Serendipity in 421, a stochastic game of life

Pierre Albarède*

November 27, 2007

Abstract

An optimal stochastic control problem is found in a popular dice game, known as 421, and set up somewhere between game theory and statistical mechanics, with emphasis on symmetries. The open loop solution corresponds to a backward induction judging program, `mean-mean`, and is related to dual Kolmogorov and Fokker-Planck equations. The closed loop solution corresponds to a backward induction policy, `mean-max`. A ratchet stratagem of `mean-max` is generalized into "cheaper" goal-driven policies, depending on three parameters: serendipity, horizon and dynamism and yielding some non-Markovian strategies. Almost all goal-driven strategies for a sample of utility functions are exactly judged. From this experiment, laws of goal-driven policy utility are inferred. Principles of meta-policy are presented and inequalities on computing and meta-computing times are proposed. In appendices, relations are established with transport theory (the Galton-Watson problem) and the indifference principle (the Buridan donkey problem).

Key words: utility, strategy, policy, indifference principle, backward induction, goal, ratchet, serendipity, horizon, dynamism, meta-policy.

JEL: C61, C63, C73.
MSC: 68T20, 90B50, 91A15, 93E20.
PACS: 05.10.Gg.

---

*http://pierre.albarede.free.fr

I Cast Dice

Mark Saric

i cast dice
the day my savior arrived
rain before and beyond the horizon
and my hands confused
suspended over the end of human wisdom
branches like the hands of the condemned
reach into the stony darkness of a deaf night sky
the calculus of redemption
in a world of hunger where the table is not yet set.

Contents

1 Introduction

The present study, initiated in [1, 2], aims primarily at giving "casual advice" [3, ch. 2] to players of the 421 game, the rules of which are explained in appendix A. Within the game, only rounds will be considered, except maybe to characterize end-of-round conditions.

A 421 round is not exactly a game (in the sense of game theory [4]) but an *optimal stochastic control* problem: a player has to optimize his present choice, with respect to a long term utility, in spite of future odds. This is the usual condition of any (operations) research. Moreover, the first player in a 421 set, unlike his fellows, has a stopping problem.

Game theory focuses on proving the existence of most useful strategies, but "usable techniques for obtaining practical answers" [5, §1.1], programs indeed, also matter. Programs will be herein loosely determined using language and equations, leaving realization for [6]; one can also call on the Curry-Howard isomorphism between proof and program [7].

Definition. *A* policy *is a program yielding exactly one strategy.*

A program, such as the policy corresponding to Zermelo theorem for the chess game [4, §11.4], may be impracticable, because of computing time and space constraints.

Many 421 round policies will be proposed and judged, to answer (not so) trivial questions, such as "Should I be driven by goals and if so how to choose them? Is it worth thinking deeper or changing my mind? If I do not reach the goal, am I still worth something?" Serendipity (the utility of not reaching a goal) suggested in [3] about the process of invention (and research), will be investigated in particular.

2 Model of 421 round

2.1 Alea

Dice are considered as particles in classical (non-quantum) statistical mechanics [8], identifying face with phase. A combination is a class of arrangements modulo permutations; for example (see appendix A for notation), the face arrangements $122, 212, 221$ make up one face combination of cardinality three, represented by 221.

Let $F \in \mathbb{N}^*$ be the dice face number (the same for all dice). Dice are modeled, in Lagrangian form, as a face combination, or, in Eulerian form, as an occupation number vector,

$$\mathbf{d} = (d_f, f \in \{1 \dots F\}) \in \mathbb{Z}^F,$$

where d_f is the number of f faces in the combination. For example, the Lagrangian combination 421 corresponds to the Eulerian vector $(1, 1, 0, 1, 0, 0)$ $(F = 6)$, $421 \equiv (1, 1, 0, 1, 0, 0)$.

$\mathbb{Z}^F$ is a partially ordered $\mathbb{Z}$-modulus (nearly a vector space). Let $\mathbf{d} \wedge \mathbf{d}'$ be the minimum of $\mathbf{d}, \mathbf{d}' \in \mathbb{Z}^F$. The partial canonic order $\leq$ on $\mathbb{Z}^F$ must be distinguished from the 421 set total order $\preceq$ (65).

The canonic basis of $\mathbb{Z}^F$ is, using Kronecker δ,

$$(\mathbf{e}_f, f \in \{1 \dots F\}),\ \mathbf{e}_f = (\delta(f', f), f' \in \{1 \dots F\}).$$

The f-brelan (see appendix A) is $3\mathbf{e}_f$.

$\mathbb{Z}^F$ is normed by

$$\|\mathbf{d}\| = \sum_{f=1}^{F} |d_f|.$$

The norm of an Eulerian vector is the number of dice it models. For $D \in \mathbb{N}$, the positive[1] ball and sphere of radius D are

$$B^+(D) = \{\mathbf{d} \in \mathbb{Z}^F, \mathbf{d} \geq 0, \|\mathbf{d}\| \leq D\},$$
$$\partial B^+(D) = \{\mathbf{d} \in \mathbb{Z}^F, \mathbf{d} \geq 0, \|\mathbf{d}\| = D\}.$$

Dice are ***assumed*** discernible, independent of each other and unloaded, so that the probability of obtaining the Eulerian vector $\mathbf{d}$ after casting dice once is $p(\mathbf{d})$, given by the multinomial formula: using vector power and factorial forms,

$$\sum_{\mathbf{d} \in \partial B^+(D)} p(\mathbf{d}) = 1,\ \mathbf{p} = \frac{1}{F}(1 \dots 1) \in \mathbb{Q}^F,$$

$$\forall(\mathbf{d} \in \mathbb{Z}^F, \mathbf{d} \geq 0),\ p(\mathbf{d}) = \mathbf{p}^{\mathbf{d}} \frac{\|\mathbf{d}\|!}{\mathbf{d}!} \in \mathbb{Q}^+. \tag{1}$$

For example, when casting two dice, the probability of raising $21 \equiv (1, 1, 0, 0, 0, 0)$ is twice the probability of raising $11 \equiv (2, 0, 0, 0, 0, 0)$.

---

[1] The present convention is that 0 be both positive and negative; accordingly, “as much as” is both more and less than, present is both past and future…

2.2 Fate

Definition. Fate *is an infinite sequence,*

$$(\mathbf{d}_j \in B^+(D), j \in \frac{1}{2}\mathbb{N}),$$

where integer and non-integer dates *index respectively* states *and* events*. A state consists in the dice that have been* pushed away *from the dice board; an event consists in the dice that have just been cast.*

One ***assumes*** that

- all the D_j dice that have not been *pushed away* at date j must be cast at date $j+1/2$ (2),
- any dice that has just been cast can be pushed away (3),
- the initial state is null and fate is only "virtually" infinite (4),

$$\forall j \in \mathbb{N},\ D_j = D - \|\mathbf{d}_j\|,\ \mathbf{d}_{j+1/2} \in \partial B^+(D_j), \tag{2}$$
$$\forall j \in \mathbb{N},\ 0 \leq \mathbf{d}_{j+1} - \mathbf{d}_j \leq \mathbf{d}_{j+1/2}, \tag{3}$$
$$\mathbf{d}_0 = \mathbf{0},\ \exists j \in \mathbb{N},\ D_j = 0. \tag{4}$$

Definition. *For every fate, the* cast number *and* effective fate *are respectively*

$$J_1 = \min(\{j \in \mathbb{N}, D_j = 0\}),\ (\mathbf{d}_j, j \in \frac{1}{2}\{0 \dots 2J_1\}). \tag{5}$$

J_1 exists because of (4); J_1 is ***assumed*** bounded for all fates and let J be its maximum.

$(\mathbf{d}_j, j \in \mathbb{N})$ increases in $B^+(D)$ from the origin to the boundary, while $(D_j, j \in \mathbb{N})$ decreases from D to 0:

$$\begin{aligned} &\mathbf{d}_0 \leq \mathbf{d}_1 \dots \mathbf{d}_{J_1-1} < \mathbf{d}_{J_1} \in \partial B^+(D), \\ &D = D_0 \geq D_1 \dots D_{J_1-1} > D_{J_1} = 0. \end{aligned} \tag{6}$$

Effective events (components of effective fate) are non-null; fate effectively ends at date J_1, where after all events are null and state is constant. Here are many equivalent final conditions:

$$\begin{aligned} &\forall j \in \mathbb{N}^*,\ (j \geq J_1 \ \Leftrightarrow\ D_j = 0 \ \Leftrightarrow\ \mathbf{d}_j - \mathbf{d}_{j-1} = \mathbf{d}_{j-1/2} \\ &\quad \Leftrightarrow\ (\forall (k \in \mathbb{N}, k \geq j),\ D_k = 0,\ \mathbf{d}_{k+1/2} = \mathbf{0},\ \mathbf{d}_k = \mathbf{d}_j \in \partial B^+(D))). \end{aligned} \tag{7}$$

Definition. *A* utility function *is a function which associates to every fate its final utility* $u(\ldots \mathbf{d}_{J_1})$.

Utility function appears in the control problem as a functional parameter, modeling the outer world, as completely as possible.[2]

A rational player (whose computing time and space are bounded) can actually compute only rational or infinite final utilities. Therefore, one ***assumes***

$$u(\ldots \mathbf{d}_{J_1}) \in \mathbb{Q} \cup \{-\infty\}. \tag{8}$$

Moreover, one ***assumes***

$$u(\ldots \mathbf{d}_{J_1} \ldots) = u(\ldots \mathbf{d}_{J_1}). \tag{9}$$

Computing fate final utility is called *judging*.

For all next players in the same 421 set, J must be first player cast number and this is obtained not with special rules but by ***assuming*** that fates ending early are infinitely harmful, so that no rational next player will ever let them happen:

$$\text{next players: } \forall j \in \{1 \ldots J-1\},\ (D_{j-1} > D_j = 0\ \Rightarrow\ u(\ldots \mathbf{d}_j) = -\infty). \tag{10}$$

Thus, all players do fit in the same model, except maybe for the values of J and the utility function.

Fate is ***assumed*** to be a causal stochastic chain. *Causal* means that the historic sequence (not fate itself)

$$((\mathbf{d}_0), (\mathbf{d}_0, \mathbf{d}_{1/2}), (\mathbf{d}_0, \mathbf{d}_{1/2}, \mathbf{d}_1) \ldots) \tag{11}$$

is a *Markovian* stochastic chain; Markovian stochastic chain means that each chain component is a random variable, the probability of which only depends on the previous chain component.

For $(j, \mathbf{d}) \in \mathbb{N}/2 \times B^+(D)$, let $P(\ldots \mathbf{d}_j, \mathbf{d})$ be the probability, knowing history until date j $(\ldots \mathbf{d}_j)$, that $\mathbf{d}_{j+1/2} = \mathbf{d}$. P appears as a (mixed) strategy, also bearing providential probabilities. From (1, 2),

$$\forall (j, \mathbf{d}) \in \mathbb{N} \times B^+(D),\ P(\ldots \mathbf{d}_j, \mathbf{d}) = \delta(D_j, \|\mathbf{d}\|)\, p(\mathbf{d}) \in \mathbb{Q}^+. \tag{12}$$

---

[2] One does not always know in practice when and where a game exactly ends, as it may be a sub-game of a larger game or a player can play many games simultaneously. For example, a tennis game is embedded in a set, embedded in a match, embedded in a tournament, embedded in a ranking system, and a tennis player may also play, say, the stock market.

For some utility function u_e, a policy A yields the strategy $P = A(u_e)$.

The set of all effective fates appears as a tree, alternatively branched by providence and player and carrying final utilities (as fruits). In [6], computing space constraints lead to a Markovian fate tree format (illustrated by fig. 1)), with only one node per dated state, which prevents remembering history; also, instead of infinite utility (10), distinct fate trees are used for first and next players.

