

Прикладная Криптография: Симметричные криптосистемы Псевдослучайные функции

Макаров Артём
МИФИ 2018

PRP и PRF

Пусть функция $F: K \times X \rightarrow Y$ определена на (K, X, Y) .

Тогда F – **псевдослучайная функция (PRF)**, если существует эффективный алгоритм, вычисляющий $F(k, m)$, $k \in K, x \in X$.

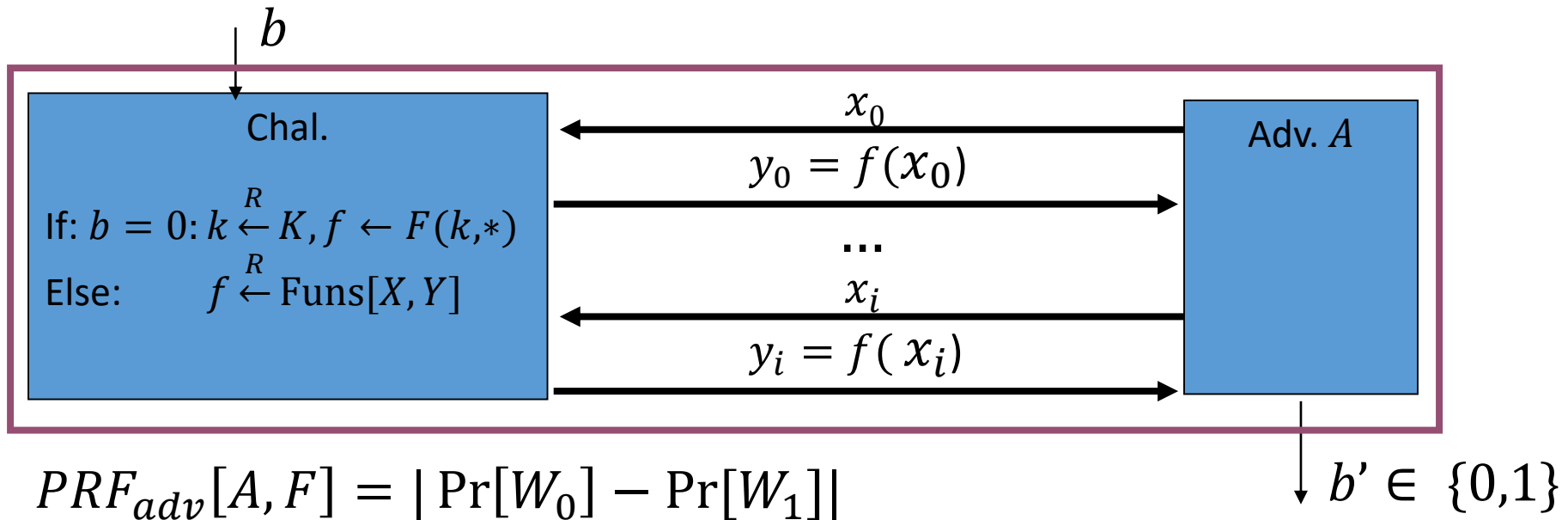
Пусть функция $E: K \times X \rightarrow X$ определена на (K, X) .

Тогда E – **псевдослучайная подстановка (PRP)**, если

- Существует эффективный алгоритм вычисляющий $E(k, x)$. $k \in K, x \in X$
- Функция $f_k = E(k, *)$ – подстановка.

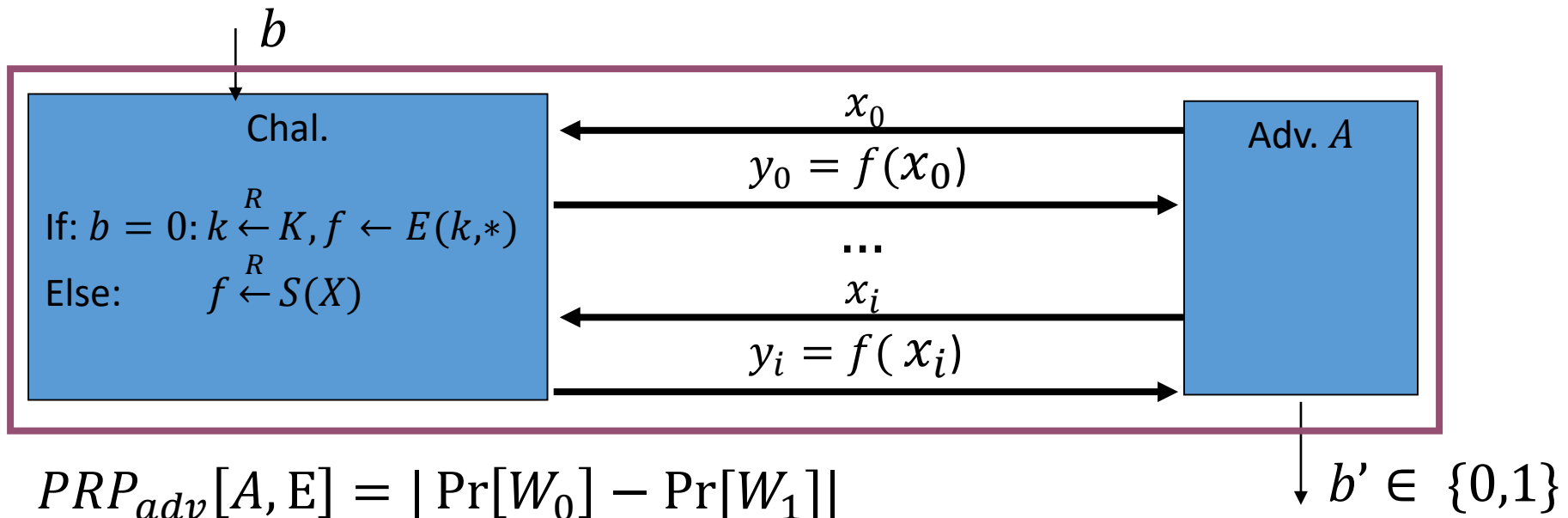
Стойкая PRF

PRF F , определённая на (K, X) , называется стойкой PRF, если $\forall A: A$ – эффективный алгоритм в игре на стойкость PRF величина $PRF_{adv}[A, F] \leq \epsilon$, где ϵ – пренебрежимо малая величина.



Стойкая PRP

PRP E , определённая на (K, X) , называется стойкой PRP, если $\forall A$: A – эффективный алгоритм в игре на стойкость PRP величина $PRP_{adv}[A, E] \leq \epsilon$, где ϵ – пренебрежимо малая величина.



Является ли PRP PRF?

- Является ли любая PRP также PRF?
 - Да, любая эффективная подстановка является эффективной функцией
- Является ли любая стойкая PRP стойкой PRF?
 - Нет!

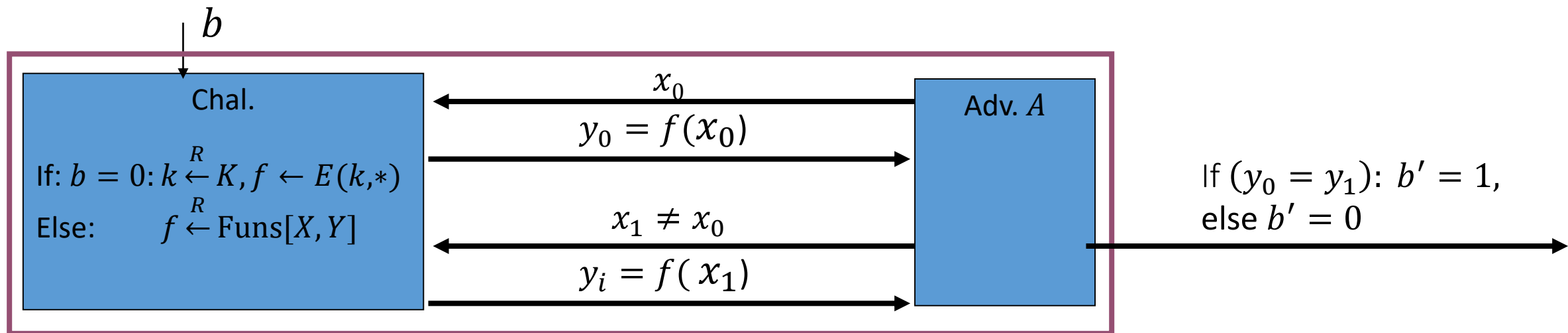
Пусть $E: K \times X \rightarrow X$ – стойкая PRP, $|X| = N$. Очевидно, что E – PRF на (K, X, X) .

