Information-Theoretic Cryptography

(Extended Abstract)

Ueli Maurer*

Department of Computer Science
Swiss Federal Institute of Technology (ETH), Zurich
CH-8092 Zurich, Switzerland,
`maurer@inf.ethz.ch`

Abstract. We discuss several applications of information theory in cryptography, both for unconditional and for computational security. Unconditionally-secure secrecy, authentication, and key agreement are reviewed. It is argued that unconditional security can practically be achieved by exploiting the fact that cryptography takes place in a physical world in which, for instance due to noise, nobody can have complete information about the state of a system.
The general concept of an information-theoretic cryptographic primitive is proposed which covers many previously considered primitives like oblivious transfer, noisy channels, and multi-party computation. Many results in information-theoretic cryptography can be phrased as reductions among such primitives We also propose the concept of a generalized random oracle which answers more general queries than the evaluation of a random function. They have applications in proofs of the computational security of certain cryptographic schemes.
This extended abstract summarizes in an informal and non-technical way some of the material presented in the author's lecture at Crypto '99.

Key words. Information theory, unconditional security, conditional independence, information-theoretic primitive, generalized random oracle.

1 Introduction

Historically, information theory and cryptography are closely intertwined, although the latter is a much older discipline. Shannon's foundation of information theory [40] was motivated in part by his work on secrecy coding during the second world war, and it may be for this reason that his work was not de-classified until 1948 when his seminal paper was published. His 1949 companion paper on the communication theory of secrecy systems [39] was, like Diffie and Hellman's later discovery of public-key cryptography [19], a key paper in the transition of cryptography from an art to a science.

There are two types of cryptographic security. The security of a cryptographic system can rely either on the computational infeasibility of breaking it (computational security), or on the theoretical impossibility of breaking it, even

* Supported in part by the Swiss National Science Foundation, grant no. 20-42105.94.

using infinite computing power (information-theoretic or unconditional security). Because no computational problem has been proved to be computationally difficult for a reasonable model of computation, the computational security of every practical cryptographic system relies on an unproven intractability assumption. In contrast, information-theoretically secure systems rely on no such assumptions, but they rely on an assumption about the probabilistic behavior of the universe, for instance of a noisy channel or a quantum measurement. However, even computationally-secure systems rely on such assumptions, at least the tacitly made assumption that random keys can be generated and that they are independent of an adversary's entire *a priori* knowledge.

While information-theoretic security is stronger than computational security, it is usually less practical. In fact, Shannon's proof that perfect secrecy requires a secret key of the same length as the plaintext is often taken as evidence that unconditional security can never be practical. However, this precipitate jump to conclusions should be reconsidered: in contrast to Shannon's model, in which his result holds, cryptography takes place in a physical world (every communication channel is based on a physical process) in which nobody can have complete information about the state of a system, for instance due to noise or theoretical limitations of quantum physics.

Information theory has several applications in cryptography. First, it allows to prove the unconditional security of cryptographic systems. Second, it allows to prove impossibility and lower bound results on the achievability of unconditional security. Third, it is a key tool in reduction proofs showing that breaking a cryptographic system is as hard as breaking an underlying cryptographic primitive (e.g. a one-way function or a pseudo-random function).

In this extended abstract we give an overview of known applications and results of information theory in cryptography. Due to space limitations we cannot give a complete overview of the extensive literature on the subject. The treatment is informal and non-technical, emphasizing concepts and general viewpoints. In Section 2, we review some basic concepts of information theory and state two basic facts on conditional independence. In Section 3, we summarize known results on unconditional secrecy, authentication, and key agreement. In Section 4 we take a general approach to cryptographic primitives and reductions among them. The concept of generalized random oracles is sketched briefly in Section 5, followed by some conclusions.

2 Random variables, entropy, and conditional independence

Information theory, like statistics, is a mathematical theory based on probability theory.[1] In almost all applications of probability theory in cryptography one considers a *discrete random experiment* which is conceptually very simple: it is defined by a finite or countably infinite set called the *sample space*, consisting of all elementary events, and a *probability measure* assigning a non-negative real number to every elementary event, such that the sum of all these probabilities is

[1] We refer to [7] and [13] for a more detailed introduction to information theory, and to [21] for an introduction to probability theory.

equal to 1. An *event* of a discrete random experiment is a subset of the sample space, and the probability assigned to it is the sum of the probabilities of its elementary events.

A *discrete random variable* X is a mapping from the sample space to a certain range $\mathcal{X}$ and is characterized by its *probability distribution* P_X that assigns to every $x \in \mathcal{X}$ the probability $P_X(x)$ of the event that X takes on the value x.

The *entropy* of a random variable X is a real number that measures the uncertainty about the value of X when the underlying random experiment is carried out. It is defined as

$$H(X) \;=\; -\sum_x P_X(x) \log_2 P_X(x),$$

assuming here and in the sequel that terms of the form $0 \log 0$ are excluded from the summation. The particular formula will be irrelevant below, but we need certain important properties of entropy. It is easy to verify that

$$0 \;\leq\; H(X) \;\leq\; \log_2 |\mathcal{X}|$$

with equality on the left if and only if $P_X(x) = 1$ for some $x \in \mathcal{X}$ and with equality on the right (for finite $\mathcal{X}$) if and only if $P_X(x) = 1/|\mathcal{X}|$ for all $x \in \mathcal{X}$.

The deviation of the entropy $H(X)$ from its maximal value can be used as a measure of non-uniformity of the distribution P_X. While there are other such non-uniformity measures (e.g., based on Rényi entropy and min-entropy, which have some interesting applications not discussed in this paper), the significance of Shannon entropy is that it satisfies some intuitive rules (e.g., the chain rule) and that it gives the right answer to fundamental questions in communication engineering: how much can (reversible) data compression reduce the size of a message, and how many information bits per channel use can be transmitted reliably over a given noisy communication channel?

When several random variables (e.g. X, Y, Z with joint distribution P_{XYZ}) are considered, they are always defined on the same random experiment. The definition of $H(X)$ can be generalized to the definition of the joint entropy of two or more random variables. For instance, we have $H(XYZ) = -\sum_{(x,y,z)} P_{XYZ}(x,y,z) \log P_{XYZ}(x,y,z)$.

