CS 591: Formal Methods in Security and Privacy

Approximate probabilistic relational Hoare Logic

Marco Gaboardi
gaboardi@bu.edu

Alley Stoughton
stough@bu.edu

Q&A

To increase interactivity, I will ask more question to each one of you.

It is not a test, you can always answer “pass!”

Assignments

Remember that the third assignment was due yesterday. If you are still working on it, no problem, but please do let us know.

Recording

This is a reminder that we will record the class and we will post the link on Piazza.

This is also a reminder to myself to start recording!

From the previous classes

An example

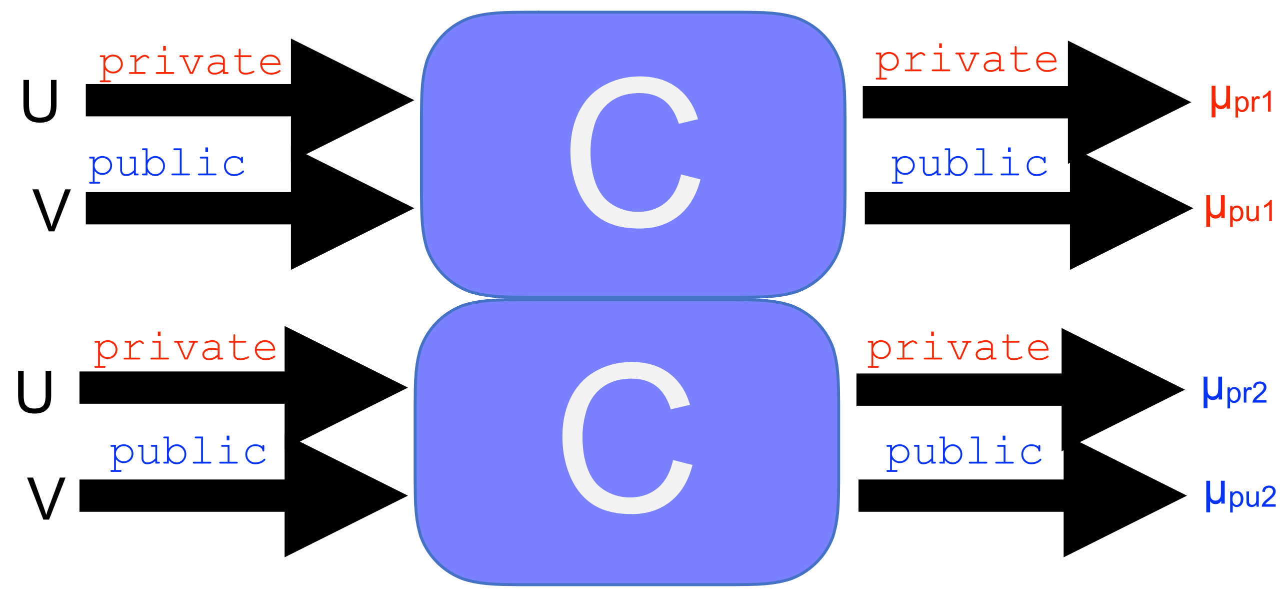
```
OneTimePad(m : private msg) : public msg
   key :=$ Uniform({0,1}^n);
   cipher := msg xor key;
   return cipher
```

Learning a ciphertext does not change any a priori knowledge about the likelihood of messages.

Probabilistic Noninterference as a Relational Property

c is probabilistically noninterferent if and only if for every $m_1 \sim_{low} m_2$:

$\{c\}_{m1} = \mu_1$ and $\{c\}_{m2} = \mu_2$ implies $\mu_1 \sim_{low} \mu_2$



Probabilistic Relational Hoare Quadruples

Precondition
(a logical formula)

Precondition
$\mathtt{Program}_1 \sim \mathtt{Program}_2$
Postcondition

$$c_1 \sim c_2 : P \Rightarrow Q$$

Probabilistic Program

Probabilistic Program

Postcondition
(a logical formula)

R-Coupling

Given two distributions $\mu_1 \in D(A)$, and $\mu_2 \in D(B)$, an R-coupling between them, for $R \subseteq A \times B$, is a joint distribution $\mu \in D(A \times B)$ such that:

1) the marginal distributions of μ are μ_1 and μ_2, respectively,
2) the support of μ is contained in R. That is, if μ`(a,b)>0, then (a,b)`$\in$R.

