Enhancing Security via Protocol Composition

Catuscia Palamidessi



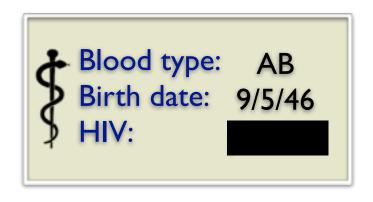


This talk in a nutshell

- Setting: Quantitative Information Flow. Inference attacks using correlation between secret observables
- Defense: The system designer can reduce the correlation secret-observables via protocol composition (typically randomized)
- Active Adversary: may interact with the system and increase the correlation secret-observables
- We formalize the interplay between defender and attacker in **Game Theory**
- Optimal strategy for composition: Saddle points / Nash equilibria. Convex analysis

Quantitative Information Flow

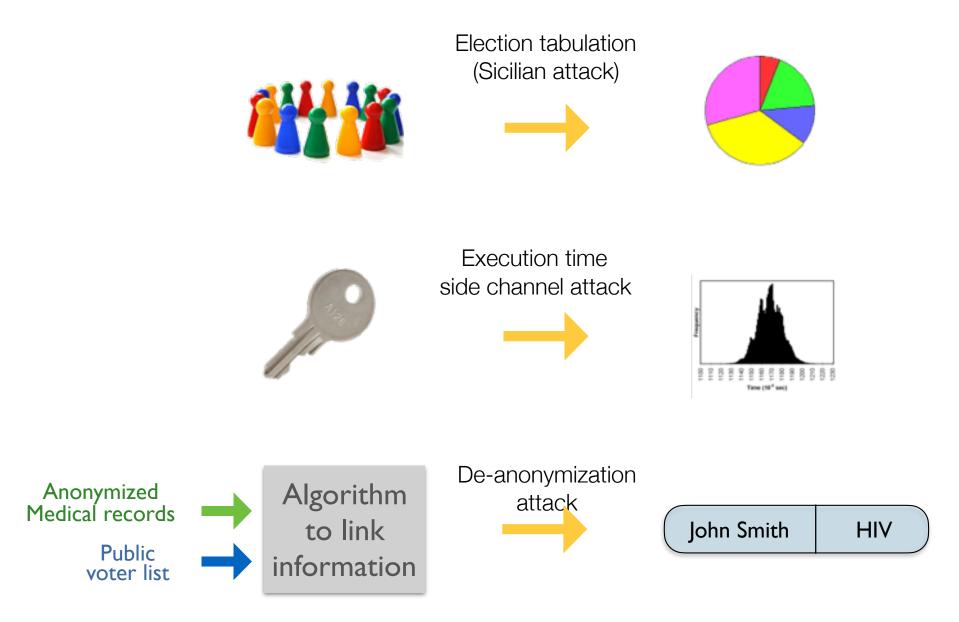
• General problem: security and privacy





- Access control and encryption are not always sufficient: systems may leak sensitive information through their correlation with information available to the adversary (observables)
- Observables: output of the system, public information, side information, physical aspects of the implementation, etc.
- QIF studies measures to assess the threats and techniques to mitigate the leakage due to correlation

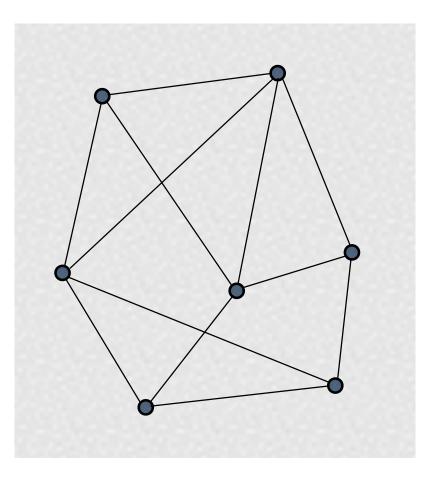
Examples of Leakage via correlated observables



Example

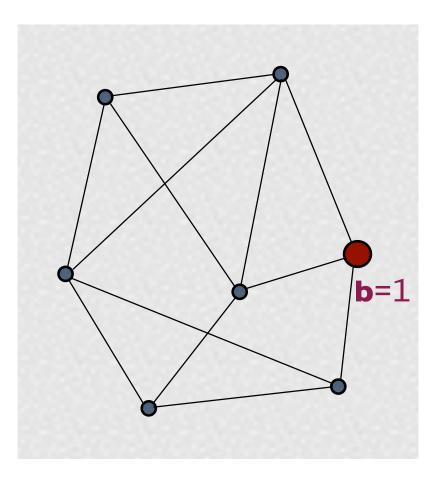
Dining Cryptographers (DC) [Chaum'88]

- A set of nodes with some communication channels (edges).
- One of the nodes (source) wants to broadcast one bit b of information
- The source (broadcaster) must remain anonymous

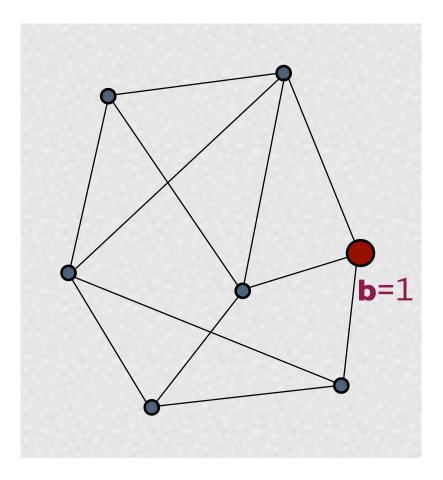


Dining Cryptographers (DC) [Chaum'88]

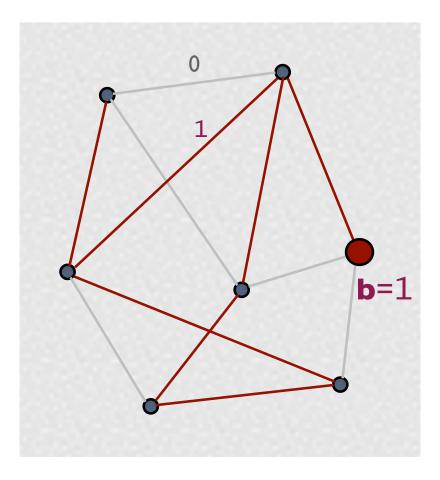
- A set of nodes with some communication channels (edges).
- One of the nodes (source) wants to broadcast one bit b of information
- The source (broadcaster) must remain anonymous



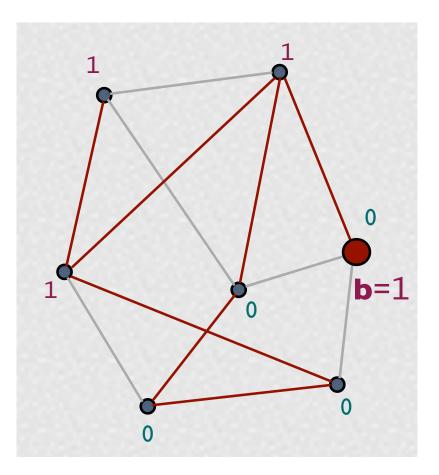
• Associate to each edge a binary coin



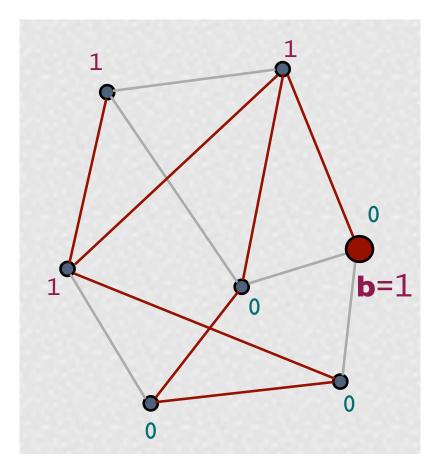
- Associate to each edge a binary coin
- Toss the coins



