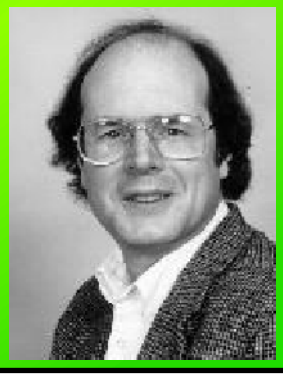
Generalized Privacy Amplification



Charles H. Bennett

IBM Research

Gilles Brassard

Université de Montréal

Claude Crépeau

École Normale Supérieure

Ueli M. Maurer

ETH Zürich

Generalized Privacy Amplification

Privacy amplification is the art of distilling highly secret shared information, perhaps for use as a cryptographic key, from a larger body of shared information that is only partially secret.

Generalized Privacy Amplification

Let Alice and Bob be given a random variable W, such as a random n–bit string, while an eavesdropper Eve learns a correlated random variable V, providing at most $t < n$ bits of information about W, i.e. $H(W|V) \geq n - t$. The details of the distribution P_{VW} are generally unknown to Alice and Bob, except that it satisfies this constraint as well as possibly some further constraints. They may or may not know P_W.

Generalized Privacy Amplification

Alice and
Bob wish to publicly choose a compression function $g : \{0,1\}^n \to \{0,1\}^r$ such that Eve's partial information on W and her complete information on g give her arbitrarily little information about $K = g(W)$, except with negligible probability (over possible choices for g).

The resulting K is virtually uniformly distributed given all Eve's information;
it can hence be used safely as a cryptographic key.

Generalized Privacy Amplification

The size r of the secret that Alice and Bob can distill depends on the kind as well as the amount of information available to Eve. Assuming that W is a random n–bit string, various possible scenarios to consider are that Eve can obtain

(1) t arbitrary bits of W,

(2) t arbitrary parity checks of W,

(3) the result of an arbitrary function mapping n–bit strings to t–bit strings, or

(4) the string W transmitted through a binary symmetric channel with bit error probability ε satisfying $h(\varepsilon) = 1 - t/n$, and hence with capacity t/n, where $h(\cdot)$ denotes the binary entropy function.

IV. Universal hashing and Rényi entropy

Definition 1 [13]**.** A class $\mathcal{G}$ of functions $\mathcal{A} \to \mathcal{B}$ is *universal*$_2$ ("universal" for short) if, for any distinct x_1 and x_2 in $\mathcal{A}$, the probability that $g(x_1) = g(x_2)$ is at most $1/|\mathcal{B}|$ when g is chosen at random from $\mathcal{G}$ according to the uniform distribution.

IV. Universal hashing and Rényi entropy

Definition 2. Let X be a random variable with alphabet $\mathcal{X}$ and distribution P_X. The *collision probability* $P_c(X)$ of X is defined as the probability that X takes on the same value twice in two independent experiments:

$$P_c(X) = \sum_{x \in \mathcal{X}} P_X(x)^2.$$

IV. Universal hashing and Rényi entropy

Definition 2.
The *Renyi entropy of order two* ("Rényi entropy" for short) of X [15], [27] is defined as the negative logarithm of the collision probability of X:

$$R(X) = -\log_2 P_c(X).$$

IV. Universal hashing and Rényi entropy

Definition 2.
For an event $\mathcal{E}$, the collision probability and the Rényi entropy of X conditioned on $\mathcal{E}$, $P_c(X|\mathcal{E})$ and $R(X|\mathcal{E})$, are defined naturally as the collision probability and the Rényi entropy, respectively, of the conditional distribution $P_{X|\mathcal{E}}$. The Rényi entropy conditioned on a random variable, $R(X|Y)$, is the expected value of the conditional Rényi entropy:

$$R(X|Y) = \sum_y P_Y(y)\, R(X|Y=y).$$

IV. Universal hashing and Rényi entropy

In order to contrast Rényi entropy with the standard entropy measure defined by Shannon, we will refer to the latter as "Shannon entropy" throughout the paper. Note that Rényi entropy (like Shannon entropy) is always positive. $R(X)$ can equivalently be expressed as $R(X) = -\log_2 E[P_X(X)]$, where $E[\cdot]$ denotes the expected value. Shannon entropy $H(X)$ can be expressed similarly as $H(X) = -E[\log_2 P_X(X)]$.