Validity of Probabilistic Hoare quadruple

We say that the quadruple $c_1 \sim c_2 : P \Rightarrow Q$ is valid if and only if for every pair of memories m_1, m_2 such that $P(m_1, m_2)$ we have:

$\{c_1\}_{m1} = \mu_1$ and $\{c_2\}_{m2} = \mu_2$ implies $Q^*(\mu_1, \mu_2)$.

Consequences of Coupling

Given the following pRHL judgment

$$\vdash c_1 \sim c_2 : \text{True} \Rightarrow Q$$

We have that:

if $Q \Rightarrow (R\langle 1\rangle \iff S\langle 2\rangle)$, then $\Pr[c_1 : R] = \Pr[c_2 : S]$

if $Q \Rightarrow (R\langle 1\rangle \Rightarrow S\langle 2\rangle)$, then $\Pr[c_1 : R] \leq \Pr[c_2 : S]$

A more realistic example

```
StreamCipher(m : private msg[n]) : public msg[n]
   pkey :=$ PRG(Uniform({0,1}^k));
   cipher := msg xor pkey;
   return cipher
```

Properties of PRG

We would like the PRG to increase the number of random bits but also to guarantee the result to be (almost) random.

We can express this as:

$$\boxed{\text{PRG}: \{0,1\}^k \rightarrow \{0,1\}^n \text{ for } n>k}$$

$$\boxed{\text{PRG}(\text{Uniform}(\{0,1\}^k) \approx \text{Uniform}(\{0,1\}^n}$$

How can we measure the similarity between the result of PRG and the uniform distribution?

Statistical distance

We say that two distributions $\mu_1, \mu_2 \in D(A)$, are at statistical distance δ if and only if:

$$\Delta(\mu 1,\mu 2)=\max_{E\subseteq A} | \mu_1(E)-\mu_2(E) | = \delta$$

For discrete distributions the statistical distance can also be characterized as:

$$\Delta(\mu 1,\mu 2)= 1/2 \sum_{a\in A} | \mu_1(a)-\mu_2(a) |$$

Properties of PRG

We would like the PRG to increase the number of random bits but also to guarantee the result to be (almost) random.

We can express this as:

$$\boxed{\text{PRG}: \{0,1\}^k \rightarrow \{0,1\}^n \text{ for } n>k}$$

$$\boxed{\Delta(\text{PRG}(\text{Uniform}(\{0,1\}^k),\text{Uniform}(\{0,1\}^n) \leq 2^{-n}}$$

In fact this is a too strong requirement - usually we require that every polynomial time adversary cannot distinguish the two distributions in statistical distance

How can we prove this secure?

```
OneTimePad(m : private msg[n])
          : public msg[n]
   key :=$ Uniform({0,1}^n);
   cipher := msg xor key;
   return cipher
```

~

```
StreamCipher(m : private msg[n])
            : public msg[n]
   pkey :=$ PRG(Uniform({0,1}^k));
   cipher := msg xor pkey;
   return cipher
```

How can we prove this secure?

```
OneTimePad(m : private msg[n])
          : public msg[n]
   key :=$ Uniform({0,1}ⁿ);
   cipher := msg xor key;
   return cipher
```

~

```
StreamCipher(m : private msg[n])
              : public msg[n]
   pkey :=$ PRG(Uniform({0,1}ᵏ));
   cipher := msg xor pkey;
   return cipher
```

m

m

How can we prove this secure?

```
OneTimePad(m : private msg[n])
          : public msg[n]
   key :=$ Uniform({0,1}ⁿ);
   cipher := msg xor key;
   return cipher
```