- Associate to each edge a binary coin
- Toss the coins
- Each node computes the binary sum of the incident edges. The source adds b. They all broadcast their results



- Associate to each edge a binary coin
- Toss the coins
- Each node computes the binary sum of the incident edges. The source adds b. They all broadcast their results
- Achievement of the goal: Compute the total binary sum: it coincides with b

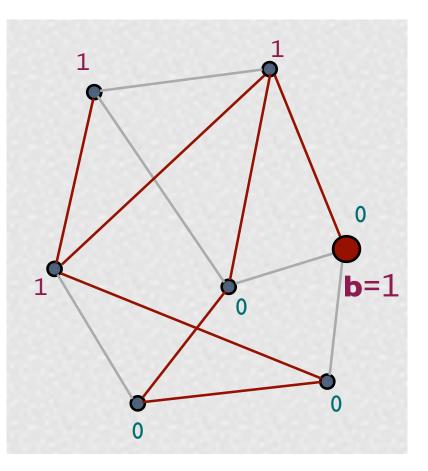


Strong anonymity (Chaum)

 If the graph is connected and the coins are fair, then for an external observer (who observes the declarations of the nodes, but cannot see the value of the coins), the protocol satisfies strong anonymity:

the *a posteriori* probability that a certain node is the source is equal to its *a priori* probability

- A priori / a posteriori = before / after observing the declarations
- Note the use of randomization to obfuscate the link between secret and observables



Anonymity of DC Nets

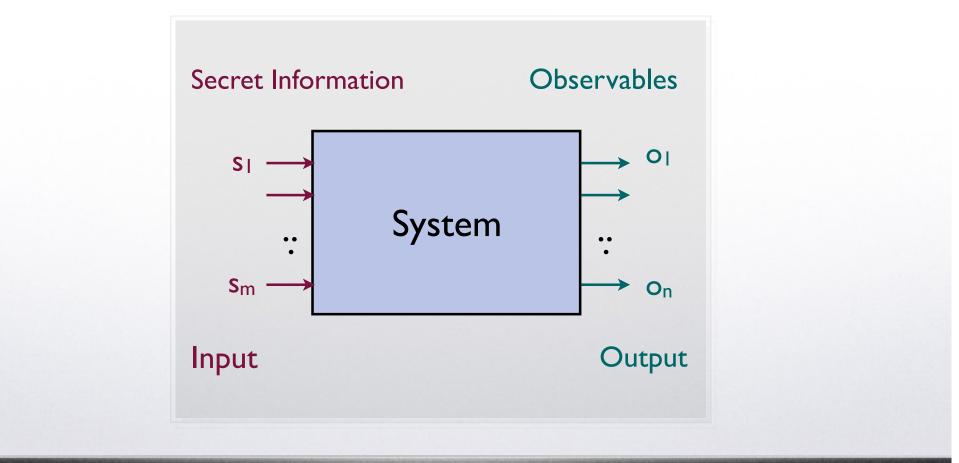
Questions:

What if the coins are biased ?

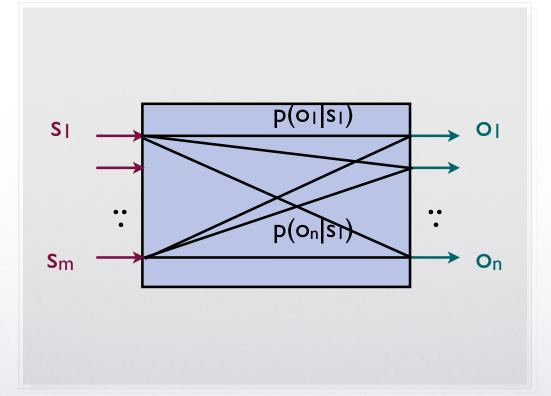
- Does the protocol still protect the anonymity of the source ? To what extent ?
- How to measure the leakage ?

The basic model:

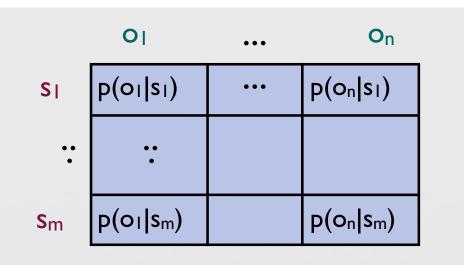
Systems = Information-Theoretic channels



Probabilistic systems are **noisy** channels: an output can correspond to different inputs, and an input can generate different outputs, according to a prob. distribution



 $p(o_j|s_i)$: the conditional probability to observe o_j given the secret s_i



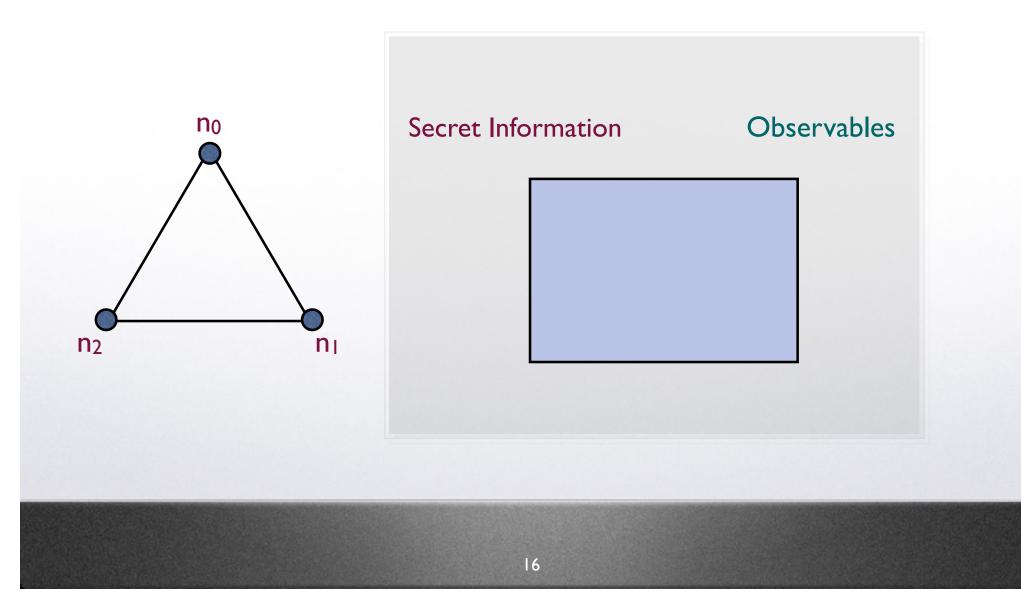
 $p(o|s) = \frac{p(o \text{ and } s)}{p(s)}$