2.3 Symmetries

2.3.1 Face permutations

Definition. $\mathbf{d}, \mathbf{d}' \in \partial B^+(D)$ *are* equivalent modulo face permutations*,* $\mathbf{d} \sim \mathbf{d}'$*, if they have the same* combination *of occupation numbers.*

Canonic representatives are chosen so as to minimize face sums in Lagrangian form, for example, $442 \sim 211$, in Eulerian form, $(0, 1, 0, 2, 0, 0) \sim (2, 1, 0, 0, 0, 0)$, for the combination of (non-null) occupation numbers $\{1, 2\}$. There are three equivalence classes in $\partial B^+(3)$:

- the class of brelans (such as 111),
- the class of sequences (to which belong 321 and 421, although 421 is not a sequence, see appendix A),
- the class of pairs (to which belong 211 and 221, although 221 is not a pair, see appendix A).

As events are not equiprobable (12) , fate trees in [6] bear event probabilities (see also fig. 1). Nevertheless, the providential probability law is *invariant modulo face permutations*, (essentially: face labels are indifferent)

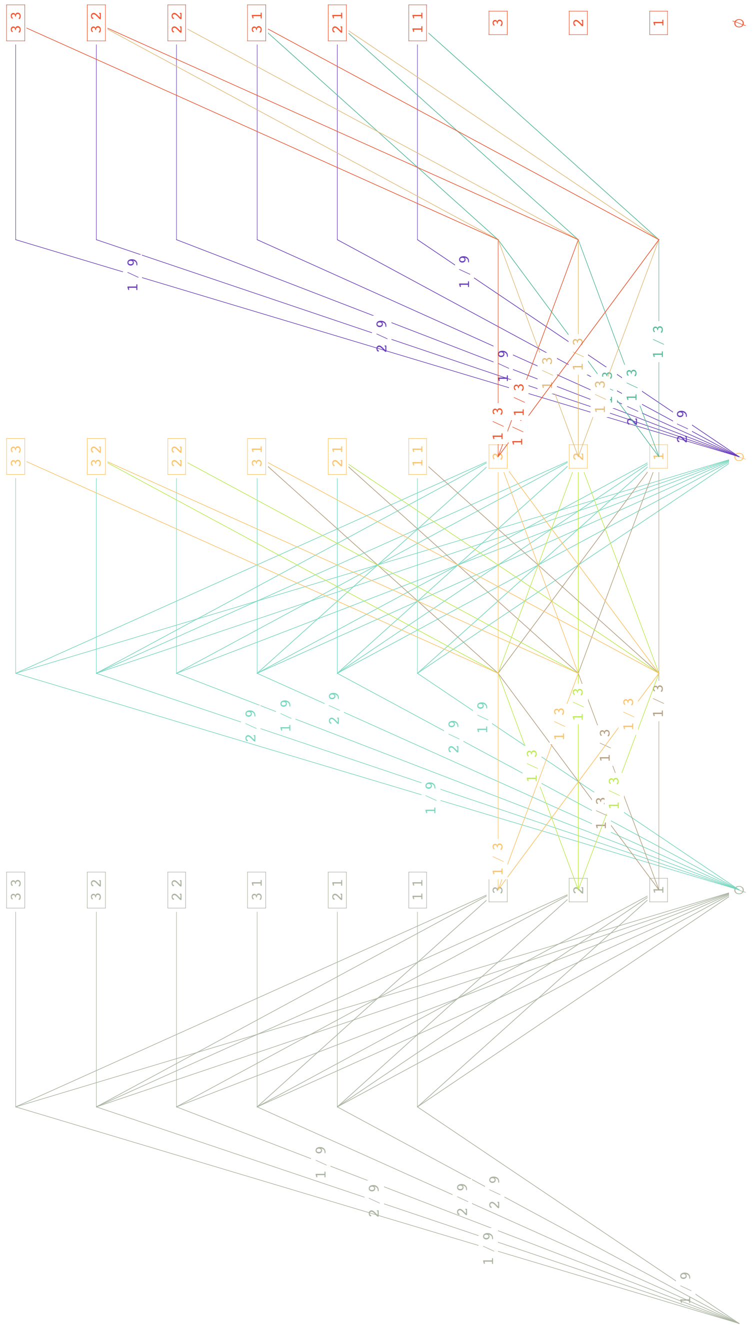
$$\mathbf{d} \sim \mathbf{d}' \;\Rightarrow\; p(\mathbf{d}) = p(\mathbf{d}').$$

Definition. *Eulerian vector couples are* equivalent modulo face permutations *if and only if they have the same combination of component couples.*

The representative of a state couple $(\mathbf{d}^*, \mathbf{d})$ is not the couple of its component state representatives, except in the diagonal case ($\mathbf{d} = \mathbf{d}^*$), as

$$\mathbf{d} \sim \mathbf{d}' \;\Leftrightarrow\; (\mathbf{d}, \mathbf{d}) \sim (\mathbf{d}', \mathbf{d}').$$

Figure 1: first player fate tree $(D, F, J) = (2, 3, 3)$ [6]



The first component of a couple representative is chosen as the representative of the first couple component; the next component is chosen so as to minimize its face sum in Lagrangian form, for example, $(421, 442) \sim (321, 211)$, in Eulerian form,

$$((1, 1, 0, 1, 0, 0), (0, 1, 0, 2, 0, 0)) \sim ((1, 1, 1, 0, 0, 0), (2, 1, 0, 0, 0, 0)),$$

for the combination of (non-null) component couples $\{(1, 0), (1, 1), (1, 2)\}$. There are 31 equivalence classes in $\partial B^+(3) \times \partial B^+(3)$, including three diagonal ones.

A policy is *covariant modulo face permutations* if its output strategy varies like its input utility function, submitted to any face permutation. In any strategy P, the providential dependence is invariant, while the player dependence need not be covariant, because he may prefer some numbers; this would be a case of "symmetry breaking" (see appendix C).

2.3.2 Self-similarity

The 421 round control problem is similar to itself, reconsidered dynamically at dated state $(j, \mathbf{d}_j) \in \{0 \dots J\} \times B^+(D)$, modulo the parameter reduction,

$$(D, J) \ \rightarrow \ (D_j, J - j) \tag{13}$$

and a corresponding utility function transformation. The parameters D, J (but not F) are thus called *dynamic*; self-similarity is much used in [6].

3 Fate as stochastic chain

This is the "open-loop" part of the 421 round control problem. Given strategy P, fate appears as a stochastic chain and Markovian methods [9, ch. 6], [10, ch. 15] and [8, ch. 15] do apply, not to fate itself, but to historic sequence (11). Dice driven by utility in $\mathbb{Z}^F$ are like particles driven by some force in Newtonian space, and "history" sounds faithfully like "hysteresis".

3.1 Kolmogorov equation on utility

From the von Neumann-Morgenstern theorem [4, ch. 27], using virtual fates,

$$\forall j \in \frac{1}{2}\mathbb{N}, \ u(\dots \mathbf{d}_j) = \sum_{\mathbf{d}_{j+1/2}} P.u(\dots \mathbf{d}_j, \mathbf{d}_{j+1/2}), \tag{14}$$

where $0 \times (-\infty)$ that may occur for next players and non-integer j (10) must be replaced by zero. Applying (14) to itself,

$$\forall j \in \mathbb{N},\ u(\ldots \mathbf{d}_j) = \sum_{\mathbf{d}_{j+1/2}} P(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2}) \sum_{\mathbf{d}_{j+1}} P.u(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2}, \mathbf{d}_{j+1}), \tag{15}$$

which corresponds to a strategy judging program (not a policy), called `mean-mean`, taking for input a utility function u_e and P and yielding all utilities, in particular, the initial utility $u(\mathbf{0})(u_e, P)$, that is also the *strategy utility*. `mean-mean` is invariant modulo (global) face permutations.

For example, (15) determines, from the final utilities

- $u_s(\mathbf{d}_j, j \in \{0 \ldots J\}) = \chi(\mathbf{d}_J \in V)$, χ denoting characteristic function: the probability of reaching $V \subset B^+(D)$;
- $u_s(\mathbf{d}_j, j \in \{0 \ldots J\}) = J_1$ (5): the average cast number;
- $u_s(\mathbf{d}_j, j \in \{0 \ldots J\}) = \delta(D_k, d)\, \chi(k \leq J_1),\ d \in \{0 \ldots D\},\ k \in \{0 \ldots J\}$: the probability that $D_k = d$ *effectively* (further determined in appendix B).

If state probabilities are rational, then all utilities also are, as shown with backward induction on (8, 12, 14). Likewise, if final utilities are binary, then all utilities are conditional probabilities of fate being useful.

For $(j, \mathbf{d}) \in \mathbb{N} \times B^+(D)$, let $\sigma(\ldots \mathbf{d}_j, \mathbf{d})$ be the probability, knowing history until date j, that $\mathbf{d}_{j+1} = \mathbf{d}$. It is decomposed over all mutually exclusive intermediary events, hence Chapman-Kolmogorov equation,

$$\sigma(\ldots \mathbf{d}_j, \mathbf{d}) = \sum_{\mathbf{d}_{j+1/2}} P(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2})\, P(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2}, \mathbf{d}). \tag{16}$$

Final utilities are ***assumed*** independent of events (17). If moreover choices are independent of events except the last one (18), then all utilities are independent of events, as shown with backward induction on (12, 14):

$$u(\mathbf{d}_j, j \in \frac{1}{2}\{0 \ldots 2k\}) = u_s(\mathbf{d}_j, j \in \{0 \ldots k\}), \tag{17}$$

$$P(\mathbf{d}_j, j \in \frac{1}{2}\{0 \ldots 2k\}) = P_s(\mathbf{d}_j, j \in \{0 \ldots k-1, k-\frac{1}{2}, k\}), \tag{18}$$

$$\sigma(\mathbf{d}_j, j \in \frac{1}{2}\{0 \ldots 2k-2, 2k\}) = \sigma_s(\mathbf{d}_j, j \in \{0 \ldots k\}). \tag{19}$$

(17) allows to factor out $\sigma(\ldots)$ in (15), hence Kolmogorov equation on utility,

$$\forall j \in \mathbb{N},\ u_s(\ldots \mathbf{d}_j) = \sum_{\mathbf{d}_{j+1}} \sigma_s(\ldots \mathbf{d}_j, \mathbf{d}_{j+1})\, u_s(\ldots \mathbf{d}_j, \mathbf{d}_{j+1}), \tag{20}$$

where, from (16, 12, 18),

$$\sigma_s(\ldots \mathbf{d}_j, \mathbf{d}_{j+1}) = \sum_{\mathbf{d}_{j+1/2}} p(\mathbf{d}_{j+1/2})\, P_s(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2}, \mathbf{d}_{j+1}). \tag{21}$$

3.2 Fokker-Plank equation on probability

For $(j, \mathbf{d}) \in \mathbb{N} \times B^+(D)$, let $\rho(j, \mathbf{d})$ be the probability that $\mathbf{d}_j = \mathbf{d}$. From (4),

$$\rho(0, \mathbf{d}) = \delta(\mathbf{d}, \mathbf{0}). \tag{22}$$

Final utilities are ***assumed*** Markovian (independent of strictly past fate). If moreover strategy is Markovian (that is, each choice only depends on last event and last state), then all utilities are Markovian, as shown with backward induction on (12, 14):

$$u_s(\ldots \mathbf{d}_j) = u_{t,s}(j, \mathbf{d}_j), \tag{23}$$

$$P(\ldots \mathbf{d}_{j-1}, \mathbf{d}_{j-1/2}, \mathbf{d}_j) = P_{t,s}(j, \mathbf{d}_j, \mathbf{d}_{j-1/2}, \mathbf{d}_j), \tag{24}$$

$$\sigma(\ldots \mathbf{d}_j, \mathbf{d}_{j+1}) = \sigma_{t,s}(j, \mathbf{d}_j, \mathbf{d}_{j+1}). \tag{25}$$

Decomposing over all mutually exclusive past states with (25) and recalling that $\sigma(\ldots)$ is a conditional probability yields Fokker-Planck equation,

$$\rho(j+1, \mathbf{d}) = \sum_{\mathbf{d}_j \in B^+(D)} \rho(j, \mathbf{d}_j)\, \sigma_{t,s}(j, \mathbf{d}_j, \mathbf{d}), \tag{26}$$

where, from (21, 24),

$$\sigma_{t,s}(j, \mathbf{d}_j, \mathbf{d}) = \sum_{\mathbf{d}_{j+1/2} \in B^+(D)} p(\mathbf{d}_{j+1/2})\, P_{t,s}(j, \mathbf{d}_j, \mathbf{d}_{j+1/2}, \mathbf{d}). \tag{27}$$

With (23), (9) becomes

$$\forall (j, \mathbf{d}) \in \{J_1 \ldots J\} \times \partial B^+(D),\ u_{t,s}(j, \mathbf{d}) = u_{t,s}(J_1, \mathbf{d}). \tag{28}$$

3.3 Duality and utility conservation

Using (23) in Kolmogorov equation (20),

$$u_{t,s}(j, \mathbf{d}) = \sum_{\mathbf{d}_{j+1} \in B^+(D)} \sigma_{t,s}(j, \mathbf{d}, \mathbf{d}_{j+1})\, u_{t,s}(j+1, \mathbf{d}_{j+1}). \tag{29}$$

(26, 29) are linear equations, adjoint to each other, respectively on probability and utility. This "duality" has useful consequences, well known for example in linear transport theory (see appendix B).

