Prequential Randomness

Vladimir Vovk[1] and Alexander Shen[2]

[1] Department of Computer Science, Royal Holloway, University of London, Egham, Surrey TW20 0EX, England, vovk@cs.rhul.ac.uk
[2] Laboratoire d'Informatique Fondamentale, Centre de Mathématiques et Informatique, 39 rue Joliot-Curie, F-13453 Marseille Cedex 13, France, alexander.shen@lif.univ-mrs.fr

Abstract. This paper studies Dawid's prequential framework from the point of view of the algorithmic theory of randomness. The main result is that two natural notions of randomness coincide. One notion is the prequential version of the standard definition due to Martin-Löf, and the other is the prequential version of the martingale definition of randomness due to Schnorr. This is another manifestation of the close relation between the two main paradigms of randomness, typicalness and unpredictability. The algorithmic theory of randomness can be stripped of the algorithms and still give meaningful results; the typicalness paradigm then corresponds to Kolmogorov's measure-theoretic probability and the unpredictability paradigm corresponds to game-theoretic probability. It is an open problem whether the main result of this paper continues to hold in the stripped version of the theory.

1 Introduction

The two fundamental paradigms of randomness are unpredictability and typicalness (see, e.g., [1], Chapter 1). Unpredictability is based on the idea of gambling: a sequence is regarded random if there is no computable way to become infinitely rich betting on its elements. Typicalness is based on the idea of measure: a sequence is regarded random if there is no computable way to specify a set of measure zero containing this sequence. The standard formal definition of unpredictability is due to Schnorr [19] and Levin [12], and the standard formal definition of typicalness is due to Martin-Löf [14]. Schnorr and Levin [19, 12] also established equivalence between unpredictability and typicalness.

The standard definitions of randomness are given relative to a stochastic mechanism supposedly generating the sequence. In this paper we ask the question whether the equivalence will still hold in the "prequential framework", i.e., when no such stochastic mechanism is assumed.

The first prequential definition of randomness was proposed by Dawid [4]. Dawid's definition, however, was based on von Mises's idea of subsequence selection rules, which Ville [22] showed to be inadequate in some important respects. Dawid ([4], Section 13.2) also gave a brief description of a prequential definition based on Ville's martingales, but did not elaborate on it. This paper provides the details of this definition, which belongs to the unpredictability paradigm. It also gives a Bayesian definition of typicalness. Its main mathematical result says that the prequential notions of unpredictability and typicalness coincide.

Prequential framework for randomness

Set $\Pi := ([0,1] \times \{0,1\})^\infty$; this will sometimes be referred to as the *prequential space*. A typical element $(p_1, y_1, p_2, y_2, \ldots)$ of Π is interpreted as the record of predictions $p_1, p_2, \ldots$ output by a forecaster and the observed outcomes $y_1, y_2, \ldots$, assumed binary. Intuitively, p_n is the forecaster's subjective probability that $y_n = 1$ after having observed $y_1, \ldots, y_{n-1}$ and taking account of all other information available at the time of issuing the forecast.

Dawid's prequential principle ([3]; it is called "M2" in [4] and "weak prequential principle" in [6, 7]) says that our evaluation of the quality of the forecasts $p_1, p_2, \ldots$ in light of the observed outcomes $y_1, y_2, \ldots$ should not depend on the forecaster's probability model (typically, a probability measure describing the forecaster's beliefs about the stochastic mechanism generating the data). The prequential principle is satisfied automatically in our framework, since the prequential space does not involve the forecaster's probability model at all.

In the following two sections we will introduce two different notions of randomness of a sequence $\pi \in \Pi$. Intuitively, $(p_1, y_1, p_2, y_2, \ldots)$ is random if the p_n are good predictions of y_n; slightly more precisely, if there is no computable way to detect inadequacy of p_n. The first definition, in Section 2, belongs to the unpredictability paradigm and the second, in Section 3, to the typicalness paradigm. The equivalence between the two definitions, which is the main result of the paper, is stated in Section 4 and proved in Section 5. The concluding Section 6 discusses directions for further research.

For understanding the intuitive meaning of our statements, the following intuitive idea of lower semicontinuity will suffice: a function $f : X \to \mathbb{R}$ is lower semicomputable if there is an algorithm that, for all $x \in X$ and $r \in \mathbb{R}$, will eventually tell us that $f(x) > r$ if this inequality is indeed true. (Lower semicomputable functions are not necessarily computable as the algorithm can work arbitrarily long.) Understanding the proofs requires precise definitions, as given in, e.g., Appendix A.

Some other notation

The set of all natural (i.e., positive integer) numbers is denoted $\mathbb{N}$, $\mathbb{N} := \{1, 2, \ldots\}$; $\overline{\mathbb{N}}$ is $\mathbb{N}$ extended by adding ∞. As always, $\mathbb{Q}$ and $\mathbb{R}$ are the sets of all rational and real numbers, respectively.

Let $\Omega := \{0,1\}^\infty$ be the set of all infinite binary sequences and $\Omega^\diamond := \{0,1\}^*$ be the set of all finite binary sequences. Set $\Pi^\diamond := ([0,1] \times \{0,1\})^*$. The empty element (sequence of length zero) of both $\Omega^\diamond$ and $\Pi^\diamond$ will be denoted $\Box$. In our applications, the elements of $\Pi^\diamond$ will be finite sequences of predictions and outcomes, and the elements of Ω and $\Omega^\diamond$ will be sequences of outcomes (infinite or finite).

For $x \in \Omega^\diamond$, let $\Gamma_x \subseteq \Omega$ be the set of all infinite continuations of x. Similarly, for $x \in \Pi^\diamond$, $\Gamma_x \subseteq \Pi$ is the set of all infinite continuations of x. For each $\omega = (y_1, y_2, \dots) \in \Omega$ and $n \in \mathbb{N}$, set $\omega^n := (y_1, \dots, y_n)$. Similarly, for each $\pi = (p_1, y_1, p_2, y_2, \dots) \in \Pi$ and $n \in \mathbb{N}$, set $\pi^n := (p_1, y_1, \dots, p_n, y_n)$.

2 Unpredictability

A *farthingale* is a function $V : \Pi^\diamond \to [0, \infty]$ satisfying

$$V(p_1, y_1, \dots, p_{n-1}, y_{n-1}) = (1 - p_n)V(p_1, y_1, \dots, p_{n-1}, y_{n-1}, p_n, 0) + p_n V(p_1, y_1, \dots, p_{n-1}, y_{n-1}, p_n, 1) \quad (1)$$

for all n and all $(p_1, y_1, p_2, y_2, \dots) \in \Pi$. If we replace "=" by "≥" in (1), we get the definition of *superfarthingales*. These are prequential versions of the standard notions of martingale and supermartingale; we are following the terminology of [7].

The value of a farthingale can be interpreted as the capital of a gambler betting according to the odds announced by the forecaster. In the case of superfarthingales, the gambler is allowed to throw away part of his capital.