Является ли PRP PRF?

См. парадокс дней рождений

Пусть $E: K \times X \rightarrow X$ – стойкая PRP, $|X| = N$. Очевидно, что E – PRF на (K, X, X) .

Рассмотрим игру на PRF. Пусть N – малая величина, такая что противник может эффективно получить полный образ произвольной функции, с областью определения в X (т.е. получить множество $\{f(x): x \in X\}$ для $f: X \rightarrow Y$).



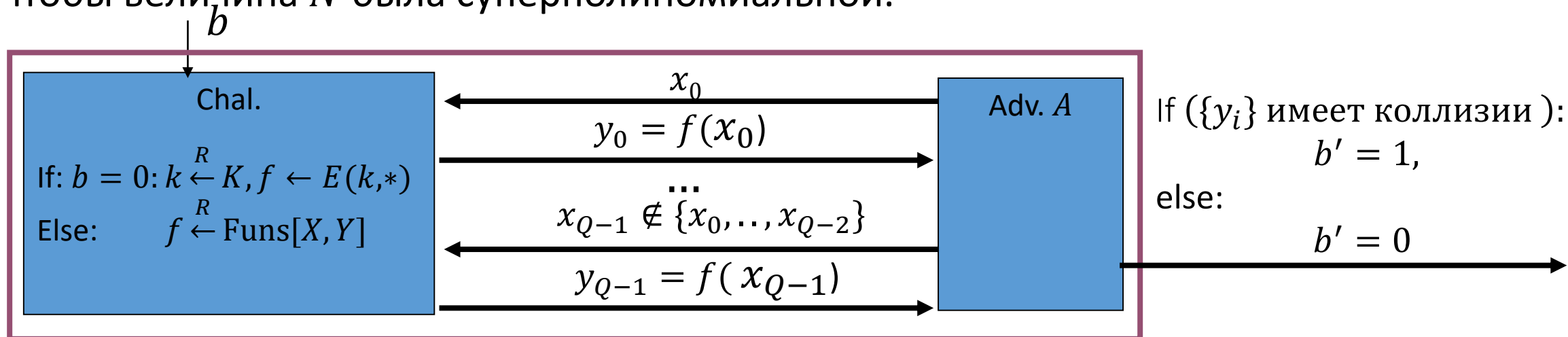
$$PRF_{adv}[A, F] = |\Pr[W_0] - \Pr[W_1]| = |0 - (1 - N!/N^N)| > 1/2$$

Является ли PRP PRF?

См. парадокс дней рождений

Какова «малость» N для осуществления атаки?

Пусть противник делает Q запросов к претенденту (оракулу), прежде чем выдать результат. Тогда для нахождения коллизии ему необходимо запросить $O(N^{1/2})$ различных сообщений, для осуществления атаки с преимуществом $\min\{Q(Q-1)/4N, 0.63\}$. Следовательно для стойкости PRP как PRF необходимо, чтобы величина N была суперполиномиальной.



$$PRF_{adv}[A, F] = |\Pr[W_0] - \Pr[W_1]| = |0 - \min\{Q(Q-1)/4N, 0.63\}|$$

PRF switching Lemma

Теорема 6.1. Пусть $E: K \times X \rightarrow X$ – стойкая PRP, $|X| = N$. Пусть A – противник в игре на стойкость PRF, делающих не более Q запросов к претенденту. Тогда

$$|PRP_{adv}[A, E] - PRP_{adv}[A, E]| \leq Q^2 / 2N$$

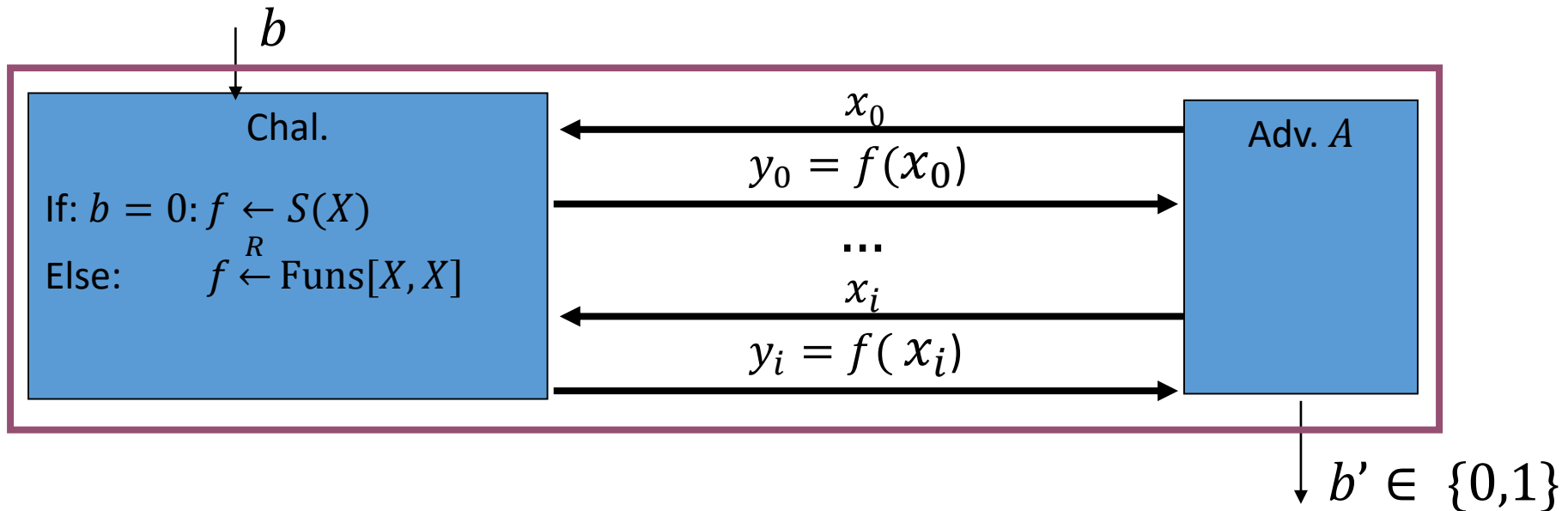
▷ Необходима вспомогательная теорема. Сформулируем игру для неё.

Игра Подстановки против Функций

Определим игру на различимость случайной подстановки от случайной функции.

Пусть W_b вероятность того что в эксперименте b $b' = 1$.