IV. Universal hashing and Rényi entropy

Lemma 2. *For every discrete probability distribution* P_X,

$$R(X) \leq H(X),$$

with equality if and only if P_X *is the uniform distribution over* $\mathcal{X}$ *or a subset of* $\mathcal{X}$. *Moreover, for every distribution* P_{XY},

$$R(X|Y) \leq H(X|Y).$$

IV. Universal hashing and Rényi entropy

At first it seems natural to extend the analogy between Rényi and Shannon entropies to the notion of information. In other words, it is tempting to define the mutual Rényi information between X and Y to be $I_R(X;Y) = R(X) - R(X|Y)$. However, this notion is not symmetric as $R(X) - R(X|Y)$ is different from $R(Y) - R(Y|X)$ in general. Moreover, $R(X) - R(X|Y)$ can be negative, as we shall see in Section VI.

IV. Universal hashing and Rényi entropy

Theorem 3. *Let X be a random variable over the alphabet $\mathcal{X}$ with probability distribution P_X and Rényi entropy $R(X)$, let G be the random variable corresponding to the random choice (with uniform distribution) of a member of a universal class of hash functions $\mathcal{X} \to \{0,1\}^r$, and let $Q = G(X)$. Then*

$$H(Q|G) \;\geq\; R(Q|G) \;\geq\; r - \log_2(1 + 2^{r-R(X)}) \;\geq\; r - \frac{2^{r-R(X)}}{\ln 2}.$$

Note that G is a random variable and the quantity $H(Q|G) = H(G(X)|G)$ is an average over all choices of the function g. It is possible that even when $R(X) \gg r$, $H(G(X)|G = g) = H(g(X))$ differs from r by a non-negligible amount for some g, but such a g can occur only with negligible probability.

IV. Universal hashing and Rényi entropy

Proof. The first inequality follows from Lemma 2. The other two inequalities are proved as follows:

$$\begin{aligned} R(G(X)|G) &= \sum_g P_G(g) R(G(X)|G=g) \\ &= \sum_g P_G(g)(-\log_2 P_c(G(X)|G=g)) \\ &\geq -\log_2 \left(\sum_g P_G(g) P_c(G(X)|G=g) \right) \end{aligned} \tag{2}$$

where the last step follows from Jensen's inequality. The sum in the last term is equal to the probability that $g(x_1) = g(x_2)$ if g is chosen at random according to P_G and x_1 and x_2 are chosen at random, independently of each other and of g, according to P_X. Therefore

IV. Universal hashing and Rényi entropy

$\Sigma_g \Pr_G(g) P_c(G(X)|G=g)$

(my mistake was to expand the sum on x instead of y=g(x).)

$=\Sigma_g \Pr_G(g) \Sigma_y (\Pr_Y(G(X)=y|G=g))^2$

$=\Sigma_g \Pr_G(g) \Sigma_y \Pr_Y[g(X_1)=y] \Pr_Y[g(X_2)=y]$

$=\Sigma_g \Pr_G(g) \Sigma_y \Pr_Y[g(X_1)=y \text{ and } g(X_2)=y]$

$=\Sigma_g \Pr_G(g) \Pr[g(X_1)=g(X_2)]$

$=\Pr[G(X_1)=G(X_2)]$

IV. Universal hashing and Rényi entropy

we have

$$\begin{aligned}
\sum_g P_G(g) P_c(G(X)|G=g) &= \text{Prob}[G(X_1)=G(X_2)] \\
&= \text{Prob}[X_1=X_2] + \\
&\quad \text{Prob}[X_1 \neq X_2] \cdot \text{Prob}[G(X_1)=G(X_2) \mid X_1 \neq X_2] \\
&\leq P_c(X) + (1-P_c(X)) \cdot 2^{-r} \\
&< 2^{-R(X)} + 2^{-r} \\
&= 2^{-r}\left(1+2^{r-R(X)}\right). \qquad (3)
\end{aligned}$$

Here the first inequality follows from the fact that the class of functions is universal and by noting that $1/|\mathcal{B}| = 2^{-r}$ according to Definition 1. The second and third inequalities of the theorem now follow immediately from (2) by taking logarithms on both sides of (3), and from the inequality $\log_2(1+y) \leq y/\ln 2$, respectively. □