The conditional probability distribution $P_{X|Y}(\cdot, y)$ of the random variable X, given the event $Y = y$, is defined by $P_{X|Y}(x,y) = P_{XY}(x,y)/P_Y(y)$ when $P_Y(y) \neq 0$. For every such $y \in \mathcal{Y}$, $P_{X|Y}(\cdot, y)$ is a probability distribution satisfying $\sum_{x \in \mathcal{X}} P_{X|Y}(x,y) = 1$. The entropy of this distribution is the conditional entropy of X, given the event $Y = y$:

$$H(X|Y = y) = -\sum_x P_{X|Y}(x,y) \log_2 P_{X|Y}(x|y).$$

The conditional uncertainty of X, given the random variable Y, is defined as the average over all y of $H(X|Y = y)$, and is not the entropy of a distribution:

$$H(X|Y) \;=\; \sum_{y \in \mathcal{Y}: P_Y(y) \neq 0} H(X|Y = y) P_Y(y).$$

One can show that additional knowledge can never increase entropy:

$$0 \;\leq\; H(X|Y) \;\leq\; H(X),$$

with equality on the left if and only if Y determines X (except when $P_Y(y) \neq 0$) and with equality on the right if and only if X and Y are statistically independent (see below).

An important rule for transforming entropies is

$$H(XY) \;=\; H(X) + H(Y|X),$$

i.e., the joint entropy about X and Y is the entropy about X plus the additional entropy about Y, given that X is known. This so-called chain rule can be used repeatedly to expand $H(X_1X_2\cdots X_N)$ as

$$H(X_1X_2\cdots X_N) \;=\; \sum_{n=1}^{N} H(X_n|X_1\cdots X_{n-1}).$$

Note that the order in which variables are extracted is arbitrary. For example,

$$\begin{aligned} H(XYZ) &= H(X) + H(Y|X) + H(Z|XY) \\ &= H(Y) + H(Z|Y) + H(X|YZ). \end{aligned}$$

The *mutual information* $I(X;Y)$ between two random variables X and Y is defined as the amount by which the uncertainty (entropy) about X is reduced by learning Y:

$$I(X;Y) \;=\; H(X) - H(X|Y).$$

The term mutual stems from the fact that, as can easily be verified, $I(X;Y) = I(Y;X) = H(Y) - H(X|Y)$. The conditional mutual information between X and Y, given the random variable Z, is defined as

$$I(X;Y|Z) = H(X|Z) - H(X|YZ).$$

We have $I(X;Y|Z) = 0$ if and only if X and Y are statistically independent when given Z.

Conditional independence is a fundamental concept in information-theoretic cryptography. Two events A and B are *statistically independent*, here denoted $[A;B]$, if $P(A \cap B) = P(A) \cdot P(B)$. In the following we will drop the symbol $\cap$ and use the shorter notation $P(A,B)$ or simply $P(AB)$ for $P(A \cap B)$. Two events A and B are conditionally independent, given the event C, denoted $[A;B|C]$, if $P(A \cap B \cap C) \cdot P(C) = P(A \cap C) \cdot P(B \cap C)$ or, in our short notation,

$$P(ABC) \cdot P(C) = P(AC) \cdot P(BC).$$

If $P(C) > 0$, this is equivalent to $P(AB|C) = P(A|C) \cdot P(B|C)$. Note that independence is symmetric, i.e. $[A;B] \Longleftrightarrow [B;A]$ and $[A;B|C] \Longleftrightarrow [B;A|C]$. Let $\overline{A}$ denote the complement of the event A. One can also show that $[A;B] \Longleftrightarrow [A;\overline{B}]$ and $[A;B] \Longleftrightarrow [\overline{A};\overline{B}]$ while $[A;BC] \Longrightarrow [A;\overline{B}C]$ is false in general.

The concept of statistical independence and this notation can be extended to a situation where any of A, B and C can be either an event or a random variable. Independence when random variables are involved means that the independence relation holds when any random variable is replaced by the event that it takes on a particular value. For instance, if A is an event and X and Y are random variables, then $[X; A|Y]$ is equivalent to $[X = x; A|Y = y]$ for all x and y.

The following theorem stated without proof implies the rules for a calculus of conditional independence and is useful for simplifying certain security proofs. It states under which condition an event or random variable can be added to an independence set. Moreover, any random variables in an independence set can be dropped, and, if accompanied in the set only by other random variables, then it can also be moved to the conditioning set.

Theorem 1. *Let S, T, U and V each be an event or a random variable (defined for the same random experiment). Then*

$$[S;T|V] \text{ and } [S;U|TV] \implies [S;TU|V].$$

If U is a random variable, then $[S;TU|V] \Longrightarrow [S;T|V]$, and if also T is a random variable, then $[S;TU|V] \Longrightarrow [S;U|TV]$

Note that if S, T, and U are events, then $[S;TU] \implies [S;T|U]$ and $[S;TU] \Longrightarrow [S;T]$ are false in general. For instance, let $P(S) = P(T) = P(U) = 0.5$, $P(ST) = P(SU) = P(TU) = 0.2$, and $P(STU) = 0.1$. Then $P(STU) = P(S)P(TU) = 0.1$ but $P(STU)P(U) = 0.05 \neq P(SU)P(TU) = 0.04$.

3 Unconditional secrecy, authenticity, and key agreement

One of the fundamental problems in cryptography is the transmission of a message M from a sender Alice to a receiver Bob such that an adversary Eve with access to the communication channel is unable to obtain information about M (secrecy). Moreover, if Eve has write-access to the channel, then Bob must not accept a fraudulent message modified or inserted by Eve (authenticity). This is achieved by Alice and Bob sharing a secret key K used together with the message to compute a ciphertext C to be transmitted over the channel.[2] The security can be either computational or information-theoretic, and we are here only interested in the latter.

3.1 Unconditional Authentication

Unconditionally secure message authentication based on a shared secret key was first considered in [24] and later in a large number of papers (e.g., see [44], [41], [42]). Another line of research is devoted to proving lower bounds on the cheating probability as a function of the entropy of the key, $H(K)$; see [33] for a discussion and generalization of these bounds. Assume that the secret key

[2] For instance, C is an encryption of M, or M together with an appended message authentication code.