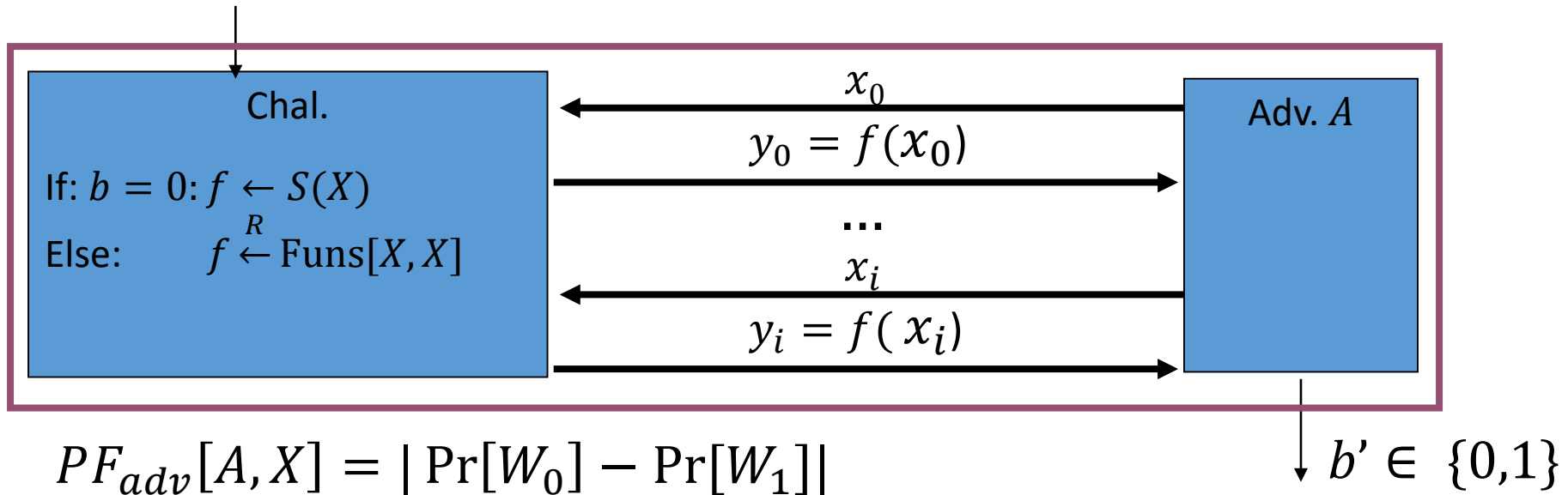
Определим $PF_{adv}[A, X] = |\Pr[W_0] - \Pr[W_1]|$



PRF switching Lemma

Теорема 6.1.1. Пусть X конечное множество, $|X| = N$. Пусть A противник в игре на различимость случайных функций и случайных подстановок. Тогда

$$PF_{adv}[A, X] \leq Q^2/2N$$



PRF switching Lemma



Теорема 6.1.1. Пусть X конечное множество, $|X| = N$. Пусть A противник в игре на различимость случайных функций и случайных подстановок. Тогда

$$PF_{adv}[A, X] \leq Q^2/2N$$

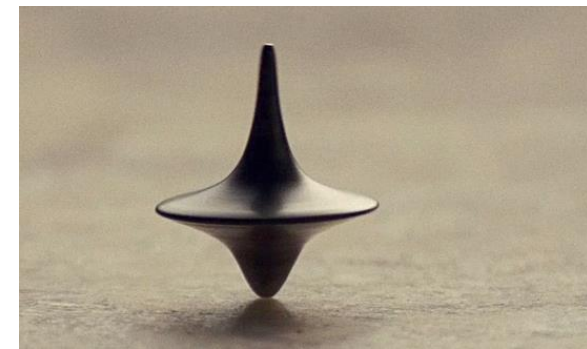
необходима ещё одна вспомогательная теорема.

Теорема 6.1.1.1. Пусть Z, W_0, W_1 события над некоторым вероятностным пространством. Пусть $W_0 \wedge !Z$ происходит тогда и только тогда когда происходит $W_1 \wedge !Z$. Тогда $|\Pr[W_0] - \Pr[W_1]| \leq \Pr[Z]$

$$\% \Pr[W_b \wedge Z] \in [0, \Pr[Z]]$$

$$|\Pr[W_0] - \Pr[W_1]| = |\Pr[W_0 \wedge Z] + \Pr[W_0 \wedge !Z] - \Pr[W_1 \wedge Z] +$$

PRF switching Lemma



Теорема 6.1.1. Пусть X конечное множество, $|X| = N$. Пусть A противник в игре на различимость случайных функций и случайных подстановок. Тогда

$$PF_{adv}[A, X] \leq Q^2/2N$$

Рассмотрим противника A в игре на различимость RP и RF. Пусть он отправляет Q различных запросов x_1, \dots, x_q .

Пусть Z событие того, что $f(x_i) = f(x_j), i \neq j$ (пара ответов оракула совпала). Если событие Z не произошло, то все величины $f(x_i)$ различны, и игры 0 и 1 идентичны. Следовательно, противник будет иметь идентичный результат в обеих играх. Следовательно можем применить Теорему 6.1.1.1. $|\Pr[W_0] - \Pr[W_1]| \leq \Pr[Z] \leq Q^2/2N$.

$\Pr[f(x_i) = f(x_j)] = 1/N$, число таких пар в игре не больше $Q^2/2$. #

PRF switching Lemma

Теорема 6.1. Пусть $E: K \times X \rightarrow X$ – стойкая PRP, $|X| = N$. Пусть A – противник в игре на стойкость PRF, делающих не более Q запросов к претенденту. Тогда

$$|PRP_{adv}[A, E] - PRP_{adv}[A, E]| \leq Q^2/2N$$

$$PF_{adv}[A, X] \leq Q^2/2N.$$

Рассмотрим игру с тремя экспериментами.



Игра на различимость RF, RP и PRP

Пусть W_b вероятность того что в эксперименте b $b' = 1$.

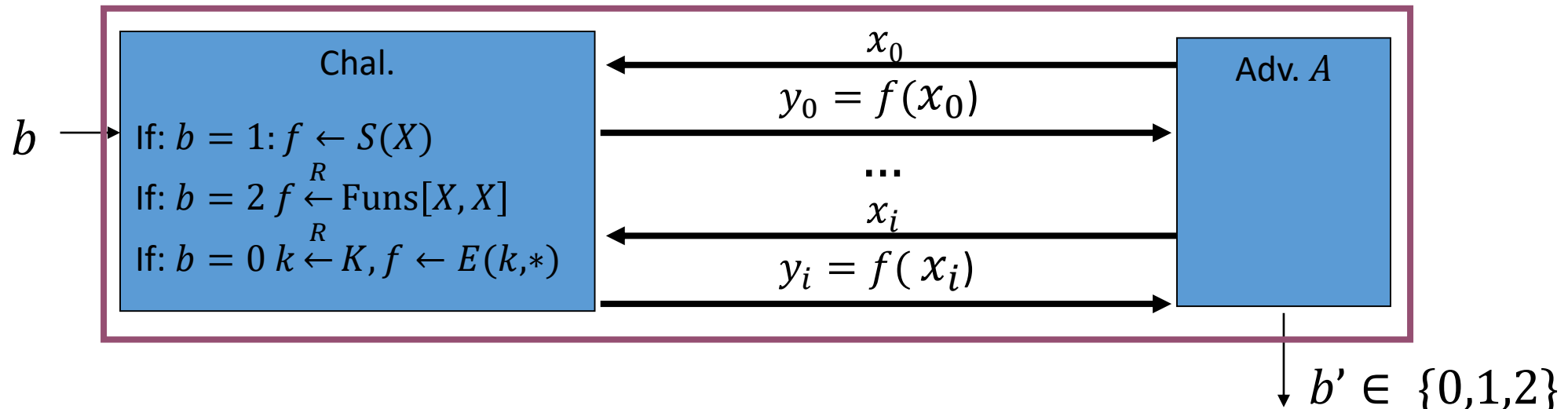
Обозначим p_b вероятность того, что в эксперименте $b \in \{0,1,2\}$ $b' = 1$.

$$|p_1 - p_0| = PRP_{adv}[A, E]$$

$$|p_2 - p_0| = PRF_{adv}[A, E]$$

$$|p_2 - p_1| = PF_{adv}[A, X] \leq Q^2/2N$$

$$|PRP_{adv}[A, E] - PRP_{adv}[A, E]| = ||p_1 - p_0| - |p_2 - p_0|| \leq |p_2 - p_1| \leq Q^2/2N \triangleleft$$



Построение PRG из PRF

Пусть F PRF на (K, X, Y) , $l \geq 1$ полиномиально ограничена, x_1, \dots, x_l различные элементы из X : $|X| \geq L$. Определим PRG G с пространством ключей K , пространством выходов Y^l для $k \in K$ следующим образом:

$$G(k) = (F(k, x_1), \dots, F(k, x_l))$$

Теорема 6.2. Если F стойкая PRF, то G стойкая PRG, причём $\forall A$ в игре на стойкость PRG $\exists B$ в игре на стойкость PRF, такой что

$$PRG_{adv}[A, G] = PRF_{adv}[B, F]$$

▷ Построим противника B . B получает от A запросы x_1, \dots, x_l , отправляет претенденту, получая ответы y_1, \dots, y_l , которые прозрачно пересылает A . Выход b' соответствует выходу алгоритма A .