~

```
StreamCipher(m : private msg[n])
             : public msg[n]
   pkey :=$ PRG(Uniform({0,1}ᵏ));
   cipher := msg xor pkey;
   return cipher
```

$$m \qquad\qquad m$$

$$\downarrow$$

$$m \oplus k$$

How can we prove this secure?

```
OneTimePad(m : private msg[n])
          : public msg[n]
   key :=$ Uniform({0,1}^n);
   cipher := msg xor key;
   return cipher
```

~

```
StreamCipher(m : private msg[n])
            : public msg[n]
   pkey :=$ PRG(Uniform({0,1}^k));
   cipher := msg xor pkey;
   return cipher
```

$$m \Big\downarrow m \oplus k$$

$$m \Big\downarrow m \oplus pk$$

How can we prove this secure?

```
OneTimePad(m : private msg[n])
          : public msg[n]
   key :=$ Uniform({0,1}^n);
   cipher := msg xor key;
   return cipher
```

~

```
StreamCipher(m : private msg[n])
             : public msg[n]
   pkey :=$ PRG(Uniform({0,1}^k));
   cipher := msg xor pkey;
   return cipher
```

$$m \longrightarrow m \oplus k \qquad\qquad m \longrightarrow m \oplus pk$$

$$\Delta(\, m \oplus k \,,\; m \oplus pk \,) \le \delta$$

Today:
approximate probabilistic noninterference

How to reason formally about this formally?

Approximate Probabilistic Relational Hoare Logic

Indistinguishability parameter

Precondition (a logical formula)

$$\vdash_{\delta} c_1 \sim c_2 : P \Rightarrow Q$$

Probabilistic Program

Probabilistic Program

Postcondition (a logical formula)

How can we define validity?

Validity of Probabilistic Hoare quadruple

We say that the quadruple $c_1 \sim c_2 : P \Rightarrow Q$ is valid if and only if for every pair of memories m_1, m_2 such that $P(m_1, m_2)$ we have: $\{c_1\}_{m1} = \mu_1$ and $\{c_2\}_{m2} = \mu_2$ implies $Q^*(\mu_1, \mu_2)$.

$\mathtt{R}$–δ–Coupling

Given two distributions $\mu_1 \in D(A)$, and $\mu_2 \in D(B)$, we have an $\mathtt{R}$-δ-coupling between them, for $\mathtt{R} \subseteq A\text{x}B$ and $0 \leq \delta \leq 1$, if there are two joint distributions $\mu_L, \mu_R \in D(A\text{x}B)$ such that:

1) $\pi_1(\mu_L) = \mu_1$ and $\pi_2(\mu_R) = \mu_2$,
2) the support of μ_L and μ_R is contained in $\mathtt{R}$. That is, if $\mu_L\mathtt{(a,b)>0}$, $\mathtt{then\ (a,b)} \in \mathtt{R}$, $\mathtt{and}$ if $\mu_R\mathtt{(a,b)>0}$, $\mathtt{then\ (a,b)} \in \mathtt{R}$.
3) $\Delta(\mu_L, \mu_R) \leq \delta$

Approximate relational lifting of a predicate

We say that two subdistributions $\mu_1 \subseteq D(A)$ and $\mu_2 \subseteq D(B)$ are in the relational δ–lifting of the relation $R \subseteq A x B$, denoted $\mu_1 \; R_\delta{*} \; \mu_2$ if and only if there exist an R-coupling between them.

Validity of approximate Probabilistic Hoare judgments

We say that the quadruple $\vdash_\delta c_1 \sim c_2 : P \Rightarrow Q$ is valid if and only if for every pair of memories m_1, m_2 such that $P(m_1, m_2)$ we have:

$\{c_1\}_{m1} = \mu_1$ and $\{c_2\}_{m2} = \mu_2$ implies $Q_\delta{}^*(\mu_1, \mu_2)$.