K is used to authenticate a message M, resulting in ciphertext C, and let p_I and p_S denote Eve's probability of successfully creating a fraudulent message (impersonation attack) and of successfully replacing a valid message by a fraudulent message (substitution attack), respectively, then the following lower bounds hold for any authentication system, for an optimal cheating strategy:

$$p_I \geq 2^{-I(C;K)}, \quad p_S \geq 2^{-H(K|C)}, \quad \text{and} \quad \max(p_I, p_S) \geq 2^{-H(K)/2}.$$

In other words, half of the key must be used for protecting against an impersonation attack, and the other half to prevent a substitution attack. These bounds can be generalized in various directions, for instance to a setting where n consecutive messages are authenticated using the same key. Then the cheating probability is lower bounded by $2^{-H(K)/(n+1)}$.

This bound can easily be achieved, when the message space is a subset of the k-bit strings, by a scheme based on polynomial evaluation (where the secret key consists of the $n+1$ coefficients of a polynomial over $GF(2^k)$ of degree n), achieving cheating probability 2^{-k}. One can show that equality in the bounds cannot be achieved for larger message spaces. However, Gemmell and Naor [23] proved that interactive protocols for authenticating a k-bit message can make more efficient use of the secret key than non-interactive protocols.

3.2 Unconditional Secrecy

It is well-known that the one-time pad [43] provides perfect secrecy (though no authenticity unless the message is redundant), where perfect secrecy is the strongest possible type of security of an encryption system and is defined as the message M and the ciphertext C being statistically independent: $I(M;C) = 0$, or $[M;C]$. Shannon [39] proved that for every perfect system, $H(K) \geq H(M)$, i.e. perfect secrecy requires an impractically large amount of secret key. A system that is perfect for every distribution P_M of the message M is called robustly perfect. The (binary) key of such a system must be at least as long as the message; hence the one-time pad is optimal. Rather than proving Shannon's bound, we look at a more general setting below from which Shannon's result follows as a special case.

In the following we assume that an insecure communication channel between Alice and Bob is available. Since we are interested in results on security and not primarily on communication efficiency, this assumption is made without loss of generality. It implies that a secure key agreement protocol implies a secure encryption scheme (e.g. using the one-time pad), and the reverse implication is trivial. Thus we can restrict our attention to key agreement.

3.3 Unconditional key agreement: impossibility results and bounds

Definition 1. A *key-agreement protocol* consists of a communication phase in which Alice and Bob alternate sending each other messages $C_1, C_2, C_3, \ldots$, where we assume that Alice sends messages $C_1, C_3, C_5, \ldots$ and Bob sends messages $C_2, C_4, C_6, \ldots$. Each message can depend on the sender's entire view of the protocol and possibly on privately generated random bits.

After the communication phase, Alice and Bob each either accepts or rejects the protocol execution, depending on whether he or she believes to be able to generate a shared secret key. If Alice accepts, she generates a key S depending on her view of the protocol.

Similarly, if Bob accepts, he generates a key S' depending on his view of the protocol. Even if a party does not accept, he or she may generate a key. $\diamond$

In the sequel we assume without loss of generality that S and S' are binary strings of length $|S| = |S'| = k$, where the goal is of course to make k as large as possible. Let t be the total number of messages and let $C^t = [C_1, \ldots, C_t]$ denote the set of exchanged messages.

Informally, a key agreement protocol is secure if the following conditions are satisfied [34, 45]:

– whenever Eve is only passive, then Alice and Bob both accept, and
– whenever one of the parties accepts, then
  - the other party has also generated a key (with or without accepting), and the two keys agree with very high probability (i.e. $P[S \neq S'] \approx 0$),
  - the key S is very close to uniformly distributed, i.e. $H(S)$ (and hence also $H(S')$) is very close to k, and
  - Eve's information about S, $I(S; C^t Z)$, given her entire knowledge, is very small (see the definition of Z below).

It appears obvious that if Alice and Bob do not share at least some partially secret information initially, they cannot generate an information-theoretically secure secret key S (i.e. $H(S) = 0$) if they can only communicate over a public channel accessible to Eve, even if this channel is authenticated.[3] This follows from inequality (1) below and implies Shannon's bound $H(K) \geq H(M)$.

In order for the key agreement problem to be interesting and relevant, we therefore need to consider a setting that takes into account the possibility that Alice and Bob each have some correlated side information about which Eve does not have complete information. Several such scenarios, practical and theoretical, have been considered. For instance, Fischer and Wright analyzed a setting in which Alice and Bob are assumed to have been dealt disjoint random deals of cards. More natural and realistic scenarios may arise from exploiting an adversary's partial uncertainty due to unavoidable noise in the communication channel or intrinsic limitations of quantum physics. We refer to [4] for a discussion of quantum key agreement. The problem of designing information-theoretically secure key agreement protocols is hence to identify practical scenarios in which the adversary's total information can be bounded, and then to design a protocol that exploits this scenario.

Such a scenario can generally be modeled by assuming that Alice, Bob, and Eve initially know random variables X, Y, and Z, respectively, which are jointly distributed according to some probability distribution P_{XYZ}.[4] These random

[3] This fact can be rephrased as follows: There exists no unconditionally-secure public-key cryptosystem or public-key distribution protocol.

[4] More generally, the distribution P_{XYZ} could be under Eve's partial control and may only partly be known to Alice and Bob, for instance in the case of a quantum transmission disturbed by Eve.

variables could be the individual received noise signals of a deep-space radio source, or the individual received noisy versions of a random bit string broadcast by a satellite at a very low signal power.

For this setting it was shown in [35] that[5]

$$H(S) \ \leq \ I(X;Y{\downarrow}Z),$$

where $I(X;Y{\downarrow}Z)$ denotes the intrinsic conditional mutual information between X and Y, given Z, which is defined as follows [35]:

$$I(X;Y{\downarrow}Z) := \inf_{P_{\overline{Z}|Z}} \left\{ I(X;Y|\overline{Z}) \ : \ P_{XY\overline{Z}} = \sum_{z \in \mathcal{Z}} P_{XYZ} \cdot P_{\overline{Z}|Z} \right\} .$$

The above inequality is a generalization of the following bound [32]:

$$H(S) \ \leq \ \min[I(X;Y), I(X;Y|Z)]. \tag{1}$$

Note that $I(X;Y)$ and $I(X;Y|Z)$ are obtained when $\overline{Z}$ is a constant or $\overline{Z} = Z$, respectively, and that $I(X;Y|Z) \geq I(X;Y)$ is possible.

