Randomness extractors

Jan Bouda

FI MU

June 4, 2010

Part I

Extracting randomness

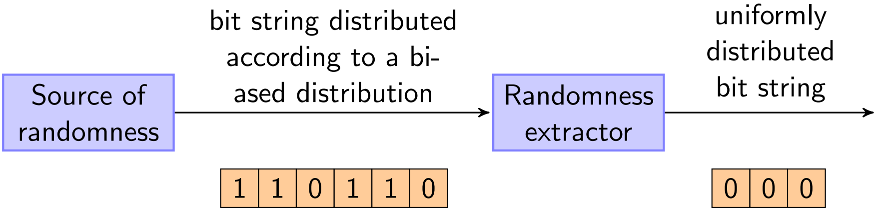
Random Numbers in Computer Science

- Random numbers are of crucial importance for a waste number of computer science applications.
- Cryptography is impossible without random numbers.
  - Cryptographic keys - encryption, authentication, digital signatures
  - Random choices in cryptographic algorithms and protocols - zero knowledge proofs
- Randomized algorithms
- Communication protocols

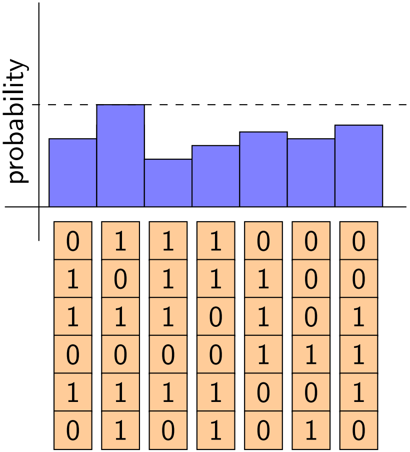
Practically all these applications

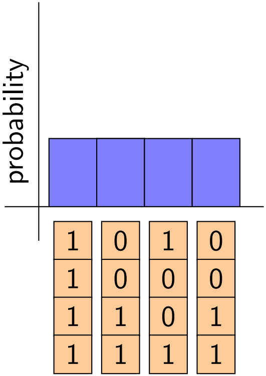
- inherently require randomness generated uniformly
- or their analysis is performed for uniform random numbers.

Randomness Extraction



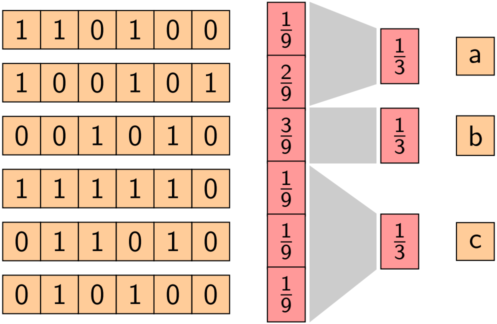
Randomness Extraction





Extraction from Know Probability Distribution

- In contrast to our requirements, most available sources of randomness generate non-uniform output.
- We have to partition the set of outputs into set of constant probability.
- Depending on the output probability distribution, this may be impossible.



Extraction from Unknown Probability Distribution

- The probability distribution of the random number generator output may vary during the computation.
- This might be due to
  - low quality of the generator design,
  - external hard-to-control effects, such as temperature,
  - or an attack of an adversary.
- Non-uniform distribution models adversary's knowledge about the outcome of a (uniform) random number generator.
- Extraction is still possible, given some limitations on the output probability distributions.

Von Neumann Extractor

- Source produces a sequence of random bits, that are generated independently according to (an unknown) a fixed probability distribution.
- On each position the source generates independently

$$0 \text{ with probability } p$$
$$1 \text{ with probability } (1-p).$$

- Von Neumann extractor divides the bit sequence into pairs and for each pair of bits it takes action depending on the value

00 outputs nothing

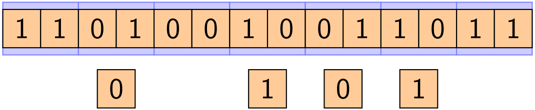
11 outputs nothing

01 outputs 0

10 outputs 1.