IV. Universal hashing and Rényi entropy

Corollary 4. *Let P_{VW} be an arbitrary probability distribution and let v be a particular value of V observed by Eve. If Eve's Rényi entropy $R(W|V=v)$ about W is known to be at least c and Alice and Bob choose $K = G(W)$ as their secret key, where G is chosen at random from a universal class of hash functions from $\mathcal{W}$ to $\{0,1\}^r$, then*

$$H(K|G, V=v) \;\geq\; r - \log_2(1+2^{r-c}) \;\geq\; r - \frac{2^{r-c}}{\ln 2}.$$

IV. Universal hashing and Rényi entropy

Thus we see that when $r < c$, Eve's entropy of the secret key K is close to maximal, i.e. her distribution of K is close to uniform. In particular, her information about K, namely $H(K) - H(K|G, V = v)$, is arbitrarily small. More precisely, her total information about K decreases exponentially in the excess compression $c - r$. It should be pointed out that Corollary 4 cannot be generalized to Rényi entropy conditioned on a random variable, i.e. both

$$H(K|GV) \;\geq\; r - \log_2\left(1 + 2^{r-R(W|V)}\right)$$

as well as the weaker inequality

$$H(K|GV) \;\geq\; r - \frac{2^{r-R(W|V)}}{\ln 2}$$

are false in general. However, if the probability is at least $1 - \delta$ that V takes on a value v satisfying $R(W|V = v) \geq c$, then we have

$$H(K|GV) \;\geq\; (1-\delta)(r - \log_2(1 + 2^{r-c})).$$

and therefore

$$I(K;GV) \;\leq\; \delta r + (1-\delta)\log_2(1 + 2^{r-c}) \leq \delta r + 2^{r-c}/\ln 2.$$

IV. Universal hashing and Rényi entropy

Corollary 5. *Let W be a random n–bit string with uniform distribution over $\{0,1\}^n$, let $V = e(W)$ for an arbitrary eavesdropping function $e : \{0,1\}^n \to \{0,1\}^t$ for some $t < n$, let $s < n - t$ be a positive safety parameter, and let $r = n - t - s$. If Alice and Bob choose $K = G(W)$ as their secret key, where G is chosen at random from a universal class of hash functions from $\{0,1\}^n$ to $\{0,1\}^r$, then Eve's expected information about the secret key K, given G and V, satisfies*

$$I(K; GV) \;\leq\; 2^{-s}/\ln 2.$$

Note that, in contrast to Corollary 4, this result is (and must be) stated as an average over the values of V. Note also that Alice's and Bob's strategy does not depend on e and hence privacy amplification works even if they have no information about e, provided they know an upper bound on t.

V. The gap between Rényi and Shannon entropy

Let us now investigate the implication of Corollary 4 in a genuine case of probabilistic information. Assume that Alice and Bob share a random n–bit string W (uniformly distributed over the n–bit strings) and that Eve can receive the output V when W is sent through a binary symmetric channel with bit error probability ε. Hence we have

$$P_{W|V=v}(w) = (1-\varepsilon)^{n-d(v,w)}\varepsilon^{d(v,w)},$$

where $d(v,w)$ is the Hamming distance between v and w. It is easy to check that Rényi entropy, like Shannon entropy, is additive for independent random variables. It follows that

$$H(W|V=v) = nh(\varepsilon) = -n(\varepsilon \log_2 \varepsilon + (1-\varepsilon)\log_2(1-\varepsilon))$$

and

$$R(W|V=v) = -n\log_2((1-\varepsilon)^2 + \varepsilon^2)$$

for all v.

V. The gap between Rényi and Shannon entropy

Example: Consider an example relevant to quantum cryptography [2]: $\varepsilon = \sin^2 22\frac{1}{2}^\circ \approx 0.15$. In this case, $R(W) - R(W|V=v) \approx 0.585\,n$ whereas $H(W) - H(W|V=v) \approx 0.399\,n$ for all v. Note that these bounds also apply to the averages over choices for v: $R(W) - R(W|V) \approx 0.585\,n$ and $H(W) - H(W|V) \approx 0.399\,n$. Observe that this eavesdropping strategy reduces Eve's Rényi entropy significantly more than it does her Shannon entropy.

