Extracting Randomness from Generalized Symbol-Fixing and Markov Sources[1]

Robert König Ueli Maurer
Department of Computer Science,
ETH Zürich, Switzerland
e-mail: {rkoenig,maurer}@inf.ethz.ch

Abstract — **We introduce a new class of realistic sources of randomness and give concrete procedures for *deterministic* extraction of almost uniform random bits from these sources. Moreover, we show how randomness can be extracted from general Markov sources. This extends the types of sources for which explicit deterministic randomness extractors are known.**

I. Generalized Symbol-Fixing Sources

A common measure of the amount of randomness contained in a source X is its min-entropy $H_\infty(X) := -\log_2(\max_x P_X(x))$. To extract randomness deterministically, a lower bound on this quantity has to be guaranteed, but it is trivial to see [1] that this condition on its own is too general to allow deterministic extraction. It is also clear that many physical sources of randomness, such as Geiger counter noise or Zener diodes, exhibit a sort of imperfectness that is only crudely modeled by a single min-entropy condition alone. In particular, the output produced by some physical device may be fixed over long periods of time and may only contain a few sudden bursts of randomness. This motivates the following definition, which generalizes the concept of oblivious bit-fixing sources [2, 3] and symbol-fixing sources [5].

Definition 1 *A* $\big(n, \{k_i\}_{i=1}^s, \Omega\big)$-generalized symbol-fixing source *is a tuple* $(X_1, \ldots, X_n)$ *of independent random variables on* Ω *subject to the following condition. There exist distinct indices* $i_1, \ldots, i_s \in [n]$ *such that* $H_\infty(X_{i_j}) \geq k_j$ *for* $j = 1, \ldots, s$. *Similarly, a* $\big(n, k, s, \Omega\big)$-generalized symbol-fixing source *is an* n*-tuple of independent random variables on* Ω *for which there exist* s *distinct indices such that* $H_\infty(X_{i_1} \cdots X_{i_s}) \geq k$.

We denote by ${}^n\mathbf{GSF}_\Omega[k_1, \ldots, k_s]$ the set of $\big(n, \{k_i\}_{i=1}^s, \Omega\big)$- and by ${}^n_s\mathbf{GSF}_\Omega[k]$ the set of $\big(n, k, s, \Omega\big)$-generalized symbol-fixing sources.

II. Deterministic Randomness Extraction

The problem of finding *deterministic* procedures for generating almost uniform random bits using only an imperfect source of randomness has been studied in various contexts in the literature. In general, one aims at finding a function $Ext : \Omega \to \mathcal{Z}$ such that for every random variable $X \in \mathcal{S}$, the random variable $Ext(X)$ is almost uniformly distributed, where $\mathcal{S}$ is a specified class of probability distributions on Ω. If $Ext(X)$ is ε-close (measured in terms of statistical distance) to the uniform distribution on $\mathcal{Z}$ for every $X \in \mathcal{S}$, the function Ext is called an $(\mathcal{S}, \varepsilon)$-extractor [4].

A typical example of such a class $\mathcal{S}$ of probability distributions is the set of so-called Chor-Goldreich sources [1], which are pairs of independent random variables, each having a certain amount of min-entropy.

We show how $({}^n\mathbf{GSF}_{\mathcal{G}}[k_1, k_2], \varepsilon)$- and $({}^n_2\mathbf{GSF}_{\mathcal{G}}[k], \varepsilon)$-extractors can be built in a generic way from group-theoretic extractors for Chor-Goldreich sources. This gives for example the following result.

Theorem 2 *Let* $p > 2$ *be a prime and let* q *be a divisor of* $p - 1$. *Let* g *be a generator of* $\mathbb{Z}_p^*$ *and let* $\log_g(\cdot)$ *denote the discrete logarithm to base* g *in* $\mathbb{Z}_p^*$. *Then for every* $\varepsilon \geq \frac{2}{p}$ *the function*

$$(x_1, \ldots, x_n) \mapsto \begin{cases} \log_g(\sum_{i=1}^n x_i \textit{ mod } p) \textit{ mod } q & \textit{if defined} \\ 0 & \textit{otherwise} \end{cases}$$

is a $({}^n_2\mathbf{GSF}_{\mathbb{Z}_p}[\log_2 p + 2\log_2(\frac{1}{\varepsilon}) + 2\log_2 q], \varepsilon)$*-extractor. Moreover, the extractor is efficiently computable if the parameters* p*,* g *and* q *are chosen appropriately.*

We show that every Markov source can be seen as an m-*independent distribution*, a concept we introduce to describe pairs of random variables having a limited amount of independence. This directly yields extractors for Markov sources, since extractors for m-independent distributions can also be built generically from extractors for Chor-Goldreich sources. We apply this to deterministic extraction of random bits produced by space-bounded samplers, a problem raised in [6] and generalize results about randomness extraction from memory-bounded sources [7].

References

[1] B. Chor and O. Goldreich, "Unbiased Bits from Sources of Weak Randomness and Probabilistic Communication Complexity", *SIAM Journal on Computing, vol. 17 (2), 1988*

[2] B. Chor, O. Goldreich, J. Hastad, J. Freidmann, S. Rudich and R. Smolensky, "The Bit Extraction Problem or t-Resilient Functions" *IEEE Symposium on Foundations of Computer Science (FOCS), 1985*

[3] A. Cohen and A. Wigderson, "Dispersers, Deterministic Amplification, and Weak Random Sources", *IEEE Symposium on Foundations of Computer Science (FOCS), 1989*

[4] Y. Dodis, "Exposure-Resilient Cryptography", *Phd thesis, Massachusetts Institute of Technology, 2000*

[5] J. Kamp and D. Zuckerman, "Deterministic Extractors for Bit-Fixing Sources and Exposure-Resilient Cryptography", *IEEE Symposium on Foundations of Computer Science (FOCS), 2003*

[6] L. Trevisan and S. P. Vadhan, "Extracting Randomness from Samplable Distributions", *IEEE Symposium on Foundations of Computer Science (FOCS), 2000*

[7] U. V. Vazirani, "Strong communication complexity or generating quasirandom sequences from two communicating semirandom sources", *Combinatorica vol. 7 (4), 1987*

[1] This work was partially supported by the Swiss National Science Foundation (SNF), project no. 20-66716.01.