of a finite number of actions prior to entering a state, and upon entering the state realizes a payoff that depends on his selected action and the state's characteristic.

In the examples below, two possible characteristics, rain or shine, are associated with every state, and the decision maker has to choose between two possible actions, take an umbrella or not. Payoffs of 1 result from visits to states with appropriately selected actions - visiting a rainy state with an umbrella and visiting a shiny state without one, and payoffs of zero result from visits with wrong selections - a rainy state without an umbrella and a shiny state with one.

The decision maker is limited to the use of simple decision rules. This limitation may be self-imposed, for example as a computational cost-reducing measure, or may be externally imposed, for example by a highly able manager passing down a simple decision rule to a subordinate who may be limited. The main issue in this note, however, is the identification of the best simple rule, and not why and how it is obtained.

The examples are restricted to decision rules that can be described by two state automata. While there is no universal acceptance of automata as the proper tool for measuring simplicity, they do serve the purpose of illustrating that there is a connection between randomization and some version of simplicity. The examples are also restricted to deterministic Markov chains. This way none of the randomized behavior can be attributed to innate uncertainty about the environment, but only to the desire to simplify it.

The first example shows that even in a relatively simple environment a two-state automaton may be an attractive simplification device. In this example, an automaton with a deterministic transition rule turns out to be optimal. In the second example, however, the unique optimal two-state automaton requires random transition rules.

There are interesting connections between simplification devices and bounded recall, as in Piccioni and Rubinstein's [1997] study of an absent-minded driver. Indeed, as shown in their paper, the optimal choices made by the absent-minded driver involve randomization. There are however some important differences. First, the limitations on the forgetful driver are exogenous and not subject to choice, whereas in Example 2 below, the *optimally selected* rule calls for randomization. Second, choices made by the forgetful driver affect the transitions of the underlying Markov process, i.e. his feasible choices and payoffs in future periods. Thus, it is not clear whether his random behavior is targeted to affect his payoff directly, or through a manipulation of his future environment. Example 2 below removes any possible confusion since the feasible choices and payoffs in any period depend exclusively on the state of the environment in that period, and this state is not influenced by the decision maker's earlier choices.

Further elaboration on the above and additional points is postponed till after the presentation of the examples.

Example 1: Long Seasons.

Consider a town where winter lasts for exactly 197 consecutive rainy days followed by a summer that lasts for exactly 168 consecutive sunny days. The mayor has to decide every morning whether or not to take an umbrella. Clearly the mayor can count the days since the summer (or winter) began, and keep track of the weather to attain a perfect average payoff of 1.

But the mayor can do almost as well by the following simple rule: "take an umbrella after rainy days and do not take one after sunny days." This simple method misses only twice a year, the first day of summer and the first day of winter, and yields an average payoff of 363/365. The following two-state automaton may describe the behavior induced by this simple rule:

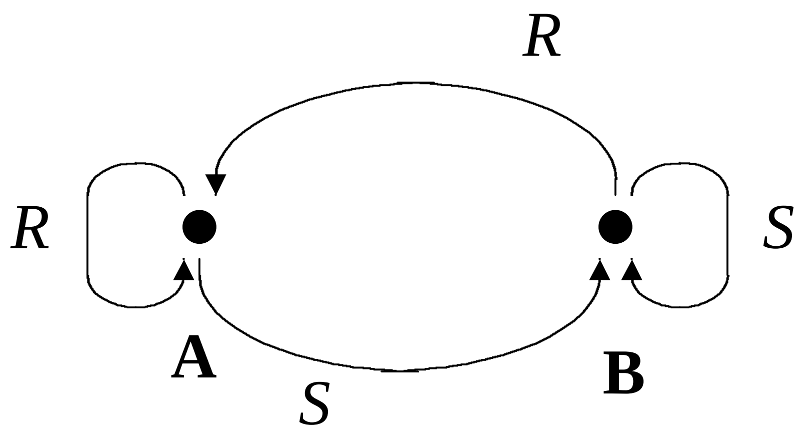

Figure 1: The optimal automaton

In this automaton, in state **A** the mayor takes an umbrella and in state **B** he does not. The daily transitions between states are determined by the last observed weather condition as described by the arrows in Figure 1.

Can the mayor do better with a probabilistic decision rule, or equivalently, a probabilistic automaton; that is, an automaton that allows both probabilistic transitions among states and probabilistic choices of actions in states?

The answer is negative. First we claim that the mayor cannot gain by choosing, in any of the states **A** or **B**, an action in a random manner. Indeed, let an optimal (possibly random) automaton be given. Let p_A be the limit of the fraction of rainy days among all days in which the automaton is in state **A**. Since the weather process is governed by a Markov chain, this limit is well defined. Since the mayor's choices do not affect the future weather conditions, if $p_A<1/2$, in state **A** the optimal action is *not* to take an umbrella, if $p_A>1/2$, in state **A** the optimal action is to *take* an umbrella, while if $p_A=1/2$, any action (deterministic or random) taken in state **A** is optimal. This, of course, is true also for state **B**.