Von Neumann Extractor

- For the aforementioned source the output is always a sequence of independent and uniformly distributed bits.



Part II

Randomness Extractors

Towards Extractor Definition

- The purpose of an extractor is to transform an input (biased) probability distribution to a probability distribution that is (close to) uniform distribution.
- Assume we have a biased distribution X on $\mathbf{X}$.
- A randomness extractor is function $e : \mathbf{X} \rightarrow \mathbf{Y}$, such that the distribution Y on $\mathbf{Y}$ induced by the distribution X, i.e.

  $$P(Y = y) = \sum_{x \in \mathbf{X}, e(x) = y} P(X = x),$$

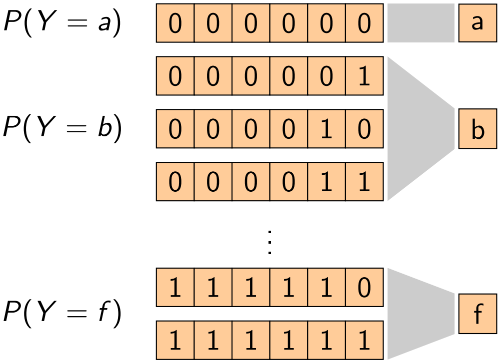
  is close (to be specified later) to the uniform distribution.
- Such an extractor has natural limitations, namely for a fixed e, and two distributions X_1 and X_2 mapped by e to uniform distribution, for each $y \in \mathbf{Y}$ it holds that

  $$\sum_{x \in \mathbf{X}, e(x) = y} P(X_1 = x) = \sum_{x \in \mathbf{X}, e(x) = y} P(X_2 = x) = P(Y = y).$$

Towards extractor definition

This means that e partitions $\mathbf{X}$ to pre-images of elements of $\mathbf{Y}$.

$$e : \mathbf{X} \to \mathbf{Y}$$



Towards extractor definition

- We may overcome this limitation by allowing a (small) auxiliary uniform input Z.
- This would give us the seeded extractor $e : \mathbf{X} \times \mathbf{Z} \rightarrow \mathbf{Y}$.
- We naturally expect that the extractor should be useful, i.e. to produce some extra randomness. We require $|Y| > |Z|$.

Trace Distance of Probability Distributions

Definition

Let X and Y be random variables defined on the same sample space $\mathcal{S}$ with probability distributions p_X and p_Y, respectively. The **trace distance** (or L_1 **distance**) of random variables X and Y is

$$d(X,Y) = \frac{1}{2}\sum_{a\in\mathcal{S}} |p_X(a) - p_Y(a)| = \max_{A\subseteq\mathcal{S}} |P(X\in A) - P(Y\in A)|. \quad (1)$$

X and Y are ϵ-**close in** L_1 iff

$$d(X,Y) \leq \epsilon. \quad (2)$$

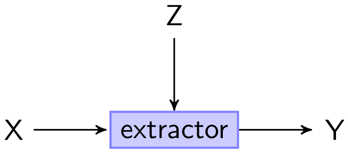
Extractor Definition

Definition

Let $\mathcal{P}(\mathbf{X})$ be the set of all probability distributions on $\mathbf{X}$, and $\mathcal{S} \subset \mathcal{P}(\mathbf{X})$. Then $e : \mathbf{X} \times \mathbf{Z} \to \mathbf{Y}$ is a $(\mathcal{S}, \epsilon)$ (seeded) **randomness extractor** iff for all $X \in \mathcal{S}$

$$d(e(X, U_Z), U_Y) \leq \epsilon, \tag{3}$$

where U_Z is the uniform distribution on $\mathbf{Z}$ and U_Y is the uniform distribution on $\mathbf{Y}$.



Part III

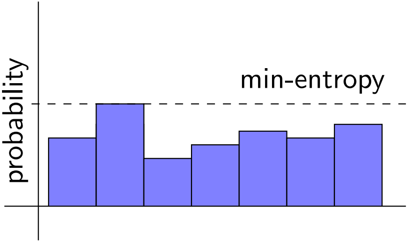
Sources of Randomness

Definition

The **min–entropy** of a probability distribution X is

$$H_\infty(X) = \min_{x \in \mathbf{X}} -\log P(X = x) = -\log \max_{x \in \mathbf{X}} P(X = x). \tag{4}$$

It is a good measure of the amount of randomness contained in the input probability distribution, as demonstrated by the next theorem.