Let $\mathcal{F}$ be the $\mathbb{Q}$-vector space of numeric applications on $B^+(D)$, Euclidean for the scalar product

$$\forall f, g \in \mathcal{F},\ \langle f, g \rangle = \sum_{\mathbf{d} \in B^+(D)} f(\mathbf{d})\, g(\mathbf{d}).$$

The matrix of the backward operator $\sigma_{t,s}(j) : u_{t,s}(j+1, .) \mapsto u_{t,s}(j, .)$ is

$$(\sigma_{t,s}(j, \mathbf{d}, \mathbf{d}'), (\mathbf{d}, \mathbf{d}') \in B^+(D) \times B^+(D)).$$

The matrix of the forward operator $\sigma_{t,s}(j, .)^\dagger : u_{t,s}(j, .) \mapsto u_{t,s}(j+1, .)$ is the transpose of the latter,

$$(\sigma_{t,s}(j, \mathbf{d}', \mathbf{d}), (\mathbf{d}, \mathbf{d}') \in B^+(D) \times B^+(D)).$$

In functional form, (26, 29) become

$$\begin{aligned} u_{t,s}(j, .) &= \sigma_{t,s}(j)(u_{t,s}(j+1, .)), \\ \sigma_{t,s}(j)^\dagger(\rho(j, .)) &= \rho(j+1, .). \end{aligned}$$

As $\sigma_{t,s}(j), \sigma_{t,s}(j)^\dagger$ are adjoint to each other, utility is spread but conserved over all possible fates:

$$\begin{aligned} \forall j \in \{0 \ldots J\},\ \langle u_{t,s}(j, .), \rho(j, .) \rangle &= \langle \sigma_{t,s}(j)(u_{t,s}(j+1, .)), \rho(j, .) \rangle \\ &= \langle u_{t,s}(j+1, .), \sigma_{t,s}(j)^\dagger(\rho(j, .)) \rangle \\ &= \langle u_{t,s}(j+1, .), \rho(j+1, .) \rangle \\ &= u_{t,s}(0, \mathbf{0}). \end{aligned} \tag{30}$$

The initial utility $u_{t,s}(0, \mathbf{0}) = u_s(\mathbf{0}) = u(\mathbf{0})$ can be computed by choosing $j \in \{0 \ldots J\}$, computing $u_{t,s}(j, .)$ backward from final utilities with (29), $\rho(j, .)$ forward from initial probabilities with (26) and at last the scalar product in the l. h. s. of (30). Although the value does not depend on j, computing time or space does and is minimum for some j.

3.4 What is really done; non-Markovian strategies

Kolmogorov and Fokker-Planck equations were presented to relate the open loop 421 round control problem with stochastic theory. In [6], utilities and probabilities are computed with a Markovian version of (15), not using virtual fate. Indeed, virtual fate and (9) are only used theoretically to avoid boundary problems and to establish Fokker-Planck equation, that cannot accommmodate first player stopped chains; in turn, first player Fokker-Planck probabilities are not effective.

The Markovian fate tree format in [6] is consistent with the Markovian condition (24), that was not formally assumed, because it excludes many strategies, possibly useful and cheap (not using much computing time or space). For example, a strategy consisting in choosing a goal and sticking to it is not Markovian and cannot be judged *directly* with `mean-mean`.

4 Playing against providence

This is the “closed-loop” part of the 421 round control problem. The optimal control condition is: for some utility function u_e, maximize strategy utility $u(\mathbf{0})(u_e, P)$ with respect to strategy P.

4.1 Backward induction policy

From the von Neumann-Morgenstern theorem [4, ch. 27],

$$\forall j \in \mathbb{N},\ u(\ldots \mathbf{d}_{j+1/2}) = \max_{\mathbf{d}_{j+1}} u(\ldots \mathbf{d}_{j+1/2}, \mathbf{d}_{j+1}). \tag{31}$$

Inserting (31) into (14),

$$u(\ldots \mathbf{d}_j) = \sum_{\mathbf{d}_{j+1/2}} P(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2}) \max_{\mathbf{d}_{j+1}} u(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2}, \mathbf{d}_{j+1}). \tag{32}$$

From (31), the set of most useful states at integer date j, knowing history until date $j + 1/2$, is

$$\mathcal{B}(\ldots \mathbf{d}_{j+1/2}) = \operatorname*{argmax}_{\mathbf{d}_{j+1}} u(\ldots \mathbf{d}_{j+1/2}, \mathbf{d}_{j+1}). \tag{33}$$

Choice is based, firstly, from the von Neumann-Morgenstern theorem, on greatest utility,

$$\forall \mathbf{d} \in B^+(D) \setminus \mathcal{B}(\ldots \mathbf{d}_{j+1/2}),\ P(\ldots \mathbf{d}_{j+1/2}, \mathbf{d}) = 0, \tag{34}$$

secondly, on ***equiprobable*** tie-breaking (see appendix C), that is completely random choice among all most useful states: card denoting cardinality,

$$\forall (j, \mathbf{d}) \in \mathbb{N} \times B^+(D),\ P(\ldots \mathbf{d}_{j+1/2}, \mathbf{d}) = \frac{\chi(\mathbf{d} \in \mathcal{B}(\ldots \mathbf{d}_{j+1/2}))}{\mathrm{card}(\mathcal{B}(\ldots \mathbf{d}_{j+1/2}))} \in \mathbb{Q}. \tag{35}$$

(34, 35) imply

$$\sum_{\mathbf{d} \in B^+(D)} P(\ldots \mathbf{d}_{j+1/2}, \mathbf{d}) = 1.$$

(32, 33, 35) correspond to a backward induction policy, called `mean-max`, after von Neumann `min-max`, taking for input a utility function u_e and yielding the complete most useful strategy `mean-max`(u_e). `mean-max` is covariant modulo face permutations.

For example, applying (32) thrice to itself, using (2, 3, 12), yields 'last judgment" `mean-max` equations, which essentially solve the 421 round control problem for all players and $J \leq 3$:

$$\begin{aligned} u(\mathbf{d}_0) &= \sum_{\mathbf{d}_{1/2}} p(\mathbf{d}_{1/2}) \ \max_{\mathbf{d}_1} u(\mathbf{d}_0, \mathbf{d}_{1/2}, \mathbf{d}_1), \\ u(\mathbf{d}_0, \mathbf{d}_{1/2}, \mathbf{d}_1) &= \sum_{\mathbf{d}_{3/2}} p(\mathbf{d}_{3/2}) \ \max_{\mathbf{d}_2} u(\mathbf{d}_0, \mathbf{d}_{1/2}, \mathbf{d}_1, \mathbf{d}_{3/2}, \mathbf{d}_2), \\ u(\mathbf{d}_0, \mathbf{d}_{1/2}, \mathbf{d}_1, \mathbf{d}_{3/2}, \mathbf{d}_2) &= \sum_{\mathbf{d}_{5/2}} p(\mathbf{d}_{5/2})\, u(\mathbf{d}_0, \mathbf{d}_{1/2}, \mathbf{d}_1, \mathbf{d}_{3/2}, \mathbf{d}_2, \mathbf{d}_{5/2}, \mathbf{d}_2 + \mathbf{d}_{5/2}). \end{aligned} \tag{36}$$

Last judgment means that, to accommodate first player, for each fate, judgement at time J_1 is virtually postponed to last time J (the same for all fates) according to (9). (10) is also needed to accomodate next players. For players avoiding infinite harm, `mean-max` utilities are rational, as shown with backward induction on (8, 12, 31, 14).

For $i \in \{1, 2\}$ and $\mathbf{d}^* \in \partial B^+(D)$, let $\hat{P}(i, J, \mathbf{d}^*)$ be the i-th player complete most useful strategy (as computed with `mean-max`), for the stationary pointwise Markovian utility function

$$(j, \mathbf{d}) \in \{1 \ldots J\} \times \partial B^+(D) \ \mapsto \ \delta(\mathbf{d}, \mathbf{d}^*). \tag{37}$$

$\hat{P}(i, J, \mathbf{d}^*)$ maximizes the probability of reaching $\mathbf{d}^*$; it is the most probably successful.

4.2 Constant-goal policies

The present section deals with a player who always tries to reach most probably exactly one *goal* $\mathbf{d}^* \in \partial B^+(D)$.

4.2.1 Bernoulli and ratchet policies

Two first player constant-goal policies are

the Bernoulli policy: to push away either no die or at once all dice; then, fate (before success) is (part of) a Bernoulli chain (a sequence of independent trials, immediately failing or succeeding).

the $\mathbf{d}^*$-ratchet policy: to push away at once all dice contributing to the goal $\mathbf{d}^*$, hence a pure strategy,

$$\forall (j, \mathbf{d}) \in \{1 \ldots J-1\} \times B^+(D),$$
$$P(\ldots \mathbf{d}_j, \mathbf{d}_{j+1/2}, \mathbf{d}) = \delta(\mathbf{d}, \mathbf{d}^* \wedge (\mathbf{d}_j + \mathbf{d}_{j+1/2})). \quad (38)$$

p decreases on $B^+(D)$ for the partial order on $\mathbb{Z}^F$ if and only if the ratchet policy (38) is most probably successful.

Actually, from (1),

$$\forall (\mathbf{d} \in B^+(F), \mathbf{d} + \mathbf{e}_1 \in B^+(F)),\ \frac{p(\mathbf{d} + \mathbf{e}_1)}{p(\mathbf{d})} = \frac{1}{F} \frac{\|\mathbf{d}\| + 1}{d_1 + 1} \leq 1,$$

p decreases on $B^+(F)$. Moreover, for $D \geq 2$,

$$\frac{1}{F} > \frac{2!}{F^2} > \frac{3!}{F^3} > \ldots \frac{D!}{F^D} \ \Leftrightarrow\ D < F.$$

It follows that the ratchet strategy is most probably successful if and only if $D \leq F$, and is the only most probably successful strategy, $\hat{P}(i, J, \mathbf{d}^*)$ indeed, if and only if $D < F$.