3.4 Unconditional key agreement by public discussion

The previous bound is an impossibility result. In order to prove constructive results about key agreement by public discussion, we need to make an explicit assumption about the distribution P_{XYZ}. A very natural assumption, which is often made in an information-theoretic context, is that the same random experiment generating X, Y, and Z is repeated independently many times. One can then define the *secret key rate* $S(X;Y||Z)$ [32] as the maximum rate (per executed random experiment) at which Alice and Bob can generate secret key, assuming (for now) an authenticated but otherwise insecure communication channel.

This rate turns out to be positive even in cases where intuition might suggest that key agreement is impossible. For instance, when a satellite broadcasts random bits and X, Y, and Z are the bits (or more generally signals) received by Alice, Bob, and Eve, respectively, then key agreement is possible under the sole condition that Eve's channel is not completely perfect, even if Alice's and Bob's channels are by orders of magnitude more noisy than Eve's channel, for instance when Alice's and Bob's bit error rates are very close to 50% (e.g. 0.499) and Eve's bit error rate is very small (e.g. 0.001).

We conjecture that the secret key rate $S(X;Y||Z)$ is positive if and only if $I(X;Y{\downarrow}Z)$ is positive, and the two quantities may even be equal. Even if the public discussion channel is *not* authenticated, key agreement is still possible. The secret key rate is even equal to $S(X;Y||Z)$ (where an authenticated channel is assumed) [34, 45], except if Eve can either generate from Z a random variable

[5] neglecting here the fact that the bound can be slightly greater if imperfect secrecy or a non-zero failure probability is tolerated.

$\tilde{Y}$ such that $P_{X\tilde{Y}} = P_{XY}$ or, symmetrically, a random variable $\tilde{X}$ such that $P_{\tilde{X}Y} = P_{XY}$. In both these cases, key agreement is impossible.

Many results on unconditionally secure key agreement were recently refined in various ways. We refer to [46] for a very good overview and to [9, 47] for detailed accounts of recent results in unconditionally-secure key agreement.

3.5 Public randomness and memory-bounded adversaries

In this section we briefly discuss two other scenarios in which Eve cannot obtain complete information, and where this can be exploited by Alice and Bob to agree on a very long unconditionally-secure secret key S (e.g. 1 Gigabyte), assuming that they share only a short secret key K (e.g. 5000 bits) initially.

Suppose that all parties, including Eve, have access to a public source of randomness (similar to a random oracle) which is too large to be read entirely by Eve in feasible time. Then Alice and Bob can access only a moderate number of random bits, selected and combined using the secret key, such that unless Eve examines a substantial fraction (e.g. one half) of the random bits (which is infeasible), she ends up having no information about the generated key [30]. More precisely, there exists an event $\mathcal{E}$ such that $[S; KW|\mathcal{E}]$, where W summarizes Eve's entire observation resulting from an adaptive access strategy. This is true even if Eve is given access to the secret key K after finishing her access phase. Moreover, for any adaptive strategy (without knowledge of K), the event $\mathcal{E}$ has probability exponentially close to 1. In other words, conditioned on this high-probability event, the scheme achieves perfect secrecy.

While the original proof assumed that Eve accesses individual bits of the source, Aumann and Rabin [37] showed that the scheme is secure even if she accesses arbitrary bits of information about the random bits, e.g. Boolean functions evaluated on all the randomizer bits.

This also motivates the following model [11]: Alice and Bob publicly exchange a random string too long to fit into Eve's memory, and use a scheme similar to that described above, based on a short initially shared secret key. It's security holds based on the sole assumption that Eve's memory capacity is bounded, without assumption about her computing power.

4 Information-theoretic primitives: a general perspective

In both complexity-theoretic and information-theoretic cryptography, an important body of research is devoted to the reduction of one primitive to another primitive, e.g. of a pseudo-random number generator to a one-way function [27] or of oblivious transfer to the existence of a noisy channel [16]. In this section we informally define a general notion of an information-theoretic cryptographic primitive, discuss the general reduction problem among such primitives and possible goals of such a reduction, and show that many results in the literature fit into this general framework.

4.1 Definition of IT-primitives

Definition 2. A (stateless) *information-theoretic cryptographic primitive* (IT-primitive or simply primitive, for short) is an abstractly defined mechanism (which can be viewed as a service offered by a trusted party) to which $n \geq 2$ players $P_1, \ldots, P_n$ have access. For every invocation of the primitive, each player P_i can provide a (secret) input X_i from a certain domain and receives a (secret) output Y_i from a certain range according to a certain (usually publicly known) conditional probability distribution $P_{Y_1,\ldots,Y_n|X_1,\ldots,X_n}$ of the outputs, given the inputs.

As described, different invocations of the primitive are independent, but more generally an IT-primitive can have an internal state: the players can provide inputs and receive outputs in consecutive rounds, where in each round the dependence of the outputs on the inputs and the current state is specified by a conditional probability distribution.

This concept encompasses as special cases a very large class of previously considered cryptographic primitives, including secure message transmission, noisy channels, all types of oblivious transfer, broadcast channels, and secure multi-party computation, as will be explained below. The concept of a secure reduction of one primitive to another primitive will be discussed later.

There are at least two different ways of defining what it means for the players to have incomplete knowledge of the distribution $P_{Y_1,\ldots,Y_n|X_1,\ldots,X_n}$.

- The distribution can be any one in a class of distributions, possibly chosen by an adversary. If such a primitive is used in a protocol, security must be guaranteed for all distributions in the class.
- The distribution is fixed, but some or all players' knowledge about the distribution may be incomplete. This can be modeled by letting each player receive an additional output summarizing his information about the distribution. This extra output can be viewed as part of the regular output and hence this case is covered by the above definition.

4.2 Examples

Let us first consider some IT-primitives for $n = 2$ players, called Alice and Bob.

