Extraction Rates of Almost Total Functionals

Douglas Cenzer
University of Florida

Joint work with Christopher Porter

CCA 2020
Bologna, ITALY
9 September 2020

Goals

1. Develop the notion of extraction rate for almost total functionals
2. Consider the canonical representation of a continuous functional
3. Study von Neumann and Levin-Kautz extraction of randoms.
4. Make connections with random number generation by Knuth and Yao
5. Calculate the extraction rate for random functionals

History

Von Neumann gave a procedure to extract an unbiased random sequence from a biased (Bernoulli) random sequence

This procedure requires two bits of input to compute one bit of output.

Levin and Kautz improved this to optimal extraction

This is related to the entropy $h(\mu)$ of a measure

Barmpalias, Brodhead, Cenzer et al developed the notion of algorithmic randomness for closed sets and continuous functionals

These functionals have natural extraction rates.

Outline

- Representation of continuous and Turing functionals

  μ-Almost total functionals

  Canonical representations

  Nowhere constant functionals

- Extraction and input/output rates

  Average rate

- Rate of block functionals

- Rate of functionals induced by DDG-trees

- Rate of Levin-Kautz conversion

- Random functionals

Continuous Functionals

A continuous functional $F : 2^\omega \to 2^\omega$ may be represented by a function $f : \{0,1\}^* \to \{0,1\}^*$ such that the following hold for all $\sigma \in \{0,1\}^*$.

(1) $\sigma_1 \sqsubset \sigma_2$ implies $f(\sigma_1) \sqsubseteq f(\sigma_2)$;

(3) For every n, there exists m such that for all $\sigma \in \{0,1\}^m$, $|f(\sigma)| \geq n$;

(3) For all $X \in 2^\omega$, $F(X) = \bigcup_n f(X\lceil n)$.

If f only satisfies (1), then f may represent a partial functional F where $F(X)$ may be a finite string.
Let $dom(\Phi) = \{X : \Phi(X) \in 2^\omega\}$.

Φ is said to be a total or partial Turing functional if f is computable

Almost Total Functionals

A partial functional $\Phi : 2^\omega \to 2^\omega$ is *μ-almost total* if $\mu(dom(\Phi)) = 1$.

$\Phi^{-1}(\tau) = \{\sigma \in 2^{<\omega} : \tau \preceq \varphi(\sigma)\ \&\ (\forall \sigma' \prec \sigma)\ \tau \not\preceq \varphi(\sigma')\}$

For $S \subseteq 2^{<\omega}$, $\Phi^{-1}(S) = \bigcup_{\tau \in S} \Phi^{-1}(\tau)$.

When $\mathcal{A}$ is a subset of 2^ω, we denote by $\Phi^{-1}(\mathcal{A})$ the set $\{X \in \mathrm{dom}(\Phi) : \Phi(X) \in \mathcal{A}\}$.

Note in particular that $\Phi^{-1}([\![\tau]\!]) = [\![\Phi^{-1}(\tau)]\!] \cap \mathrm{dom}(\Phi)$.

Lemma

A Turing functional Φ is μ-almost total if and only if $\mathrm{MLR}_\mu \subseteq dom(\Phi)$.

Canonical representation

For the *canonical representation*, $\varphi(\sigma)$ is be the longest common initial segment of all members of $\{\Phi(X) : \sigma \prec X\}$.

Φ is *nowhere constant* if for any string σ, if $[\![\sigma]\!] \subset \mathit{dom}(\Phi)$, then Φ is not constant on $[\![\sigma]\!]$.

Proposition

If Φ is a total, nowhere constant Turing functional, then the canonical representation φ of Φ is computable.

In general, the canonical representation of a partial computable functional is computable in $\emptyset'$ and need not be computable.

Extraction rates

Let $u_\varphi(X, n)$ be the least m such that $|\varphi(X\restriction m)| \geq n$.