Lemma 1. *Let $\mathcal{V}$ be the class of all non-negative lower semicomputable superfarthingales V with initial value $V(\Box) = 1$. There exists a largest superfarthingale in $\mathcal{V}$ to within a constant factor.*

Proof. Fix a universal sequence of lower semicomputable functions $f_1, f_2, \dots$ on $\Pi^\diamond$ (see Lemma 6 in Appendix A). It is easy to construct a new computable sequence of semicontinuous functions $f'_1, f'_2, \dots$ such that each of f'_l is a superfarthingale in $\mathcal{V}$ and that $f'_l = f_l$ whenever f_l is already in $\mathcal{V}$, $l \in \mathbb{N}$. Then $\sum_{l=1}^{\infty} 2^{-l} f'_l$ will be a largest superfarthingale in $\mathcal{V}$. □

Let us fix a largest superfarthingale U in $\mathcal{V}$ and call it the *universal* superfarthingale. A sequence $\pi \in \Pi$ is *unpredictable* if $U(\pi^n)$ stays bounded as $n \to \infty$. The following lemma gives an equivalent definition of unpredictable sequences.

Lemma 2. *A sequence $\pi \in \Pi$ is unpredictable if and only if $U(\pi^n)$ does not tend to infinity as $n \to \infty$.*

Proof. Following the proof of Lemma 3.1 in [21], we can construct a superfarthingale $V \in \mathcal{V}$ such that $\liminf_{n\to\infty} V(\pi^n) = \infty$ whenever $\sup_n U(\pi^n) = \infty$. (Therefore, $\liminf_{n\to\infty} U(\pi^n) = \infty$ whenever $\sup_n U(\pi^n) = \infty$.) Indeed, for each $m \in \mathbb{N}$, the function $U : \Pi^\diamond \to [0, \infty)$ defined by

$$U^m(x) := \begin{cases} 2^m & \text{if } U(y) > 2^m \text{ for some prefix } y \text{ of } x \\ U(x) & \text{otherwise} \end{cases}$$

is a superfarthingale; it is clear that it is lower semicomputable and so belongs to $\mathcal{V}$. Since $U^1, U^2, \ldots$ is a computable sequence of lower semicontinuous functions, we can set

$$V := \sum_{m=1}^{\infty} 2^{-m} U^m. \qquad \square$$

3 Typicalness

We can also adapt the standard definition of typicalness to the prequential framework. First we give an informal version of the definition.

A *forecasting system* is a function $\phi : \Omega^\diamond \to [0,1]$. Let Φ be the set of all forecasting systems. For each $\phi \in \Phi$ there exists a unique probability measure $\mathbb{P}_\phi$ on Ω such that, for each $x \in \Omega^\diamond$, $\mathbb{P}_\phi(\Gamma_{x1}) = \phi(x)\,\mathbb{P}_\phi(\Gamma_x)$. (In other words, such that $\phi(x)$ is a version of the conditional probability, according to $\mathbb{P}_\phi$, that x will be followed by 1.) The notion of a forecasting system is close to that of a probability measure on Ω: the correspondence $\phi \mapsto \mathbb{P}_\phi$ becomes an isomorphism if we only consider forecasting systems taking values in the open interval $(0,1)$ and probability measures taking positive values on the sets Γ_x, $x \in \Omega^\diamond$.

Informally, we say that a sequence $\omega \in \Omega$ is *typical* w.r.t. to a forecasting system ϕ if it is typical (i.e., random in the sense of Martin-Löf [14]) w.r.t. to $\mathbb{P}_\phi$ when ϕ is given as an oracle. We will formalize "given as an oracle" using some simplest notions of effective topology (see Appendix A). The following definition is a version of Levin's "uniform test of randomness" [13, 9].

Definition 1. *A* uniform test of typicalness *is a lower semicomputable function* $T : \Omega \times \Phi \to \overline{\mathbb{N}}$ *such that, for all* $\phi \in \Phi$ *and all* $m \in \mathbb{N}$,

$$\mathbb{P}_\phi\{\omega \in \Omega \mid T(\omega,\phi) \ge m\} \le 2^{-m}. \tag{2}$$

Intuitively, $T(\omega,\phi)$ is the number of anomalies (measured in bits, according to (2)) discovered in ω w.r.t. to ϕ. The requirement of lower semicomputability means that the anomalies have to be genuine: a discovery of anomaly can never be undone.

Lemma 3. *There exists a largest, to within an additive constant, test of typicalness. In other words, there exists a test of typicalness T such that, for any other test of typicalness T', there exists a constant C such that, for any $(\omega,\phi) \in \Omega \times \Phi$,*

$$T(\omega,\phi) \ge T'(\omega,\phi) - C.$$

Proof. The proof is similar to the standard one given by Martin-Löf [14]; it will, however, crucially depend on the compactness of Φ, as in [13]. For each set $G \subseteq \Omega \times \Phi$ and each $\phi \in \Phi$ we will use the notation

$$G[\phi] := \{\omega \in \Omega \mid (\omega,\phi) \in G\}$$

for the ϕ-*cut* of G. A convenient alternative representation of a test of typicalness T is as a computable sequence of nested open sets $G_1 \supseteq G_2 \supseteq \cdots$ in $\Omega \times \Phi$ such that

$$\mathbb{P}_\phi(G_m[\phi]) \le 2^{-m} \tag{3}$$

for all $\phi \in \Phi$ and $m \in \mathbb{N}$. It is easy to see that the representations are indeed equivalent: when given T we can set $G_m := \{(\omega, \phi) \mid T(\omega, \phi) \ge m\}$, and when given $G_1, G_2, \ldots$, we can set $T(\omega, \phi) := \max\{m \mid (\omega, \phi) \in G_m\}$. Such sequences $G_1, G_2, \ldots$ will also be referred to as *tests of typicalness* (dropping the word "uniform").

Let $G_{l,m}$ be a universal computable family of sequences of open sets (cf. Lemma 5 in Appendix A). Put $G'_{l,m} := \cap_{i=1}^m G_{l,i}$, so that $G'_{l,m}$ is a computable family of nested sequences of open sets containing all nested computable sequences of open sets. We can further "trim" each $G'_{l,m}$ to $G''_{l,m}$ so that:

- $\mathbb{P}_\phi(G''_{l,m}[\phi]) \le 2^{-m}$ for all $\phi \in \Phi$;
- $G''_{l,m} = G'_{l,m}$ whenever $\mathbb{P}_\phi(G'_{l,m}[\phi]) < 2^{-m}$ for all $\phi \in \Phi$.