For example, with

$$D = 3 < F = 6,\ J > 1,\ \mathbf{d}^* \equiv 421,\ \mathbf{d}_{1/2} \equiv 651,$$
$$p(421) < p(42),\ p(41) < p(4),\ p(21) < p(2),$$

the maximum probability of raising 42 is greater than the maximum probability of raising 421; the ratchet choice (to push away 1) is the only most probably successful. On the contrary, with

$$D = 3 > F = 2,\ \mathbf{d}^* \equiv 211,\ \mathbf{d}_{1/2} \equiv 222,$$
$$p(211) = \frac{3}{F^3} = \frac{3}{8} > p(11) = \frac{1}{F^2} = \frac{1}{4},$$

the Bernoulli choice, to replay all dice, is the only most probably successful. With

$$D = F = 2,\ \mathbf{d}^* \equiv 21,\ \mathbf{d}_{1/2} \equiv 11,\ p(21) = p(2),$$

both the Bernoulli and ratchet choices are most probably successful.

Hypothesis:

$$D < F. \tag{39}$$

This hypothesis is verified in the normal case, (according to the rules of the game)

$$(D, F, J) = (3, 6, 3).$$

Moreover, dynamically, it is eventually verified, because of (13, 6).

When some next player could reach his goal early by pushing away all dice, his most probably successful choice is to cast again exactly any one die. Unless his goal is a brelan, he has a *dilemma*, that is a choice between equally harmful states. The number of his most probably successful pure strategies is the number of distinct goal faces, at the power $J - 1$. Apart from dilemmas, next players should try to reach most probably any state preceding (in Eulerian form, for the partial order on $\mathbb{Z}^F$) the goal, which they can do by ratcheting, almost like first player.

4.2.2 Final state probabilities

A rational 421 player will be interested in the probability of player i, having at most k casts left and some *relative* goal $\mathbf{d}^*$, to *effectively* reach after exactly j casts a relative final state $\mathbf{d}$,

$$q(i, k, \mathbf{d}^*, j, \mathbf{d}),$$
$$i \in \{1, 2\},\ 0 \leq j \leq k \leq J,\ \mathbf{d}^* \in B^+(D),\ \mathbf{d} \in \partial B^+(\|\mathbf{d}^*\|),$$
$$\|\mathbf{d}^*\| < D \ \Rightarrow\ k < J, \tag{40}$$

independent of the current dated state by self-similarity (see section 2.3.2).

In [6], success probabilities are computed as `mean-max` utilities and failure probabilities are computed as `mean-mean` utilities, according to

$$q(i, J, \mathbf{d}^*, j, \mathbf{d}) = u_{t,s}(0, \mathbf{0})(\delta_{(j,\mathbf{d})}, \hat{P}(i, J, \mathbf{d}^*)) \in \mathbb{Q},$$

where $\delta_x(y) = \delta(y, x)$, or, more generally, by self-similarity,

$$\begin{aligned} & j' \leq j,\ \mathbf{d}' \leq \mathbf{d}^* \wedge \mathbf{d}, \\ & q(i, J - j', \mathbf{d}^* - \mathbf{d}', j - j', \mathbf{d} - \mathbf{d}') = u_{t,s}(j', \mathbf{d}')(\delta_{(j,\mathbf{d})}, \hat{P}(i, J, \mathbf{d}^*)) \in \mathbb{Q}. \end{aligned} \tag{41}$$

For (39), the strategy $\hat{P}(1, J, \mathbf{d}^*)$ is pure, so that $q(1, k, \mathbf{d}^*, j, \mathbf{d})$ does not depend on tie-breaking, as opposed to $q(2, k, \mathbf{d}^*, j, \mathbf{d})$ except in the diagonal case or when $\mathbf{d}^*$ is a brelan. Properties of q, s follow.

- As `mean-max` is covariant and `mean-mean` is invariant, q, s are invariant modulo (global) face permutations (42).
- An initial condition: the unique relative goal that can be reached instantly is zero and it is reached certainly (43).
- A boundary condition: zero can be a relative goal only for the present (44), as casting no die does not increment effective time.
- The sum of probabilities of all mutually exclusive final dated states is one (45).
- All dice must be pushed away at last cast, so that a 421 round ends up as a lottery (46, 1).
- First player will never stop early unless he succeeds (47); next players must not stop early (48).
- The maximum cast number does not actually affect first player success probability ($\mathbf{d} = \mathbf{d}^*$) (49).
- First and next player failure probabilities ($\mathbf{d} \neq \mathbf{d}^*$) are identical, unless the goal and the final state have exactly one face in common (because only dilemmas make a difference).

$$\begin{aligned}(\mathbf{d}^*, \mathbf{d}) \sim (\mathbf{d}^{*\prime}, \mathbf{d}') &\Rightarrow q(i, k, \mathbf{d}^*, j, \mathbf{d}) = q(i, k, \mathbf{d}^{*\prime}, j, \mathbf{d}'), \\ \mathbf{d} \sim \mathbf{d}' &\Rightarrow \quad s(i, k, \mathbf{d}) = s(i, k, \mathbf{d}').\end{aligned} \tag{42}$$

$$q(i, k, \mathbf{d}^*, 0, \mathbf{d}) = \delta(\mathbf{d}, \mathbf{0}), \tag{43}$$

$$q(i, k, \mathbf{0}, j, \mathbf{0}) = \delta(k, 0), \tag{44}$$

$$\sum_{\mathbf{d} \in \partial B^+(\|\mathbf{d}^*\|)} \sum_{j=0}^{k} q(i, k, \mathbf{d}^*, j, \mathbf{d}) = 1, \tag{45}$$

$$q(i, 1, \mathbf{d}^*, 1, \mathbf{d}) = p(\mathbf{d}), \tag{46}$$

$$q(1, k, \mathbf{d}^*, j, \mathbf{d}) = 0, \; j < k, \; \mathbf{d}^* \neq \mathbf{d}, \tag{47}$$

$$q(2, k, \mathbf{d}^*, j, \mathbf{d}) = 0, \; j < k, \tag{48}$$

$$q(1, k, \mathbf{d}, j, \mathbf{d}) = q(1, j, \mathbf{d}, j, \mathbf{d}). \tag{49}$$

As first player has more freedom than his fellows, the success probability after at most k casts,

$$s(i, k, \mathbf{d}) = \sum_{j=0}^{k} q(i, k, \mathbf{d}, j, \mathbf{d}), \tag{50}$$

obeys

$$s(1, k, \mathbf{d}) \geq s(2, k, \mathbf{d}) = q(2, k, \mathbf{d}, k, \mathbf{d}) \geq q(1, k, \mathbf{d}, k, \mathbf{d}),$$

where the first and second inequalities are strict for $k > 1$ (and $\mathbf{d} \neq 0$).

Probabilities are tabulated on domains restricted in time by (49), in space by invariance modulo face permutations (42). Further restriction is possible using property 4.2.2.

Success probabilities: tab. 1. In every cell stands a column of $q(i, j, \mathbf{d}, j, \mathbf{d})$, j increasing from top to bottom, and, right to it for first player only, a column of $s(1, j, \mathbf{d})$, partial sums of the latter (50, 49). The line header is the canonic representative of $\mathbf{d}$.

Failure probabilities: tab. 2, 3, 4, 5, 7, 8, 9. Every cell shows, at left, a relative goal $\mathbf{d}^*$, pointing downward to a failing relative final state $\mathbf{d}$; at right, a column of $q(i, j, \mathbf{d}^*, j, \mathbf{d})$, j increasing from top to bottom. The line header is the goal canonic representative and the column header is the final state canonic representative.

For example, what are the probabilities, for the goal $\mathbf{d}^* \equiv 641$, of reaching $\mathbf{d} \equiv 655$, after exactly one, two or three casts? The canonic representatives of the goal $\mathbf{d}^*$ and the failing final state $\mathbf{d}$ are, respectively, 321 and 211. The former points, for first player, to tab. 5, whence the latter points to the second column. The canonic representative of $(\mathbf{d}^*, \mathbf{d})$, $(321, 441)$, points further to the third row, where the three requested probabilities stand. For exercise, what is the probability of first player reaching 221 (nénette) at last cast while aiming at 421? (Answer: approximately 0.040.)

Table 1: success probabilities [6]

	First player		Next players
$\boxed{1}$	$\frac{1}{6} \simeq 0.167$		$\frac{1}{6} \simeq 0.167$
	$\frac{5}{36} \simeq 0.139$	$\frac{11}{36} \simeq 0.306$	$\frac{1}{6} \simeq 0.167$
$\boxed{1\ 1}$	$\frac{1}{36} \simeq 0.028$		$\frac{1}{36} \simeq 0.028$
	$\frac{85}{1296} \simeq 0.066$	$\frac{121}{1296} \simeq 0.093$	$\frac{91}{1296} \simeq 0.070$
$\boxed{2\ 1}$	$\frac{1}{18} \simeq 0.056$		$\frac{1}{18} \simeq 0.056$
	$\frac{35}{324} \simeq 0.108$	$\frac{53}{324} \simeq 0.164$	$\frac{19}{162} \simeq 0.117$
$\boxed{1\ 1\ 1}$	$\frac{1}{216} \simeq 0.005$		$\frac{1}{216} \simeq 0.005$
	$\frac{1115}{46656} \simeq 0.024$	$\frac{1331}{46656} \simeq 0.029$	$\frac{1151}{46656} \simeq 0.025$
	$\frac{466075}{10077696} \simeq 0.046$	$\frac{753571}{10077696} \simeq 0.075$	$\frac{513991}{10077696} \simeq 0.051$
$\boxed{2\ 1\ 1}$	$\frac{1}{72} \simeq 0.014$		$\frac{1}{72} \simeq 0.014$
	$\frac{143}{2592} \simeq 0.055$	$\frac{179}{2592} \simeq 0.069$	$\frac{149}{2592} \simeq 0.057$
	$\frac{23681}{279936} \simeq 0.085$	$\frac{43013}{279936} \simeq 0.154$	$\frac{26903}{279936} \simeq 0.096$
$\boxed{3\ 2\ 1}$	$\frac{1}{36} \simeq 0.028$		$\frac{1}{36} \simeq 0.028$
	$\frac{227}{2592} \simeq 0.088$	$\frac{299}{2592} \simeq 0.115$	$\frac{239}{2592} \simeq 0.092$
	$\frac{21043}{186624} \simeq 0.113$	$\frac{42571}{186624} \simeq 0.228$	$\frac{24631}{186624} \simeq 0.132$

Table 2: first player failure probabilities ($\|\mathbf{d}^*\| < 3$) [6]

	$\boxed{1}$
$\boxed{1}$	$\boxed{1} \downarrow \boxed{2}$: $\frac{1}{6} \simeq 0.167$ $\frac{5}{36} \simeq 0.139$

	$\boxed{1\ 1}$	$\boxed{2\ 1}$
$\boxed{1\ 1}$	$\boxed{1\ 1} \downarrow \boxed{2\ 2}$: $\frac{1}{36} \simeq 0.028$ $\frac{25}{1296} \simeq 0.019$	$\boxed{1\ 1} \downarrow \boxed{2\ 1}$: $\frac{1}{18} \simeq 0.056$ $\frac{55}{648} \simeq 0.085$ $\boxed{1\ 1} \downarrow \boxed{3\ 2}$: $\frac{1}{18} \simeq 0.056$ $\frac{25}{648} \simeq 0.039$

	$\boxed{1\ 1}$	$\boxed{2\ 1}$
$\boxed{2\ 1}$	$\boxed{2\ 1} \downarrow \boxed{1\ 1}$: $\frac{1}{36} \simeq 0.028$ $\frac{35}{648} \simeq 0.054$ $\boxed{2\ 1} \downarrow \boxed{3\ 3}$: $\frac{1}{36} \simeq 0.028$ $\frac{1}{81} \simeq 0.012$	$\boxed{2\ 1} \downarrow \boxed{3\ 1}$: $\frac{1}{18} \simeq 0.056$ $\frac{43}{648} \simeq 0.066$ $\boxed{2\ 1} \downarrow \boxed{4\ 3}$: $\frac{1}{18} \simeq 0.056$ $\frac{2}{81} \simeq 0.025$

Table 3: brelan goal first player failure probabilities [6]