The *extraction rate* of $Y = \Phi(X)$ from X is $\frac{n}{u_\varphi(X,n)}$

Alternatively, the *φ-output/input ratio* is $OI_\varphi(\sigma) = \frac{|\varphi(\sigma)|}{|\sigma|}$

Lemma

For any partial (computable) Turing functional Φ with representation φ and any X such that $\Phi(X) = Y$ is defined,

$$\lim_{n\to\infty} \frac{|\varphi(X\restriction n)|}{n} = \lim_{m\to\infty} \frac{m}{u_\varphi(X, m)},$$

provided that both limits exists.

Let $OI_\varphi(X)$ for $\lim_{n\to\infty} OI_\varphi(X\restriction n)$ if this limit exists.

Average rate

Let μ be a measure on 2^ω; $\mu(\sigma)$ denotes $\mu([\![\sigma]\!])$.

The *average* Φ*-output/input ratio* for strings of length n is

$$Avg(\Phi, \mu, n) = \sum_{\sigma \in 2^n} \mu(\sigma) OI_\Phi(\sigma) = \frac{1}{n} \sum_{\sigma \in 2^n} \mu(\sigma) |\varphi(\sigma)|,$$

where φ is the canonical representation of Φ.
This is also the μ-average value of $F_n(X) = OI_\Phi(X \upharpoonright n)$ for $X \in 2^\omega$

$$Avg(\Phi, \mu, n) = \int_{2^\omega} F_n(X) d\mu(X).$$

Extraction rate of a functional

The μ-extraction rate of Φ is

$$Rate(\Phi, \mu) = \limsup_{n\to\infty} Avg(\Phi, \mu, n).$$

Lemma

Suppose for some $c \in \omega$ and all $\sigma \in 2^{<\omega}$, $|\varphi(\sigma)| \leq c|\sigma|$ and that

$$\lim_{n\to\infty} \frac{|\varphi(X\restriction n)|}{n} = r$$

for μ-almost every $X \in 2^\omega$. Then $Rate(\Phi, \mu) = r$.

We will present several examples of functionals Φ for which

(i) $\lim_{n\to\infty} Avg(\Phi, \mu, n)$ exists (for an appropriate μ), and

(ii) $OI_\Phi(X) = \lim_{n\to\infty} OI_\Phi(X\restriction n) = Rate(\Phi, \mu)$ for all sufficiently μ-random X.

Block functionals

φ is an *n-block map* if for any string $\sigma = \sigma_1 ^\frown \ldots ^\frown \sigma_k ^\frown \tau$, where $|\sigma_i| = n$ for $i = 1, \ldots, k$ and $|\tau| < k$,

$$\varphi(\sigma) = \varphi(\sigma_1) ^\frown \ldots ^\frown \varphi(\sigma_k).$$

φ is *non-trivial* if $|\varphi(\sigma)| > 0$ for some $\sigma \in 2^n$.

Φ is an *n-block functional* if it has an n-block representation.

Block maps have been studied by Elias, Peres,Pae-Il and others

von Neumann's extractor Φ is a 2-block functional defined by $\varphi(01) = 0$; $\varphi(10) = 1$; $\varphi(00) = \epsilon = \varphi(11)$

Bernoulli measures

The Bernoulli measure μ_p, with $0 < p < 1$, on 2^ω is given by $\mu(\sigma) = p^t(1-p)^{n-t}$, where $t = |\{i : \sigma(i) = 0\}|$, where $|\sigma| = n$

An *n-step Bernoulli measure* is a Bernoulli measure on $(2^n)^\omega$ is similarly induced by the measures $\mu(\sigma)$ for σ of length n, so that if $\sigma = \sigma_1 ^\frown \ldots ^\frown \sigma_k$ with each $|\sigma_i| = n$,

$$\mu(\sigma) = \mu(\sigma_1) \cdot \ldots \cdot \mu(\sigma_k)$$

μ is *positive* if $\mu(\sigma) > 0$ for all $\sigma \in \{0,1\}^n$.