V. The gap between Rényi and Shannon entropy

It is instructive to contrast the described eavesdropping scenario with the following eavesdropping strategy, also relevant to quantum cryptography. If Eve could obtain $\tilde{V}$ consisting of $n/2$ arbitrary bits of W of her choice, this would reduce her Rényi entropy by $R(W) - R(W|\tilde{V} = \tilde{v}) = 0.5\, n$ bits and would also reduce her Shannon entropy by $H(W) - H(W|\tilde{V} = \tilde{v}) = 0.5\, n$ for all $\tilde{v} \in \{0,1\}^{n/2}$. Note that in these examples we have $H(W|V = v) = H(W|V)$ and $R(W|V = v) = R(W|V)$ for all v, and $H(W|\tilde{V} = \tilde{v}) = R(W|\tilde{V} = \tilde{v}) = H(W|\tilde{V}) = R(W|\tilde{V})$ for all $\tilde{v}$. Therefore

$$R(W) - R(W|V) \;>\; R(W) - R(W|\tilde{V})$$

whereas

$$H(W) - H(W|V) \;<\; H(W) - H(W|\tilde{V}).$$

VI. Auxiliary random variables

Our goal is to leave Eve with negligible Shannon information about the secret key. By virtue of Lemma 2 we know that this can be accomplished by making her Rényi entropy close to maximal, but this may be overkill. To illustrate this, consider a random variable W chosen with uniform distribution over the n–bit strings, and let Eve's distribution $P_{W|V=v}$ over the n–bit strings be defined by

$$P_{W|V=v}(w) \;=\; \begin{cases} 2^{-n/4} & \text{if } w = v \\ \dfrac{1-2^{-n/4}}{2^n-1} & \text{otherwise.} \end{cases}$$

Although $P_c(W|V=v) > (2^{-n/4})^2 = 2^{-n/2}$ and hence $R(W|V=v) < n/2$ is far from maximal, it is straightforward to check that $H(W|V=v) > n(1-2^{-n/4})$ and hence Eve has an exponentially small amount of information about W. Therefore $K = W$ can be used directly as secret key, with no need to sacrifice more than half the key length to privacy amplification, as Corollary 4 would suggest.

VI. Auxiliary random variables

This example also illustrates a counter-intuitive property of Rényi entropy, which we are going to exploit in the sequel. Unlike Shannon entropy, Rényi entropy can increase when it is conditioned on a random variable. In other words,

$$R(X|Y) > R(X)$$

is possible. In the previous example, consider an oracle who gives Eve for free the random variable U defined by $U = 0$ if $w = v$ and $U = 1$ otherwise. With probability $2^{-n/4}$, this gives Eve complete information about W, and her Rényi entropy falls from roughly $n/2$ bits to 0 bits: $R(W|U = 0, V = v) = 0$. However, with overwhelming complementary probability $1 - 2^{-n/4}$, we have $R(W|U = 1, V = v) = \log_2(2^n - 1)$ and hence

$$\begin{aligned} R(W|U, V = v) &= (1 - 2^{-n/4}) \log_2(2^n - 1) \\ &\geq (1 - 2^{-n/4})(n - 2^{1-n}) \\ &> n - n2^{-n/4} - 2^{1-n} \end{aligned}$$

provided $n \geq 1$, which is approximately twice as large as the unconditioned Rényi entropy $R(W|V = v)$.

VI. Auxiliary random variables

Corollary 7. *Let W, V, G and K be defined as in Corollary 4 and let U be another random variable, jointly distributed with W and V according to some distribution P_{UVW} for which the marginal distribution of $[V,W]$ coincides with P_{VW}. Then*

$$H(K|G,V=v) \;\geq\; r - \sum_u P_{U|V}(u,v)\cdot\log_2\left(1+2^{r-R(W|U=u,V=v)}\right)$$

and

$$H(K|GV) \;\geq\; r - \sum_{u,v} P_{UV}(u,v)\cdot\log_2\left(1+2^{r-R(W|U=u,V=v)}\right).$$

In particular, $H(K|G,V=v)$ and $H(K|GV)$ are lower bounded by the maximum of the respective right-hand sides over choices of P_{UVW} consistent with the given P_{VW}.