Remark 1: The above argument is general and shows that in any Markovian environment with transitions not affected by the decision maker actions, ***optimal automata*** may be restricted to ***use deterministically chosen actions***. If randomization can improve payoffs, it must be done in the transition rules. Note that this argument did *not* assume that the payoff is independent of time; as long as the payoff is independent of previous decisions, the above argument holds.

Back to the example, we now notice that in an optimal automaton of size 2, in one state (say, state **A**) the action is to take an umbrella, and in the other (state **B**) the action is not to take an umbrella. For otherwise the mayor will either always take an umbrella and get an average payoff of only 197/365, or will never take one and get an average payoff of only 168/365.

Finally, we claim that in this example the mayor cannot gain by using stochastic transitions. Assume that p is the probability of transiting to state **B** after observing rain in state **A**, and q is the probability of transiting to state **A** after observing sun in state **B** (see Figure 2). Note that we consider only two of the four transitions.

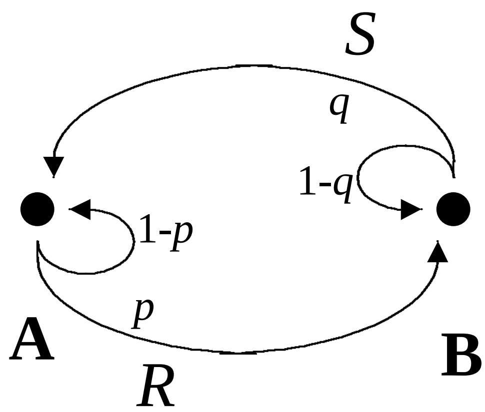

Figure 2

We first prove that since the automaton is optimal, $p = q = 0$. We then show that the other two transitions are also deterministic.

Let {1, 2, …,197,198, …,365} be the days of the year, where {1,2, …,197} correspond to winter, and {198,199, …,365} correspond to summer. Divide the days into pairs as follows: {365,1},{2,3},{4,5}, …,{194,195},{197,198}, …,{363,364} (one day, 196, is not taken into account). We shall count the expected number of misses of the automaton. It is easily verified that in the pair {365,1} it misses at least once with probability at least 1-q: if in day 365 the automaton is in state **A**, it misses in that day (and maybe also in day 1), whereas if it is in state **B**, it misses in day 1 with probability 1-q. Similarly, in each of the pairs {2,3},{4,5}, …,{194,195} it misses at least once with probability at least p, in the pair {197,198} it misses at least once with probability at least 1-p, and in each of the pairs {199,200}, …,{363,364} it misses at least once with probability at least q. Thus, unless $p = q = 0$, the average number of misses is strictly more than 2.

It follows that the other two transitions are also deterministic: if the automaton is in state **A** and it is shiny, then it must be the first day of summer, hence in an optimal automaton the next state is **B**. Similar argument shows that if the automaton is in state **B** and it is rainy, the new state should be **A**.

Example 2: Short Seasons.

We now consider a town where the weather is a cycle of length three, repeating the pattern rainy, rainy, shiny, rainy, rainy, shiny,… . Restricting the mayor to simple decision rules that are representable by automata of size 2, we will see that deterministic automata yield a maximal payoff of 2/3, whereas randomizing automata can yield 3/4.

Continuing with the same notations as in Example 1, it is easy to see that the best the mayor can do now using a deterministic automaton of size 2 is 2/3, which he can by always taking an umbrella. Indeed, if the transition from state **A** (where he takes an umbrella) after a rain is to state **B** (where he does not take an umbrella), he misses at the second rainy day, while if it is to stay at **A**, he misses at the shiny day.

What is the best that the mayor can do using a randomizing automaton? To do better than the deterministic automata, the optimal randomizing automaton must miss on average strictly less than once every cycle. Does this condition restrict some of its transitions?

We first argue that in such an optimal automaton, after a shiny day the automaton moves to state **A**. We then argue that if the automaton is in state **B** and it is a rainy day, then the automaton *remains* at state **B**. We thus reduce the complexity of such an optimal automaton from four unknown transitions to one unknown transition.

Let π(**A**) be the expected average payoff in one cycle (rainy, rainy, shiny) conditioned on the automaton starting the cycle in state **A**, and let π(**B**) be the expected average payoff in one cycle conditioned on the automaton starting the cycle in state **B**. Note that π(**B**) $\leq$ 2/3, since if the state of the automaton at the beginning of the cycle is **B**, it misses at that day. Moreover, the average payoff of the mayor is a weighted average of π(**A**) and π(**B**). It follows that π(**A**) > 2/3 $\geq$ π(**B**).