This induces a measure μ on 2^ω

Proposition

Suppose μ is a positive n-step Bernoulli measure on 2^ω and Φ is a non-trivial n-block functional Φ. Then Φ is μ-almost total.

Example of von Neumann extraction

Recall $\varphi(00) = (0)$ and $\varphi(10) = 1$; $\varphi(00) = \varphi(11) = \epsilon$

This is in fact the canonical representation

(01) and (10) each occur with probability $p(1-p)$ for the Bernoulli measure μ_p

Let $A(n) = \sum_{\sigma \in 2^n} \mu(\sigma)\varphi(\sigma)/n$

Then $A(2) = 2p(1-p) \cdot 1/2 = p(1-p)$.

In general, $A(2n) = p(1-p)$ and $A(2n+1) = \frac{2n}{2n+1} p(1-p)$.

Thus $Rate(\varphi, \mu_p) = p(1-p)$.

Average rates for n-step Bernoulli functionals

Theorem

Let μ be a positive n-step Bernoulli measure and Φ a non-trivial n-block functional. Then $Rate(\Phi, \mu) = Avg(\Phi, \mu, n)$

Sketch: $Avg(\Phi, \mu, n) = \sum_{\sigma \in 2^n} \mu(\sigma)|\varphi(\sigma)|$
is the expected length of $\varphi(\sigma)$

Since blocks are independent,
$Avg(\Phi, \mu, nk) = \sum_{\sigma \in 2^{nk}} \mu(\sigma)|\varphi(\sigma)| = Avg(\Phi, \mu, n)$
so

$$Rate(\Phi, \mu) = \lim_{k \to \infty} Avg(\Phi, \mu, nk) = Avg(\Phi, \mu, n).$$

Theorem on n-block Functionals

Theorem

Let μ be a computable, positive n-step Bernoulli measure, and let $X \in 2^\omega$ be . Then for every non-trivial n-block functional Φ and every μ-Schnorr random X

$$\lim_{n\to\infty} \frac{|\varphi(X\upharpoonright n)|}{n} = \mathit{Rate}(\Phi, \mu).$$

Here is an outline of the proof

Ergodic Transformations

Lemma

Let μ be a measure on 2^ω and let $T : 2^\omega \to 2^\omega$ be μ-invariant. Then T is ergodic if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}[\![\sigma]\!] \cap [\![\tau]\!]) = \mu(\sigma)\mu(\tau)$$

for all $\sigma, \tau \in 2^{<\omega}$.

Lemma

The n-shift on 2^ω is ergodic with respect to an n-step Bernoulli measure.

Effective Birkhoff's Ergodic Theorem

Theorem (Franklin-Towsner 2014)

Let μ be a computable measure on 2^ω and let $T : 2^\omega \to 2^\omega$ be a computable, μ-invariant, ergodic transformation. Then for any bounded computable function Φ and any μ-Schnorr random $X \in 2^\omega$,

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} \Phi(T^i(X)) = \int F d\mu.$$

Let μ be an n-step Bernoulli measure and let T be the n-shift.

Define $F(X) = \dfrac{|\varphi(X\upharpoonright n)|}{n}$. Then

$$\int F d\mu = \sum_{\sigma\in 2^n} \mu(\sigma)\varphi(\sigma) = \sum_{\sigma\in 2^n} \mu(\sigma)\frac{|\varphi(\sigma)|}{n} = Avg(F,\mu,n) = Rate(F,\mu)$$