	1 1 1	2 1 1	3 2 1
1 1 1	1 1 1 $\downarrow$ 2 2 2 $\frac{1}{216} \simeq 0.005$ $\frac{125}{46656} \simeq 0.003$ $\frac{15625}{10077696} \simeq 0.002$	1 1 1 $\downarrow$ 2 1 1 $\frac{1}{72} \simeq 0.014$ $\frac{605}{15552} \simeq 0.039$ $\frac{207025}{3359232} \simeq 0.062$ 1 1 1 $\downarrow$ 2 2 1 $\frac{1}{72} \simeq 0.014$ $\frac{275}{15552} \simeq 0.018$ $\frac{56875}{3359232} \simeq 0.017$ 1 1 1 $\downarrow$ 3 2 2 $\frac{1}{72} \simeq 0.014$ $\frac{125}{15552} \simeq 0.008$ $\frac{15625}{3359232} \simeq 0.005$	1 1 1 $\downarrow$ 3 2 1 $\frac{1}{36} \simeq 0.028$ $\frac{275}{7776} \simeq 0.035$ $\frac{56875}{1679616} \simeq 0.034$ 1 1 1 $\downarrow$ 4 3 2 $\frac{1}{36} \simeq 0.028$ $\frac{125}{7776} \simeq 0.016$ $\frac{15625}{1679616} \simeq 0.009$

Table 4: pair goal first player failure probabilities [6]

2 1 1	1 1 1	2 1 1 → 1 1 1	$\frac{1}{216} \simeq 0.005$	$\frac{205}{11664} \simeq 0.018$	$\frac{16105}{629856} \simeq 0.026$
		2 1 1 → 2 2 2	$\frac{1}{216} \simeq 0.005$	$\frac{215}{23328} \simeq 0.009$	$\frac{20605}{2519424} \simeq 0.008$
		2 1 1 → 3 3 3	$\frac{1}{216} \simeq 0.005$	$\frac{1}{729} \simeq 0.001$	$\frac{8}{19683} \simeq 4.\times 10^{-4}$
	2 1 1	2 1 1 → 3 1 1	$\frac{1}{72} \simeq 0.014$	$\frac{103}{3888} \simeq 0.026$	$\frac{6499}{209952} \simeq 0.031$
		2 1 1 → 2 2 1	$\frac{1}{72} \simeq 0.014$	$\frac{91}{1944} \simeq 0.047$	$\frac{28219}{419904} \simeq 0.067$
		2 1 1 → 3 3 1	$\frac{1}{72} \simeq 0.014$	$\frac{5}{486} \simeq 0.010$	$\frac{38}{6561} \simeq 0.006$
		2 1 1 → 3 2 2	$\frac{1}{72} \simeq 0.014$	$\frac{77}{3888} \simeq 0.020$	$\frac{7039}{419904} \simeq 0.017$
		2 1 1 → 3 3 2	$\frac{1}{72} \simeq 0.014$	$\frac{31}{2592} \simeq 0.012$	$\frac{839}{93312} \simeq 0.009$
		2 1 1 → 4 3 3	$\frac{1}{72} \simeq 0.014$	$\frac{1}{243} \simeq 0.004$	$\frac{8}{6561} \simeq 0.001$
	3 2 1	2 1 1 → 3 2 1	$\frac{1}{36} \simeq 0.028$	$\frac{37}{648} \simeq 0.057$	$\frac{10217}{139968} \simeq 0.073$
		2 1 1 → 4 3 1	$\frac{1}{36} \simeq 0.028$	$\frac{5}{243} \simeq 0.021$	$\frac{76}{6561} \simeq 0.012$
		2 1 1 → 4 3 2	$\frac{1}{36} \simeq 0.028$	$\frac{31}{1296} \simeq 0.024$	$\frac{839}{46656} \simeq 0.018$
		2 1 1 → 5 4 3	$\frac{1}{36} \simeq 0.028$	$\frac{2}{243} \simeq 0.008$	$\frac{16}{6561} \simeq 0.002$

Table 5: sequence goal first player failure probabilities [6]

3 2 1	1 1 1	3 2 1 → 1 1 1	$\frac{1}{216} \simeq 0.005$ $\frac{83}{15552} \simeq 0.005$ $\frac{3115}{1119744} \simeq 0.003$
		3 2 1 → 4 4 4	$\frac{1}{216} \simeq 0.005$ $\frac{1}{1728} \simeq 6. \times 10^{-4}$ $\frac{1}{13824} \simeq 7. \times 10^{-5}$
	2 1 1	3 2 1 → 2 1 1	$\frac{1}{72} \simeq 0.014$ $\frac{179}{5184} \simeq 0.035$ $\frac{15067}{373248} \simeq 0.040$
		3 2 1 → 4 1 1	$\frac{1}{72} \simeq 0.014$ $\frac{175}{15552} \simeq 0.011$ $\frac{6311}{1119744} \simeq 0.006$
		3 2 1 → 4 4 1	$\frac{1}{72} \simeq 0.014$ $\frac{101}{15552} \simeq 0.006$ $\frac{3277}{1119744} \simeq 0.003$
		3 2 1 → 5 4 4	$\frac{1}{72} \simeq 0.014$ $\frac{1}{576} \simeq 0.002$ $\frac{1}{4608} \simeq 2. \times 10^{-4}$
	3 2 1	3 2 1 → 4 2 1	$\frac{1}{36} \simeq 0.028$ $\frac{319}{7776} \simeq 0.041$ $\frac{24239}{559872} \simeq 0.043$
		3 2 1 → 5 4 1	$\frac{1}{36} \simeq 0.028$ $\frac{101}{7776} \simeq 0.013$ $\frac{3277}{559872} \simeq 0.006$
		3 2 1 → 6 5 4	$\frac{1}{36} \simeq 0.028$ $\frac{1}{288} \simeq 0.003$ $\frac{1}{2304} \simeq 4. \times 10^{-4}$

Table 6: next player failure probabilities ($\|\mathbf{d}^*\| < 3$) [6]

	$\boxed{1}$
$\boxed{1}$	$\boxed{1} \downarrow \boxed{2}$ $\frac{1}{6} \simeq 0.167$ $\frac{1}{6} \simeq 0.167$

	$\boxed{1\ 1}$	$\boxed{2\ 1}$
$\boxed{1\ 1}$	$\boxed{1\ 1} \downarrow \boxed{2\ 2}$ $\frac{1}{36} \simeq 0.028$ $\frac{25}{1296} \simeq 0.019$	$\boxed{1\ 1} \downarrow \boxed{2\ 1}$ $\frac{1}{18} \simeq 0.056$ $\frac{29}{324} \simeq 0.090$ $\boxed{1\ 1} \downarrow \boxed{3\ 2}$ $\frac{1}{18} \simeq 0.056$ $\frac{25}{648} \simeq 0.039$

	$\boxed{1\ 1}$	$\boxed{2\ 1}$
$\boxed{2\ 1}$	$\boxed{2\ 1} \downarrow \boxed{1\ 1}$ $\frac{1}{36} \simeq 0.028$ $\frac{19}{324} \simeq 0.059$ $\boxed{2\ 1} \downarrow \boxed{3\ 3}$ $\frac{1}{36} \simeq 0.028$ $\frac{1}{81} \simeq 0.012$	$\boxed{2\ 1} \downarrow \boxed{3\ 1}$ $\frac{1}{18} \simeq 0.056$ $\frac{23}{324} \simeq 0.071$ $\boxed{2\ 1} \downarrow \boxed{4\ 3}$ $\frac{1}{18} \simeq 0.056$ $\frac{2}{81} \simeq 0.025$

Table 7: brelan goal next player failure probabilities [6]

	1 1 1	2 1 1	3 2 1
1 1 1	1 1 1 → 2 2 2 $\frac{1}{216} \simeq 0.005$ $\frac{125}{46656} \simeq 0.003$ $\frac{15625}{10077696} \simeq 0.002$	1 1 1 → 2 1 1 $\frac{1}{72} \simeq 0.014$ $\frac{617}{15552} \simeq 0.040$ $\frac{222997}{3359232} \simeq 0.066$	1 1 1 → 3 2 1 $\frac{1}{36} \simeq 0.028$ $\frac{275}{7776} \simeq 0.035$ $\frac{56875}{1679616} \simeq 0.034$
		1 1 1 → 2 2 1 $\frac{1}{72} \simeq 0.014$ $\frac{275}{15552} \simeq 0.018$ $\frac{56875}{3359232} \simeq 0.017$	1 1 1 → 4 3 2 $\frac{1}{36} \simeq 0.028$ $\frac{125}{7776} \simeq 0.016$ $\frac{15625}{1679616} \simeq 0.009$
		1 1 1 → 3 2 2 $\frac{1}{72} \simeq 0.014$ $\frac{125}{15552} \simeq 0.008$ $\frac{15625}{3359232} \simeq 0.005$	

Table 8: pair goal next player failure probabilities [6]

2 1 1				
1 1 1	2 1 1 → 1 1 1	$\frac{1}{216} \simeq 0.005$	$\frac{437}{23328} \simeq 0.019$	$\frac{18751}{629856} \simeq 0.030$
	2 1 1 → 2 2 2	$\frac{1}{216} \simeq 0.005$	$\frac{215}{23328} \simeq 0.009$	$\frac{20605}{2519424} \simeq 0.008$
	2 1 1 → 3 3 3	$\frac{1}{216} \simeq 0.005$	$\frac{1}{729} \simeq 0.001$	$\frac{8}{19683} \simeq 4. \times 10^{-4}$
2 1 1	2 1 1 → 3 1 1	$\frac{1}{72} \simeq 0.014$	$\frac{215}{7776} \simeq 0.028$	$\frac{7381}{209952} \simeq 0.035$
	2 1 1 → 2 2 1	$\frac{1}{72} \simeq 0.014$	$\frac{373}{7776} \simeq 0.048$	$\frac{3911}{52488} \simeq 0.075$
	2 1 1 → 3 3 1	$\frac{1}{72} \simeq 0.014$	$\frac{5}{486} \simeq 0.010$	$\frac{38}{6561} \simeq 0.006$
	2 1 1 → 3 2 2	$\frac{1}{72} \simeq 0.014$	$\frac{77}{3888} \simeq 0.020$	$\frac{7039}{419904} \simeq 0.017$
	2 1 1 → 3 3 2	$\frac{1}{72} \simeq 0.014$	$\frac{31}{2592} \simeq 0.012$	$\frac{839}{93312} \simeq 0.009$
	2 1 1 → 4 3 3	$\frac{1}{72} \simeq 0.014$	$\frac{1}{243} \simeq 0.004$	$\frac{8}{6561} \simeq 0.001$
3 2 1	2 1 1 → 3 2 1	$\frac{1}{36} \simeq 0.028$	$\frac{151}{2592} \simeq 0.058$	$\frac{1405}{17496} \simeq 0.080$
	2 1 1 → 4 3 1	$\frac{1}{36} \simeq 0.028$	$\frac{5}{243} \simeq 0.021$	$\frac{76}{6561} \simeq 0.012$
	2 1 1 → 4 3 2	$\frac{1}{36} \simeq 0.028$	$\frac{31}{1296} \simeq 0.024$	$\frac{839}{46656} \simeq 0.018$
	2 1 1 → 5 4 3	$\frac{1}{36} \simeq 0.028$	$\frac{2}{243} \simeq 0.008$	$\frac{16}{6561} \simeq 0.002$

Table 9: sequence goal next player failure probabilities [6]