- A noisy channel from Alice to Bob: Alice can provide an input X from a certain domain (e.g. a bit) and Bob receives the output Y generated according to the conditional distribution $P_{Y|X}$. For instance, in a binary symmetric channel with error rate ϵ, $P_{Y|X}(y,x) = \epsilon$ if $x \neq y$ and $P_{Y|X}(y,x) = 1 - \epsilon$ if $x = y$. The (γ, δ)-unfair noisy channel of [18] is a binary symmetric channel with error probability chosen by the adversary in the interval $[\gamma, \delta]$.
- Oblivious transfer (OT) of any type is a classical example of an IT-primitive. In standard 1-out-of-2 OT, Alice chooses as input two bits (or bit strings), Bob chooses as input a selection bit, and Bob learns as output the corresponding bit (or string), while Alice learns nothing. In the generalized OT of Brassard and Crépeau [8], Bob can choose to receive any binary function

of the two input bits, and this can be generalized further [8, 10] to allow Bob to specify any channel over which he receives the two input bits, with the only constraint that his uncertainty (entropy) about the two input bits be at least γ, for some $0 < \gamma < 2$.
- Damgård, Kilian, and Salvail [18] introduced a more general two-party primitive which they called weak oblivious transfer (WOT). In this model, also Alice receives an output that gives her partial information about Bob's choice.
- A commitment scheme is a primitive with state: Alice inputs a value (which is kept as the state). Later, upon initiation of an opening phase, Alice chooses (with a binary input) whether or not she agrees to open the commitment, and Bob's output is the committed value or a dummy value, respectively. This primitive can also be defined with respect to several players (see below).

Damgård et al. [18] also introduced a two-player primitive called weak generic transfer (WGT) that is similar to the two-player case of our general IT-primitive. However, the models differ in the following way: In WGT, cheating by one of the players is modeled by an additional output which the player receives only when cheating, but not otherwise. Passive cheating means that the player collects this information, without deviating from the protocol, and active cheating means that the player can take this extra information into account during the execution of the protocol. In contrast, cheating is in our definition not considered as part of the primitive, but as misbehavior to be protected against in a protocol in which the primitive is used. The possible assumption that a player receives additional information when cheating can be phrased as a security condition of a protocol in which the primitive is used.

Next we consider primitives for $n = 3$ players which we call Alice, Bob, and Eve, and where it is known in advance that Alice and Bob are honest while Eve is a cheating adversary.

- Key agreement between Alice and Bob: The players have no input[6] and receive outputs Y_A, Y_B, and Y_E, respectively. The specification of the primitive is that Y_A is chosen uniformly at random from the key space, that $Y_B = Y_A$, and that Y_E is some dummy output that is independent of Y_A.
- A noisy random source: there are again no inputs, but the outputs Y_A, Y_B, and Y_E are selected according to a general distribution. This corresponds to the key agreement scenario discussed in the previous section, where the random variables Y_A, Y_B, and Y_E were called X, Y, and Z and generated according to some distribution P_{XYZ}. The distribution of a random deal of cards [22] is also a special case.
- Wyner's wire-tap channel [48] and the generalization due to Csiszár and Körner [17]: A symbol sent by Alice is received by Bob and Eve over two dependent noisy channels.
- Secure message transmission also fits into this framework: Alice's input is received by Bob, Eve receives no output. If the channel is authenticated but not confidential, then Eve also receives the output.

[6] Invocation of the primitive could actually be considered as a special type of input.

- A quantum channel from Alice to Bob can also be modeled as an IT-primitive if the eavesdropper is forced to measure each quantum state independently. For modeling the possibility that Eve could perform general quantum computations on the quantum states, our IT-primitive can be generalized to the quantum world.

We now describe some IT-primitives for general n:

- A broadcast channel: A designated sender provides an input which is received (consistently) by all other $n-1$ players.
- Secure function evaluation for an agreed function: each player provides a secret input and receives the output of a function evaluated on all inputs. The players' output functions can be different.
- Secure multi-party computation [26, 3, 12] among a set of $n \geq 2$ players: here the primitive keeps state. This corresponds to the general paradigm of simulating a trusted party.
- A random oracle can also be interpreted in this framework.

4.3 Reductions among IT-primitives

The general reduction problem can be phrased as follows: assuming the availability of one primitive G (more generally, several primitives $G_1, \ldots, G_s$), can one construct primitive H, even if some of the players cheat, where the type of tolerable cheating must be specified. One can distinguish many types of cheating. Three well-defined cases are active cheater(s) who deviate from the protocol in an arbitrary manner, passive cheaters who follow the protocol but record all information in order to violate other players' privacy, and fail-corrupted cheaters who may stop executing the protocol at any time. Cheating players are usually modeled by assuming a central adversary who may corrupt some of the players.

One generally assumes without essential loss of generality or applicability that insecure communication channels between any pair of entities are available.

Such information-theoretic reductions among primitives are interesting for at least two reasons. First, if a certain primitive exists or can be assumed to exist in nature (e.g. a noisy channel), then it can be used to build practical unconditionally-secure protocols. Second, if a primitive can be realized or approximated cryptographically (e.g. oblivious transfer), then one can construct computationally-secure cryptographic protocols with a well-defined isolated complexity assumption. The relevance of a reduction result depends on at least the following criteria:

- whether the resulting primitive is useful in applications,
- whether the assumed primitive has a natural realization in the physical world, or can efficiently be realized by cryptographic means,
- the efficiency of the reduction, for instance the number of invocations of primitive G needed to realize one occurrence of H?
- the assumption about cheating (e.g. less than $n/3$ cheaters), and
- which additional assumptions are made (e.g., availability of an authenticated channel between certain or all pairs of players).

Informally, a reduction of one primitive to another is secure against an adversary with some specified corruption power if the adversary can do nothing more in the protocol than what he could do in an idealized implementation of the primitive, except possibly with exponentially small probability.