Proof Continued

Now let X be μ-Schnorr random. Then

$$\frac{1}{k}\sum_{i=0}^{k-1} F(T^i(X)) = \frac{1}{k}\sum_{i=0}^{k-1} \frac{|\varphi(T^i(X)\upharpoonright n)|}{n}$$

$$= \frac{1}{nk}\sum_{i=0}^{k-1} \left|\varphi\left(X\upharpoonright[ni, n(i+1)\right)\right| = \frac{|\varphi(X\upharpoonright nk)|}{nk},$$

where the last equality follows from the fact that φ is an n-block map. Then

$$\lim_{n\to\infty} \frac{|\varphi(X\upharpoonright n)|}{n} = \lim_{k\to\infty} \frac{|\varphi(X\upharpoonright nk)|}{nk}$$

$$= \lim_{k\to\infty} \frac{1}{k}\sum_{i=0}^{k-1} F(T^i(X)) = \int f d\mu = Rate(\Phi, \mu),$$

Functionals induced by DDG-trees

Discrete Distribution Generating trees were defined by Knuth and Yao as follows:

A DDG-tree is a tree $S \subseteq 2^{<\omega}$ with terminal nodes $D(S) \subset S$, together with a labelling function $\ell_S : D(S) \to A = \{a_1, \ldots, a_k\}$ which induces a discrete probability distribution on A, by setting $p_i = \sum_{\ell_S(\tau)=a_i} 2^{-|\tau|}$.

We assume that the set $[S]$ of infinite paths through S has measure 0, so that $\sum_{i=1}^{k} p_i = 1$

A DDG-tree S defines a partial functional $F_T : 2^\omega \to A$ so that $F_T(X) = \ell(X \restriction n)$, where $X \restriction n \in D(S)$, for $X \notin [S]$.

Average time of extraction

Knuth and Yao define the average running time of randomness extraction by a DDG-tree S to be

$$AvgRT(S) = \sum_{i \in \omega} i \cdot \lambda([\![D(S) \cap 2^i]\!]).$$

That is, $AvgRT(S)$ is the average number of input bits needed to produce a single output bit.

The functional F_S may be used to define an almost total map $\Phi_S : 2^\omega \to 2^\omega$ with representation φ_S as follows.
Any string $\sigma = \sigma_1 ^\frown \ldots ^\frown \sigma_k$, where $\sigma_1, \ldots, \sigma_{k-1} \in D(S)$ and $\sigma_k \notin D(S)$, where this decomposition is unique, as $D(S)$ is prefix-free. Then we set

$$\varphi_S(\sigma) = \varphi_S(\sigma_1) ^\frown \ldots ^\frown \varphi_S(\sigma_{k-1}) ^\frown \varphi_S(\sigma_k) = \ell_S(\sigma_1) ^\frown \ldots ^\frown \ell_S(\sigma_{k-1})$$

The Tree-Shift

$T_S(X) = Y$, where $X = \sigma^\frown Y$ and $\sigma \in D(S)$, when such σ exists.

Lemma

If S is a tree with $\lambda([\![D(S)]\!]) = 1$, then the tree-shift T_S is λ-invariant and ergodic.

Birkhoff's Ergodic Theorem, Version 2

Our result will depend on the following effective version of Birkhoff's Ergodic Theorem, due to Gács, Hoyrup and Rojas.

Theorem (Effective Birkhoff's Ergodic Theorem, version 2)

Let μ be a computable measure on 2^ω and let $T : 2^\omega \to 2^\omega$ be an a.e. computable, μ-invariant, ergodic transformation. Then for any a.e. computable function F that is effectively integrable and any Schnorr random $X \in 2^\omega$,

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=0}^{k-1} F(T^i(X)) = \int F d\mu.$$

Extraction rate of Φ_S

Theorem

Let $X \in 2^\omega$ be Schnorr random. Then for every computable DDG-tree S, we have

$$\lim_{n\to\infty} \frac{|\varphi_S(X\upharpoonright n)|}{n} = \frac{1}{AvgRT(S)}.$$

Finally, by integrating over the Schnorr randoms, we have

Corollary

$$Rate(\Phi_S, \lambda) = \frac{1}{Avg(RT)(S)}.$$

Levin-Kautz conversion

Extraction from of ν-random sequence from a μ-random sequence, for computable measures μ and ν has been studied independently by Levin (1970), Kautz (1991), Schnorr and Fuchs(1977) and Knuth and Yao (1976).
It is the main idea behind the data compression technique known as *arithmetic coding*.