Min-entropy Bounds Extractor Output

Theorem

Let X be a random variable with image $\mathbf{X} = \{0,1\}^n$ satisfying $H_\infty(X) \leq k-1$ for some $k \in \mathbb{N}$. Then there no $(\{X\}, 0)$ extractor with $\mathbf{Z} = \{0,1\}^d$ and $\mathbf{Y} = \{0,1\}^m$ such that $m \geq k+d$.

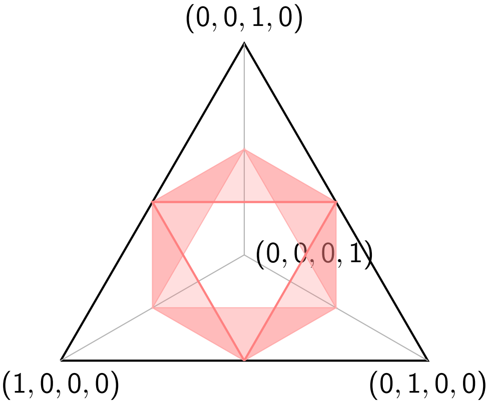
Proof.

The fact that $H_\infty(X) \leq k-1$ implies that there is some element x such that $P(X=x) \geq 2^{-(k-1)}$. Therefore, for any auxiliary input $z \in \mathbf{Z}$, the probability of the corresponding output $e(x,z)$ is at least $2^{-(k-1)}2^{-d} = 2^{-(k+d-1)} > 2^{-m}$ and therefore the output probability distribution is not uniform and its distance from the uniform distribution is bounded by the min-entropy of the input. □

Previous theorem shows us that the gain of randomness extraction is limited by the min–entropy of the source distribution.

Min-Entropy Source

A first example of an extractable set of probability distributions is the min-entropy source. We define the source with min-entropy k as $\mathcal{S} \subset \mathcal{P}(\mathbf{X})$ such that $\forall X \in \mathcal{S}\ H_\infty(X) \geq k$.



Min-Entropy Extractor

Definition

The function $e : \mathbf{X} \times \mathbf{Z} \rightarrow \mathbf{Y}$ is a (k, ϵ) (seeded) **randomness extractor** iff for all X with $H_\infty(X)$ it holds that

$$d(e(X, U_Z), U_Y) \leq \epsilon, \quad (5)$$

where U_Z is the uniform distribution on $\mathbf{Z}$ and U_Y is the uniform distribution on $\mathbf{Y}$.

- Extractor is non-trivial if it extracts more randomness than it consumes as the auxiliary input, i.e. $|\mathbf{Y}| > |\mathbf{Z}|$.
- We want to extract as much randomness as possible, i.e. $|\mathbf{Y}| \dashrightarrow 2^k |\mathbf{Z}|$.

Importance of Seed in Min-Entropy Extractor

Theorem

There is no function $e : \{0,1\}^n \to \{0,1\}$ *giving a single random bit (uniform distribution on* $\{0,1\}$*) as an output for any input random variable* X *on* n*–bit strings satisfying* $H_\infty(X) \geq n - 1$*.*

Intuitively, an input distribution with min–entropy at least $n - 1$ contains much more randomness than necessary to obtain a single random bit.

Proof.