Indeed, let $G'_{l,m} = \cup\{U_k \mid (l, m, k) \in A\}$ be the representation of $G'_{l,m}$ as the union of basic sets. Set $H_K := \cup\{U_k \mid (l, m, k) \in A, k \le K\}$, so that $H_1, H_2, \ldots$ is a non-decreasing sequence of simple sets whose union is $G'_{l,m}$. Remember that, by (15), $\overline{H_K} \subseteq G'_{l,m}$. We may "quarantine" new H_K until they are "cleared", i.e.,

$$\forall \phi \in \Phi : \mathbb{P}_\phi(\overline{H_K}[\phi]) < 2^{-m} \tag{4}$$

is established. The open set $G''_{l,m}$ is defined as the union of the H_K that are cleared.

Let us check that condition (4) can indeed be eventually established by a computable procedure when it is satisfied. Suppose (4) is satisfied. The set

$$S := \left\{\phi \in \Phi \mid \mathbb{P}_\phi(\overline{H_K}[\phi]) < 2^{-m}\right\}$$

is effectively open, so that we can effectively generate a sequence of basic sets $U'_k \subseteq \Phi$ whose union is S. By the compactness of Φ, already a finite number of U'_k will cover S when $S = \Phi$, and so (4) can be established in a computable manner.

Therefore, we can list all tests of typicalness, in the following sense: there is a computable sequence $(G''_{l,m})_{m=1}^\infty$, $l = 1, 2, \ldots$, of tests of typicalness that contains all "strict" tests of typicalness (i.e., those satisfying the required inequality with "$<$" instead of "$\le$"). To obtain a largest test of typicalness G_m, it suffices to set

$$G_m := \cup_{l=1}^\infty G_{l,m+l}.$$

Indeed, the computability of the sequence of open sets G_m is obvious,

$$\mathbb{P}_\phi(G_m[\phi]) \le \sum_{l=1}^\infty \mathbb{P}_\phi(G_{l,m+l}[\phi]) \le \sum_{l=1}^\infty 2^{-m-l} = 2^{-m}, \quad \forall \phi \in \Phi, \forall m \in \mathbb{N},$$

and, for each $l \in \mathbb{N}$,

$$T(\omega, \phi) = \max\{m \mid (\omega, \phi) \in G_m\} \geq \max\{m \mid (\omega, \phi) \in G_{l,m+l}\} = T_l(\omega, \phi) - l,$$

where T is the functional representation of the test $(G_m)_{m=1}^{\infty}$ and T_l is the functional representation of the test $(G_{l,m})_{m=1}^{\infty}$. $\square$

Let us fix a largest uniform test of typicalness T and call it the *universal test of typicalness*. A sequence $\omega \in \Omega$ is said to be *typical w.r.t. to* $\phi \in \Phi$ if $T(\omega, \phi) < \infty$.

Definition 2. *We say that* $\pi = (p_1, y_1, p_2, y_2, \ldots) \in \Pi$ *is* typical *if there exists a forecasting system* ϕ *such that* $(y_1, y_2, \ldots)$ *is typical w.r.t. to* ϕ *and* ϕ agrees *with* π*, in the sense that* $p_n = \phi(y_1, \ldots, y_{n-1})$ *for all* n.

4 Main result

Theorem 1. *A sequence* $\pi \in \Pi$ *is unpredictable if and only if it is typical.*

This theorem will be proved in the next section. The proof will be based on Levin's [13] ideas (see also [9]).

The philosophical significance of Theorem 1 is that it establishes the equivalence of the purely prequential and Bayesian viewpoints in the framework of the algorithmic theory of randomness. The definition of typicalness is Bayesian, in that the forecaster is modelled as a coherent decision maker, computing his forecasts by conditioning a probability measure; rejecting the forecasts is the same as rejecting all probability measures that could have produced those forecasts. The definition of unpredictability is purely prequential, in that it does not see any probability measures behind the forecasts; the latter are used for testing directly.

A simple corollary of Theorem 1 is the following observation:

Corollary 1. *Let* ϕ *be a computable forecasting system. A binary sequence* $(y_1, y_2, \ldots)$ *is random w.r.t. to* $\mathbb{P}_\phi$ *in the sense of Martin-Löf if and only if the sequence* $(p_1, y_1, p_2, y_2, \ldots)$ *is unpredictable (equivalently, typical), where* $p_n := \phi(y_1, \ldots, y_{n-1})$*,* $n \in \mathbb{N}$.

Therefore, the prequential notions of unpredictability and typicalness generalize Martin-Löf's notion of randomness.

Remark 1. Notice that we have never assumed that the past observations $y_1, \ldots, y_{n-1}$ are the only information available to the forecaster when choosing the prediction p_n for the next outcome y_n. The forecaster is allowed to (and typically does) use all kinds of "side information" in addition to the past observations. It is easy to extend all our definitions and results to the case where some of this side information, x_n, is also known to the gambler. (As in [4], Section 9.) As an example, the definition of a superfarthingale, (1), becomes

$$
\begin{aligned}
V(x_1, p_1, y_1, \ldots, x_{n-1}, p_{n-1}, y_{n-1}) \\
&= (1 - p_n) V(x_1, p_1, y_1, \ldots, x_{n-1}, p_{n-1}, y_{n-1}, x_n, p_n, 0) \\
&\qquad + p_n V(x_1, p_1, y_1, \ldots, x_{n-1}, p_{n-1}, y_{n-1}, x_n, p_n, 1).
\end{aligned}
$$

Remark 2. Since we do not record side information in the main part of this paper, the forecasting systems that we consider are never assumed computable: even if the forecaster computes each forecast from the past outcomes and the side information, typically the forecast cannot be computed from the past outcomes alone. It is not even obvious that the notion of a forecasting system ϕ as we defined it (a function of past outcomes) is meaningful. It involves the following controversial picture along the lines of Pearl's "local surgeries". To elicit the value of the function ϕ on a binary sequence $y_1, \ldots, y_n$, we act as follows. First we wait until the nature produces the first piece of side information x_1 and, in response, the forecaster produces p_1. Then we perform a "local surgery" replacing the nature's outcome by y_1 (if it is different from y_1). Now the nature produces x_2 and the forecaster produces p_2. Another local surgery replaces the outcome by y_2. Etc. Finally, the forecaster produces p_n, which is taken to be the value of ϕ on $y_1, \ldots, y_n$. According to Theorem 1, this philosophically questionable approach (see, e.g., Section 4 of Pearl's response in [18] and Dawid's and Shafer's critique of Pearl's use of counterfactuals) leads to the same notion of randomness as the philosophically immaculate approach of Section 2.