Theorem (Levin, Kautz)

For any computable measures μ and ν on $\mathcal{P}(2^\omega)$, if $X \in \mathrm{MLR}_\mu$ and X is not computable, then there is some $Y \in \mathrm{MLR}_\nu$ such that $X \equiv_T Y$.

The Levin-Kautz Procedure

We define for computable measures μ and ν, an almost total functional $\Phi_{\mu\to\nu}$ that transforms μ-randomness into ν-randomness.

For non-computable $X \in \mathrm{MLR}_\mu$ and $Y \in \mathrm{MLR}_\nu$ such that $\Phi_{\mu\to\nu}(X) = Y$,

(i) $(\Phi_{\mu\to\nu} \circ \Phi_{\nu\to\mu})(X) = X$, and
(ii) $(\Phi_{\nu\to\mu} \circ \Phi_{\mu\to\nu})(Y) = Y$,

so that $X \equiv_T Y$.

A measure μ on 2^ω is *strongly positive* if there is some $\delta \in (0, \frac{1}{2})$ such that for every $\sigma \in 2^{<\omega}$, $\mu(\sigma 0 \mid \sigma) \in [\delta, 1-\delta]$.

Effective Shannon-McMillan-Breimann Theorem

To prove our result, we use an effective version of the Shannon-McMillan-Breimann theorem due to Hoyrup (2012)

Theorem (Hoyrup)

Let μ be a computable, shift-invariant, ergodic measure on 2^ω. Then for every μ-Martin-Löf random sequence $X \in 2^\omega$,

$$\lim_{n\to\infty} \frac{K(X\upharpoonright n)}{n} = \lim_{n\to\infty} \frac{-\log \mu(X\upharpoonright n)}{n} = h(\mu).$$

The extraction rate for Levin-Kautz conversion

We prove the following effective, pointwise version of the classical result of Uyematsu and Kanaya

Theorem

Let μ and ν be computable, shift-invariant, ergodic measures that are strongly positive. Then for every non-computable $A \in \mathrm{MLR}_\mu$,

$$OI_{\Phi_{\mu\to\nu}}(A) = \frac{h(\mu)}{h(\nu)}.$$

In particular, in the case that $\nu = \lambda$, we have $OI_{\Phi_{\mu\to\lambda}}(A) = h(\mu)$.

Random functionals

Random continuous functions on 2^ω were introduced by Barmpalias, Cenzer et al and studied further by the authors.

An arbitrary continuous function $F : 2^\omega \to 2^\omega$ may be given by a *labeled* binary tree, that is, by a function $\ell : 2^{<\omega} \to \{0, 1, B\}$.

Then $F(X)$ is obtained by deleting the B's from the sequence $(\ell(X\restriction i) : i \in \omega)$ and $f(\sigma)$ is similarly defined.

A measure on the space $\mathcal{C}$ of continuous functions is given by assigning probabilities q, r, s to the event that $\ell(\sigma)$ is 0, 1, or B.

Let p be the probability that $\ell(\sigma)$ is either 0 or 1.

Then the expected number of strings in $\{0, 1\}^n$ with $\ell(\sigma) \in \{0, 1\}$ is $p \cdot 2^n$.

Are random functionals total

The following was shown by the authors in 2015

Theorem

Let μ be a measure on $\mathcal{C}$ with p be the probability that $\ell(\sigma)$ is either 0 or 1. Then the probability that a given functional is total is one if $p \geq \frac{1}{2}$ and is zero otherwise.

We will show later that for $p > \frac{1}{2}$, random functionals are still *almost* total.