3 2 1

1 1 1	3 2 1 ↓ 1 1 1	$\frac{1}{216} \simeq 0.005$ $\frac{83}{15552} \simeq 0.005$ $\frac{3115}{1119744} \simeq 0.003$
	3 2 1 ↓ 4 4 4	$\frac{1}{216} \simeq 0.005$ $\frac{1}{1728} \simeq 6. \times 10^{-4}$ $\frac{1}{13824} \simeq 7. \times 10^{-5}$
2 1 1	3 2 1 ↓ 2 1 1	$\frac{1}{72} \simeq 0.014$ $\frac{187}{5184} \simeq 0.036$ $\frac{17459}{373248} \simeq 0.047$
	3 2 1 ↓ 4 1 1	$\frac{1}{72} \simeq 0.014$ $\frac{175}{15552} \simeq 0.011$ $\frac{6311}{1119744} \simeq 0.006$
	3 2 1 ↓ 4 4 1	$\frac{1}{72} \simeq 0.014$ $\frac{101}{15552} \simeq 0.006$ $\frac{3277}{1119744} \simeq 0.003$
	3 2 1 ↓ 5 4 4	$\frac{1}{72} \simeq 0.014$ $\frac{1}{576} \simeq 0.002$ $\frac{1}{4608} \simeq 2. \times 10^{-4}$
3 2 1	3 2 1 ↓ 4 2 1	$\frac{1}{36} \simeq 0.028$ $\frac{331}{7776} \simeq 0.043$ $\frac{27827}{559872} \simeq 0.050$
	3 2 1 ↓ 5 4 1	$\frac{1}{36} \simeq 0.028$ $\frac{101}{7776} \simeq 0.013$ $\frac{3277}{559872} \simeq 0.006$
	3 2 1 ↓ 6 5 4	$\frac{1}{36} \simeq 0.028$ $\frac{1}{288} \simeq 0.003$ $\frac{1}{2304} \simeq 4. \times 10^{-4}$

4.3 Goal-driven policies

A 421 round is considered for some Markovian utility function: abbreviating $u_{t,s} \to u$,

$$(j, \mathbf{d}) \in \{1 \dots J\} \times \partial B^+(D) \ \mapsto\ u(j, \mathbf{d}).$$

One may attempt to find the most useful goals by looking at the utility function. For the utility function (37), $\mathbf{d}^*$ is the only most useful goal; for a flat utility function, all goals are most useful; the problem is when the utility function is somehow “between peaked and flat”.

Any Markovian utility function is a linear combination of pointwise Markovian utility functions, each leading (through `mean-max`) to one most probably successful strategy. Unfortunately, in general, no “mixture” or “superposition” whatsoever of the latter strategies makes up a most useful strategy.

The criterion for goal choice should not be final utility, but initial utility, taking into account final state probabilities (40), intervening as a “logic of science” [11].

4.3.1 Serendipity, horizon and dynamism

Definition. *For* $(j, \mathbf{d}) \in \{0 \dots J\} \times B^+(D)$, *and*

$$V = \partial B^+(D - \|\mathbf{d}\|),$$

the non-serendipitous and serendipitous *goal-driven Markovian utilities of dated state* $(j, \mathbf{d})$ *are, respectively,*

$$u^*_N(j, \mathbf{d}) = \max_{\mathbf{d}^* \in V} \sum_{k=0}^{J-j} \qquad q(i, J-j, \mathbf{d}^*, k, \mathbf{d}^*)\, u(j+k, \mathbf{d}+\mathbf{d}^*), \tag{51}$$

$$u^*_Y(j, \mathbf{d}) = \max_{\mathbf{d}^* \in V} \sum_{k=0}^{J-j} \sum_{\mathbf{d}' \in V} q(i, J-j, \mathbf{d}^*, k, \mathbf{d}')\, u(j+k, \mathbf{d}+\mathbf{d}'). \tag{52}$$

Serendipity in (52) means that even the utility of failure ($\mathbf{d}' \neq \mathbf{d}^*$) is taken into account, as opposed to (51).

Goal-driven utility obeys

- $u^*_N \leq u^*_Y$;
- from (44),

$$\forall (j, \mathbf{d}) \in \{1 \dots J\} \times \partial B^+(D),\ u^*_N(j, \mathbf{d}) = u^*_Y(j, \mathbf{d}) = u(j, \mathbf{d}); \tag{53}$$

- from (32, 46),

$$\forall \mathbf{d} \in B^+(D),\ u^*_Y(J-1, \mathbf{d}) = u(J-1, \mathbf{d}); \tag{54}$$

- more generally, if the relative strategy after dated state $(j, \mathbf{d})$ is constant-goal, then its relative goal maximizes (52) and

$$u^*_Y(j, \mathbf{d}) = u(j, \mathbf{d}); \tag{55}$$

- for stationary utility function, averaging is eliminated:

$$u^*_N(j, \mathbf{d}) = \max_{\mathbf{d}^* \in \partial B^+(D - \|\mathbf{d}\|)} s(i, J-j, \mathbf{d}^*)\, u(\_, \mathbf{d} + \mathbf{d}^*). \tag{56}$$

Goal-driven utility is used for policy design, as follows. For $(j, \mathbf{d}) \in \{0 \dots J\} \times B^+(D),\ h \in \mathbb{N}, j + h \leq J$, the relative fate tree after dated state $(j, \mathbf{d})$ is cut off at depth h, where utility is replaced by goal-driven utility. Hence a *h-horizon* relative control problem, solved by `mean-max`. The partial strategy computed in this way is most useful if goal-driven utility equals horizon utility, for example, if $j + h = J$ (53), or, with serendipity, if $j + h \geq J - 1$ (54) or if the complete most useful strategy after every horizon state happens to be constant-goal (55).

Null horizon implies equiprobably choosing exactly one most useful goal, before any event, and trying to reach it most probably.

Unit horizon implies equiprobably choosing the first state, remarkably without averaging, according to

$$\mathbf{d}_1 \in \operatorname*{argmax}_{\mathbf{d} \in B^+(D),\ \mathbf{d} \leq \mathbf{d}_{1/2}} u^*_s(1, \mathbf{d}),\ s \in \{N, Y\}. \tag{57}$$

h-horizon can be used once initially, while the last $J - h$ strategy levels are computed from goals chosen at the horizon, or repeatedly, according to dynamic programming [12], belief revision [13] or cybernetic feedback (of fate on strategy). Such a goal-driven policy is called *dynamic*.

As a non-full combination state can be consistent with many goals, goal choice, as opposed to state choice, can be delayed, possibly in a useful way.

Definition. *When two relative goals maximize goal-driven utility in (51) or (52), a player behaves with* duplicity *if he chooses his goals only after, and depending on, the next event.*

For example, in the chess game, double attacks are based on duplicity.

In [6], goal-driven policies, depending on serendipity, horizon and dynamism, but ***without duplicity***, are realized, with equiprobable tie-breaking, so that they are covariant, as opposed to the policies realized previously in [2].

Serendipity is intrinsically Boolean (false or true). $h \geq J - 1$ with serendipity is a (probably slower) variation on `mean-max` (53) and will not be further considered; $h = J - 1$ without serendipity ***will not be considered*** either, for brevity. The normal case $J = 3$ makes horizon binary and dynamism Boolean (whether or not to revise at the second and last choice). Hence a workable array of eight strictly reduced goal-driven policies, besides `mean-max` and the completely random or "monkey" policy.

The resulting goal-driven strategies are judged in [6], if possible, using `mean-mean`, for a sample of utility functions, consisting of sums of stationary pointwise functions, with or without common faces; the token transfer function (appendix A, tab. 13) and completely pseudo-random functions. Some interesting difficulties occur:

- Some strategies based on goal memories are non-Markovian and treated implicitely as successive alternatives between partial Markovian strategies.
- Strategy judgment space occasionally explodes, by multiplication of goal ties. Explosion is contained by pruning, that is elimination of dominated strategies [4], like so-called von Neumann $\alpha - \beta$.
- In turn, some dominated strategies, stepping out of pruned trees, cannot be judged.

4.3.2 Fuzzy utility functions

Definition. *A utility function is* fuzzy *(depending on the player) if every strategy it leads to, through any strictly reduced horizon goal-driven policy, becomes strictly more useful with serendipity, whatever horizon and dynamism.*

Fuzziness values are gathered in tab. 10. Question marks reflect the difficulties discussed at the end of section 4.3.1.

Table 10: utility function fuzziness [6]

	First player	Next players
δ[{1, 2, 3}, #2] &	False	False
δ[{1, 2, 3}, #2] + δ[{4, 5, 6}, #2] &	False	False
δ[{1, 2, 3}, #2] + δ[{2, 3, 4}, #2] &	False	False
δ[{1, 1, 1}, #2] + δ[{2, 2, 2}, #2] &	False	False
δ[{1, 2, 3}, #2] + δ[{3, 4, 5}, #2] + δ[{1, 5, 6}, #2] &	False	True
δ[{1, 1, 1}, #2] + δ[{2, 2, 2}, #2] + δ[{1, 1, 3}, #2] + δ[{2, 2, 3}, #2] &	False	False
transfer[3, 6][#2] &	False	True
Max[#2] &	?	?
Total[#2] &	True	True
Times @@ (#1 - 1 &) /@ #2 &	False	False
Times @@ #2 &	False	False
randomUtility[3, 6, 3, 280865]	True	True
randomUtility[3, 6, 3, 28086]	True	False
randomUtility[3, 6, 3, 2808]	True	True
randomUtility[3, 6, 3, 280]	True	True
randomUtility[3, 6, 3, 28]	False	True
randomUtility[3, 6, 3, 2]	True	True

5 Meta-policy or politics

5.1 Policy utility

Definition. *For some policy A and for some utility function u_e, let U, U_0, U_1 be the respective utilities of the strategy $A(u_e)$, the monkey strategy and the complete most useful strategy. The* utility of the policy A *for the utility function u_e is*

$$\overline{U} = \frac{U - U_0}{U_1 - U_0}. \tag{58}$$

Explicitly, using a notation already introduced p. 12,

$$\begin{aligned} U &= u(\mathbf{0})(u_e, A(u_e)), \\ U_0 &= u(\mathbf{0})(u_e, \texttt{mean-max}(0)), \\ U_1 &= u(\mathbf{0})(u_e, \texttt{mean-max}(u_e)). \end{aligned}$$

- $\overline{U} \leq 1$ and for $\overline{U} = 1$, the policy yields a most useful strategy.
- $\overline{U} \lesseqgtr 0$ and for $\overline{U} \leq 0$, the policy yields a harmful strategy.

Policy utilities, for sample utility functions, are computed, if possible, and gathered in tab. 11, 12.

5.2 Stratagem

Definition. *A* stratagem *of a policy A, valid for a set of utility functions $\mathcal{U}$, is a policy B, such that, for some utility function $u_e \in \mathcal{U}$, the strategy $B(u_e)$ is more useful than the strategy $A(u_e)$ but checking $u_e \in \mathcal{U}$ and applying B use less time and space than only applying A: T, S denoting computing time and space,*[3]

$$\begin{aligned} \forall u_e \in \mathcal{U},\ u(\mathbf{0})(u_e, B(u_e)) &\geq u(\mathbf{0})(u_e, A(u_e)), \\ (T, S)(u_e \in \mathcal{U}, B(u_e)) &\leq (T, S)(A(u_e)). \end{aligned}$$

Using a stratagem means taking a logically justified short-cut, illustrated by fig. 2. The reverse process may define panic.

---

[3] One must be careful that $x = y$ does not imply $T(x) = T(y)$; accordingly, in *Mathematica*, T (`Timing` actually) has the `HoldAll` attribute [14].