Remark 3. There is a feature of our definition of typicalness which some readers might find counterintuitive. Consider a sequence $(p_1, y_1, p_2, y_2, \ldots) \in \Pi$ generated from the Markov chain $\mathbb{P}_{\phi_{\mathrm{Mark}}}$, where

$$
\phi_{\mathrm{Mark}}(y_1, \ldots, y_n) := \begin{cases} 1/3 & \text{if } n > 0 \text{ and } y_n = 0 \\ 2/3 & \text{if } n > 0 \text{ and } y_n = 1 \\ 1/2 & \text{if } n = 0 \end{cases} \tag{5}
$$

for all $(y_1, \ldots, y_n) \in \Omega^{\diamond}$. Therefore, $p_1 = 1/2$, y_n is obtained by tossing a biased coin so that $y_n = 1$ with probability p_n, $n = 1, 2, \ldots$, and $p_n = 1/3$ or $p_n = 2/3$ according to whether $y_{n-1} = 0$ or $y_{n-1} = 1$, $n = 2, 3, \ldots$. When will the universal test of typicalness declare such a sequence untypical? Of course, this will always happen if the test is given the sequence of forecasts $(p_1, p_2, \ldots)$: the sequence of outcomes $(y_1, y_2, \ldots)$ is computable from $(p_1, p_2, \ldots)$ (y_1 is determined by p_2, y_2 is determined by p_3, etc.) but the set $\{(y_1, y_2, \ldots)\}$ has measure zero. According to our definition of typicalness at the end of Section 3, for the sequence $(p_1, y_1, p_2, y_2, \ldots)$ to be typical there should *exist* a forecasting system ϕ such that $p_1, p_2, \ldots$ are produced by ϕ along the sequence $(y_1, y_2, \ldots)$ and such that $(y_1, y_2, \ldots)$ is typical w.r.t. to ϕ. This works because we can choose a ϕ that "hides" $(p_1, p_2, \ldots)$ very well: the Markov forecasting system (5) will do so. The quantifier "exist" is essential: it cannot be replaced by "for all". In particular, we cannot take as ϕ Dawid's ([5], Section 7.3) *prequential model*

$$
\phi_{\mathrm{preq}}(x) := p_n, \quad \forall n \in \mathbb{N}, \forall x \in \Omega^{n-1}.
$$

The problem is that the prequential model is too honest and does not try to hide the p_n at all; consequently, $T((y_1, y_2, \ldots), \phi_{\mathrm{preq}}) = \infty$, where T is the universal test of typicalness. This is a manifestation of the fact that the universal test of typicalness violates Dawid's strong prequential principle [6, 7], which recommends, roughly, the following procedure for testing agreement between $(p_1, p_2, \ldots)$ and $(y_1, y_2, \ldots)$. Choose a test of typicalness $T(\omega, \phi)$ that depends on (ω, ϕ) only via $(\omega, \phi(\omega))$, where $\phi(\omega)$ is the sequence of forecasts made by ϕ on ω (cf. (7) in the proof of Theorem 1 below). There might be some restrictions (regarded as mild in a given context) on the class of forecasting systems ϕ considered. Now the disagreement between $(p_1, p_2, \ldots)$ and $(y_1, y_2, \ldots)$ can be defined as $T(\omega, \phi)$ being large, where ϕ is any forecasting system that agrees with $(p_1, y_1, p_2, y_2, \ldots)$. It does not matter which ϕ is chosen; e.g., the prequential model will do.

5 Proof of Theorem 1 and Corollary 1

The proof of the theorem will depend on a fundamental result called Ville's inequality. Let ϕ be a forecasting system. A *martingale* w.r. to ϕ is a function $V : \Omega^\diamond \to [0, \infty]$ satisfying

$$V(x) = (1 - \phi(x))V(x, 0) + \phi(x)V(x, 1) \tag{6}$$

for all $x \in \Omega^\diamond$. If we replace "=" by "$\geq$" (respectively, "$\leq$") in (6), we get the definition of a *supermartingale* (respectively, *submartingale*) w.r. to ϕ.

Proposition 1 (Ville's inequality, [22], p. 100). *If ϕ is a forecasting system, V is a non-negative supermartingale w.r. to ϕ with initial value $V(\Box) = 1$, and $C > 0$,*

$$\mathbb{P}_\phi \left\{ \omega \in \Omega \mid \sup_n V(\omega^n) \geq C \right\} \leq \frac{1}{C}.$$

Fix $\pi \in \Pi$.

Part "if" of Theorem 1

Suppose π is not unpredictable. Then $\pi \in G_m$ for all $m \in \mathbb{N}$, where

$$G_m := \{\pi \in \Pi \mid \exists n : U(\pi^n) > 2^m\}$$

and U is the universal superfarthingale.

We will not distinguish between $(p_1, y_1, p_2, y_2, \ldots) \in \Pi$ and the pair of sequences $((p_1, p_2, \ldots), (y_1, y_2, \ldots)) \in [0, 1]^\infty \times \Omega$. For $\phi \in \Phi$ and $\omega = (y_1, y_2, \ldots) \in \Omega$ we set

$$\phi(\omega) := (\phi(\Box), \phi(y_1), \phi(y_1, y_2), \ldots) \in [0, 1]^\infty. \tag{7}$$

The mapping $(\omega, \phi) \mapsto \phi(\omega)$ from $\Omega \times \Phi$ to $[0, 1]^\infty$ is continuous. Therefore, the mapping $(\omega, \phi) \mapsto (\phi(\omega), \omega)$ from $\Omega \times \Phi$ to Π is also continuous. Therefore, the set

$$G'_m := \{(\omega, \phi) \mid (\phi(\omega), \omega) \in G_m\}$$

is open.

Let us check that G'_m is a test of typicalness. The computability requirement is obvious. Fix $m \in \mathbb{N}$ and $\phi \in \Phi$. To check (3), i.e., $\mathbb{P}_\phi(G'_m[\phi]) \le 2^{-m}$ in the current notation, notice that the function $U^\phi : \Omega^\diamond \to [0,\infty]$ defined by

$$U^\phi(y_1,\ldots,y_n) := U(\phi(\Box), y_1, \phi(y_1), y_2, \ldots, \phi(y_1,\ldots,y_{n-1}), y_n)$$

is a martingale w.r.t. to ϕ. Now Ville's inequality implies

$$\begin{aligned}\mathbb{P}_\phi(G'_m[\phi]) = \mathbb{P}_\phi \{\omega \in \Omega \mid (\omega,\phi) \in G'_m\} = \mathbb{P}_\phi \{\omega \in \Omega \mid (\phi(\omega),\omega) \in G_m\} \\ = \mathbb{P}_\phi \left\{\omega \in \Omega \mid \exists n : U^\phi(\omega^n) > 2^m\right\} \le 2^{-m}, \quad \forall \phi \in \Phi.\end{aligned}$$

Suppose π is not only unpredictable but also typical. Then there exists $\phi \in \Phi$ such that $\pi = (\phi(\omega),\omega)$ for some ω typical w.r.t. to ϕ. Since $\pi \in G_m$, we have $(\omega,\phi) \in G'_m$; since this is true for each $m \in \mathbb{N}$, ω is not typical w.r.t. to ϕ, and so we have arrived at a contradiction.