Note that π(**A**) and π(**B**) are independent of the transitions after a shiny day, but these transitions do influence the probability that the first state in a cycle is **A** (or **B**). Since π(**A**) > π(**B**), in an optimal automaton after a shiny day the automaton moves to state **A**, so that the average payoff is equal to π(**A**).

Next we argue that this implies that if the automaton is in state **B** and it is a rainy day, the automaton remains in state **B**. Indeed, if such a case arises then it must be the second rainy day, hence tomorrow will be a shiny day.

Thus, the optimal randomized automaton has the following transitions:

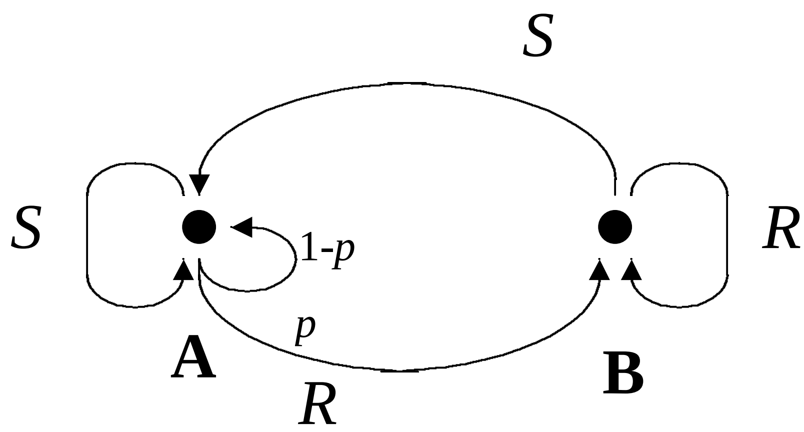

Figure 3: The transitions of an optimal automaton

where p is the probability to move to state **B** if the current state is **A** and it is rainy.

Since after a shiny day the automaton will be in state **A**, the probability of success in the first rainy day is 1, the probability of success in the second rainy day is 1-p, and the probability of success in the sunny day is $1-(1-p)^2 = 2p-p^2$. In particular, the expected average payoff is $(2+p-p^2)/3$, which is maximized at p=1/2, and gives an average payoff 3/4.

Additional Comments:

1. <u>On randomization, flexibility and bounded recall:</u> At first glance, it seems surprising that a decision maker would choose to randomize in a one-person decision problem. Under the conditions of Kuhn's [1953] theorem, any randomizing strategy of an extensive form game can be written as a convex combination of pure strategies, with payoffs being linear in the convex combinations. Thus, no random strategy could do better than all pure strategies. Figure 4 helps clear the situation.

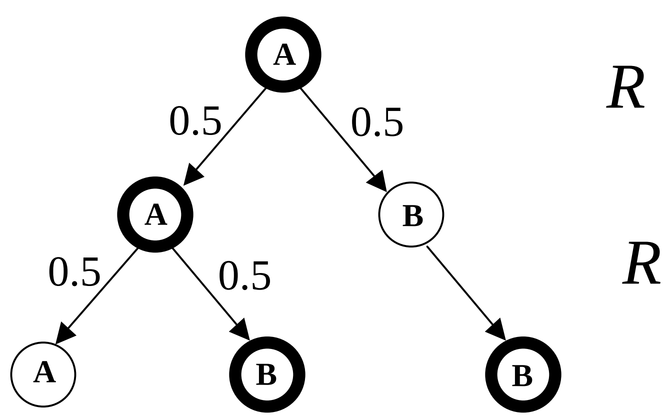

Figure 4

For one complete cycle that starts after a shiny day, the graph describes the probability tree of the optimal strategy in Example 2. It gives the paths that can occur in the cycle.

The state of the automaton at the beginning of the cycle is **A**. Then, a signal *R* is received: with probability 0.5 the new state is **A**, and with probability 0.5 it is **B**. Again a signal *R* is received, and a new state is chosen. A bold circle means that at that stage the action chosen by the automaton is correct, and a thin circle means it is incorrect.

Note that any path yields an average payoff at least 2/3. The path **AAA** can be represented by an automaton that prescribes always taking an umbrella, and the path **ABB** can be represented by an automaton that prescribes taking an umbrella after a shiny day, and not taking an umbrella after a rainy day. These two automata are deterministic, and yield average payoff 2/3. The middle path, **AAB**, yields average payoff 1, but alas, cannot be generated by a deterministic automaton of size 2.

Thus, with probability 0.25, the randomizing automaton adds to the decision maker a behavior pattern not possible with deterministic automata of size 2. Or, in other words, flexibility not possible otherwise. This is exactly were our gain came from.