CTR, вспомним 2 теоремы

Теорема 2.4. Пусть $G: S \rightarrow \{0,1\}^n$ стойкий генератор (PRG).

Тогда поточный шифр E определённый с использованием G семантически стойкий, т.е. $\forall A$: A – противник в игре на семантическую стойкость, \exists противник B в игре на стойкость PRG (различимость):

$$Adv_{ss}[A, E] \leq 2 * Adv_{PRG}[B, G]$$

Теорема 6.1. Пусть $E: K \times X \rightarrow X$ – стойкая PRP, $|X| = N$. Пусть A – противник в игре на стойкость PRF, делающих не более Q запросов к претенденту. Тогда

$$|PRP_{adv}[A, E] - PRF_{adv}[A, E]| \leq Q^2 / 2N$$

CTR

Формально опишем полученный шифр.

$E' = (E', D')$ определён на $(K, X^{\leq l}, X^{\leq l})$, где l – полиномиально ограниченная, $l \leq N$. Пусть x_1, \dots, x_l элементы из X . Обозначим $x_i = (i - 1)_n$ - двоичное n битное представление числа $i - 1$.

Для $k \in K, m \in X^{\leq l}, v = |m|$

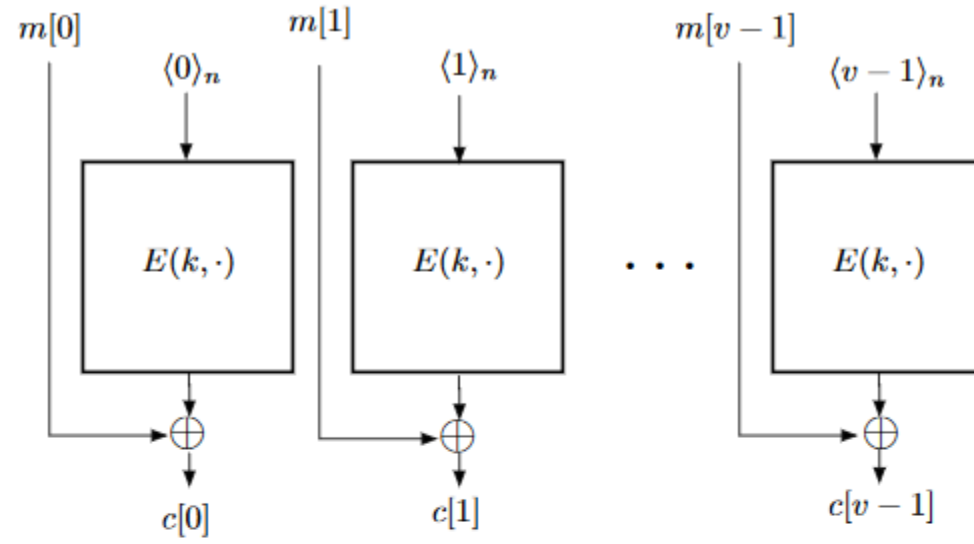
$$E'(k, m) = (E(k, (0)_n) \oplus m[0], \dots, E(k, (v - 1)_n) \oplus m[v - 1])$$

Для $k \in K, c \in X^{\leq l}, c = |c|$

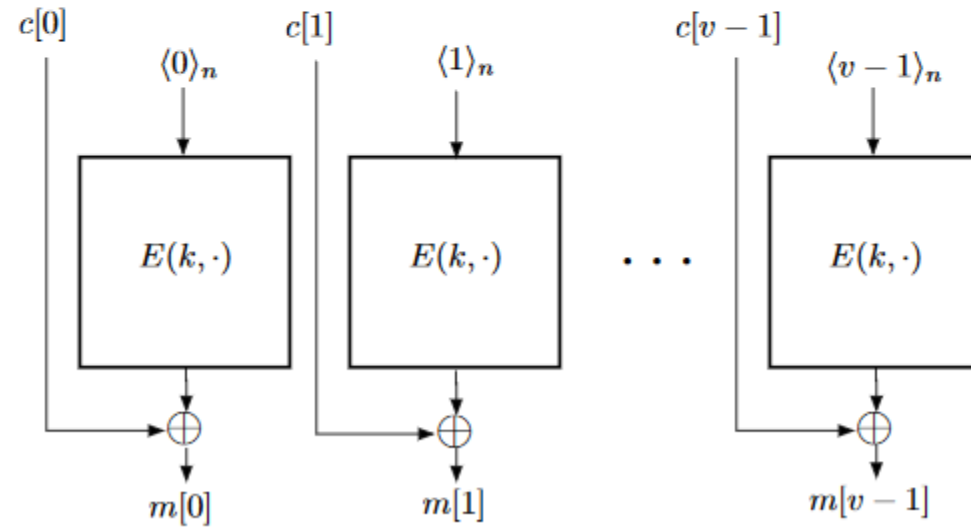
$$D'(k, c) = (E(k, (0)_n) \oplus c[0], \dots, E(k, (v - 1)_n) \oplus c[v - 1])$$

Полученный шифр называется детерминированным CTR режимом для блочного шифра E .

CTR



(a) encryption



(b) decryption

CTR

Теорема 6.2 позволяет построить семантически стойкий шифр с помощью стойкого блочного шифра. Пусть $E = (E, D)$ стойкий блочный шифр на (K, X) , $X = \{0,1\}^n$, $N = 2^n$ - суперполиномиальная.

По **Теореме 6.1** функция зашифрования блочного шифра E является стойкой PRF на (K, X, X) . Используя **Теорему 6.2** получаем стойкий PRG, и используя **Теорему 2.4** (стойкий генератор даёт семантически стойкий поточный шифр) получаем семантически стойкий шифр.

CTR

Теорема 6.3. Если $E = (E, D)$ стойкий блочный шифр, то E' на $(K, X^{\leq l}, X^{\leq l})$ введённый ранее – семантически стойкий шифр, причём $\forall A$ – противников в игре на стойкость блочного шифра (стойкость PRP) $\exists B$ в игре на семантическую стойкость, причём

$$SS_{adv}[A, E'] = 2 * PRP_{adv}[B, E] + l^2 / N$$

▷ Используя **Теорему 6.1** получаем PRF из PRP (добавляется слагаемое l^2 / N), используя **Теорему 6.3** получаем PRG из PRF, используя **Теорему 2.4** получаем семантически стойкий шифр из PRG (множитель 2) ◁

CTR

- $SS_{adv}[A, E'] = 2 * PRP_{adv}[B, E] + l^2/N$
- Стойкий для сообщений произвольной длины
- Стойкость на больших сообщениях убывает квадратично быстро (из за слагаемого l^2/N , стойкость зависит от размера блока (N)).