VI. Auxiliary random variables

Theorem 8. *For all positive ε and γ, there exists a positive α such that if Alice and Bob share a random n–bit string W, which Eve receives through a binary symmetric channel with bit error probability ε, and if they apply privacy amplification with universal hashing to obtain an r–bit string K where $r = \lfloor (h(\varepsilon) - \gamma)n \rfloor$, then for all sufficiently large n, Eve's expected information about K is at most $2^{-\alpha n}$ bits.*

VI. Auxiliary random variables

Consider the auxiliary random variable $U = d(W, v)$ consisting of the Hamming distance between W and the particular value v known to Eve. Given $U = u$, all $\binom{n}{u}$ strings w at distance u from v are equally likely candidates for W, i.e.

$$R(W|U = u, V = v) \;=\; \log_2 \binom{n}{u}.$$

For any λ between 0 and 1, we have

$$\binom{n}{\lambda n} \geq \frac{2^{nh(\lambda)}}{\sqrt{8n\lambda(1-\lambda)}} \geq \frac{2^{nh(\lambda)}}{\sqrt{2n}}. \tag{4}$$

The left inequality is proved as Lemma 7 in [22, p. 309] and the second inequality follows from $8\lambda(1-\lambda) \leq 2$.

VI. Auxiliary random variables

Consider now an arbitrarily small positive constant $\delta \leq \varepsilon$. By the law of large numbers we have

$$(\varepsilon - \delta)n < d(W, v) < (\varepsilon + \delta)n,$$

and hence also

$$\binom{n}{u} > \binom{n}{(\varepsilon - \delta)n},$$

except with probability exponentially small in n. Hence inequality (4) implies that

$$\binom{n}{u} > \frac{2^{nh(\varepsilon-\delta)}}{\sqrt{2n}},$$

and hence that

$$R(W|U = u, V = v) \;>\; nh(\varepsilon - \delta) - \log_2 \sqrt{2n},$$

except with probability exponentially small in n.

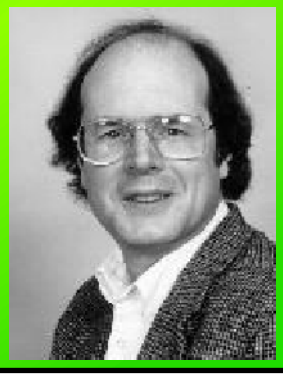
VI. Auxiliary random variables

Consider now any fixed $\gamma > 0$ and let
$\delta > 0$ be so that $h(\varepsilon - \delta) > h(\varepsilon) - \gamma/3$. We conclude that

$$R(W|U = u, V = v) \; > \; (h(\varepsilon) - \gamma/2)n$$

for all sufficiently large n, except with probability exponentially small in n. This exponentially small probability cannot contribute more than an exponentially small amount of Shannon information for Eve.

Generalized Privacy Amplification



Charles H. Bennett

IBM Research

Gilles Brassard

Université de Montréal

Claude Crépeau

École Normale Supérieure

Ueli M. Maurer

ETH Zürich

How good is this bound ???