A channel is characterized by its matrix: the array of conditional probabilities

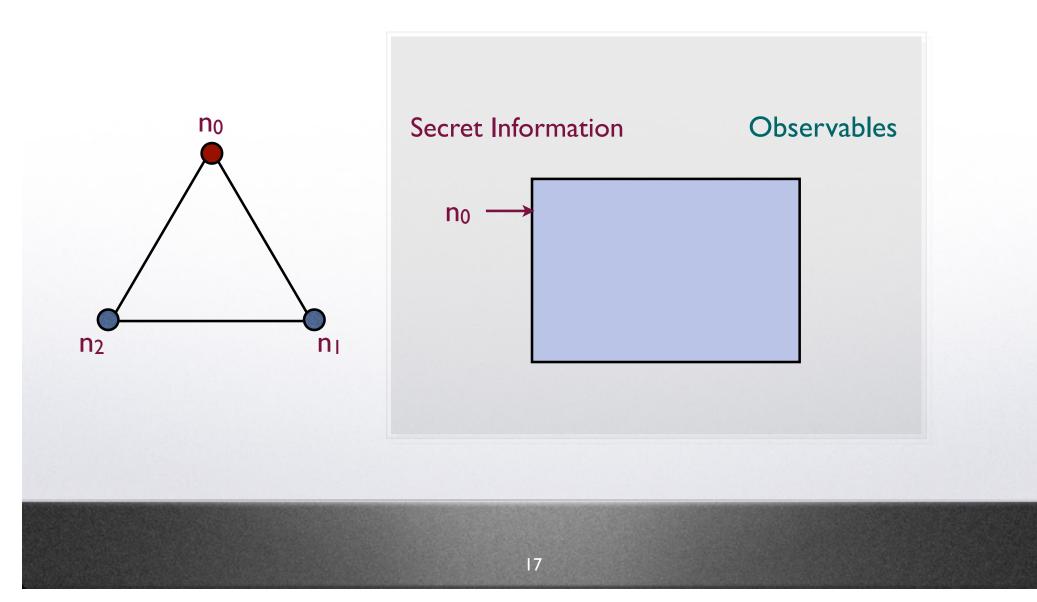
In a information-theoretic channel these conditional probabilities are independent from the input distribution