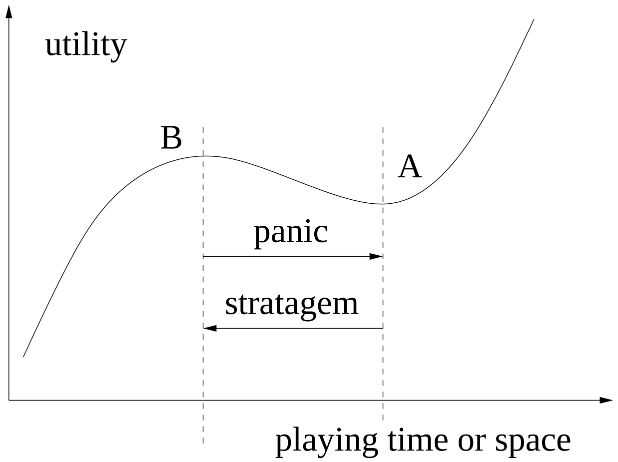
Table 11: first player policy utilities [6]

		serendipity → False		serendipity → True	
		dyn → False	dyn → True	dyn → False	dyn → True
δ[{1, 2, 3}, #2] &	h → 0	1	1	1	1
	h → 1	1	1	1	1
δ[{1, 2, 3}, #2] + δ[{4, 5, 6}, #2] &	h → 0	0.684	0.684	0.684	0.684
	h → 1	0.980	0.980	0.980	0.980
δ[{1, 2, 3}, #2] + δ[{2, 3, 4}, #2] &	h → 0	0.806	0.806	0.806	0.806
	h → 1	0.911	0.980	0.911	0.980
δ[{1, 1, 1}, #2] + δ[{2, 2, 2}, #2] &	h → 0	0.605	0.605	0.605	0.605
	h → 1	0.946	1	0.946	1
δ[{1, 2, 3}, #2] + δ[{3, 4, 5}, #2] + δ[{1, 5, 6}, #2] &	h → 0	0.599	0.599	0.599	0.599
	h → 1	0.946	0.960	0.950	0.974
δ[{1, 1, 1}, #2] + δ[{2, 2, 2}, #2] + δ[{1, 1, 3}, #2] + δ[{2, 2, 3}, #2] &	h → 0	0.587	0.587	0.587	0.587
	h → 1	0.865	1	0.865	1
transfer[3, 6][#2] &	h → 0	0.833	0.833	0.833	0.840
	h → 1	0.840	0.828	0.920	0.993
Max[#2] &	h → 0	0.684	?	1	?
	h → 1	0.050	0.050	1	1
Total[#2] &	h → 0	0.768	0.829	0.857	0.893
	h → 1	0.004	0.004	0.915	0.996
Times @@ (#1 - 1 &) /@ #2 &	h → 0	0.778	0.815	0.778	0.819
	h → 1	0.739	0.706	0.857	0.992
Times @@ #2 &	h → 0	0.798	0.835	0.798	0.838
	h → 1	0.507	0.433	0.871	0.993
randomUtility[3, 6, 3, 280865]	h → 0	-0.095	-0.044	0.277	0.303
	h → 1	0.411	0.507	0.697	0.858
randomUtility[3, 6, 3, 28086]	h → 0	0.336	0.348	0.437	0.461
	h → 1	0.620	0.561	0.877	0.974
randomUtility[3, 6, 3, 2808]	h → 0	0.442	0.440	0.479	0.504
	h → 1	0.484	0.531	0.786	0.965
randomUtility[3, 6, 3, 280]	h → 0	0.395	0.395	0.444	0.452
	h → 1	0.397	0.386	0.742	0.926
randomUtility[3, 6, 3, 28]	h → 0	0.400	0.339	0.400	0.406
	h → 1	0.436	0.465	0.761	0.944
randomUtility[3, 6, 3, 2]	h → 0	0.325	0.325	0.338	0.340
	h → 1	0.520	0.516	0.846	0.938

Table 12: next player policy utilities [6]

		serendipity → False		serendipity → True	
		dyn → False	dyn → True	dyn → False	dyn → True
δ[{1, 2, 3}, #2] &	h → 0	1	1	1	1
	h → 1	1	1	1	1
δ[{1, 2, 3}, #2] + δ[{4, 5, 6}, #2] &	h → 0	0.754	0.754	0.754	0.754
	h → 1	0.951	0.951	0.951	0.951
δ[{1, 2, 3}, #2] + δ[{2, 3, 4}, #2] &	h → 0	0.846	0.846	0.846	0.846
	h → 1	0.792	0.801	0.918	1
δ[{1, 1, 1}, #2] + δ[{2, 2, 2}, #2] &	h → 0	0.636	0.636	0.636	0.642
	h → 1	0.930	1	0.933	1
δ[{1, 2, 3}, #2] + δ[{3, 4, 5}, #2] + δ[{1, 5, 6}, #2] &	h → 0	0.756	0.756	0.821	0.821
	h → 1	0.916	0.928	0.944	1
δ[{1, 1, 1}, #2] + δ[{2, 2, 2}, #2] + δ[{1, 1, 3}, #2] + δ[{2, 2, 3}, #2] &	h → 0	0.646	0.646	0.646	0.699
	h → 1	0.876	0.938	0.937	0.987
transfer[3, 6][#2] &	h → 0	0.730	0.730	0.835	0.881
	h → 1	0.718	0.735	0.936	0.998
Max[#2] &	h → 0	0.610	?	1	1
	h → 1	-0.032	-0.032	1	1
Total[#2] &	h → 0	0.805	0.850	0.912	0.945
	h → 1	0.488	0.487	0.955	1
Times @@ (#1 - 1 &) /@ #2 &	h → 0	0.893	0.931	0.893	0.931
	h → 1	0.867	0.882	0.943	1
Times @@ #2 &	h → 0	0.897	0.930	0.897	0.934
	h → 1	0.872	0.858	0.946	1
randomUtility[3, 6, 3, 280865]	h → 0	-0.101	-0.080	0.724	0.770
	h → 1	0.433	0.433	0.929	0.982
randomUtility[3, 6, 3, 28086]	h → 0	0.891	0.891	0.891	0.898
	h → 1	0.516	0.516	0.945	1
randomUtility[3, 6, 3, 2808]	h → 0	0.379	0.454	0.875	0.887
	h → 1	0.349	0.349	0.935	0.990
randomUtility[3, 6, 3, 280]	h → 0	0.659	0.614	0.743	0.821
	h → 1	0.319	0.319	0.925	0.974
randomUtility[3, 6, 3, 28]	h → 0	0.031	0.031	0.773	0.828
	h → 1	0.369	0.369	0.955	0.984
randomUtility[3, 6, 3, 2]	h → 0	0.250	0.250	0.790	0.827
	h → 1	0.224	0.224	0.914	0.988

Figure 2: stratagem and panic



In the Nim game, a stratagem of all policies exists, based on a congruence property [4, §1.3]. In a 421 round, as shown in section 4.2.1, the ratchet policy is a stratagem of `mean-max`, valid for stationary pointwise Markovian utility functions. Fate tree pruning [6] may be a stratagem of `mean-max` (or some goal-driven policies), just as von Neumann $\alpha - \beta$ is a stratagem of `min-max`.

5.3 Free utility

Definition. *A* meta-policy *or* politics *is a program yielding exactly one policy.*

Definition. *The* free utility *of a policy A for some utility function u_e is $\overline{U} - c$, where $\overline{U}$ is the utility of A and c is the negative utility of computing $A(u_e)$.*

A meta-policy consists in maximizing free utility on a finite policy set, such as the one considered in section 4.3.1. For $c = 0$, that is, without computing constraints, `mean-max` would be most (freely) useful. In queuing problems, for example, $\overline{U} = 0$ (results have no utility for the queuing system) and only c remains to be minimized.

For c depending linearly on computing time and space, free utility is a Legendre transform, like free energy in thermodynamics [8],

$$c = \lambda T + \mu S,\ \ \lambda, \mu \geq 0,$$

where the Lagrange multipliers λ, μ are exchange rates of utility to computing time and space.

For λ or μ large enough, even the monkey policy will be more freely useful than `mean-max`. For $\lambda > 0, \mu = 0$ and arbitrarily large J, bounded horizon goal-driven policies are asymptotically more freely useful than `mean-max`. Proof: according to (51, 52, 47, 48), bounded horizon goal-driven strategy computing time is quadratic in J, hence negligible before `mean-max` exponential.

Strategy computing time or space can be reduced,

- quite generally, but exclusively (not together), by compiling functions or sharing data, without touching output;
- by lexicographic tie-breaking (as in [2]) instead of equiprobable tie-breaking (utility is conserved but not covariance);
- among goal-driven policies, by not using serendipity, reducing horizon or not using dynamism (covariance is conserved but not generally utility).

In order to obtain more precise results on free utility, strategy computing time and space should be more carefully defined and computed. This is left for future work.

5.4 Meta-computing time and space

Definition. *For some policy A and some utility function u_e, the* meta-computing time *(or space) of the strategy $A(u_e)$ is the computing time (or space) of the free utility of A for u_e,*

$$(T', S')(A(u_e)) = (T, S)((\overline{U} - c)(A(u_e))).$$

Meta-computing time or space must not be confused with computing time or space. In particular, to avoid self-reference, free utility cannot depend on meta-computing time or space.[4]

---

[4] The computing time or space of a program cannot be an input of the same program.

In [6], meta-computing times are more easily available than computing times. Meta-computing time increases, when computing time is expected to decrease, at least in the following cases.

- Among equiprobable tie-breaking goal-driven policies, such as A, B, the `mean-max` strategy is expected to have a greatest computing time, is its own meta-policy and has a lowest meta-computing time [6]:

  $$T(A(u_e)) \leq T(\texttt{mean-max}(u_e)) = T'(\texttt{mean-max}(u_e)) \leq T'(B(u_e)). \quad (59)$$

  Reciprocally, if, for every utility function u_e, $T'(A(u_e)) \approx T(A(u_e))$, then A must be a variation of `mean-max`, $A \approx \texttt{mean-max}$.

- Among equiprobable tie-breaking Markovian policies, the monkey strategy has a lowest computing time and a greatest meta-computing time.

- Reducing horizon in goal-driven policies reduces computing time, but leads to goal ties, which increase meta-computing time and space.

It is therefore postulated that the faster the policy, the slower its judgment, in inverse reason:

$$(T.T')(A(u_e)) \geq T(\texttt{mean-max}(u_e))^2, \quad (60)$$

which is consistent with (59).

6 Conclusion

6.1 Laws of goal-driven policy utility

Laws of goal-driven policy utility ***without computing constraints*** ($c = 0$) are inferred[5] from tab. 10, 11, 12.

1. Some utility functions are fuzzy, in particular most pseudo-random utility functions.
2. Some utility functions are non-fuzzy, in particular, most sums of stationary δ functions.

---

[5] More precisely, existential laws are deduced and universal laws are induced.

3. Serendipity is almost always useful (though not strictly).
4. Serendipity makes horizon and dynamism useful.
5. No serendipity makes horizon or dynamism often strictly harmful, especially for fuzzy utility functions.
6. For non-fuzzy utility functions, horizon and dynamism are often more useful than serendipity.

Few people are aware (even in ordinary life) that strictly reduced horizon goal-driven policies do not yield in general most useful strategies and that serendipity is useful in goal-driven policies (law 3).

Law 5, combined with section 5.3, suggests as a non-serendipitous stratagem to reduce horizon or not to use dynamism. (This should be confirmed by a theorem, the hypotheses of which should be checkable in less time and space than horizon or dynamism would use in the first place.)

6.2 Advice to 421 players

A player, for whom arithmetic is more difficult than maximizing, permuting and reading tables, can choose a few goals depending on final utilities weighted by any-time success probabilities, found in tab. 1. This defines a heuristic stationary utility function, hopefully non-fuzzy for law 2 and forgiving lack of serendipity for law 6. Let G be the set of chosen goals.

A non-serendipitous unit-horizon goal-driven policy is advised. The first choice obeys (57), where horizon utilities are obtained from (53, 56):

$$\forall(\mathbf{d} \in B^+(3)\, \mathbf{d} < \mathbf{d}_{1/2}), \tag{61}$$

$$\begin{aligned} u_N^*(1, \mathbf{d}) &= \max_{\mathbf{d}^* \in \partial B^+(3-\|\mathbf{d}\|)} \sum_{k=1}^{2} q(i, 2, \mathbf{d}^*, k, \mathbf{d}^*)\, u(1+k, \mathbf{d}+\mathbf{d}^*) & (62)\\ &= \max_{\mathbf{d}^* \in \partial B^+(3-\|\mathbf{d}\|)} s(i, 2, \mathbf{d}^*)\, u(\_, \mathbf{d}+\mathbf{d}^*) & (63)\\ &= \max_{\mathbf{d}^* \in G,\ \mathbf{d}^* \geq \mathbf{d}} s(i, 2, \mathbf{d}^* - \mathbf{d})\, u(\_, \mathbf{d}^*). & (64) \end{aligned}$$

One of the maximizing goals in (64) is equiprobably chosen and the second (and last) decision is made by ratcheting towards this goal (38).