It is also easy to see why the conclusion of Kuhn's theorem does not hold. Kuhn's decomposition of the optimal strategy in that example involves two pure strategies (automata) of two states, and one pure strategy of three states, which is not permissible.

In terms of Kuhn's assumptions, requiring a decision maker to use strategies describable by automata with bounded number of states, forces him to have imperfect recall - the automaton only "knows" what state it is in but not how it got there.

Imperfect recall is also present in the absent-minded driver example of Piccioni and Rubinstein mentioned earlier. Indeed, there too one obtains an optimal strategy that calls for randomization. It is important to note, however, that the absent-minded driver deals with an environment that is not separable across periods. His chosen action in one period does change the set of possible actions, and payoffs, available to him in the next period. This is not a minor difference. For example, the general observation made in Example 1, that one only needs to randomize on the transitions of the automata and not the selected actions, no longer holds. Indeed, the absent-minded driver does randomize over his selected actions (to exit or not at various decision nodes). Using a completely separable environment, Example 2 shows that simplicity, flexibility and randomization remain tide together even in the *most elementary* environments.

2. <u>On the type of simplification device:</u> There is no universal agreement on the proper way of measuring complexity, or simplicity, of decision rules. In the language of this paper it would be hard to agree on the appropriate notion of a simplification device.

Since the environments of the decision maker above are Markov chains, a seemingly natural formulation of simple decision rules is to describe the decision maker as a low-state Markov chain, rather than automaton. In other words, in the examples above he would be disallowed from using the observed weather condition as input into the transition rules. But this is an artificial restriction, since the decision maker does observe the weather, yet it drastically changes the measure of complexity as can be observed by considering the following cases in Example 1 with the long seasons.

In Example 1, one can show, using the periodic decomposition of Markov chains (see, e.g., Feller [1960], chapter XV.7) and since $365 = 73 \times 5$, which are both primes, that a Markov chain of size strictly less than 365 can hit at most $3/5 \times 365 = 219$ times every year. If, in that example, every fourth year was a leap year, one would need a much more complicated rule (and many more states in the Markov chain) to stay synchronized with the weather. On the other hand, the same automaton of size 2 that was used in Example 1 still misses only twice a year and seems quite satisfactory. This same problem would become hopelessly severe when the underlying Markov process is probabilistic.

But the fact that the observed weather should be an allowed input does not mean that the decision making rule should be described as an automaton. For example, one could argue that it should be described by a Turing machine with a possible limitation on the computation time. These issues are too difficult to be resolved here, but the fact that the examples above deal with automata of only two states should be reassuring.

Automata, in comparison with Turing machines, exaggerate the complexity of decision rules up. For example counting to 365, which requires 365 states with an automaton, is a simple matter with a Turing machine. In other words, decision rules describable by two state automata should be judged simple by most measures. Thus, at a minimum, this note establishes a connection between randomization and strong version of simplicity.

3. <u>What is the optimal automaton of size 2 when the weather has a cycle of length four</u>, repeating the pattern rainy, rainy, rainy, shiny, rainy, rainy, rainy, shiny, … ? Clearly, the deterministic simple rule that always takes an umbrella yields an average payoff 3/4. It turns out, though the calculations are more tedious, that no random automaton of size 2 can outperform this simple rule.

The above examples raise a large number of general open questions. For example,

a) When is optimality obtained by a random (rather than a deterministic) automaton?
b) Can one bound the performance of the optimal automaton (with an exogenous bound on the number of states)?
c) Can one bound the improvement by which the optimal random automaton will outperform the optimal deterministic automaton?

4. <u>The complexity of randomization:</u> The discussion above suggested that the performance of simple decision rules may be improved through randomization. It ignored however, the cost and complexity of the randomization process itself. This may be the case if the randomization is done in one's mind, but not if the randomization is

done by the use of some costly device. Does randomization improve performance even when its cost is taken into consideration seems like an interesting open question.

Bibliography

[1974] Aumann R.J., “Subjectivity and Correlation in Randomized Strategies,” *Journal of Mathematical Economics*, **1**, 67-96.

[1960] Feller W., “An Introduction to Probability Theory and Its Applications”, Volume I, second edition, John Wiley & Sons.

[1953] Kuhn, H.W., “Extensive Games and the Problem of Information,” in H. W. Kuhn and A. W. Tucker, eds., *Contributions to the Theory of Games* I, Princeton University Press, 193-216.

[1928] von Neumann J., “Zur Theorie der Gesellschaftspiele,” *Mathmatische Annalen*, **100**, 295-320.

[1997] Piccioni M. and A. Rubinstein, “On the Interpretation of Decision Problems With Imperfect Recall,” *Games and Economic Behavior*, 20, No. 1, 3-25.

[1980] Rabin, M.O., “Probabilistic Algorithm for Testing Primality,” *Journal of Number Theory*, **2**, 128-138