Corollary 1

To see that a sequence $\omega = (y_1, y_2, \ldots) \in \Omega$ is random w.r.t. to $\mathbb{P}_\phi$ in the sense of Martin-Löf if and only if $\pi := (p_1, y_1, p_2, y_2, \ldots)$ is unpredictable, where $p_n := \phi(y_1,\ldots,y_{n-1})$, $n \in \mathbb{N}$, notice that the unpredictability of π is equivalent to Schnorr and Levin's reformulation of Martin-Löf randomness of ω w.r.t. to $\mathbb{P}_\phi$ (i.e., to the universal lower semicomputable supermartingale w.r.t. to ϕ being bounded on ω^n).

Part "only if" of Theorem 1

Let $G_m = \cup\{U_k \mid (m,k) \in A\}$ be a representation of the universal test of typicalness via basic sets, with $A \subseteq \mathbb{N}^2$ a recursively enumerable set. Without loss of generality we can assume that each basic set U_k in this representation has the form $\Gamma_c \times \left\{\phi \in \Phi \mid a(x) < \phi(x) < b(x), \forall x \in \Omega^{\le n}\right\}$ for some $c \in \Omega^n$, $a, b : \Omega^{\le n} \to \mathbb{Q}$, and $n \in \mathbb{N}$. Define G'_m to be the set of all $(p_1, y_1, p_2, y_2, \ldots) \in \Pi$ such that $((y_1, y_2, \ldots), \phi) \in G_m$ for all ϕ that agree with $(p_1, y_1, p_2, y_2, \ldots)$.

The compactness of Φ easily implies that each set $G'_m \subseteq \Pi$ is open. Indeed, suppose $\pi = (p_1, y_1, p_2, y_2, \ldots) \in G'_m$. For each $\phi \in \Phi$, either ϕ disagrees with π or $((y_1, y_2, \ldots), \phi) \in G_m$. In both cases there is a neighbourhood O'_ϕ of π and a neighbourhood O''_ϕ of ϕ such that either all elements of O''_ϕ disagree with all elements of O'_ϕ or $((y'_1, y'_2, \ldots), \phi') \in G_m$ for all $(p'_1, y'_1, p'_2, y'_2, \ldots) \in O'_\phi$ and all $\phi' \in O''_\phi$. Since Φ is compact, there is a finite set $\phi_1, \ldots, \phi_J$ such that $\cup_{j=1}^J O''_{\phi_j} = \Phi$. We can see that the neighbourhood $\cap_{j=1}^J O'_{\phi_j}$ of π is a subset of G'_m.

The same argument shows that the G'_m form a computable sequence of open sets. Let us show that there exists a non-negative superfarthingale V_m with initial value 2^{-m} or less that eventually exceeds 1 on each sequence in G'_m. (In this sense G'_m form a prequential test of typicalness.)

Let $G'_m = \cup\{U_k \mid (m,k) \in A\}$ be a representation of G'_m via basic sets, where $A \subseteq \mathbb{N}^2$ is a recursively enumerable set. Let $A = \cup_{i=1}^{\infty} A_i$ be a representation of A as the union of a computable sequence $\emptyset \subset A_1 \subseteq A_2 \subseteq \cdots$ of finite sets. Fix an m. For each $i \in \mathbb{N}$, define a superfarthingale W_i as follows. Let N be so large that, for all $x \in \Pi^N$ and $(m,k) \in A_i$, either $\Gamma_x \subseteq U_k$ or $\Gamma_x \cap U_k = \emptyset$. (For example, we can set N to the largest n_k in (14) over k such that $(m,k) \in A_i$.) For $n \geq N$ and $x \in \Pi^n$, set

$$W_i(x) := \begin{cases} 1 & \text{if } \Gamma_x \subseteq U_k \text{ for some } k \text{ with } (m,k) \in A_i \\ 0 & \text{otherwise.} \end{cases}$$

After that proceed by backward induction. If $W_i(x)$ is already defined for $x \in \Pi^n$, $n = N, N-1, \ldots, 1$, set, for each $x \in \Pi^{n-1}$,

$$W_i(x) := \sup_{p \in [0,1]} \bigl((1-p)W_i(x,p,0) + pW_i(x,p,1)\bigr). \tag{8}$$

It is clear that W_i is a superfarthingale that does not depend on the choice of N.

We will need to establish several properties of W_i. First, it is lower semicontinuous. Indeed, there is an N (e.g., the largest n_k in (14) over $(m,k) \in A_i$) such that $W_i(x)$ is lower semicontinuous when x is restricted to Π^n with $n \geq N$. (It will even be lower semicomputable when x is restricted to $\Pi^{\geq N}$.) And the operation sup preserves lower semicontinuity:

Lemma 4. *If a function $f : X \times Y \to \mathbb{R}$ defined on the product of topological spaces X and Y is lower semicontinuous, then the function $x \in X \mapsto g(x) := \sup_{y \in Y} f(x,y)$ is also lower semicontinuous.*

Proof. It suffices to notice that, for each $c \in \mathbb{R}$, $\{x \mid g(x) > c\} = \{x \mid \exists y : f(x,y) > c\}$, and projections of open sets are open.

The lower semicontinuity of W_i implies its lower semicomputability: indeed, we can restrict p to $\mathbb{Q} \cap [0,1]$ in (8).

Let us check that $W_i(\Box) \leq 2^{-m}$. Suppose that, on the contrary, $W_i(\Box) > 2^{-m}$. Construct a forecasting system ϕ as follows. For each $x \in \Omega^n$, $n = 0, 1, \ldots, N-1$, choose $\phi(x)$ such that

$$\begin{aligned}(1-\phi(x))&W_i(x,\phi(x),0) + \phi(x)W_i(x,\phi(x),1) \\ &\geq \sup_{p \in [0,1]} \bigl((1-p)W_i(x,p,0) + pW_i(x,p,1)\bigr) - \epsilon/N = W_i(x) - \epsilon/N,\end{aligned}$$

where $\epsilon > 0$ satisfies $W_i(\Box) > 2^{-m} + \epsilon$. For each $x \in \Omega^{\geq N}$, define $\phi(x)$ arbitrarily, say $\phi(x) := 0$. Since $(\phi(\omega), \omega) \notin G'_m$ for all $\omega \notin G_m[\phi]$, we have $W_i^{\phi}(\omega^N) = 0$ for all $\omega \notin G_m[\phi]$. Combining the fact that

$$\left\{\omega \mid W_i^{\phi}(\omega^N) = 1\right\} \subseteq G_m[\phi]$$

with the fact that the function $x \in \Omega^\diamond \mapsto S(x) := W_i^\phi(x) + \epsilon n/N$, where n is the length of x, is a submartingale w.r. to ϕ, we obtain

$$\mathbb{P}_\phi(G_m[\phi]) \ge \mathbb{P}_\phi\left\{\omega \mid W_i^\phi(\omega^N) = 1\right\} = \mathbb{E}_\phi W_i^\phi(\omega^N) = \mathbb{E}_\phi(S(\omega^N) - \epsilon) \\ \ge S(\square) - \epsilon = W_i^\phi(\square) - \epsilon > 2^{-m}, \quad (9)$$

where $\mathbb{E}_\phi$ stands for the expectation of a function of $\omega \in \Omega$ w.r. to $\mathbb{P}_\phi$. The inequality between the extreme terms of (9) fails by the definition of a test of typicalness.