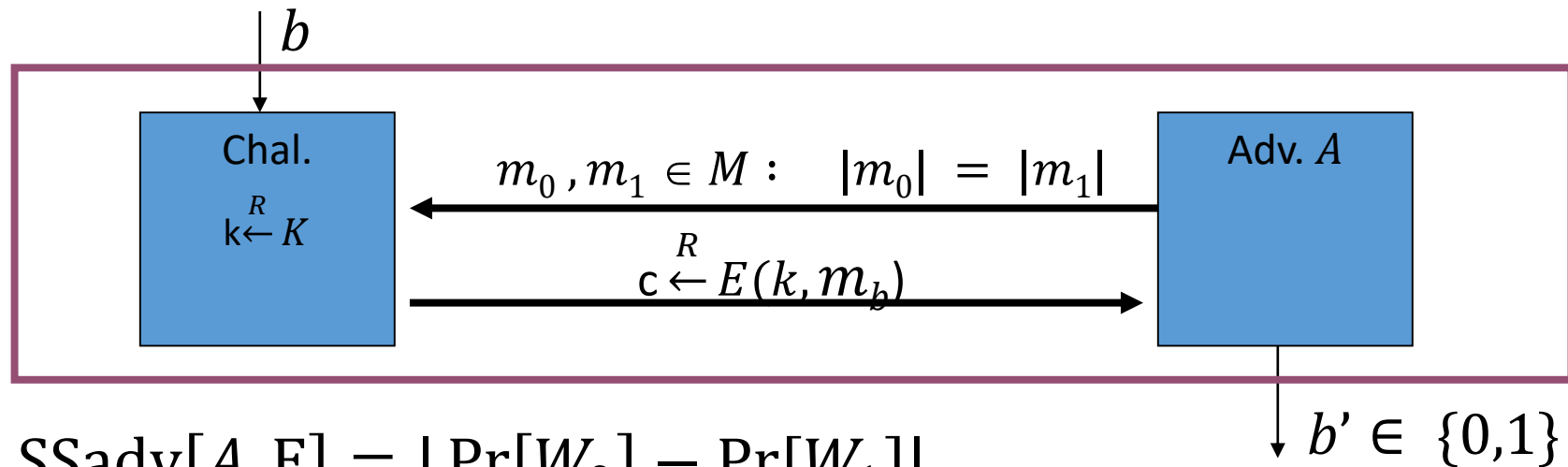
Рассмотрим атаку - m_0 - сообщение из l нулевых блоков, m_1 - сообщение из l случайных блоков. При шифровании сообщения m_0 шифртекст не будет содержать повторяющихся блоков. При шифровании m_1 вероятность получить повторяющиеся блоки $\min\{l(l-1)/4N, 0.63\}$. Т.е. можно построить алгоритм A в игре на семантическую стойкость для $l \sim N^{1/2}$.

Многоразовое использование ключей

- Семантическая стойкость – ослабленная версия абсолютной стойкости, позволяющая описывать стойкость шифров, для которых энтропия ключа меньше энтропии множества открытых текстов.
- Семантически стойкие шифры позволяют использовать короткие ключи.

Многократное использование ключей

До этого момента мы рассматривали ситуации, когда ключ использовался претендентом только один раз. Т.е. мы моделировали одноразовое использование ключа.



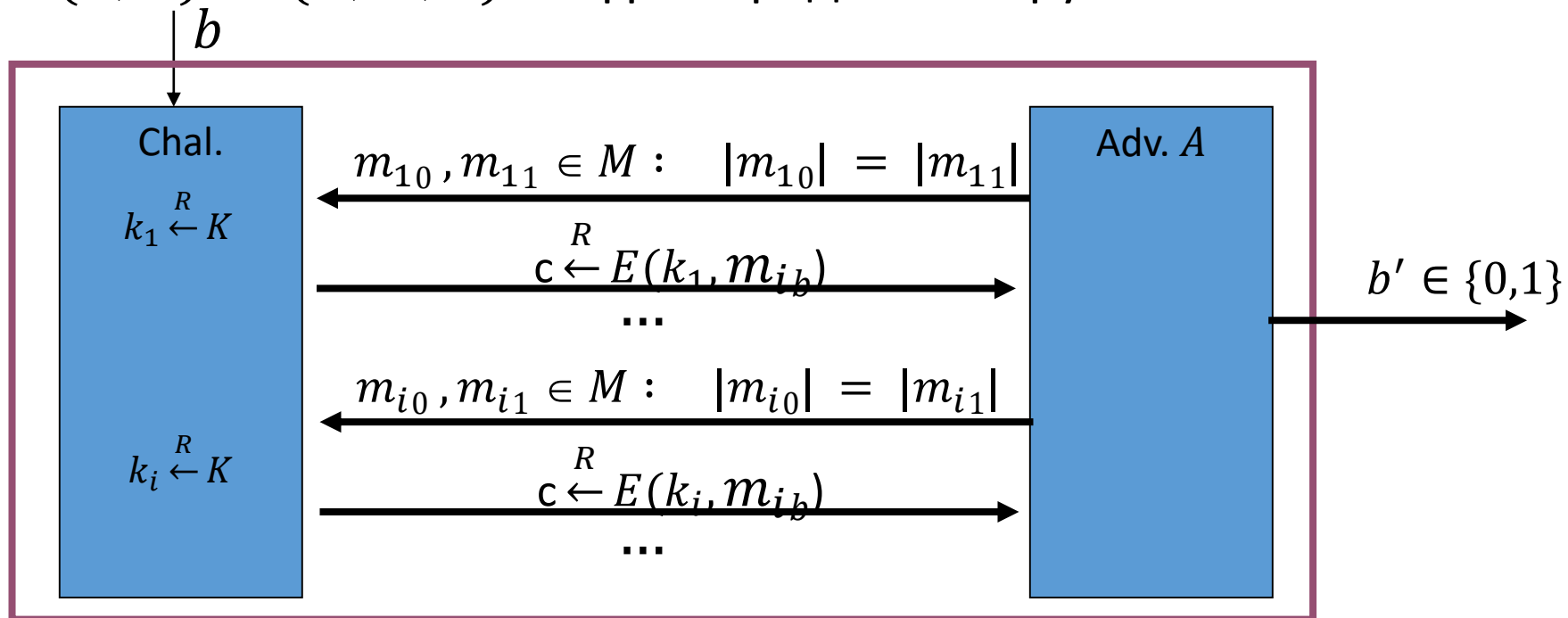
$$\text{SSadv}[A, E] = |\Pr[W_0] - \Pr[W_1]|$$

Но мы хотели бы иметь возможность использовать ключи для шифрования множества сообщений.

Использование множества ключей

- Будет ли семантически стойким шифр, если для шифрования множества сообщений будут использоваться различные случайные независимые ключи?

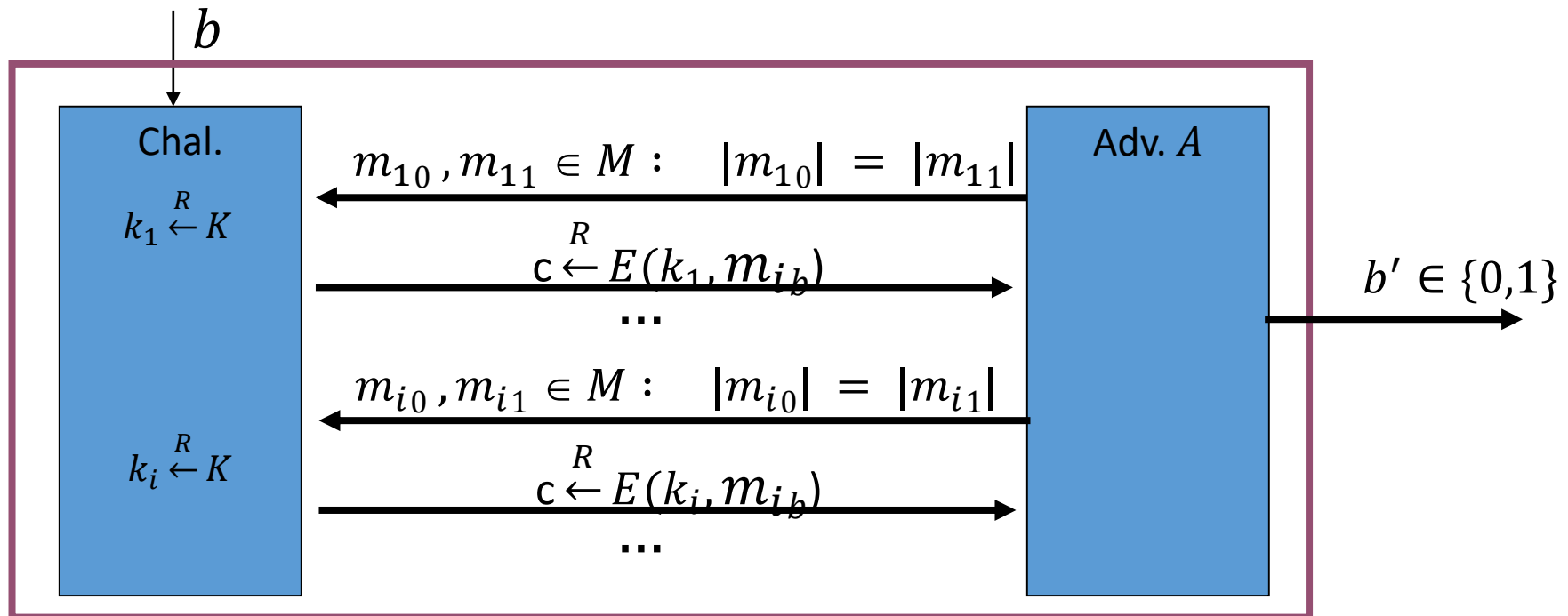
Пусть $E = (E, D)$ на (K, M, C) шифр. Определим игру