Many results in cryptography can be phrased as a reduction among primitives. Some of them were mentioned above, and a few are listed below:

- Many reduction results for oblivious transfer (e.g. [8, 10, 14–16, 18, 20]).
- Secret-key agreement by public discussion from noisy channels as discussed in the previous section can be interpreted as the reduction of key agreement to a certain type of noisy source.
- Privacy amplification [6, 5], an important sub-protocol in unconditional key agreement, can be interpreted as the reduction of key agreement to a setting in which Alice and Bob share the same string, but where Eve has some arbitrary unknown type information about the string with the only constraint being an upper bound on the total amount of information.
- Byzantine agreement protocols (e.g., see [28]) can be interpreted as the reduction of the broadcast primitive to the primitive of bilateral authenticated channels, assuming active cheating by less than $n/3$ of the players.
- The commitment primitive can be defined for an arbitrary number n of players, one of which commits to an input. Secret sharing can be interpreted as a reduction of such a commitment primitive to the primitive of bilateral secure communication channels, assuming only passive cheating by a non-qualified set of players. Verifiable secret sharing is like secret sharing, but security is guaranteed with respect to active cheaters (e.g. less than $n/3$).
- The results of Ben-Or, Goldwasser and Wigderson [3] and Chaum, Crépeau, and Damgård [12] can be interpreted as the reduction of the primitive secure multi-party computation to the primitive bilateral secure communication channels, assuming active cheating by less than $n/3$ of the players. If also the broadcast primitive is assumed, then less than $n/2$ cheaters can be tolerated [38].

4.4 General transfer primitives

A two-player primitive covering various types of oblivious transfer as special cases can be defined as follows: Alice inputs a random variable X, and Bob can select (by the random variable $C \in \{1, \ldots, n\}$) to receive any one of n random variables $Y_1, \ldots, Y_n$ defined by a conditional distribution $P_{Y_1,\ldots,Y_n|X}$, such that for all (or for certain) distributions P_X Alice learns nothing about C and Bob learns nothing about X beyond what he learns from Y_C.

We refer to an (α, β)-transfer as any transfer of the described type for which for at least one distribution P_X (e.g., the uniform distribution) we have $H(X) = \alpha$ and $H(Y_i) \leq \beta$ for $1 \leq i \leq n$, and assuming the natural condition that X is determined from $Y_1, \ldots, Y_n$, i.e. that X contains no information that is irrelevant in the transfer. A transfer is said to hide at most γ bits if for all distributions P_X and for $1 \leq i \leq n$ we have $H(X|Y_i) \leq \gamma$.

For example, in 1-out-of-n OT of l-bit strings, X is the concatenation of the n input strings, which are $Y_1, \ldots, Y_n$. Such an OT is an (ln, l)-transfer and can easily be shown to hide at most $(n-1)l$ bits. Motivated by Dodis' and Micali's [20] lower bound on reducing weak 1-out-of-N OT of L-bit strings to 1-out-of-n OT of l-bit strings we prove the following more general theorem.

Theorem 2. *The reduction of any (α, β)-transfer to any transfer that hides at most γ bits requires at least $(\alpha - \beta)/\gamma$ invocations of the latter.*

Proof. Let X be the input and Y be the output of the (α, β)-transfer to be realized, and let T be the entire communication taking place during the reduction protocol over the standard communication channel. Let k be the number of invocations of the second transfer, let $U_1, \ldots, U_k$ and $V_1, \ldots, V_k$ be the corresponding k inputs and outputs, and let $U^k = [U_1, \ldots, U_k]$ and $V^k = [V_1, \ldots, V_k]$. Then we have $H(X|V^kT) \geq \alpha - \beta$ (for at least one distribution P_X) because Bob must not learn more than β bits about X, and $H(X|U^kT) = 0$ because unless Alice enters all information about X into the protocol, she will learn something about C. We expand $H(XU^k|V^kT)$ in two different ways:

$$H(XU^k|V^kT) = H(U^k|V^kT) + H(X|U^kV^kT) = H(X|V^kT) + H(U^k|XV^kT),$$

and observe that $H(X|U^kV^kT) \leq H(X|U^kT) = 0$ and $H(U^k|XV^kT) \geq 0$. Applying repeatedly the chain rule and the fact the further conditioning cannot increase entropy, we obtain $\alpha - \beta \leq H(X|V^kT) \leq H(U^k|V^kT) \leq H(U^k|V^k) = \sum_{j=1}^{k} H(U_j|V^k, U_1 \cdots U_{j-1}) \leq \sum_{j=1}^{k} H(U_j|V_j) \leq k\gamma$, and the theorem follows.

5 Generalized random oracles

In this section we briefly sketch the definition of a new general concept, which we call generalized random oracles for lack of a perhaps better name, and describe some applications and constructions.

One motivation for introducing this concept is the fact that many proofs of the computational security of a cryptographic system (e.g. a MAC scheme or a pseudo-random permutation) based on a pseudo-random function (PRF) [25] rely on information-theoretic arguments, although this is not always made explicit in the proofs.

An adversary's attack is modeled as a usually adaptive algorithm for performing certain query operations to the system.[7] The proof of the computational security of such a system consists of two steps: 1) a purely information-theoretic proof that the attack cannot succeed in an idealized setting where the PRF is replaced by a random function, and 2) the simple observation that if the random function is replaced by a PRF, then any efficient successful attack algorithm could be converted into an efficient distinguisher for the PRF, thus contradicting the underlying intractability assumption.

For instance, the Luby-Rackoff construction of a $2n$-bit to $2n$-bit pseudo-random permutation generator [29] (involving three pseudo-random functions

[7] For example, for the case of a MAC, two query operations are allowed: evaluation of the MAC for a chosen message, and verification of a message-MAC pair (yielding a binary output) [2].

from n bits to n bits, combined in a three-round Feistel construction) can be proved secure by showing that no adaptive algorithm querying the permutation for less than a certain (super-polynomial) number of arguments (actually $2^{n/2}$) can distinguish it from a truly random permutation with non-negligible advantage. If the three random functions are replaced by PRFs, the construction is computationally indistinguishable from a random permutation.

We sketch a general approach that allows to simplify and generalize many of the security proofs given in the literature by interpreting them as the proof of indistinguishability of two particular types of generalized random oracles. Some of the proofs can be obtained from a set of simpler information-theoretic arguments which can be formulated as results of independent interest and can serve as a toolkit for new constructions. Some of the proofs that can be revisited are those for the Luby-Rackoff construction [29] and generalizations thereof [31, 36], and the analysis of the CBC MAC [1] and the XOR MAC [2].