Define $V_m(x) := \sup_i W_i(x)$, $x \in \Pi^\diamond$, to be the limit of the non-decreasing sequence of superfarthingales W_i. It is clear that V_m is also a superfarthingale and $V_m(\square) \le 2^{-m}$. Set $V := \sum_{m=1}^\infty V_m$; this is a lower semicomputable superfarthingale with initial value $V(\square) \le 1$ (so that $V \in \mathcal{V}$ if we redefine $V(\square) := 1$).

Now it is easy to finish the proof of the theorem. Suppose that π is not typical. Then $\pi \in G'_m$ for all $m \in \mathbb{N}$. Then $V(\pi^n) \to \infty$ as $n \to \infty$, and so π is not unpredictable.

6 Conclusion

In this section we will discuss some open problems and directions of further research (for further open problems and background information, see [17]).

Quantitative prequential randomness

Let U be the universal superfarthingale. Let us denote its binary logarithm $\log U(p_1, y_1, \ldots, p_n, y_n)$ by $\mathrm{DU}(p_1, y_1, \ldots, p_n, y_n)$ and call it the *deficiency of unpredictability* of $(p_1, y_1, \ldots, p_n, y_n)$. For $\pi \in \Pi$, set $\mathrm{DU}(\pi) := \sup_n \mathrm{DU}(\pi^n)$.

Let T be the universal test of typicalness. Let us denote its value $T(\omega, \phi)$ by $\mathrm{DT}(\omega, \phi)$ and call it the *deficiency of typicalness*. The deficiency of typicalness $\mathrm{DT}(\pi)$ of $\pi \in \Pi$ can now be defined as $\inf_\phi \mathrm{DT}(\omega, \phi)$, ϕ ranging over the forecasting systems that agree with π. We can further define $\mathrm{DT}(p_1, y_1, \ldots, p_n, y_n)$ as the infimum of $\mathrm{DT}(\pi)$, $\pi \in \Pi$ ranging over all continuations of $(p_1, y_1, \ldots, p_n, y_n)$.

In this paper we have demonstrated the equivalence of unpredictability and typicalness in the prequential setting: $\mathrm{DU}(\pi) < \infty$ if and only if $\mathrm{DT}(\pi) < \infty$. A natural next step is to study inequalities between the deficiency of unpredictability and the deficiency of typicalness. The arguments of this paper show that $\mathrm{DU} \approx \mathrm{DT}$. It would be interesting to establish optimal explicit inequalities.

Stripped algorithmic theory of randomness

The non-algorithmic counterpart of the notion of randomness is Cournot's principle (see, e.g., [20]): a data sequence is not random if it belongs to a pre-specified event of small probability. Therefore, the non-algorithmic counterpart of equivalence of various notions of randomness is the closeness of various notions of probability. As the notion of randomness branches into unpredictability and typicalness, the notion of probability branches into game-theoretic probability and measure-theoretic probability. Unfortunately, the equivalence of unpredictability and typicalness does not translate automatically into equivalence between game-theoretic and measure-theoretic probability. In this subsection we will give some definitions, and in the next will state an open problem.

Measure-theoretic probability, as formalized by Kolmogorov [11], is standard. The game-theoretic approach to probability is as old as measure-theoretic (see, e.g., von Mises [16] and Ville [22]) but game-theoretic probability was formalized only recently [23, 7, 21]. Game-theoretic probability can be introduced as either upper or lower probability; its natural home is the prequential framework.

If E is a *prequential event* (i.e., a subset of Π), the *upper probability* of E is

$$\overline{\mathbb{P}}(E) := \inf \left\{ \epsilon : \exists V : V(\square) = \epsilon \text{ and } \forall \pi \in \Pi : \limsup_n V(\pi^n) \geq 1 \right\}, \qquad (10)$$

where V ranges over the non-negative farthingales. It is clear that nothing changes if lim sup is replaced by sup or lim inf (we can always stop when 1 is reached) and/or if we allow V to range over the non-negative superfarthingales. The *lower probability* of E is defined as

$$\underline{\mathbb{P}}(E) := 1 - \overline{\mathbb{P}}(E^c),$$

where E^c is the complement of E. The *exact probability* of E exists if $\overline{\mathbb{P}}(E) = \underline{\mathbb{P}}(E)$ and is equal to this common value.

Open problem

Consider the situation where the forecaster's prediction y_n is restricted to a finite set $F_n \subseteq [0, 1]$ (the most interesting case is where F_n is a grid of points in $[0, 1]$ becoming more and more dense as $n \to \infty$; alternatively, we could consider $F_n := \mathbb{Q} \cap [0, 1]$ instead of finite F_n). The *sample space*, denoted

$$\Pi := \prod_{n=1}^{\infty} (F_n \times \{0, 1\}), \qquad (11)$$

is the set of all possible sequences of predictions and actual outcomes.

For each set $E \subseteq \Pi$, we can define its upper probability $\overline{\mathbb{P}}(E)$ as before, by (10). We can also give the following "dual" definition, in the spirit of [10], Section 10.2. For each forecasting system ϕ, let $\mathbb{P}^{\phi}$ be the probability distribution on Π corresponding to the following process: p_1 is chosen according to ϕ ($p_1 := \phi(\square)$), then y_1 is chosen according to the Bernoulli distribution with parameter p_1, then p_2 is chosen according to ϕ ($p_2 := \phi(y_1)$), then y_2 is chosen according to the Bernoulli distribution with parameter p_2, etc. Set

$$\mathbb{P}^{*}(E) := \sup_{\phi} \mathbb{P}^{\phi}(E).$$

We are mostly interested in the case of a Borel E, in which case the notation $\mathbb{P}^{\phi}(E)$ is unambiguous. In general, $\mathbb{P}^{\phi}(E)$ can be understood to be the outer measure of E.

Question 1. Suppose E is Borel. Is it true that $\overline{\mathbb{P}}(E) = \mathbb{P}^{*}(E)$?

The argument of this paper shows that for open E the answer is "yes". As a next step, it would be interesting to find the answer for closed E and other E in the low classes of the Borel hierarchy (such as Σ_2^0 and Π_2^0).

Natural extensions of Question 1 are:

- What is the answer when the upper probabilities are replaced by upper expectations?
- Can anything be said for $F_n = [0, 1]$, $\forall n$? (Because of measurability issues, this question might be less clear-cut and, therefore, less interesting than the case of finite F_n: cf. [21], pp. 168–169.)

Randomized forecasting systems

In this paper we only considered deterministic forecasting systems. It would be interesting to extend Theorem 1 and its future quantitative and non-algorithmic versions (as discussed in the previous subsections) to the case of randomized forecasting systems.