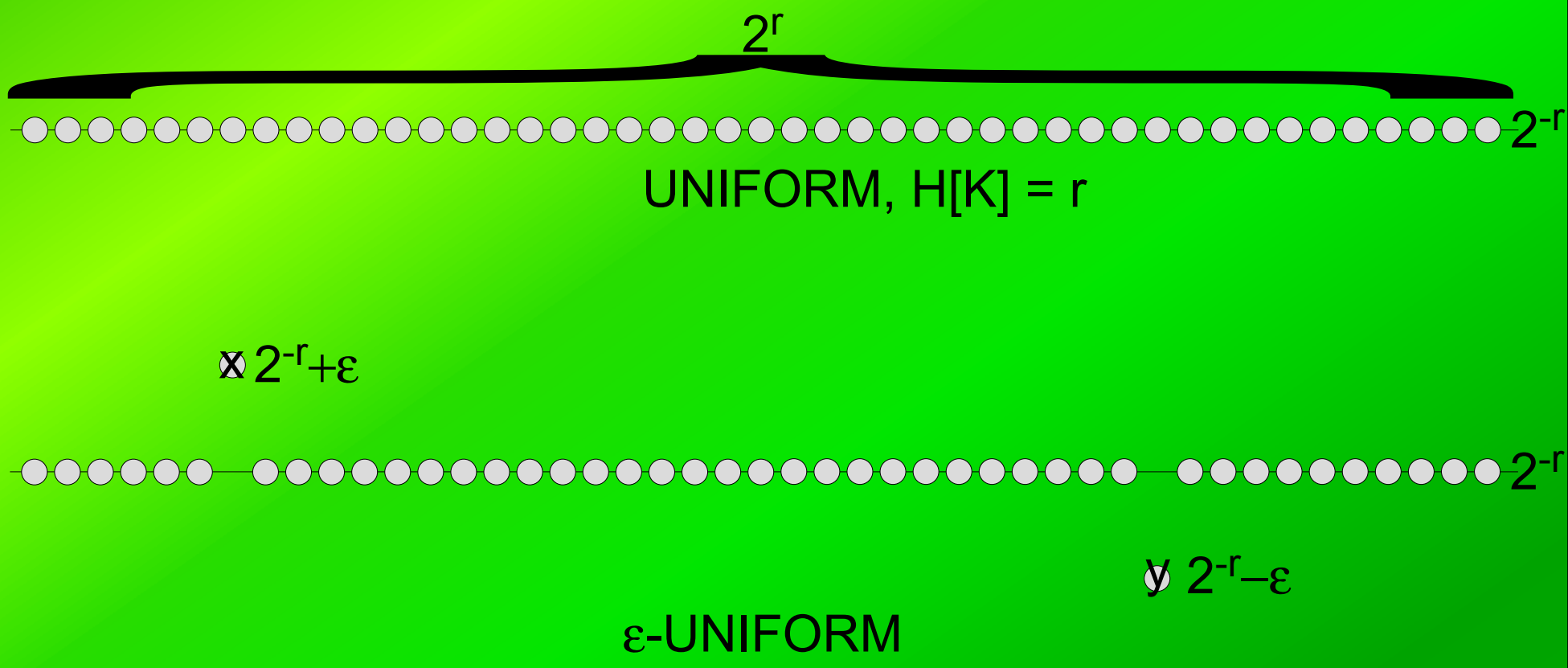
Corollary 4. *Let P_{VW} be an arbitrary probability distribution and let v be a particular value of V observed by Eve. If Eve's Rényi entropy $R(W|V=v)$ about W is known to be at least c and Alice and Bob choose $K = G(W)$ as their secret key, where G is chosen at random from a universal class of hash functions from $\mathcal{W}$ to $\{0,1\}^r$, then*

$$H(K|G, V=v) \;\geq\; r - \log_2(1+2^{r-c}) \;\geq\; r - \frac{2^{r-c}}{\ln 2}.$$

How good is this as a One-Time Pad ???

$$H(K|G, V = v) \geq r - \frac{2^{r-c}}{\ln 2}$$

Definition 6.17 (Perfect Secrecy) *Let P, K, C be random variables describing the uncertainty about the plaintext, ciphertext and key of a cryptosystem $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$. We say that a cryptosystem has perfect secrecy if* $\forall_{p \in \mathcal{P}, c \in \mathcal{C}} \Pr_{P,C}(p|c) = \Pr_P(p)$


2^r
2^{-r}
UNIFORM, H[K] = r
x $2^{-r}+\varepsilon$
2^{-r}
y $2^{-r}-\varepsilon$
ε-UNIFORM

$\exists_{r.v.}$ P={x (1/2),y (1/2)} such that H[P]=1 and

$Pr[P=x|C=0] = Pr[K=x] = 2^{-r}+2^{-\gamma r}$

$Pr[P=x|C\in C_x=x(+)K_x] = Pr[K\in K_x] = 2^{r-1}(2^{-r}+2^{-\gamma r}) = 1/2 + 2^{(1-\gamma)r-1}$

$Pr[P=x|C\in C_y=x(+)K_y] = Pr[K\in K_y] = 2^{r-1}(2^{-r}-2^{-\gamma r}) = 1/2 - 2^{(1-\gamma)r-1}$

$H[P|C] = H[P|C\in C_x] = H[P|C\in C_y] = h(1/2 - 2^{(1-\gamma)r-1}, 1/2 + 2^{(1-\gamma)r-1})$

$H[P|C] \approx$