This means that we can model systems abstracting from the input distribution

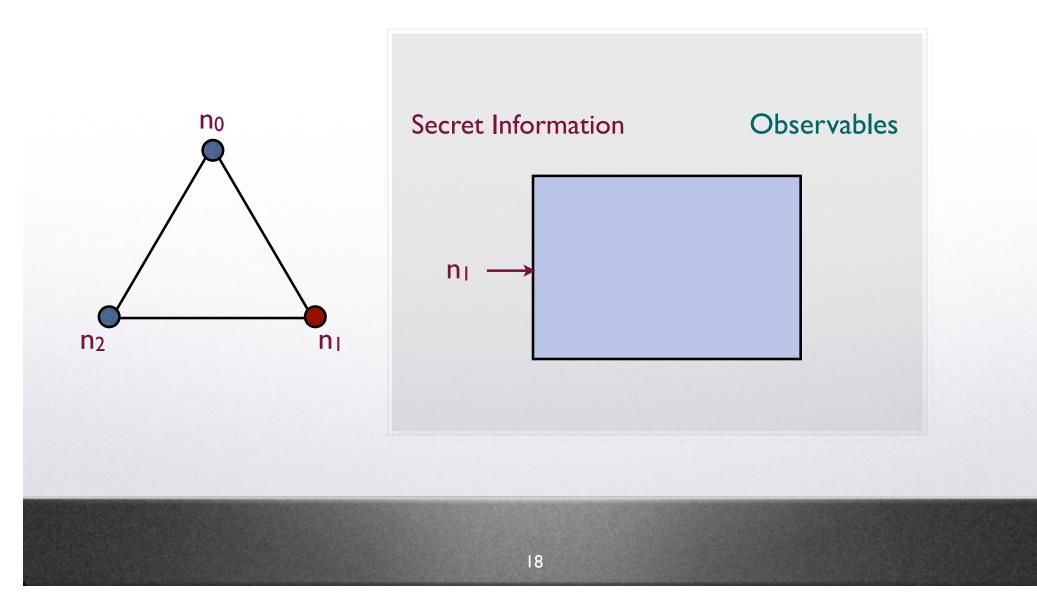




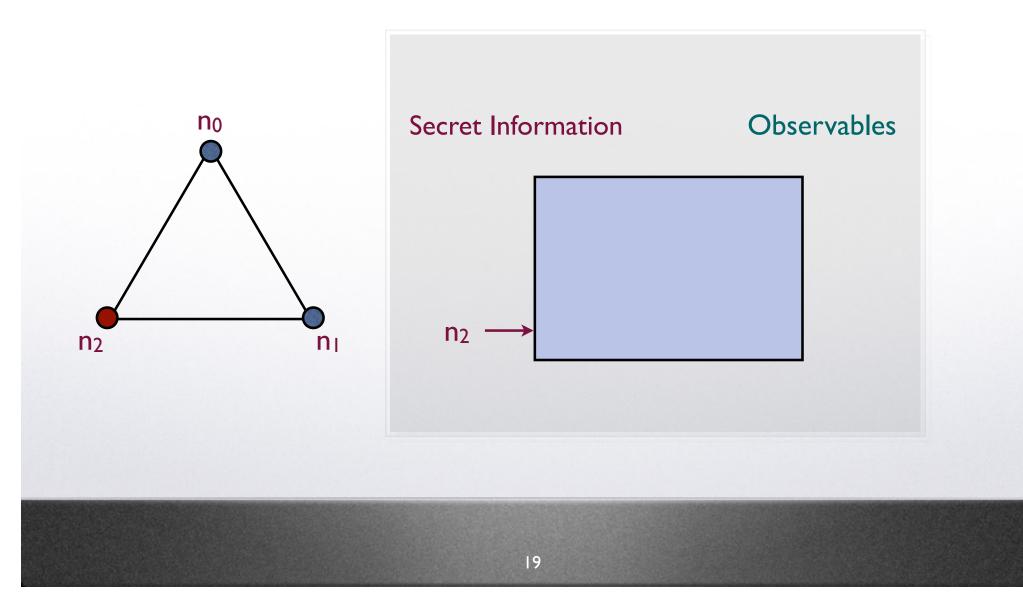


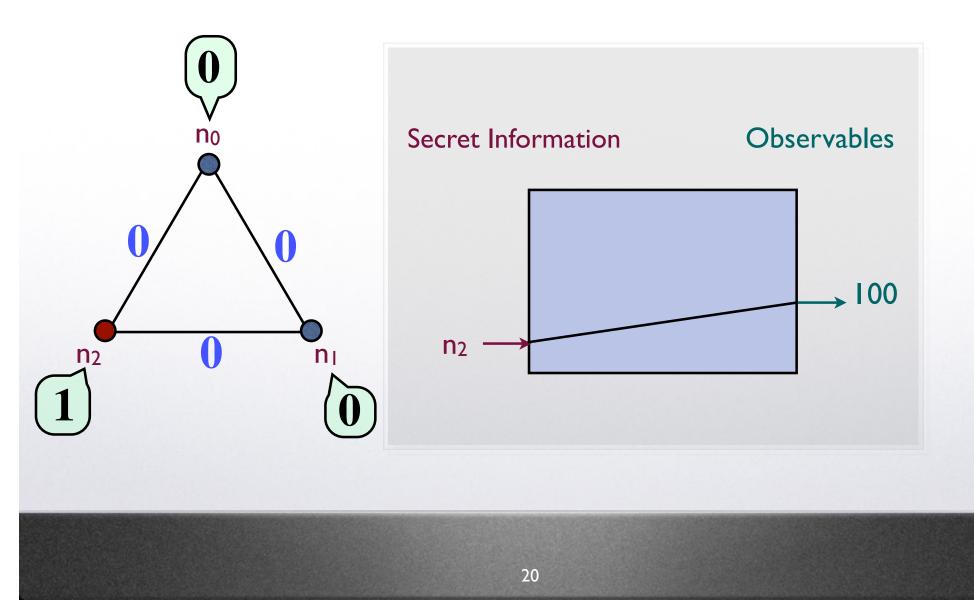




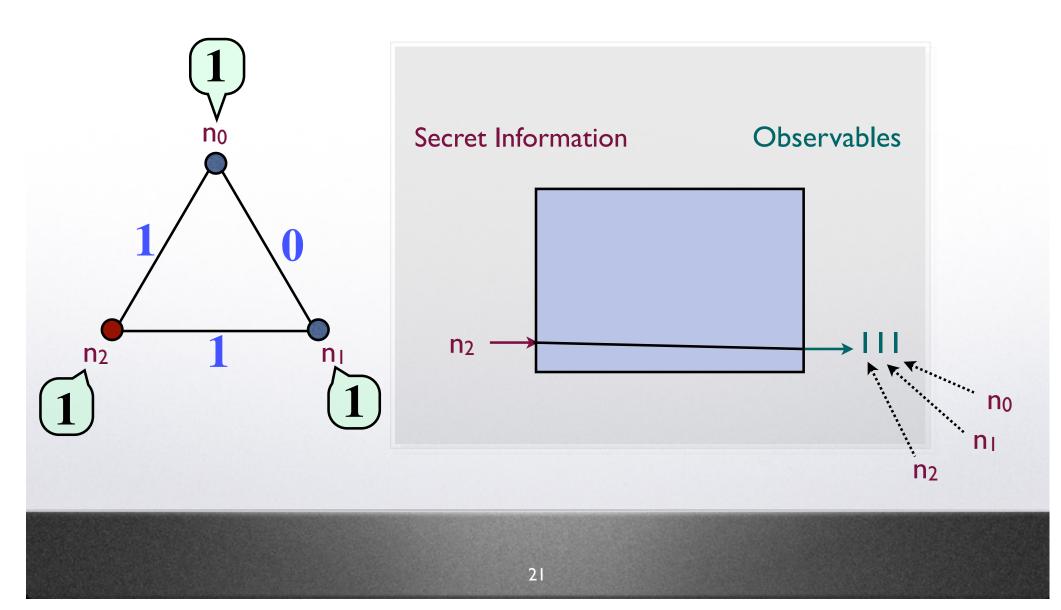




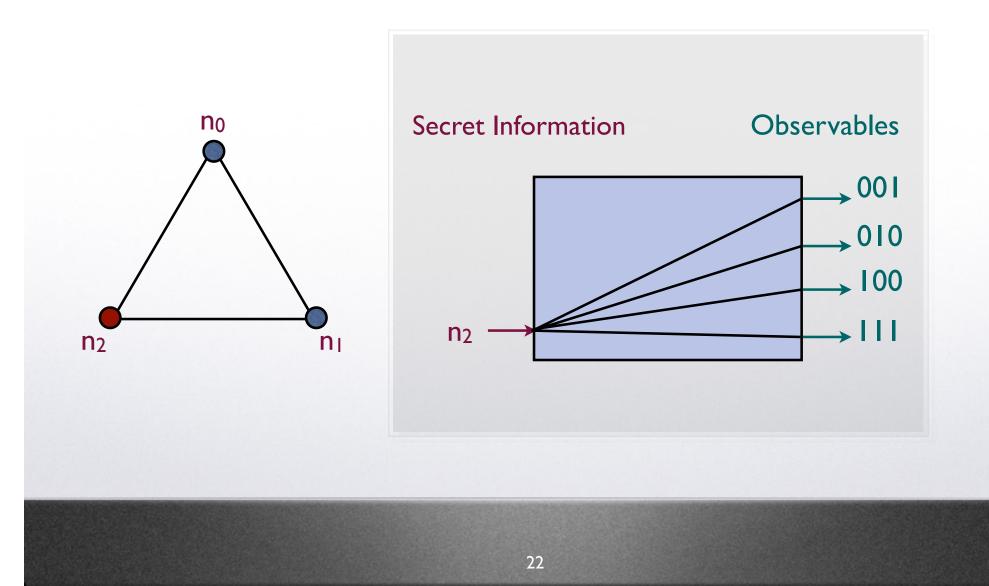




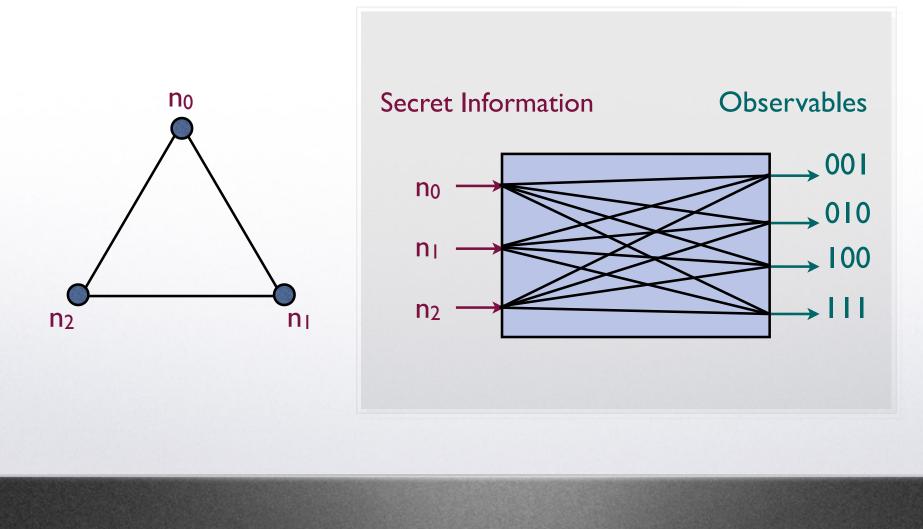














	001	010	100	111
n ₀	1⁄4	1⁄4	1⁄4	1⁄4
nı	1⁄4	1⁄4	1⁄4	1⁄4
n ₂	1⁄4	1⁄4	1⁄4	1⁄4

	001	010	100	111
n ₀	1⁄3	² /9	² / ₉	² / ₉
nı	² /9	1⁄3	² /9	² /9
n ₂	2/9	² / ₉	1⁄3	² / ₉

fair coins: $Pr(0) = Pr(1) = \frac{1}{2}$

strong anonymity

biased coins: $Pr(0) = \frac{2}{3}$, $Pr(1) = \frac{1}{3}$ The source is more likely to declare 1 than 0

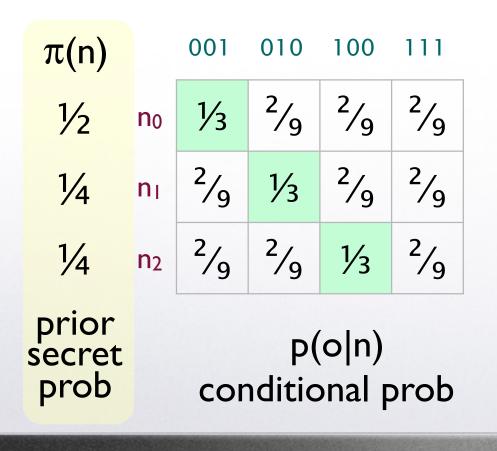
Quantitative Information Flow

- Intuitively, the leakage is the (probabilistic) information that the adversary gains about the secret through the observables
- Each observable changes the prior probability distribution on the secret values into a posterior probability distribution according to the Bayes theorem (Bayesian update)
- In average, the information content (about the actual secret value) of the posterior probability is more than or equal to the one of the prior