Acknowledgements

The main result of this paper is motivated by Question 1, which has been asked independently by several people, including Glenn Shafer and, more recently, Sasha Shen. Shen's version, where F_n (cf. (11)) are finite, is especially interesting; I am grateful to him for reviving my interest in this problem. This work was partially supported by EPSRC (grant EP/F002998/1).

References

1. Laurent Bienvenu. *Caractérisations de l'aléatoire par les jeux: impredictibilité et stochasticité*. PhD thesis, Université de Provence, 2008.
2. Patrick Billingsley. *Convergence of Probability Measures*. Wiley, New York, 1968.
3. A. Philip Dawid. Statistical theory: the prequential approach. *Journal of the Royal Statistical Society* A, 147:278–292, 1984.
4. A. Philip Dawid. Calibration-based empirical probability (with discussion). *Annals of Statistics*, 13:1251–1285, 1985.
5. A. Philip Dawid. Fisherian inference in likelihood and prequential frames of reference (with discussion). *Journal of the Royal Statistical Society* B, 53:79–109, 1991.
6. A. Philip Dawid. Prequential analysis. In Samuel Kotz, Campbell B. Read, and David L. Banks, editors, *Encyclopedia of Statistical Sciences*, volume 1 (update), pages 464–470. Wiley, New York, 1997.

7. A. Philip Dawid and Vladimir Vovk. Prequential probability: principles and properties. *Bernoulli*, 5:125–162, 1999.
8. Ryszard Engelking. *General Topology.* Heldermann, Berlin, second edition, 1989.
9. Peter Gács. Uniform test of algorithmic randomness over a general space. *Theoretical Computer Science*, 341:91–137, 2005.
10. Peter J. Huber. *Robust Statistics.* Wiley, New York, 1981.
11. Andrei N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer, Berlin, 1933. English translation: *Foundations of the Theory of Probability.* Chelsea, New York, 1950.
12. Leonid A. Levin. On the notion of a random sequence. *Soviet Mathematics Doklady*, 14:1413–1416, 1973.
13. Leonid A. Levin. Uniform tests of randomness. *Soviet Mathematics Doklady*, 17:337–340, 1976. The Russian original in: Доклады АН СССР 227(1), 1976.
14. Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.
15. Per Martin-Löf. *Notes on Constructive Mathematics.* Almqvist & Wiksell, Stockholm, 1970.
16. Richard von Mises. Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 5:52–99, 1919.
17. On-line prediction wiki, http://onlineprediction.net.
18. Judea Pearl. Causal diagrams for empirical research (with discussion). *Biometrika*, 82:669–710, 1995.
19. Claus P. Schnorr. *Zufälligkeit und Wahrscheinlichkeit.* Springer, Berlin, 1971.
20. Glenn Shafer. From Cournot's principle to market efficiency. In Jean-Philippe Touffut, editor, *Augustin Cournot: Modelling Economics*, pages 55–95. Edward Elgar, 2007. Can be downloaded from http://probabilityandfinance.com, Working Paper 15.
21. Glenn Shafer and Vladimir Vovk. *Probability and Finance: It's Only a Game!* Wiley, New York, 2001.
22. Jean Ville. *Etude critique de la notion de collectif.* Gauthier-Villars, Paris, 1939.
23. Vladimir Vovk. A logic of probability, with application to the foundations of statistics (with discussion). *Journal of the Royal Statistical Society* B, 55:317–351, 1993.

A Effective topology

In this section we will give definitions of various notions connected with computability in topological spaces, mainly following Martin-Löf [15] (see also [9], Appendix C.2). The details of the definitions become important only in the proofs. We will use the terminology of Engelking [8].

In this appendix and in some proofs in the main part of the paper we will be using the following notation for $n \in \mathbb{N}$: $\Omega^n := \{0,1\}^n$ is the set of all finite binary sequences of length n; $\Omega^{\le n}$ is the set of all finite binary sequences of length at most n; $\Pi^n := ([0,1] \times \{0,1\})^n$; $\Pi^{\ge n} := \cup_{i=n}^{\infty}([0,1] \times \{0,1\})^i$.

An *effective topological space* is a second-countable topological space with a fixed numbering $(U_k)_{k=1}^{\infty}$ of its countable base. In other words, an effective topological space is a triple $(X, \mathcal{O}, (U_k)_{k=1}^{\infty})$, where $(X, \mathcal{O})$ is a topological space and $(U_k)_{k=1}^{\infty}$ is a numbering of its countable base. The family $(U_k)_{k=1}^{\infty}$ is called the *effective base* of the effective topological space, and its elements are called *basic sets*. Finite unions of basic sets are called *simple sets*. We do not distinguish between two effective topological spaces $(X, \mathcal{O}, (U_k)_{k=1}^\infty)$ and $(X', \mathcal{O}', (U'_k)_{k=1}^\infty)$ if $(X, \mathcal{O}) = (X', \mathcal{O}')$ and there exists a computable bijection $f : \mathbb{N} \to \mathbb{N}$ such that $U'_k = U_{f(k)}$ for all k.

Example 1 ($\mathbb{N}$). The usual discrete topology on $\mathbb{N}$ has as its base the set of all singletons $\{k\}$, $k \in \mathbb{N}$. They can serve as the effective base, $U_k := \{k\}$.

Example 2 ($\overline{\mathbb{N}}$). The effective base of $\overline{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$ consists of both the singletons $\{k\}$ and the sets $\{k, k+1, \ldots, \infty\}$. Set $U_{2k-1} := \{k\}$ and $U_{2k} := \{k, k+1, \ldots, \infty\}$, $k \in \mathbb{N}$.

Example 3 ($\mathbb{R}$). The topology on $\mathbb{R}$ has as its base the set of all intervals (a, b), $a < b$. To make $\mathbb{R}$ into an effective topological space, fix a computable enumeration (a_k, b_k), $k = 1, 2, \ldots$, of all intervals with rational end-points, and take $U_k := (a_k, b_k)$ as the effective base.

Example 4 (Ω). The topology on $\Omega := \{0, 1\}^\infty$ is the usual product topology, which makes Ω a compact topological space. To make it into an effective topological space, fix a computable bijection $f : \mathbb{N} \to \Omega^\diamond$ and take $U_k := \Gamma_{f(k)}$ as the effective base.

Example 5 (Φ). The basic sets in Φ (the set of all forecasting systems) have the form

$$\left\{\phi \in \Phi \mid a(x) < \phi(x) < b(x), \forall x \in \Omega^{\leq n}\right\} \tag{12}$$

for some $n \in \mathbb{N}$ and $a, b : \Omega^{\leq n} \to \mathbb{Q}$. Let (n_k, a_k, b_k), $k = 1, 2, \ldots$, be a computable enumeration of all such triples (n, a, b). Set U_k to (12) with $(n, a, b) := (n_k, a_k, b_k)$.