Использование множества ключей

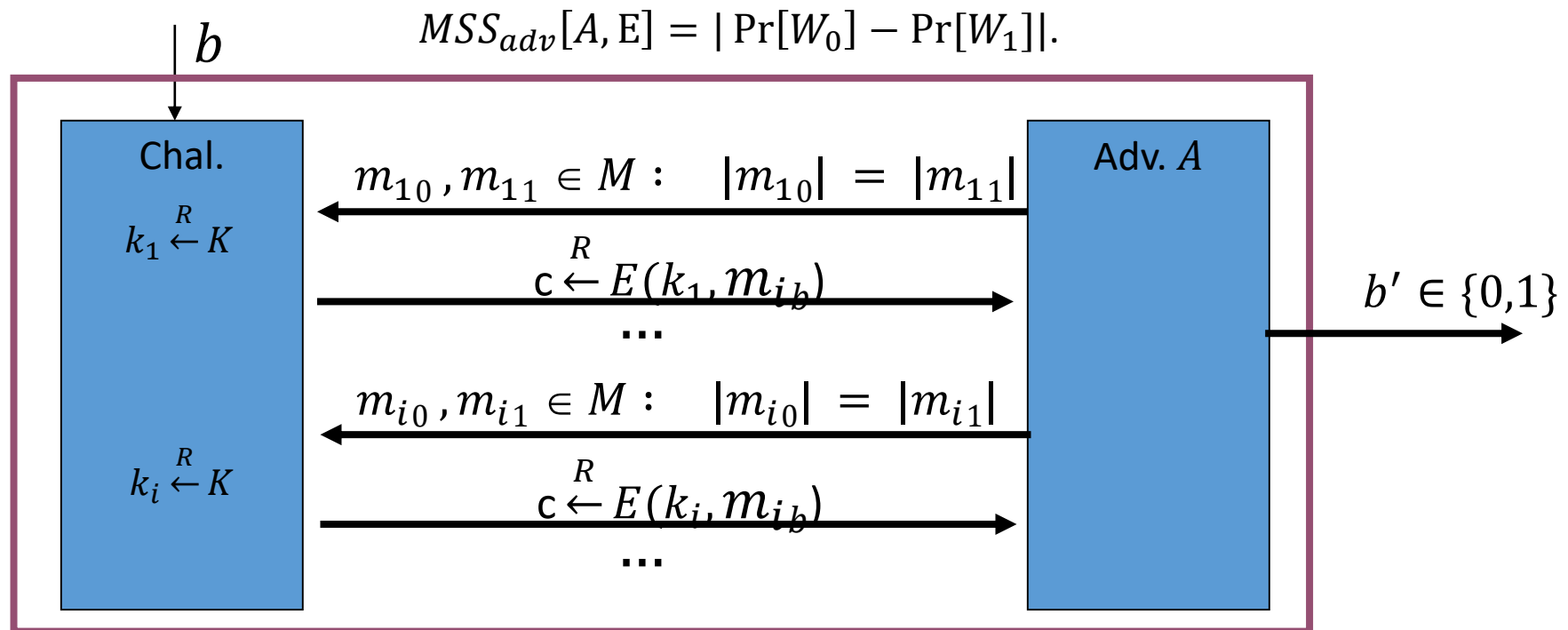
Пусть W_b событие того что в игре b $b' = 1$.

Введём $MSS_{adv}[A, E] = |\Pr[W_0] - \Pr[W_1]|$.



Использование множества ключей

Шифр E называется семантически стойким при использовании множества ключей, если величина $MSS_{adv}[A, E] \leq \epsilon$, где ϵ – пренебрежимо малая.

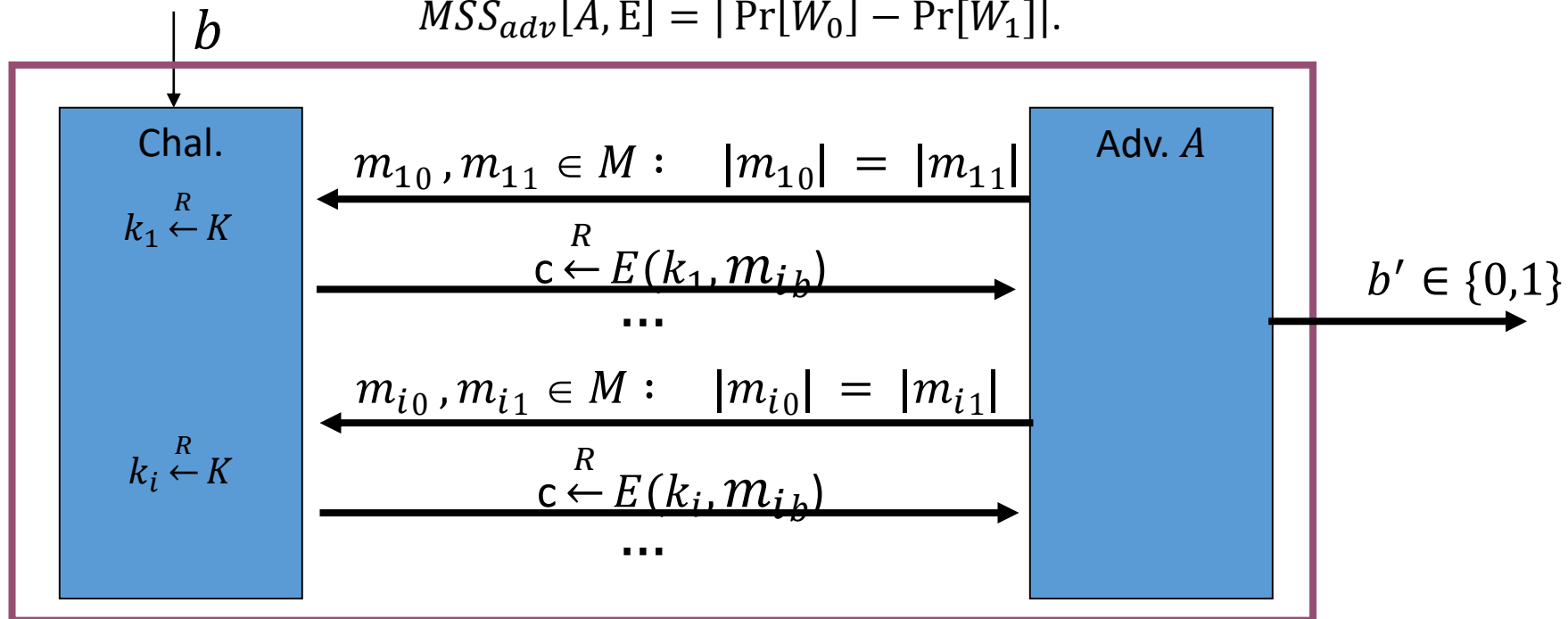


Использование множества ключей

Альтернативное определение преимущества – вероятность угадывания эксперимента противником. Обозначим $MSS_{adv}^*[A, E]$.