Definition 3. A *generalized random oracle (GRO)* is characterized by 1) a set of query operations, each of which takes as input an argument from a certain domain and outputs a corresponding value in a certain range, and 2) a random experiment for which each elementary event in the sample space is a complete set of answers to all possible queries, with some probability distribution over the sample space.

A more operational interpretation of a GRO may help: In many cases a GRO is constructed using an array of k (usually k is exponential in a security parameter) independent random bits. The sample space hence consists of 2^k equally probable elementary events, and each access operation consists of (efficiently) evaluating a certain function involving the k bits.

The simplest form of a GRO is a random function from n bits to 1 bit (hence $k = 2^n$). The (single) query operation evaluates the random function for a given argument. A generalization is obtained by allowing other types of queries, e.g. arbitrary linear combinations of the bits. Allowing outputs of general size (e.g. also n bits as in [29]) entails no essential generalization, except that a new type of binary query exists: for a given input/output pair, do they match?

The concept of locally random functions proposed in [31] also fits into this framework: these are efficient constructions of GROs with $k \ll 2^n$ and a single query operation $\{0,1\}^n \to \{0,1\}$, and which are indistinguishable from a random function for any algorithm accessing them less than, but close to k times.

Definition 4. Two GRO's $\mathcal{A}$ and $\mathcal{B}$ have *compatible access operations* if each operation for $\mathcal{A}$ is compatible with an operation for $\mathcal{B}$ in the sense of having the same input domain and output range. Informally, two GRO's $\mathcal{A}$ and $\mathcal{B}$ with compatible access operations are *perfectly (statistically) indistinguishable* for a given number of access operations of each type if no adaptive algorithm that can access the GROs in the specified manner has different (non-negligibly different) probability of outputting 1 when the queries are answered using $\mathcal{A}$ or $\mathcal{B}$.

Note that there is no restriction on the computing power of the distinguishing algorithm; hence a GRO is an information-theoretic rather than a complexity-theoretic concept.

6 Conclusions

The three main points addressed in this paper are:

– In cryptography and more generally in computer science one generally considers only digital operations. However, all processes in the real world, including computation and communication, are physical processes involving noise and other uncertainty factors. We propose to further investigate and exploit this fact to achieve unconditional security in cryptography.
– A general definition of an information-theoretic cryptographic primitive was proposed which encompasses many primitives previously proposed in the literature and leads to new research problems on the reduction between such primitives.
– A generalized definition of a random oracle has been proposed which has applications for security proofs in complexity-theoretic cryptography.

Acknowledgments

It is impossible to list all the people with whom I have had stimulating discussions on topics related to information-theoretic cryptography. Among them, I would like to thank the following people with whom I have had the pleasure to collaborate on related topics: Charles Bennett, Gilles Brassard, Christian Cachin, Ronald Cramer, Claude Crépeau, Ivan Damgård, Matthias Fitzi, Martin Gander, Martin Hirt, Olaf Keller, and Stefan Wolf. Above all, I am grateful to Jim Massey who introduced me to information theory and cryptography. I also thank Michael Wiener and the CRYPTO 99 program committee for inviting me to give this lecture.

References

1. M. Bellare, J. Kilian, and P. Rogaway, The security of the cipher block chaining message authentication code, *Advances in Cryptology - CRYPTO '94*, Lecture Notes in Computer Science, vol. 839, Springer-Verlag, 1995.
2. M. Bellare, J. Guérin, and P. Rogaway, The security of the cipher block chaining message authentication code, *Advances in Cryptology - CRYPTO '95*, Lecture Notes in Computer Science, vol. 963, Springer-Verlag, 1994.
3. M. Ben-Or, S. Goldwasser, and A. Wigderson, Completeness theorems for non-cryptographic fault-tolerant distributed computation, In *Proc. 20th ACM Symposium on the Theory of Computing (STOC)*, pp. 1–10, 1988.
4. C.H. Bennett, F. Bessette, G. Brassard, L. Salvail, and J. Smolin, Experimental quantum cryptography, *Journal of Cryptology*, vol. 5, no. 1, pp. 3–28, Springer-Verlag, 1992.
5. C.H. Bennett, G. Brassard, C. Crépeau, and U.M. Maurer, Generalized privacy amplification, *IEEE Transactions on Information Theory*, vol. 41, no. 6, pp. 1915–1923, 1995.
6. C.H. Bennett, G. Brassard, and J.-M. Robert, Privacy amplification by public discussion, *SIAM Journal on Computing*, vol. 17, pp. 210–229, 1988.
7. R.E. Blahut, *Principles and practice of information theory*, Addison-Wesley Publishing Company, 1988.