Example of R-δ-Coupling

μ_1

00	0.25
01	0.25
10	0.25
11	0.25

μ_2

00	0.20
01	0.25
10	0.25
11	0.30

Example of R-δ-Coupling

μ_1

OO	0.25
O1	0.25
1O	0.25
11	0.25

$$R(a,b) = \{a=b\}$$

μ_2

OO	0.20
O1	0.25
1O	0.25
11	0.30

Example of R-δ-Coupling

μ_1

00	0.25
01	0.25
10	0.25
11	0.25

$$R(a,b) = \{a=b\}$$

μ_2

00	0.20
01	0.25
10	0.25
11	0.30

μ_L	00	01	10	11
00	0.25			
01		0.25		
10			0.25	
11				0.25

μ_R	00	01	10	11
00	0.20			
01		0.25		
10			0.25	
11				0.30

Example of R-δ-Coupling

μ_1

00	0.25
01	0.25
10	0.25
11	0.25

$$R(a,b) = \{a=b\}$$

μ_2

00	0.20
01	0.25
10	0.25
11	0.30

μ_L	00	01	10	11
00	0.25			
01		0.25		
10			0.25	
11				0.25

μ_R	00	01	10	11
00	0.20			
01		0.25		
10			0.25	
11				0.30

$$\Delta(\mu_L, \mu_R) = 0.05$$

Example of R-δ-Coupling

μ_1

OO	0.2
O1	0.25
1O	0.25
11	0.3

μ_2

OO	0
O1	0.40
1O	0
11	0.6

Example of R-δ-Coupling

μ_1

00	0.2
01	0.25
10	0.25
11	0.3

$$R(a,b) = \{a \leq b\}$$

μ_2

00	0
01	0.40
10	0
11	0.6

Example of R-δ-Coupling

μ_1

00	0.2
01	0.25
10	0.25
11	0.3

$$R(a,b) = \{a \leq b\}$$

μ_2

00	0
01	0.40
10	0
11	0.6

μ_L	00	01	10	11
00		0.20		
01		0.25		
10				0.25
11				0.30

μ_R	00	01	10	11
00		0.20		
01		0.20		
10				0.3
11				0.3

Example of R-δ-Coupling

μ_1

00	0.2
01	0.25
10	0.25
11	0.3

$$R(a,b) = \{a \leq b\}$$

μ_2

00	0
01	0.40
10	0
11	0.6

μ_L	00	01	10	11
00		0.20		
01		0.25		
10				0.25
11				0.30

μ_R	00	01	10	11
00		0.20		
01		0.20		
10				0.3
11				0.3

$$\Delta(\mu_L, \mu_R) = 0.05$$

A more realistic example

```
StreamCipher(m : private msg[n]) : public msg[n]
   pkey :=$ PRG(Uniform({0,1}^k));
   cipher := msg xor pkey;
   return cipher
```

Example of R-δ-Coupling

μ_1

00	0.25
01	0.25
10	0.25
11	0.25

μ_2

00	0
01	0
10	0.5
11	0.5

Example of R-δ-Coupling

μ_1

00	0.25
01	0.25
10	0.25
11	0.25

$$R(a,b) = \{a=b\}$$

μ_2

00	0
01	0
10	0.5
11	0.5

Example of R-δ-Coupling

μ_1

00	0.25
01	0.25
10	0.25
11	0.25

μ_2

00	0
01	0
10	0.5
11	0.5

$$R(a,b) = \{a=b\}$$

μ_L	00	01	10	11
00	0.25			
01		0.25		
10			0.25	
11				0.25

μ_R	00	01	10	11
00	0			
01		0		
10			0.5	
11				0.5

Example of R-δ-Coupling

μ_1

00	0.25
01	0.25
10	0.25
11	0.25

$$R(a,b) = \{a=b\}$$

μ_2

00	0
01	0
10	0.5
11	0.5

μ_L	00	01	10	11
00	0.25			
01		0.25		
10			0.25	
11				0.25

μ_R	00	01	10	11
00	0			
01		0		
10			0.5	
11				0.5

$$\Delta(\mu_L, \mu_R) = 0.5$$