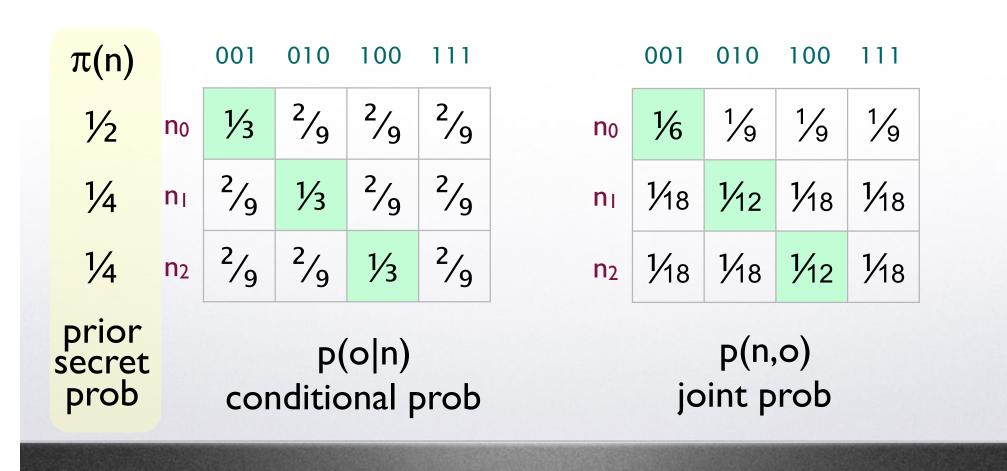








27



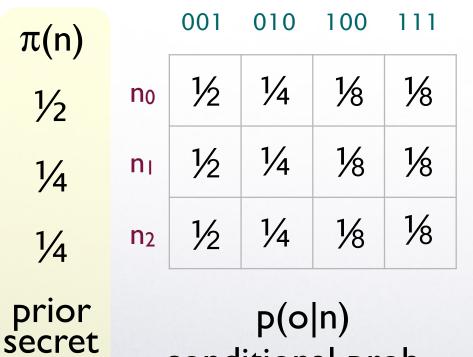
								p(o)	⁵ ⁄18	1⁄4	1⁄4	2 _{⁄9}	obs prob
	π(n)		001	010	100	111			001	010	100	111	P. 00
	1⁄2	n ₀	1⁄3	² /9	² / ₉	² / ₉		n ₀	1⁄6	1⁄9	1/9	1⁄9	
	1⁄4	nı	² / ₉	1⁄3	² /9	² / ₉		nı	1⁄18	1⁄12	1⁄18	1⁄18	
	1⁄4	n ₂	² /9	² /9	1⁄3	² / ₉		n ₂	1⁄16	1⁄18	1⁄12	1⁄18	
•	prior secret prob		p(o n) conditional prob					p(n,o) joint prob					

Bayesian update: prior \Rightarrow posterior $p(n|o) = \frac{p(n,o)}{p(o)}$ obs 5⁄18 1⁄4 100 1⁄4 2_{⁄9} p(o) prob **π(n|**001) 001 010 010 111 100 111 $\frac{1}{3}$ $\frac{2}{9}$ $\frac{2}{9}$ $^{2}/_{9}$ $\frac{1}{9}$ 3⁄5 1⁄6 $\frac{1}{9}$ $\frac{1}{9}$ n₀ n₀ $\frac{1}{3}$ $\frac{2}{9}$ $^{2}/_{9}$ $^{2}/_{9}$ 1/18 1/12 1/18 $\frac{1}{5}$ 1/18 nı nı $^{2}/_{9}$ 2/9 $^{2}/_{9}$ 1⁄3 1/5 1/18 1/18 1/12 1/18 n₂ n₂ post p(n,o) p(o|n)secret prob joint prob conditional prob