Dynamism, for its uncertain utility in this case, and its certain computing time and space increase, is not advised. Duplicity (defined p. 32) may offer additional utility for cheap but this is left for future work.

6.3 Serendipitous findings

Politics is not an exact science:

- The utility function or the actual game boundaries may be uncompletely known (see p. 8).
- Goal-driven policy utility may depend greatly on utility function (tab. 11, 12), in a way that can hardly be predicted, except for the laws of section 6.1. A statistical approach in this matter is rather delicate.
- Strictly reduced horizon goal-driven policies replace `mean-max` when it is not practicable. But then, from (59) (and possibly a similar argument on computing space) `mean-mean` is neither practicable: surrogate policies are thus needed, paradoxically, when they cannot be judged.

Strategy judgment (with `mean-mean`) explains a point of law theory: strategy utility is computed anti-causally (backward) from final utilities, themselves computed causally (forward) from history; anti-causality is the way of philosophers' "natural law"; causality in the way of judges' "positive law". Both are interwoven in a dual system.

The 421 game has many psychological features. Some of them depend specifically on the rules of the game:

- next player dilemmas;
- Markovian utility function;
- (39) gives an incentive to build (by ratcheting) rather than to fool around, like *The Cicada and the Ant*;
- the best and worst final states, 421 and 221, differ by only one face, so that one risks the worst by trying to reach the best.

The above features may be anthropically justified, as the game would not have become popular, if it did not reflect people's concern.

Other psychological features seem to emerge more generally from the quest for utility under computing constraints: goal, serendipity, horizon, dynamism, duplicity, fuzziness, stratagem and panic. But do people use policies or meta-policies, consciously or unconsciously, individually or collectively? Is the von Neumann-Morgenstern under computing constraints enough to make a "theory of mind"? At least one point is confirmed, the utility of utility.

A Rules of 421

[15] The hardware consists of three dice and eleven tokens, initially in a pot, and a dice board. There are two or more players who can always see dice and tokens. They must not communicate on their strategies.

In the first part of the game, the charge, players get tokens from the pot. In the second and last part of the game, the discharge, players get tokens from each other. The winners are the players who get no token during the charge or the first player who gets rid of his tokens during the discharge.

Charge or discharge is a sequence of sets. A set is a sequence of rounds, when each player alone at his turn handles the dice.[6]

A face combination is conventionally noted as a decreasing sequence like 421, without delimiters. For $f \in \{1 \ldots 6\}$, the f-"brelan" is fff; for $f \in \{2 \ldots 6\}$, the f-pair is $f11$; the sequences are $654, 543, 432, 321$; "nénette" is 221.

The first player casts dice up to thrice and next players in the same set must cast exactly as many times. After casting, a player pushes some dice away from the board, where they remain still for the rest of the round, and casts again any remaining dice.

All end of round face combinations are ordered as follows: $\succ$ meaning "better than",

$$\begin{aligned} &421 \succ 111 \succ 611 \succ 666 \succ 511 \succ 555 \succ 411 \succ 444 \succ 311 \\ &\qquad \succ 333 \succ 211 \succ 222 \succ 654 \succ 543 \succ 432 \succ 321 \succ 665 \succ \ldots 221, \end{aligned} \tag{65}$$

where ellipsed combinations are ordered like the natural numbers they look like, for example, $664 \succ 663$.

At end of set, the last[7] player with the worst combination gets a number of tokens, according to tab. 13, at charge, from the pot or, at discharge, from the first[8] player with the best combination. When the pot is empty, the charge is over; when any player cannot provide, the discharge is over.

---

[6] The order of players matters and remains to be defined.

[7] "Last": a proposal for tie-breaking.

[8] "First": a proposal for tie-breaking.

Table 13: token transfer function, $f \in \{2 \ldots 6\}$

best combination	number of token transfered
421	10
111	7
f-brelan	f
f-pair	f
sequence	2
other	1

B Transport theory, Galton-Watson problem

Identifying face, presence probability, transition probability and utility respectively with phase, flux, cross section and importance, (26, 29) appear as a linear transport equation and its adjoint equation [16]. Moreover, casting and pushing away are similar to diffusion and capture.

Indeed, linear transport theory is about branching processes, from neutron chain reaction to surname transmission (the Galton-Watson problem), as is the point of [17]. A 421 round can be discussed in genealogical terms: each die is considered as an individual, dying after being cast, with no child (when it is pushed away) or exactly one child (itself reborn), so that the population D_j (2) is decreasing (6).

In Galton-Watson problem, each individual's offspring must be independent of others'. For $\mathbf{d}^* \in B^+(D)$, according to the $\mathbf{d}^*$-ratchet policy (38), childless dice have their faces in $\mathbf{d}^*$, but not the converse in general. For example, with $\mathbf{d}^* \equiv 221, \mathbf{d}_{1/2} \equiv 211, J > 1$, the two dice facing up 1 have correlated offsprings: either has a child if and only if the other has none. Nevertheless, dice always have mutually independent offspring if and only if $\mathbf{d}^*$ is a brelan and the player is first, which is ***assumed***.

Following step by step [9, §6.2], the probability law of $D_j, j \in \{0 \ldots \text{J-1}\}$, will be determined. For $j > 0$, the D_{j-1} dice to be cast at date $j - 1/2$ are indexed by $d \in \{1 \ldots D_{j-1}\}$; let Z_d be the child number of die d,

$$Z_d \in \{0, 1\},\ D_j = \sum_{d=1}^{D_{j-1}} Z_d. \tag{66}$$

The Z_d are independent random variables, all with the same stationary probability law q and generating function g,

$$q_z = \mathcal{P}(Z_d = z),\ q_0 = \frac{1}{F},\ q_1 = 1 - q_0,\ g(z) = \langle z^{Z_d} \rangle = q_0 + q_1 z.$$

From (66), and a well-known property of generating functions, the generating function of D_j, knowing D_{j-1}, is

$$\langle z^{D_j} | D_{j-1} = d \rangle = g(z)^d.$$

Let g_j be the generating function of D_j,

$$g_j(z) = \langle z^{D_j} \rangle = \sum_{d=0}^{D} \mathcal{P}(D_{j-1} = d)\, \langle z^{D_j} | D_{j-1} = d \rangle = \sum_{d=0}^{D} \mathcal{P}(D_{j-1} = d)\, g(z)^d,$$
$$g_j(z) = g_{j-1}(g(z)).$$

From (6), $D_0 = D$, $g_0(z) = z^D$, and, as shown by induction,

$$g_j = g_0 \circ g^{\circ j} = (g^{\circ j})^D = (1 - q_1^j + q_1^j z)^D,$$

so that D_j follows a binomial probability law,

$$\forall j \in \{1 \dots J-1\},\ \mathcal{P}(D_j = d) = \frac{D!}{d!\,(D-d)!}\, q_1^{j\,d}\, (1 - q_1^j)^{D-d},$$

which could have been directly deduced by noticing that each die dies, when it is pushed away, or lives, with the probability q_1^j after j casts, independently of other dice.

C Indifference principle, Buridan donkey problem

Equiprobable tie-breaking (35) follows from the indifference principle [18], discussed in [19, ch. 4, § 13] and summarized as follows [19, ch. 7, § 3]:

> According to the Principle of Indifference, 'cases' are held to be equi-probable when there is no reason for preferring any one to any other, when there is nothing, as with Buridan ass, to determine the mind in any one direction.

In statistical mechanics [8], equiprobability follows from maximizing missing information (entropy) without constraints. Occasionally, equiprobability may be justified by mechanical symmetry.

Actually, using probabilities implies some kind of ignorance, paid by uncertainty on probabilities themselves (or probabilities on probabilities); the arbitrary and paradoxical characters of probabilities are summarized in [20, ch. 1, § 5]:

> La définition complète de la probabilité est donc une sorte de pétition de principe: comment reconnaître que tous les cas sont équiprobables ?
>
> (The complete definition of probability is thus a kind of petition of principle: how to recognize that all cases are equiprobable?)

Back to 421, the principle of indifference is shaken. Firstly, lexicographic tie-breaking is a rational alternative to equiprobable tie-breaking [2, 6]. Secondly, considering providence as a player, providential tie-breaking appears to be non-equiprobable (12) (see also p. 9). One may object that providence differs from a player by its lack of consciousness. But what is consciousness?

Random dice face combination non-equiprobability is usually explained by dice discernibility: arrangements, not combinations, are equiprobable. However, discernibility has its own limit, the Gibbs paradox [8].

The principle of indifference is wrong, it must be replaced by statistical experiment.

One can argue that no tie-breaking problem exists practically, because utilities are real numbers, that will always differ slightly, at least for fluctuations or errors. However, this is not at all the present stand.

The tie-breaking problem, epitomized by Buridan donkey story, exists, at least in normal world, where one must choose, but has essentially no unique rational solution and many irrational solutions [21].

References

[1] Pierre Albarède. 421 : un jeu stochastique. In J. B. Hiriart-Urruty et al., editors, *Huitièmes journées du groupe MODE de la SMAI*, Toulouse, March 2000. Univ. Paul Sabatier MIP. `http://mip.ups-tlse.fr/MODE2000`.

[2] Pierre Albarède. Fate stochastic management and policy benchmark in 421, a popular game. 2000. `http://arxiv.org/abs/math.OC/0007129`.

[3] Norbert Wiener. *Invention: the care and feeding of ideas.* MIT Press, 1993.

[4] Prajit K. Dutta. *Strategies and games - Theory and Practice.* MIT Press, 1999.

[5] Rufus Isaacs. *Differential games A mathematical theory with applications to warfare and pursuit, control and optimization.* Wiley, 1965.

[6] Pierre Albarède. Control policies for 421, a stochastic game of life. 2005. `http://pierre.albarede.free.fr`.

[7] Rémi Barrère. *Mathematica, calcul formel et programmation symbolique pour l'informatique scientifique.* Vuibert, 2002.

[8] Frederick Reif. *Fundamentals of statistical and thermal physics.* Mc Graw Hill, 1988.

[9] Emanuel Parzen. *Stochastic process.* Holden day, San Francisco, 1962.

[10] Athanasios Papoulis. *Probability, random variables and stochastic processes.* Mc Graw Hill, 1965.

[11] E. T. Jaynes. *Probability theory, the logic of science.* Cambridge Univ. Press, 2003.

[12] Richard Bellman and Robert Kalaba. *Dynamic programming and modern control theory.* Academic Press, 1965.

[13] Patrick Fabiani. *Représentation dynamique de l'incertain et stratégie de perception pour un système autonome en environnement évolutif.* PhD thesis, École Nationale Supérieure de l'Aéronautique et de l'Espace, Toulouse, 1996.

[14] Stephen Wolfram. *The Mathematica book.* Wolfram Media, 5-th edition, 1998-2003.

[15] Éditions Jeujura. *Règles de jeux de société*, 1999.

[16] Kenneth M. Case and Paul Zweifel. *Linear Transport Theory*. Addison-Wesley, 1967.

[17] T. E. Harris. *The theory of branching processes*. Springer, 1963.

[18] Eric W. Weisstein. Principle of insufficient reason. `http://mathworld.wolfram.com/PrincipleofInsufficientReason.html`.

[19] John Maynard Keynes. *A treatise on probability*. Macmillan, 1921.

[20] Henri Poincaré. *Calcul des probabilités*. Jacques Gabay, 1912–1987.

[21] Franck Cochoy. *Une sociologie du packaging ou l'âne de Buridan face au marché*. Presses Univ. de France, 2002.

.oOo.