8. G. Brassard and C. Crépeau, Oblivious transfer and privacy amplification, *Advances in Cryptology - EUROCRYPT '97*, Lecture Notes in Computer Science, vol. 1233, pp. 334–345, Springer-Verlag, 1997.
9. C. Cachin, *Entropy measures and unconditional security in cryptography*, Ph. D. Thesis, ETH Zürich, Hartung-Gorre Verlag, Konstanz, 1997.
10. ——, On the foundation of oblivious transfer, *Advances in Cryptology - EUROCRYPT '98*, Lecture Notes in Computer Science, vol. 1403, pp. 361–374, Springer-Verlag, 1998.
11. C. Cachin and U. M. Maurer, Unconditional security against memory-bounded adversaries, *Advances in Cryptology - CRYPTO '97*, Lecture Notes in Computer Science, vol. 1294, pp. 292–306, Springer-Verlag, 1997.
12. D. Chaum, C. Crépeau, and I. Damgård, Multiparty unconditionally secure protocols, In *Proc. 20th ACM Symposium on the Theory of Computing (STOC)*, pages 11–19, 1988.
13. T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley Series in Telecommunications, 1992.
14. C. Crépeau, Equivalence between two flavours of oblivious transfer, *Advances in Cryptology - CRYPTO '87*, Lecture Notes in Computer Science, pp. 350–354, Springer-Verlag, 1988.
15. ——, Efficient cryptographic protocols based on noisy channels, *Advances in Cryptology - EUROCRYPT '97*, Lecture Notes in Computer Science, vol. 1233, pp. 306–317, Springer-Verlag, 1997.
16. C. Crépeau and J. Kilian, Achieving oblivious transfer using weakened security assumptions, *29th Symposium on Foundations of Computer Science*, pp. 42–52, IEEE, 1988.
17. I. Csiszár and J. Körner, Broadcast channels with confidential messages, *IEEE Transactions on Information Theory*, vol. IT-24, pp. 339–348, 1978.
18. I. Damgård, J. Kilian, and L. Salvail, On the (im)possibility of basing oblivious transfer and bit commitment on weakened security assumptions, *Advances in Cryptology - EUROCRYPT '99*, Lecture Notes in Computer Science, vol. 1592, pp. 56–73, Springer-Verlag, 1999.
19. W. Diffie and M. E. Hellman, New directions in cryptography, *IEEE Transactions on Information Theory*, vol. 22, no. 6, pp. 644–654, 1976.
20. Y. Dodis and S. Micali, Lower bounds for oblivious transfer reductions, *Advances in Cryptology - EUROCRYPT '99*, Lecture Notes in Computer Science, vol. 1592, pp. 42–55, Springer-Verlag, 1999.
21. W. Feller, *An introduction to probability theory and its applications*, 3rd edition, vol. 1, Wiley International, 1968.
22. M. J. Fischer and R. N. Wright, Bounds on secret key exchange using a random deal of cards, *Journal of Cryptology*, vol. 9, no. 2, pp. 71–99, Springer-Verlag, 1996.
23. P. Gemmell and M. Naor, Codes for interactive authentication, *Advances in Cryptology - CRYPTO '93*, Lecture Notes in Computer Science, vol. 773, pp. 355–367, Springer-Verlag, 1993.
24. E. N. Gilbert, F. J. MacWilliams, and N. J. A. Sloane, Codes which detect deception, *Bell Syst. Tech. J.*, vol. 53, No. 3, 1974, pp. 405–424.
25. O. Goldreich, S. Goldwasser, and S. Micali, How to construct random functions, *Journal of the ACM*, vol. 33, no. 4, pp. 210–217, 1986.
26. O. Goldreich, S. Micali, and A. Wigderson. How to play any mental game — a completeness theorem for protocols with honest majority. In *Proc. 19th ACM Symposium on the Theory of Computing (STOC)*, pp. 218–229, 1987.
27. J. Håstad, R. Impagliazzo, L. Levin, and M. Luby, Construction of a pseudorandom generator from any one-way function, Technical Report no. 91-068, ICSI, Berkeley, CA, 1991.

28. L. Lamport, R. Shostak, and M. Pease, The Byzantine generals problem, *ACM Transactions on Programming Languages and Systems*, vol. 4, pp. 382–401, 1982.
29. M. Luby and C. Rackoff, How to construct pseudorandom permutations from pseudorandom functions, *SIAM Journal on Computing*, vol. 17, no. 2, pp. 373–386, 1988.
30. U. M. Maurer, Conditionally-perfect secrecy and a provably-secure randomized cipher, *Journal of Cryptology*, vol. 5, pp. 53–66, Springer-Verlag, 1992.
31. ——, A simplified and generalized treatment of Luby-Rackoff pseudo-random permutation generators, *Advances in Cryptology - EUROCRYPT '92*, Lecture Notes in Computer Science, vol. 658, pp. 239–255, Springer-Verlag, 1992.
32. ——, Secret key agreement by public discussion from common information, *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 733–742, 1993.
33. ——, A unified and generalized treatment of authentication theory, *Proceedings 13th Symp. on Theoretical Aspects of Computer Science - STACS '96*, Lecture Notes in Computer Science, vol. 1046, pp. 387–398, Springer-Verlag, 1996.
34. ——, Information-theoretically secure secret-key agreement by NOT authenticated public discussion, *Advances in Cryptology - EUROCRYPT '97*, Lecture Notes in Computer Science, vol. 1233, pp. 209–225, Springer-Verlag, 1997.
35. U. M. Maurer and S. Wolf, Unconditionally secure key agreement and the intrinsic conditional information, *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 499–514, 1999.
36. M. Naor and O. Reingold, On the construction of pseudorandom permutations: Luby-Rackoff revisited, *Journal of Cryptology*, vol. 12, no. 1, pp. 29–66, 1999.
37. M. O. Rabin, personal communication, 1998.
38. T. Rabin and M. Ben-Or, Verifiable secret sharing and multiparty protocols with honest majority, *Proc. 21st ACM Symposium on the Theory of Computing (STOC)*, pp. 73–85, 1989.
39. C. E. Shannon, Communication theory of secrecy systems, *Bell System Technical Journal*, vol. 28, pp. 656–715, 1949.
40. ——, A mathematical theory of communication, *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, 1948.
41. G. J. Simmons, A survey of information authentication, *Proceedings of the IEEE*, vol. 76, pp. 603–620, 1988.
42. D. R. Stinson, Universal hashing and authentication codes, *Advances in Cryptology - CRYPTO '91*, Lecture Notes in Computer Science, vol. 576, pp. 74–85, Springer-Verlag, 1992.
43. G. S. Vernam, Cipher printing telegraph systems for secret wire and radio telegraphic communications, *Journal of the American Institute for Electrical Engineers*, vol. 55, pp. 109–115, 1926.
44. M. N. Wegman and J. L. Carter, New hash functions and their use in authentication and set equality, *Journal of Computer and System Sciences*, vol. 22, pp. 265-279, 1981.
45. S. Wolf, Strong security against active attacks in information-theoretic secret-key agreement, *Advances in Cryptology - ASIACRYPT '98*, Lecture Notes in Computer Science, vol. 1514, pp. 405–419, Springer-Verlag, 1998.
46. ——, Unconditional security in cryptography, *Proceedings of Summer School in Cryptology and Data Security*, July 1998, Aarhus, Denmark, Lecture Notes in Computer Science, vol. 1561, pp. 217–250, Springer-Verlag, 1999.
47. ——, Information-theoretically and unconditionally secure key agreement in cryptography, Ph.D. Thesis no. 13138, ETH Zurich, 1999.
48. A. D. Wyner, The wire-tap channel, *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.