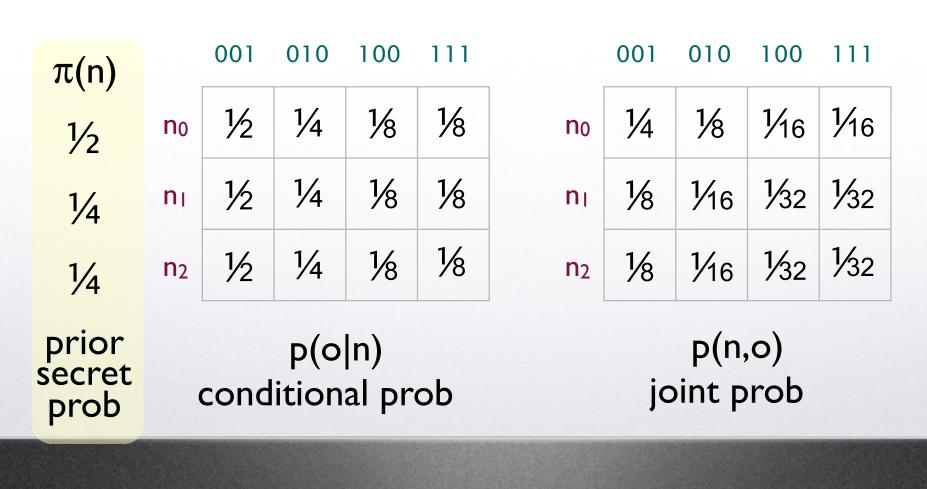
30

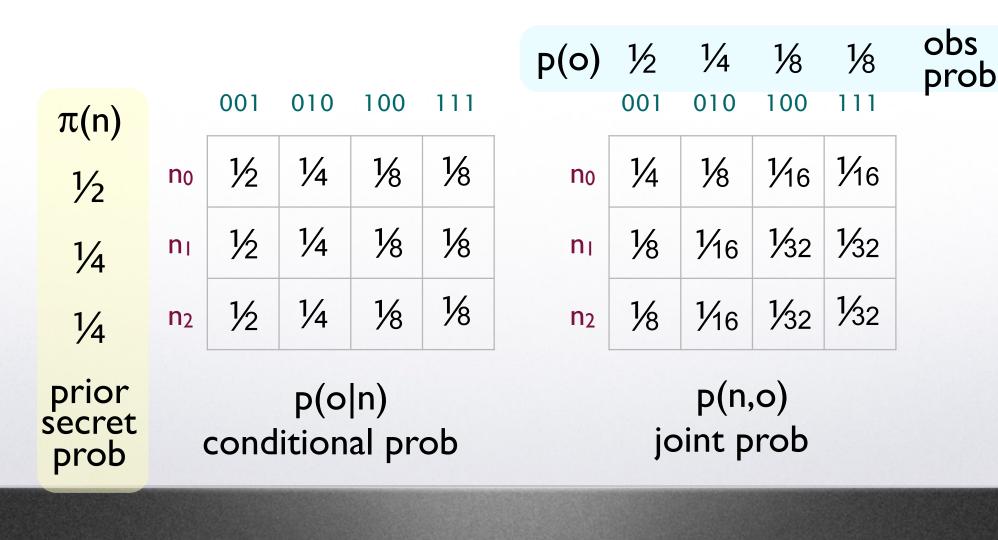


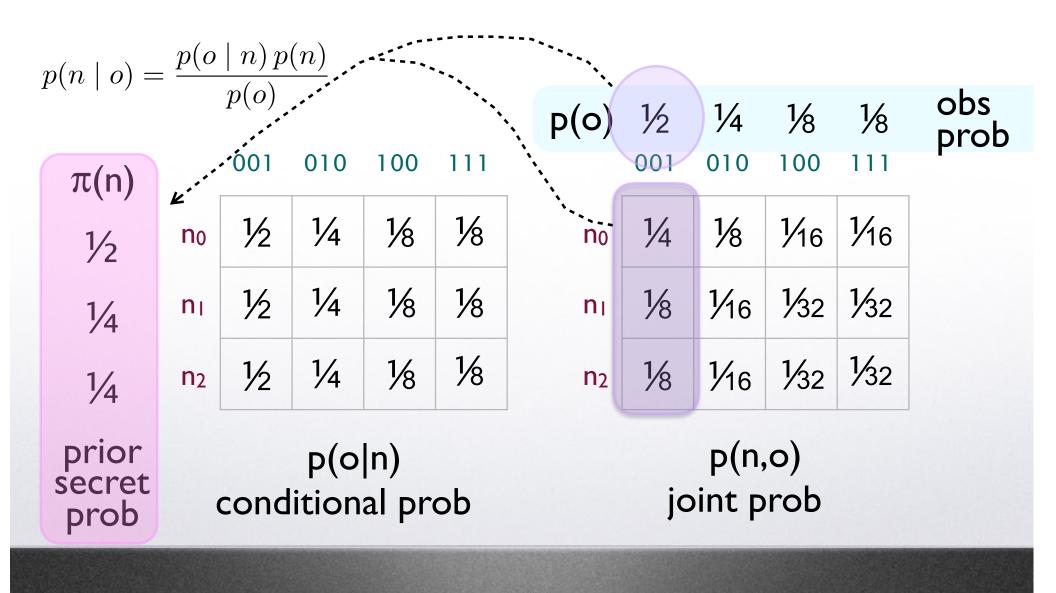
32



conditional prob prob





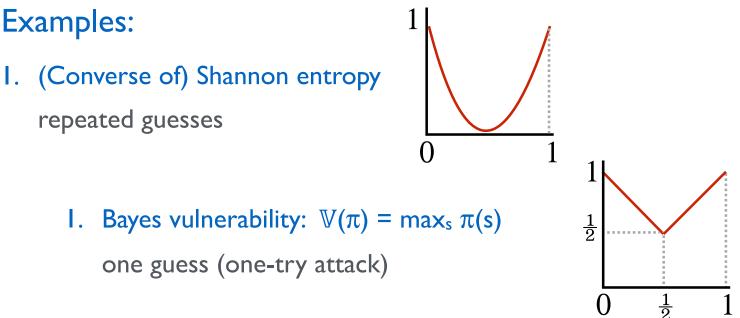


35

Formal measures of leakage

Vulnerability $\mathbb{V}(\pi)$ of a secret with prior π

- Represents the "expected damage" the adversary can inflict by making his best guess about the secret value
- The exact definition depends on the operational model of adversary
- Common feature: $\mathbb V$ is convex on π
- Convexity is a consequence of Data Processing Inequality



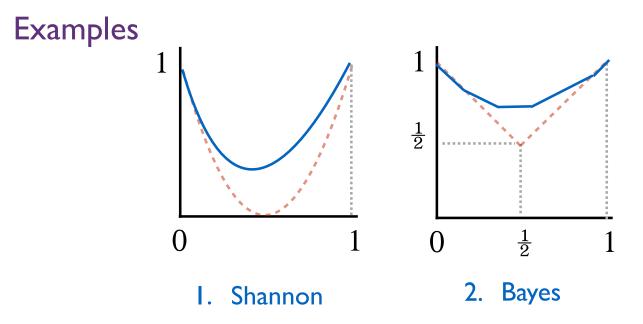
Formal measures of leakage

Posterior Vulnerability $\mathbb{V}(\pi, \mathbb{C})$ of a secret with prior π observed through a channel \mathbb{C}

• Vulnerability of the secret after the adversary observes the output: the expected vulnerability of the posterior distributions

$$\mathbb{V}(\pi, \mathbb{C}) = \sum_{o} p(o) \ \mathbb{V}(p(\cdot | o))$$

• Convex in π



$\mathbb{V}(\pi, \mathbb{C})$ is also convex in \mathbb{C}

The defender may lower the vulnerability by randomly combining different channels

Important: random protocols can always be seen as a random combination of deterministic protocols

$\mathbb{V}(\pi, \mathbb{C})$ is also convex in \mathbb{C}

The defender may lower the vulnerability by randomly combining different channels

Important: random protocols can always be seen as a random combination of deterministic protocols

Example: DC with 2 nodes, 2 biased coins

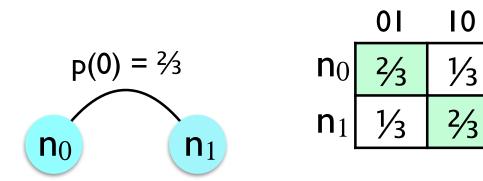


$\mathbb{V}(\pi, \mathbb{C})$ is also convex in \mathbb{C}

The defender may lower the vulnerability by randomly combining different channels

Important: random protocols can always be seen as a random combination of deterministic protocols

Example: DC with 2 nodes, 2 biased coins

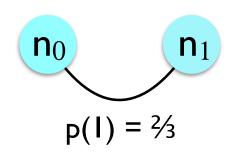


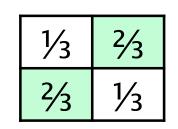
$\mathbb{V}(\pi, \mathbb{C})$ is also convex in \mathbb{C}

The defender may lower the vulnerability by randomly combining different channels

Important: random protocols can always be seen as a random combination of deterministic protocols

Example: DC with 2 nodes, 2 biased coins



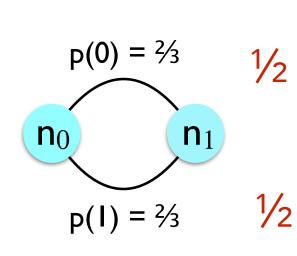


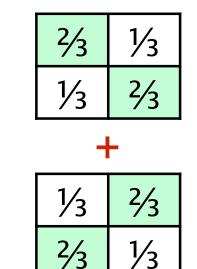
$\mathbb{V}(\pi, \mathbb{C})$ is also convex in \mathbb{C}

The defender may lower the vulnerability by randomly combining different channels

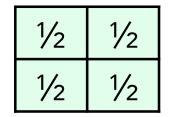
Important: random protocols can always be seen as a random combination of deterministic protocols

Example: DC with 2 nodes, 2 biased coins





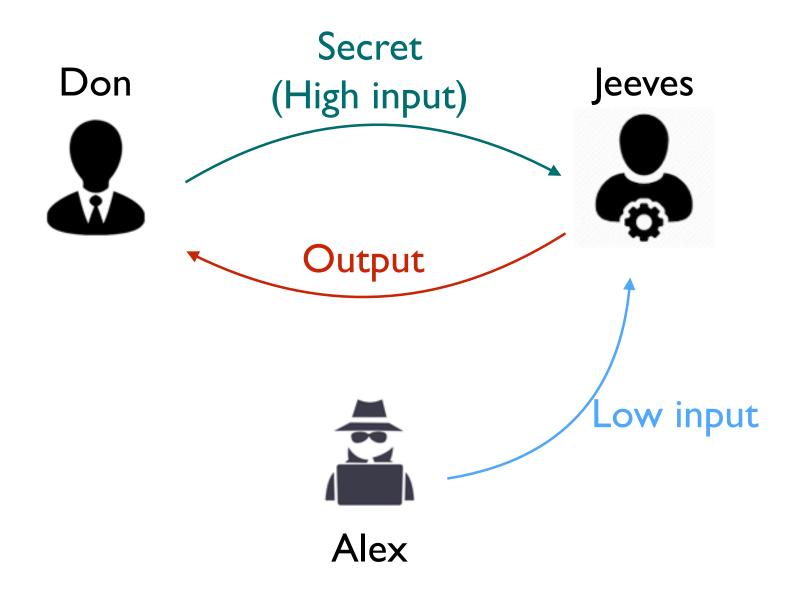
The completely opaque channel



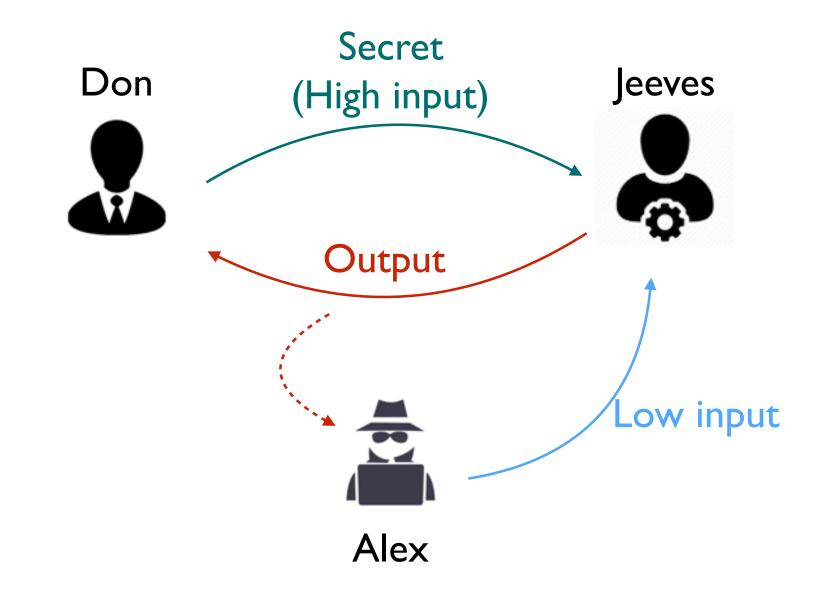
Active adversary

- The adversary may interfere with the system
 - For instance, in the DC, the adversary may control one or more coins
 - More typically, the adversary influences the system by changing the value of some inputs (low inputs)
- As a result, the adversary may change the channel matrix as well