$$MSS_{adv}[A, E] = 2 * MSS_{adv}^*[A, E]$$

$$MSS_{adv}[A, E] = |\Pr[W_0] - \Pr[W_1]|.$$



Использование множества ключей

Теорема 6.4. Пусть E семантически стойкий на (K, M, C) . Тогда он семантически стойкий при использовании множества ключей, причём $\forall A$ в игре на семантическую стойкость при использовании множества ключей, использующий не более Q запросов к претенденту, $\exists B$ в игре на семантическую стойкость такой что

$$MSS_{adv}[A, E] = Q * SS_{adv}[B, E]$$

▷ доказательство основано на использовании гибридных игр, аналогично

Теореме 3.1. В эксперименте 0 игры MSS претендент шифрует сообщения $m_{10}, m_{20}, \dots, m_{Q0}$. Для шифрования сообщения m_{10} используется ключ k_1 . Так как шифр семантически стойкий можем заменить шифрование m_{10} на шифрование m_{11} и противник не заметит разницы. В итоге производя Q таких модификаций мы получим эксперимент 1 игры MSS◁

Многоразовое использование ключей

- Описанная ранее семантическая стойкость с использованием множества ключей требует уникального случайного ключа для каждого нового зашифровываемого сообщения.
- Можно ли построить шифр так, чтобы на одном фиксированном ключе можно было зашифровать множество сообщений?
- Вводится понятие многоразовой семантической стойкости, т.е. семантической стойкости, при которой ключ используется более одного раза для зашифрования сообщения.

Многоразовое использование ключей

- Является ли одноразово семантически стойкий шифр многократно семантически стойким?
 - Нет. Пример. При использовании поточного шифра необходима уникальность ключа $s \in S$. При повторении ключа получаем двухразовый блокнот, который не является семантически стойким, так как позволяет восстановить исходные сообщения.
- Нужно новое определение - Необходим аналог семантической стойкости, но при многократном использовании ключа.

Многоразовое использование ключей

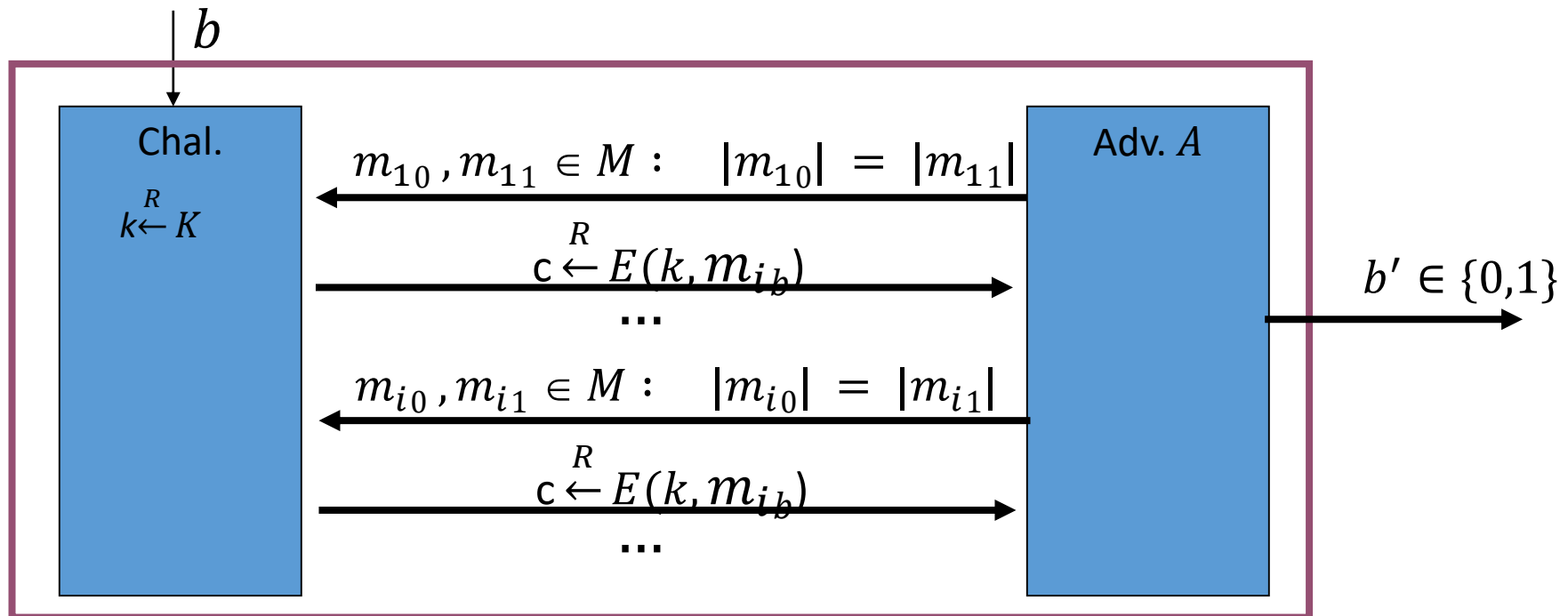
- Попробуем выдвинуть необходимые требования к шифру, семантически стойкому при многократном использовании ключей.
- Поточный поточные шифры не подходят, как видели ранее. Не подойдут и любые **детерминированные** шифры, т.е. такие, которые при фиксированном ключе на данном шифртексте дают одинаковый выход.
 - Если противник знает, что $E(k, x) = c$ и шифр детерминированный, то он может отличить сообщения $m_0 = x, m_1 \neq x$ по их шифртекстам.
- Следовательно, шифр должен быть **вероятностным**, т.е. дающим разные шифртексты на фиксированном ключе для указанного сообщения. Это **необходимое** условие.

CRA

- CRA (атака по выбранному открытому тексту, атака по парам открытый текст – шифртекст).
- Возможности противника – получить шифртексты для произвольных открытых текстов при фиксированном ключе.
- Цель противника – атака на семантическую стойкость.
- Рассмотрим игру

CPA

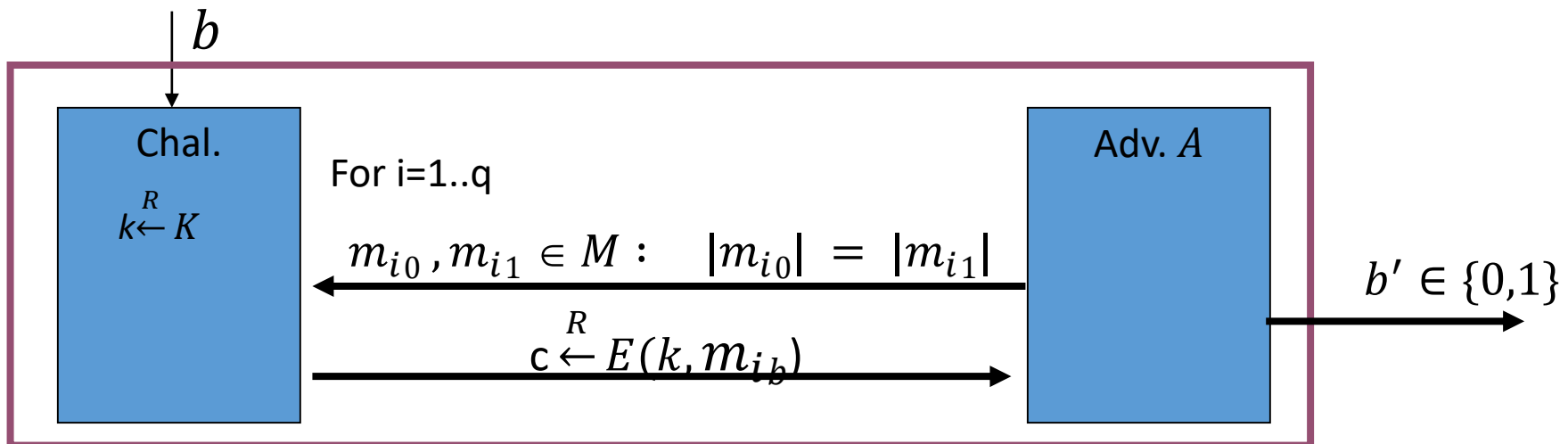
Введём игру аналогично игре при использовании множества ключей семантически стойким шифром, но фиксируя ключ. Определим $CPA_{adv}[A, E] = |\Pr[W_0] - \Pr[W_1]|$.