Example 6 (Π). The topology on the prequential space Π is the standard product topology of $[0, 1] \times \{0, 1\} \times [0, 1] \times \{0, 1\} \times \cdots$. The basic sets are

$$\left\{(p_1, y_1, p_2, y_2, \ldots) \in \Pi \mid a_1 < p_1 < b_1, y_1 = c_1, \ldots, a_n < p_n < b_n, y_n = c_n\right\} \tag{13}$$

where n ranges over $\mathbb{N}$, $a_i, b_i \in \mathbb{Q}$, and $c_i \in \{0, 1\}$, $i = 1, \ldots, n$. Let

$$(n_k, a_{1,k}, b_{1,k}, c_{1,k}, \ldots, a_{n_k,k}, b_{n_k,k}, c_{n_k,k}) \tag{14}$$

be a computable enumeration of all such sequences $(n, a_1, b_1, c_1, \ldots, a_n, b_n, c_n)$. We can define U_k as (13) with (14) in place of $(n, a_1, b_1, c_1, \ldots, a_n, b_n, c_n)$.

Let X' and X'' be two effective topological spaces with effective bases $(U'_k)_{k=1}^\infty$ and $(U''_k)_{k=1}^\infty$, respectively. The Cartesian product of X' and X'' is the product of the topological spaces X' and X'' equipped with the effective base $(U_k)_{k=1}^\infty$, where $U_{f(k',k'')} := U'_{k'} \times U''_{k''}$ and $f : \mathbb{N}^2 \to \mathbb{N}$ is a fixed computable bijection. We will be particularly interested in the product $\Omega \times \Phi$; sometimes we will need products of more than two spaces, such as $\Omega \times \Phi \times \mathbb{R} := (\Omega \times \Phi) \times \mathbb{R}$.

Let X be an effective topological space with effective base $(U_k)_{k=1}^{\infty}$. As described in the previous paragraph, we define the structure of an effective topological space on the power set X^n, $n \in \mathbb{N}$; let the effective base in X^n be $(U_k^n)_{k=1}^{\infty}$. For $n = 0$, X^0 is the trivial one-element effective topological space with all $U_k = X^0$, $k \in \mathbb{N}$. The set X^* of all finite sequences of elements of X is equipped with the topology of the direct sum of X^n, $n \geq 0$. An effective base in it can be defined by $U_{f(n,k)} := U_k^n$, where $f : (\mathbb{N} \cup \{0\}) \times \mathbb{N} \to \mathbb{N}$ is a computable bijection.

Let X be a fixed effective topological space with effective base $(U_k)_{k=1}^{\infty}$. An open set $G \subseteq X$ is said to be *effectively open* if it can be represented in the form $\cup\{U_k \mid k \in A\}$ for a recursively enumerable set $A \subseteq \mathbb{N}$. For any effectively open set G we only consider its representations $\cup\{U_k \mid k \in A\}$ such that

$$\overline{U_k} \subseteq G; \tag{15}$$

it is clear that this can be done without loss of generality. A *computable sequence of open sets* is a sequence of open sets $G_1, G_2, \ldots$ such that there exists a recursively enumerable set $A \subseteq \mathbb{N}^2$ satisfying $G_m = \cup\{U_k \mid (m,k) \in A\}$ for all $m \in \mathbb{N}$. A *computable family of sequences of open sets* is a family $(G_{l,m})$, $l, m \in \mathbb{N}$, of sequences of open sets such that there exists a recursively enumerable set $A \subseteq \mathbb{N}^3$ satisfying $G_{l,m} = \cup\{U_k \mid (l,m,k) \in A\}$ for all l, m. The existence of a universal Turing machine immediately implies

Lemma 5. *There exists a computable family $(G_{l,m})$ of sequences of open sets such that for any computable sequence G'_m of open sets there exists $l \in \mathbb{N}$ such that $G'_m = G_{l,m}$ for all $m \in \mathbb{N}$.*

Any computable family of sequences of open sets satisfying the condition in Lemma 5 will be called a *universal computable family of sequences of open sets.*

A function $f : X \to \mathbb{R}$ is called *lower semicomputable* if the set $\{(x,r) \mid x \in X, r \in \mathbb{R}, f(x) > r\}$ is effectively open in $X \times \mathbb{R}$. Similarly, a function $f : X \to \overline{\mathbb{N}}$ is *lower semicomputable* if the set $\{(x,r) \mid x \in X, r \in \mathbb{N}, f(x) \geq r\}$ is effectively open in $X \times \mathbb{N}$. A sequence $f_1, f_2, \ldots$ of lower semicomputable functions $f_l : X \to \mathbb{R}$ is called *computable* if the set $\{(l,x,r) \mid x \in X, r \in \mathbb{R}, f_l(x) > r\}$ is effectively open in $\mathbb{N} \times X \times \mathbb{R}$. The existence of a universal Turing machine also implies

Lemma 6. *There exists a computable sequence $f_1, f_2, \ldots$ of lower semicomputable functions that contains every lower semicomputable function.*

Any computable sequence of lower semicomputable functions satisfying the condition in Lemma 6 will be called a *universal computable sequence of lower semicomputable functions.*

A function $f : X \to \mathbb{R}$ is called *computable* if both f and $-f$ are lower semicomputable. It is easy to see that the analogue of Lemma 6 does not hold for computable functions.

Weak topology

In this subsection we will establish a connection between the topology of Example 5 and the weak topology on the set of probability measures on Ω ([2], Appendix III; [9], Appendix C.2).

A *computable numbering* $(V_k)_{k=1}^{\infty}$ of the family of all simple sets in an effective topological space with effective base $(U_k)_{k=1}^{\infty}$ is defined as

$$V_{f(k_1,\dots,k_n)} := U_{k_1} \cup \dots \cup U_{k_n},$$

where $f : \mathbb{N}^* \to \mathbb{N}$ is a computable bijection and $(k_1, \dots, k_n)$ ranges over $\mathbb{N}^*$.

Lemma 7. *Let $(V_k)_{k=1}^{\infty}$ be a computable numbering of the simple sets in Ω. The sequence of functions $\phi \in \Phi \mapsto \mathbb{P}_{\phi}(V_k)$, $k = 1, 2, \dots$, is a computable sequence of lower semicomputable functions.*

Proof. Suppose that $\mathbb{P}_{\phi}(V_k) > r$ for some $\phi \in \Phi$, $k \in \mathbb{N}$, and $r \in \mathbb{R}$. We are required to show that there is a computable way to eventually find basic neighbourhoods of ϕ and r such that $\mathbb{P}_{\phi'}(V_k) > r'$ holds for all ϕ' and r' in the neighbourhoods. The last statement follows from the computability of the basic arithmetic operations ($+$ and $\times$). □

Since in the space Ω the complement of each simple set is again simple, we have the following corollary.

Corollary 2. *Let $(V_k)_{k=1}^{\infty}$ be a computable numbering of the simple sets in Ω. The function $(\phi, k) \mapsto \mathbb{P}_{\phi}(V_k)$ is a computable function on $\Phi \times \mathbb{N}$.*