Example: the two millionaires problem



Example: the two millionaires problem



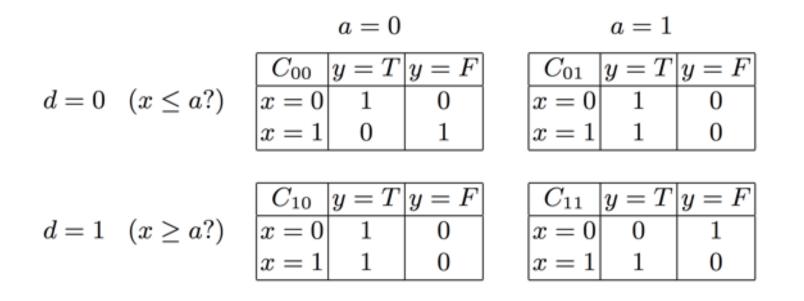
Reducing the vulnerability

Jeeves can run two different programs, both serving the purpose.

Don sends to Jeeves a bit **d** indicating which program he should run

Program 0 High Input: $x \in \{0, 1\}$ Low Input: $a \in \{0, 1\}$ Output: $y \in \{T, F\}$ return $x \leq a$ Program 1

High Input: $x \in \{0, 1\}$ Low Input: $a \in \{0, 1\}$ Output: $y \in \{T, F\}$ return $x \ge a$ Depending on the choices a (adversary) and d (defender), we get the following channel matrices:



This can be modeled as a 0-sum game, where the actions a and d are the pure strategies, and the payoff is the leakage (or equivalently, the posterior vulnerability).

The adversary wants to maximize the vulnerability, while the defender wants to minimize it

$$a = 0 \qquad a = 1$$

$$d = 0 \quad (x \le a?) \qquad \boxed{\begin{array}{c} C_{00} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 0 \quad 1 \end{array}} \qquad \boxed{\begin{array}{c} C_{01} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}}$$

$$d = 1 \quad (x \ge a?) \qquad \boxed{\begin{array}{c} C_{10} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} \qquad \boxed{\begin{array}{c} C_{11} \quad y = T \mid y = F \\ x = 0 \quad 0 \quad 1 \\ x = 1 \quad 1 \quad 0 \end{array}}$$

$$a = 0 \qquad a = 1$$

$$d = 0 \quad (x \le a?) \qquad \boxed{\begin{array}{c} C_{00} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 0 \quad 1 \end{array}} \qquad \boxed{\begin{array}{c} C_{01} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}}$$

$$d = 1 \quad (x \ge a?) \qquad \boxed{\begin{array}{c} C_{10} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} \qquad \boxed{\begin{array}{c} C_{11} \quad y = T \mid y = F \\ x = 0 \quad 0 \quad 1 \\ x = 1 \quad 1 \quad 0 \end{array}}$$

$$a = 0 \qquad a = 1$$

$$d = 0 \quad (x \le a?) \qquad \boxed{\begin{array}{c} C_{00} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 0 \quad 1 \end{array}} 1 \qquad \boxed{\begin{array}{c} C_{01} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2$$

$$d = 1 \quad (x \ge a?) \qquad \boxed{\begin{array}{c} C_{10} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2 \qquad \boxed{\begin{array}{c} C_{11} \quad y = T \mid y = F \\ x = 0 \quad 0 \quad 1 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2$$

$$a = 0 \qquad a = 1$$

$$d = 0 \quad (x \le a?) \qquad \boxed{\begin{array}{c} C_{00} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 0 \quad 1 \end{array}} 1 \qquad \boxed{\begin{array}{c} C_{01} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2$$

$$d = 1 \quad (x \ge a?) \qquad \boxed{\begin{array}{c} C_{10} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2 \qquad \boxed{\begin{array}{c} C_{11} \quad y = T \mid y = F \\ x = 0 \quad 0 \quad 1 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2$$

Payoff table

$$egin{array}{|c|c|c|c|} \mathbb{V} & a = 0 & a = 1 \\ \hline d = 0 & 1 & \frac{1}{2} \\ \hline d = 1 & \frac{1}{2} & 1 \\ \end{array}$$

$$a = 0 \qquad a = 1$$

$$d = 0 \quad (x \le a?) \qquad \boxed{\begin{array}{c} C_{00} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 0 \quad 1 \end{array}} 1 \qquad \boxed{\begin{array}{c} C_{01} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2$$

$$d = 1 \quad (x \ge a?) \qquad \boxed{\begin{array}{c} C_{10} \quad y = T \mid y = F \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2 \qquad \boxed{\begin{array}{c} C_{11} \quad y = T \mid y = F \\ x = 0 \quad 0 \quad 1 \\ x = 1 \quad 1 \quad 0 \end{array}} 1/2$$

Payoff table

V	a = 0	a = 1
d = 0	1	1/2
d = 1	1/2	1

Similar to the game of the matching pennies

Nash Equilibrium: we have a NE when neither player has any interest to change his strategy unilaterally

Nash Equilibrium: we have a NE when neither player has any interest to change his strategy unilaterally

Usually there is no pure NE, but there is always a **mixed NE**

Nash Equilibrium: we have a NE when neither player has any interest to change his strategy unilaterally

Usually there is no pure NE, but there is always a **mixed NE**

\mathbb{V}	a = 0	a = 1
d = 0	1	1/2
d = 1	1/2	1

Nash Equilibrium: we have a NE when neither player has any interest to change his strategy unilaterally

Usually there is no pure NE, but there is always a **mixed NE**

	\mathbb{V}	a = 0	a = 1
Þ	d = 0	1	1/2
1 -p	d = 1	1/2	1

Nash Equilibrium: we have a NE when neither player has any interest to change his strategy unilaterally

Usually there is no pure NE, but there is always a **mixed NE**

$$q \quad 1-q$$

$$\boxed{\mathbb{V} \quad a = 0 \quad a = 1}$$

$$p \quad d = 0 \quad 1 \quad \frac{1}{2}$$

$$1-p \quad d = 1 \quad \frac{1}{2} \quad 1$$

Nash Equilibrium: we have a NE when neither player has any interest to change his strategy unilaterally