For every function e there is a bit $b \in \{0,1\}$ such that $|\{x \in \{0,1\}^n | e(x) = b\}| \geq 2^{n-1}$ since there are 2^n inputs in the domain of e. Let us consider a random variable X uniformly distributed on the set $\{x \in \{0,1\}^n | e(x) = b\} \subset \{0,1\}^n$. Such a random variable obeys $H_\infty(X) \geq n - 1$ and yet the output distribution $e(X)$ is constant, i.e. $P(e(X) = b) = 1$. □

Part IV

Carter-Wegman Hashing

Universal hashing

Definition

Let A and B be sets such that $|A| > |B|$. A family H of hash functions $h : A \to B$ is k-**universal** iff for any $x_1, x_2, \ldots, x_k \in A$ and a hash function $h \in H$ randomly and uniformly chosen from H it holds that

$$\mathcal{P}(h(x_1) = h(x_2) = \cdots = h(x_k)) \leq \frac{1}{|B|^{k-1}}. \tag{6}$$

Definition

Let A and B be sets such that $|A| > |B|$. A family H of hash functions $h : A \to B$ is **strongly** k-**universal** iff for any $x_1 \neq x_2 \neq \cdots \neq x_k \in A$, any $y_1, y_2, \ldots, y_k \in B$ and a hash function $h \in H$ randomly and uniformly chosen from H it holds that

$$\mathcal{P}(h(x_1) = y_1 \wedge h(x_2) = y_2 \ldots h(x_k) = y_k) \leq \frac{1}{|B|^k}. \tag{7}$$

Universal Hashing: Example

Let $A = \{0, 1, \ldots, m-1\}$ and $B = \{0, 1, \ldots, n-1\}$ with $m \geq n$. Let $p \geq m$ be some prime. Consider the class of hash functions

$$h_{a,b}(x) = ((ax + b) \bmod p) \bmod n. \tag{8}$$

Let

$$H = \{h_{a,b} | 1 \leq a \leq p-1, 0 \leq b \leq p\}, \tag{9}$$

stressing that $a \neq 0$.

Theorem

H is 2-universal.

Universal Hashing: Example

Proof.

We count the number of functions from H for which two fixed and distinct elements x_1 and x_2 from A collide. $x_1 \neq x_2$ implies

$$ax_1 + b \not\equiv ax_2 + b \pmod{p},$$

since the opposite occurs only if $a(x_1 - x_2) \equiv 0 \pmod{p}$. However, we know that neither $a \equiv 0 \pmod{p}$ nor $x_1 - x_2 \equiv 0 \pmod{p}$, what implies the equation.

Fixing x_1 and x_2, for every pair $u \neq v \in B$ there exists exactly one pair a, b such that $ax_1 + b \equiv u \pmod{p}$ and $ax_2 + b \equiv v \pmod{p}$. □

Universal Hashing: Example

Proof.

Solving the system of two linear equations we obtain the unique solution

$$a = \frac{v-u}{x_2-x_1} \bmod p \tag{10}$$

$$b = u - ax_1 \bmod p. \tag{11}$$

Since there is exactly one hash function for each pair (a, b), we have there is exactly one hash function in H such that

$$ax_1 + b \equiv u \pmod{p} \text{ and } ax_2 + b \equiv v \pmod{p}.$$

We have that the number of collisions equals to the number of pairs (u, v) from $\{0, \ldots, p-1\}$ satisfying $u \neq v$ and $u \equiv v \pmod{n}$. For each choice of u there are at most $\lceil p/n \rceil - 1$ possible values of v. □

Universal Hashing: Example

Proof.

Together we have that there are at most

$$p(\lceil p/n \rceil - 1) \leq p\left(\frac{p+(n-1)}{n} - \frac{n}{n}\right) = \frac{p(p-1)}{n}.$$

such pairs. Therefore, the collision probability is

$$P(h_{a,b}(x_1) = h_{a,b}(x_2)) \leq \frac{p(p-1)/n}{p(p-1)} = \frac{1}{n}.$$

□

Part V

Extractors for Min-entropy Sources

Min-Entropy Strong Extractor

Definition

The function $e : \mathbf{X} \times \mathbf{Z} \rightarrow \mathbf{Y}$ is a (k, ϵ) (seeded) **strong randomness extractor** iff for all X with $H_\infty(X)$ it holds that

$$d\big([U_z, e(X, U_Z)], [U_Z, U_Y]\big) \leq \epsilon, \qquad (12)$$

where U_Z is the uniform distribution on $\mathbf{Z}$ and U_Y is the uniform distribution on $\mathbf{Y}$.