The extraction rate for a random functional

First we consider the random representation f of a function F
Later we look at the canonical representation of such functions

Using the law of large numbers, we show

Theorem
Let $F : 2^\omega \to 2^\omega$ be a μ-random continuous function, where μ gives probability p that $\ell(\sigma) \in \{0, 1\}$ for the labelling function ℓ. Then

$$Rate(F, \lambda) = p.$$

Random Inputs

Theorem

Let $F : 2^\omega \to 2^\omega$ is a μ-random continuous function, where μ gives probability p that $\ell(\sigma) \in \{0, 1\}$ for the labelling function ℓ. Then for any $A \in 2^\omega$ that Martin-Löf random relative to F,

$$\lim_{n\to\infty} OI_f(A\upharpoonright n) = Rate(F, \lambda) = p.$$

It follows that random functionals are defined for all random inputs. So we have the following:

Corollary

Every μ- random continuous functional F is almost total.

Canonical Representations

The final result shows that the canonical rate equals the rate

Theorem

Let μ be any measure on the space of continuous functions, let F be a μ-random continuous functional. and let $A \in 2^\omega$ be random relative to F. Then $lim_{n\to\infty} \frac{|\varphi_F(A\upharpoonright n)|}{|f(n)|} = 1$.

Proof for Online Functions

F is *online* if $s = 0$; we may assume $q \leq r < 1$.

Theorem

Let F be a random continuous online functional with representation f and canonical representation φ and let $A \in 2^\omega$ be random relative to f. Then $\lim_{n\to\infty} \frac{|\varphi_F(A\restriction n)|}{n} = 1$.

Proof: Note that $|f(\sigma)| = |\sigma| \leq |\varphi(\sigma)|$ for all σ

Lemma

For sufficiently large k and all strings σ, the probability p_k that $|\varphi(\sigma)| \geq |\sigma| + k$ is $\leq 2^{-k}$.

Proof of Lemma

$|\varphi(\sigma)| \geq |\sigma| + k$ if and only if there exists $\rho \in \{0,1\}^k$ such that, for all $\tau \in \{0,1\}^k$, $f(\sigma^\frown\tau) = f(\sigma)^\frown\rho$.

This means that for any $i < k$, and all τ of length i, the 2^i outputs $f(\sigma^\frown\tau)$ are all the same. They are either all 0, which has probability q^{2^i}, or all 1, which has probability r^{2^i}. Therefore

$$p_k \leq (q^2 + r^2)(q^4 + r^4) \cdot \cdots \cdot (q^{2^k} + r^{2^k}).$$

Since $q \leq r$, we have

$$p_k \leq 2r^2(2r^4) \cdot \cdots \cdot (2r^{2^k}) \leq 2^k r^{2^k}$$

Let $r = 2^{-t}$ for some $t > 0$ and k large enough so $t \cdot 2^k > 2k$. Then

$$2^k r^{2^k} = 2^k 2^{-t2^k} = 2^{k - t2^k} < 2^{-k}.$$

Proof of Theorem

Let K be large enough so that for $k \geq K$, $p_k \leq 2^{-k}$.
Fix A random relative to F, so F is random relative to A. Let

$$V_n = \{F \in \mathcal{C} : (\exists m > n)|\varphi(A\upharpoonright m)| \geq m + K + n\}.$$

Then $\lambda(V_n) \leq 2^{-K-n} < 2^{-n}$, by the Lemma,

so this is an A- Martin-Löf test.

Thus for some n, $F \notin V_n$.

This implies that for all $m > n$, $|\varphi(A\upharpoonright m)| < m + K + n$.

Since we have $m \leq |\varphi(A\upharpoonright m)|$ for all m,

$lim_{m\to\infty} \frac{|\varphi(A\upharpoonright m)|}{m} = 1$, as desired.

Ongoing Research

Extraction rate for real continuous functions

This is related to the modulus of convergence

There are connections with differentiability

The End

Thank You