CPA

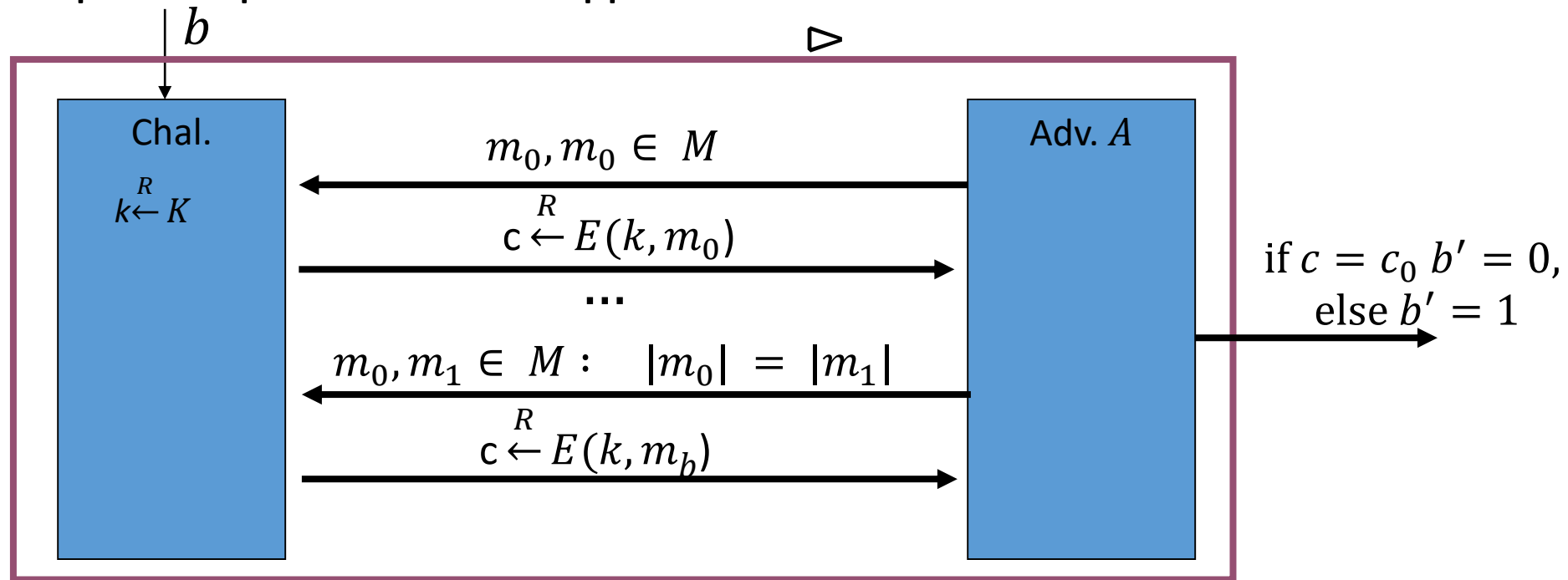
Шифр E называется стойким к атаке по выбранным открытым текстам (CPA стойким), если величина $CPA_{adv}[A, E] \leq \epsilon$, где ϵ – пренебрежимо малая.

Заметим, что противник может получить $c = E(k, m)$ отправив претенденту $m_{j0} = m_{i1} = m$.



CPA

- Детерминированные шифры не CPA стойкие

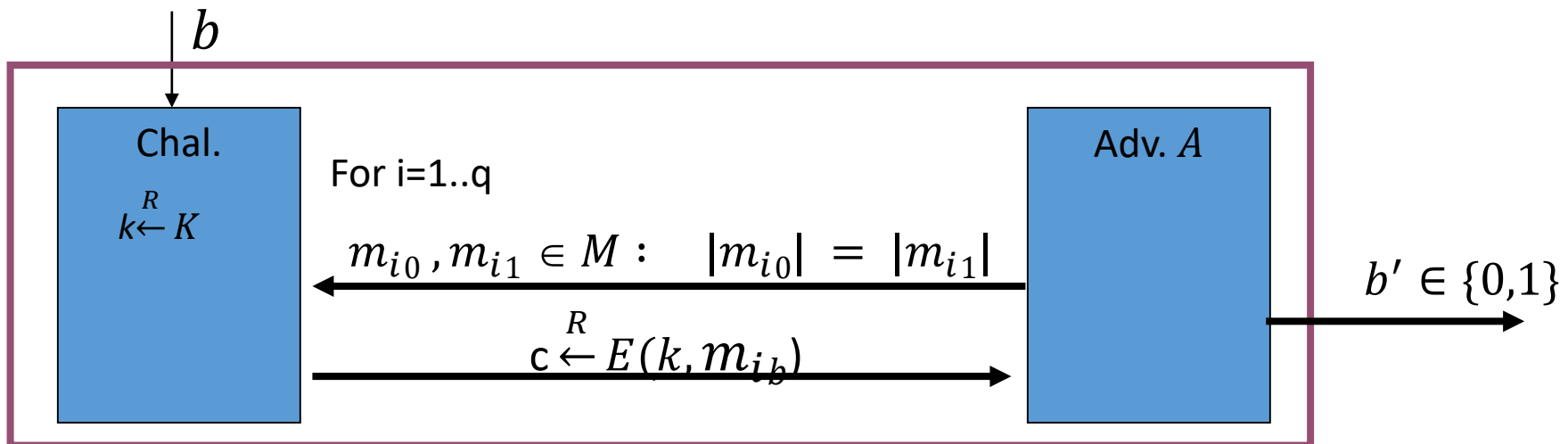


- Следовательно необходимо использовать вероятностные алгоритмы

CPA

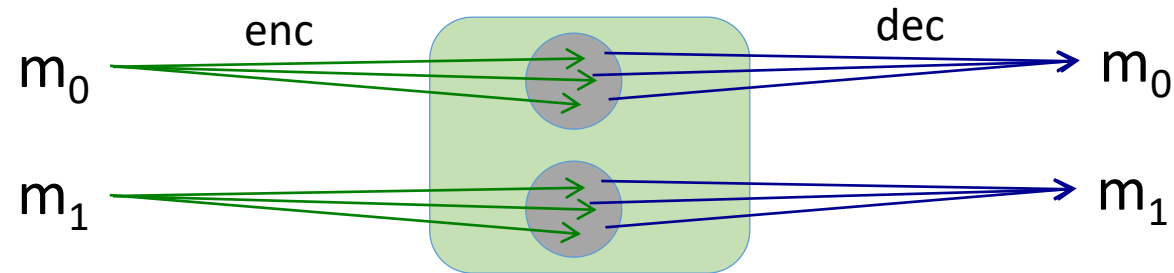
Альтернативное определение преимущества – вероятность угадывания эксперимента противником. Обозначим $CPA_{adv}^*[A, E]$.

$$CPA_{adv}[A, E] = 2 * CPA_{adv}^*[A, E]$$



Вероятностное шифрование

- Как показано ранее, для СРА стойкости необходима «рандомизация» шифртекстов
- Подход 1 – рандомизация функции зашифрования



- Зашифрование одного и того же сообщения даст разные шифртексты
- Необходим внешний источник энтропии
- Шифртексты всегда длиннее открытых текстов, так как необходимо также передать энтропию, необходимую для восстановления открытого текста

Вероятностное шифрование

- Подход 2 – использование уникальных, неповторяющихся величин (nonce)
- $m \rightarrow E(k, *, n) \rightarrow c \rightarrow D(k, *, m) \rightarrow m$
- Nonce должна быть уникально для каждого сообщения, пара (nonce, key) не должна повторяться при жизни ключа.
- В качестве nonce можно использовать счётчик или случайные величины
- Nonce может не пересылаться в явном виде, обе стороны могут синхронно обновлять его.
- Не любое использование nonce даёт стойкие схемы!