Usually there is no pure NE, but there is always a **mixed NE**

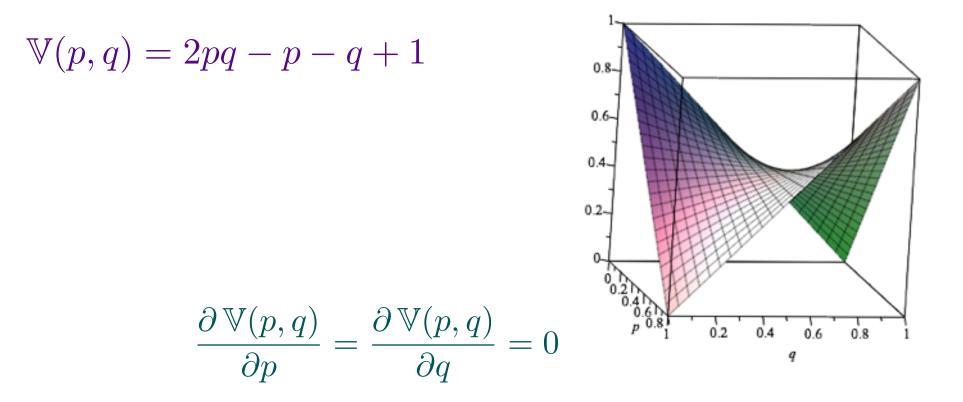
$$q \quad 1-q$$

$$\boxed{\mathbb{V} \quad a = 0 \quad a = 1}$$

$$p \quad d = 0 \quad 1 \quad \frac{1}{2}$$

$$1-p \quad d = 1 \quad \frac{1}{2} \quad 1$$

$$\mathbb{V}(p,q) = 1pq + \frac{1}{2}(1-p)q + \frac{1}{2}p(1-q) + 1(1-p)(1-q)$$
$$= 2pq - p - q + 1$$

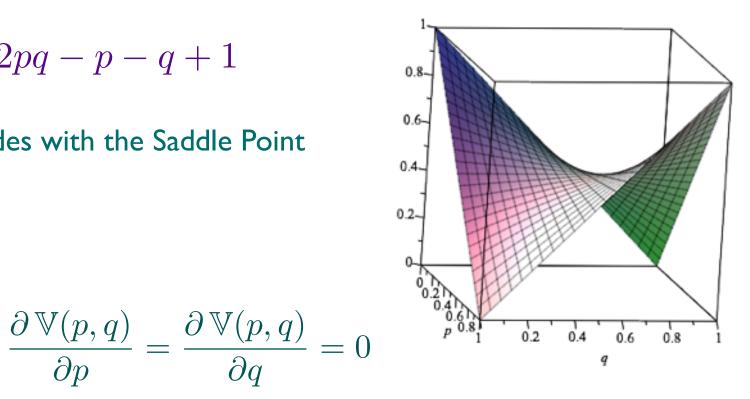


Van Neumann Th: If f(p,q) is convex in p and concave in q, then

$$\min_{p} \max_{q} f(p,q) = \max_{q} \min_{p} f(p,q)$$

$$\mathbb{V}(p,q) = 2pq - p - q + 1$$

The NE coincides with the Saddle Point



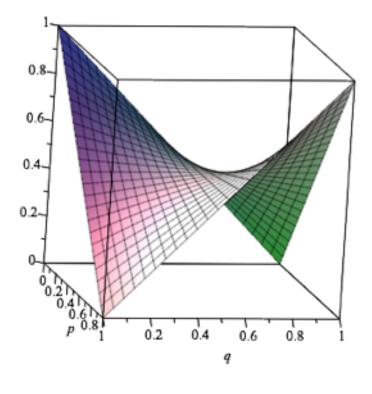
Van Neumann Th: If f(p,q) is convex in p and concave in q, then

$$\min_{p} \max_{q} f(p,q) = \max_{q} \min_{p} f(p,q)$$

$$\mathbb{V}(p,q) = 2pq - p - q + 1$$

The NE coincides with the Saddle Point

When the partial derivatives exist, the Saddle Point can be computed by imposing $\frac{\partial \mathbb{V}(p,q)}{\partial p} = \frac{\partial \mathbb{V}(p,q)}{\partial q} = 0$



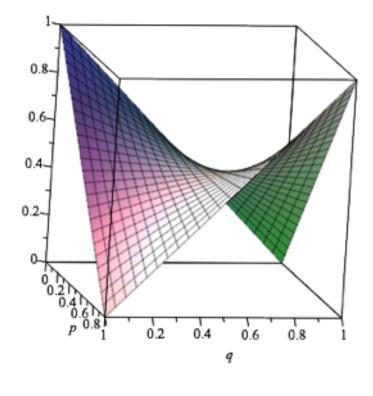
Van Neumann Th: If f(p,q) is convex in p and concave in q, then

$$\min_{p} \max_{q} f(p,q) = \max_{q} \min_{p} f(p,q)$$

$$\mathbb{V}(p,q) = 2pq - p - q + 1$$

The NE coincides with the Saddle Point

When the partial derivatives exist, the Saddle Point can be computed by imposing $\frac{\partial \mathbb{V}(p,q)}{\partial p} = \frac{\partial \mathbb{V}(p,q)}{\partial q} = 0$



In the example, the Saddle Point is for p = q = 1/2

Van Neumann Th: If f(p,q) is convex in p and concave in q, then

$$\min_{p} \max_{q} f(p,q) = \max_{q} \min_{p} f(p,q)$$

Non-standard games

Assume now that Don wants to know the binary sum

Again, Jeeves has two programs and Don sends to Jeeves a bit d indicating which program he should run

Program 0

High Input: $x \in \{0, 1\}$ Low Input: $a \in \{0, 1\}$ Output: $y \in \{0, 1\}$ return $x \oplus a$ Program 1

High Input: $x \in \{0, 1\}$ Low Input: $a \in \{0, 1\}$ Output: $y \in \{0, 1\}$ return $x \oplus a \oplus 1$

Corresponding channel matrices:

$$a = 0 \qquad a = 1$$

$$d = 0 \quad (x \oplus a) \qquad \boxed{\begin{array}{c} C_{00} \quad y = 0 \quad y = 1 \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 0 \quad 1 \end{array}} \qquad \boxed{\begin{array}{c} C_{01} \quad y = 0 \quad y = 1 \\ x = 0 \quad 0 \quad 1 \\ x = 1 \quad 1 \quad 0 \end{array}}$$

$$d = 1 \quad (x \oplus a \oplus 1) \qquad \boxed{\begin{array}{c} C_{10} \quad y = 0 \quad y = 1 \\ x = 0 \quad 0 \quad 1 \\ x = 1 \quad 1 \quad 0 \end{array}} \qquad \boxed{\begin{array}{c} C_{11} \quad y = 0 \quad y = 1 \\ x = 0 \quad 1 \quad 0 \\ x = 1 \quad 0 \quad 1 \end{array}}$$

Payoff table

V	a = 0	a = 1
d = 0	1	1
d = 1	1	1

From standard game theory the utility would be I and all strategies would be equivalent

However, in the case of our games, the saddle point is $(\frac{1}{2}, \frac{1}{2})$, and the Utility is 1/2

In fact, if the probability of d = 0 is p (strategy of the defender), from the point of view of the adversary the channels are as follows

Clearly, the optimal strategy of the defender is for p = 1/2, which gives perfectly opaque channels whatever action the attacker chooses

Explanation

- The reason why we get different results than in standard game theory is because the standard utility function is defined as expectation, hence it is affine on the strategies of both players
- In contrast, our games are **convex** on the strategy of the defender (and affine in that of the attacker)
- Van Neumann minimax theorem is still applicable
- Unfortunately, in general the partial derivatives do not exist
- However the saddle point can still be computed by convex analysis

Conclusion

- Probabilistic composition of protocols can be useful to mitigate the Information leakage
- If the attacker is active, then the attacker also has interest to use a probabilistic strategy
- We can model the interplay defender-attacker in Game Theory
- The games are non-standard, but the optimal strategies still exist and can be computed by convex analysis

Future work

- Explore the relation with risk-adverse players
- The goals of the adversary and of the defender may be different => non 0-sum games
- Both adversary and defender may have multiple goals => multiple utility games
- Repeated attacks (repeated runs of the protocol)
 => repeated games
- Other kinds of interaction => Simultaneous vs alternate games
- Develop the theory of protocol composition (choice and sequential composition)

Thank you !