- The advantage of the strong extractor is that the output is close to the uniform distribution even if the value of U_Z is known.
- Next we will show how to implement a strong extractor using Wegman-Carter hashing.

Min-Entropy Extractor

Theorem

Let X be a random variable defined on $\mathbf{X} = \{0,1\}^n$ with min-entropy $H_\infty(X) \geq k$, $H = \{h | h : \{0,1\}^n \to \{0,1\}^{k-2e}\}$ be a universal$_2$ class of hash functions. Let $x \in_R \mathbf{X}$ be randomly chosen from $\mathbf{X}$ according to X and h be randomly and uniformly chosen from H. Then the distribution of $(h, h(x))$ is 2^{-e} close to the uniform distribution in the trace distance, i.e. application of a function randomly chosen from H is a $(k, 2^{-e})$ strong randomness extractor.

Min-Entropy Extractor

Theorem

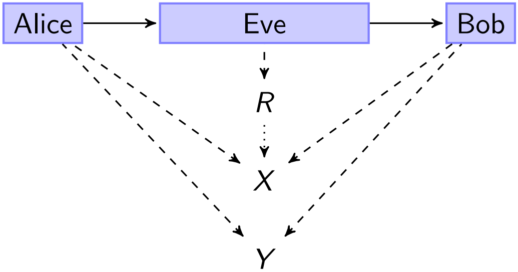
Let $X_1, X_2, \ldots, X_l$ be independent identically distributed random variables each defined on $\mathbf{X} = \{0,1\}^n$ with min-entropy $H_\infty(X) \geq k$, $H = \{h | h : \{0,1\}^n \rightarrow \{0,1\}^{k-2e}\}$ be a universal$_2$ class of hash functions. Let $x_i \in_R \mathbf{X}$ be randomly chosen from $\mathbf{X}$ according to X_i and h be randomly and uniformly chosen from H. Then the distribution of $(h, h(x_1), \ldots, h(x_l))$ is $l\,2^{-e}$ close to the uniform distribution in the trace distance, i.e. l repeated applications of a fixed function randomly chosen from H is a $(k, l\,2^{-e})$ strong randomness extractor.

Part VI

Privacy Amplification

Initial Situation

- Alice sends an information to Bob via a channel that can be (partially) observed by Eve.
- After the communication the information between Alice and Bob is perfectly preserved, described by a random variable X.
- Eve has a partial knowledge of X represented by a random variable R.
- Alice and Bob want to extract a shorter shared information Y, such that E contains no information about Y.



Eliminating Eve

Assuming Eve know's the value of R to be r, her knowledge about X is the conditional probability distribution

$$P(X = x | R = r).$$

- Alice and Bob agree publicly on a strong extractor $e : \mathbf{X} \times \mathbf{Z} \rightarrow \mathbf{Y}$.
- Alice sends $x \in \mathbf{X}$ to Bob (Eve learns partial information $r \in \mathbf{R}$).
- Alice chooses randomly and uniformly $z \in \mathbf{Z}$ and sends it via public and authenticated channel to Bob (Eve learns it).
- Both Alice and Bob compute $y = e(x, z)$.

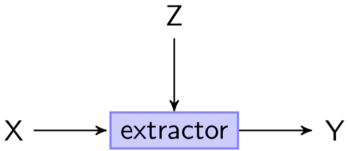
Eve has no information about y, her prediction is (almost) uniform distribution over $\mathbf{Y}$.

Eliminating Eve

This is possible thanks to the properties of the strong extractor:

$$d\big([U_z, e(X, U_Z)], [U_Z, U_Y]\big) \leq \epsilon, \qquad (13)$$

We have to evaluate the min-entropy of the conditional probability distribution.



Part VII

More Extractors

Other Types of Sources and Generalized Extractors

- von Neumann sources: independence, fixed/limited bias
- Santha-Vazirani sources: possibly dependent, limited bias
- independent sources
  - one source vs. multi-source point of view
  - blenders
- bit-fixing sources
  - cryptographic application
  - model e.g. adversary's knowledge

Condensers - increase of min-entropy.

Thank You for Your Attention!