Defendable Security in Interaction Protocols

Wojtek Jamroga, University of Luxembourg
(joint work with Matthijs Melissen and Henning Schnoor)

Strategic Reasoning @ Grenoble, 5/04/2014

Outline

Outline

Introduction

- **Interaction protocols** are ubiquitous in IT systems
- As soon as two machines communicate, a protocol is required
- Goal of a protocol: set of outcomes that are preferred
  Outcomes = runs
- Protocol is correct wrt its goal iff the goal is achieved in all runs where a predefined subset of players follows the protocol

Introduction

- **Interaction protocols** are ubiquitous in IT systems
- As soon as two machines communicate, a protocol is required
- Goal of a protocol: set of outcomes that are preferred
  Outcomes = runs
- Protocol is correct wrt its goal iff the goal is achieved in all runs where a predefined subset of players follows the protocol
- Too strong? What if the goal can be violated only by **irrational** actions of “adversaries”?

Introduction

- **Interaction protocols** are ubiquitous in IT systems
- As soon as two machines communicate, a protocol is required
- Goal of a protocol: set of outcomes that are preferred
  Outcomes = runs
- Protocol is correct wrt its goal iff the goal is achieved in all runs where a predefined subset of players follows the protocol
- Too strong? What if the goal can be violated only by **irrational** actions of “adversaries”?
- Too weak? What if the goal can be achieved only through **irrational** actions of “defenders”?

Introduction

- Incentive-based correctness: correct iff the goal holds in all runs that execute **rational** strategies [Dodis and Rabin 2007]
- For this, we need an idea of: incentives of participants, rationality of choices given incentives

  $\rightsquigarrow$ Game-theoretic view

Introduction

- Incentive-based correctness: correct iff the goal holds in all runs that execute **rational** strategies [Dodis and Rabin 2007]
- For this, we need an idea of: incentives of participants, rationality of choices given incentives

  $\rightsquigarrow$ Game-theoretic view
- **What if the actual incentives are unknown?**

Introduction

- Incentive-based correctness: correct iff the goal holds in all runs that execute **rational** strategies [Dodis and Rabin 2007]
- For this, we need an idea of: incentives of participants, rationality of choices given incentives

  $\rightsquigarrow$ Game-theoretic view
- **What if the actual incentives are unknown?**
- Idea: assume that a subset of participants ("defenders") is in favor of the goal
- ...and check if the protocol is correct for every distribution of incentives consistent with the assumption

Introduction

Contribution of this paper:

1. We generalize the idea of Dodis and Rabin

   $\rightsquigarrow$ Game-theoretic correctness parameterized with incentives of participants and a notion of rationality

Introduction

Contribution of this paper:

1. We generalize the idea of Dodis and Rabin

   $\rightsquigarrow$ Game-theoretic correctness parameterized with utilities of participants and a solution concept

Introduction

Contribution of this paper:

1. We generalize the idea of Dodis and Rabin

   $\rightsquigarrow$ Game-theoretic correctness parameterized with utilities of participants and a solution concept

2. The concept of **defendability of a protocol** by a given set of defenders

Introduction

Contribution of this paper:

1. We generalize the idea of Dodis and Rabin

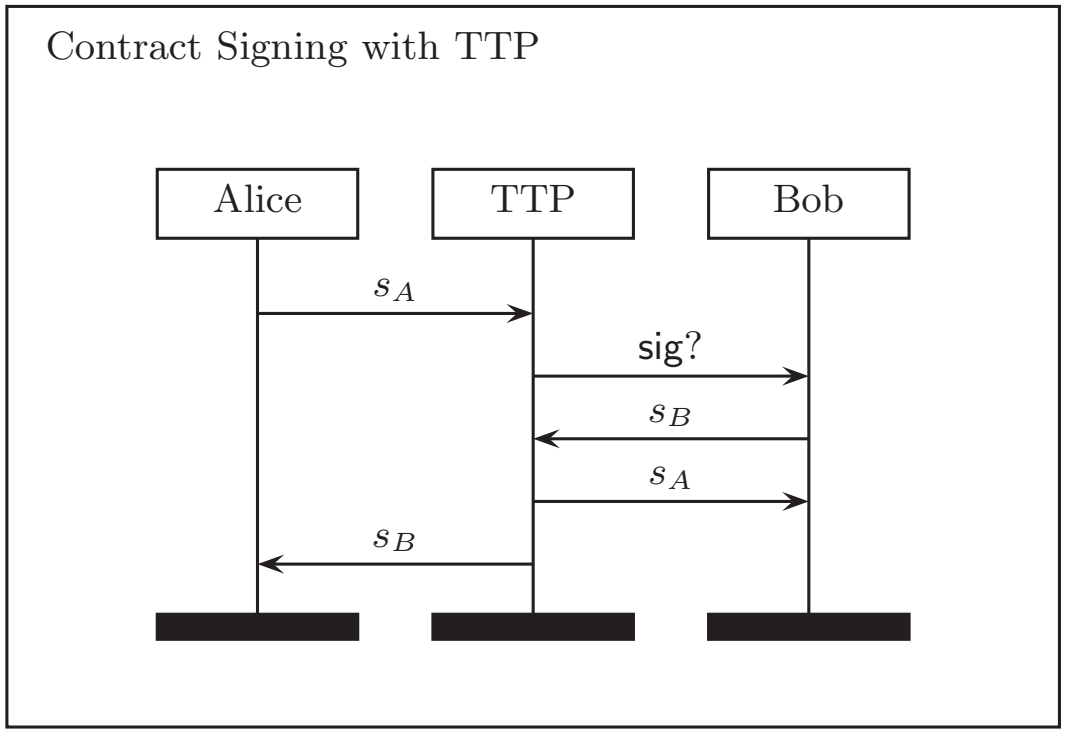
   $\rightsquigarrow$ Game-theoretic correctness parameterized with utilities of participants and a solution concept

2. The concept of **defendability of a protocol** by a given set of defenders

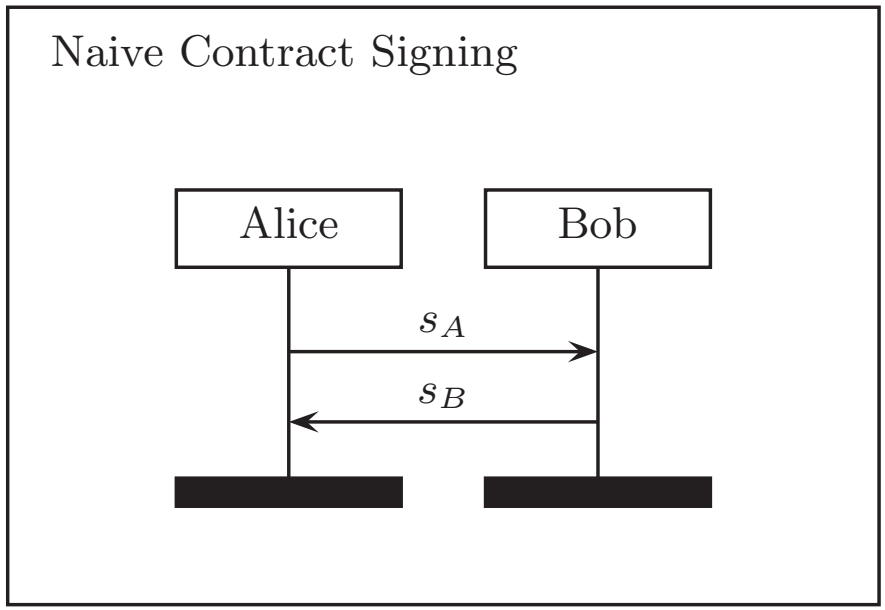
3. ...Plus some theorems :-)

Outline

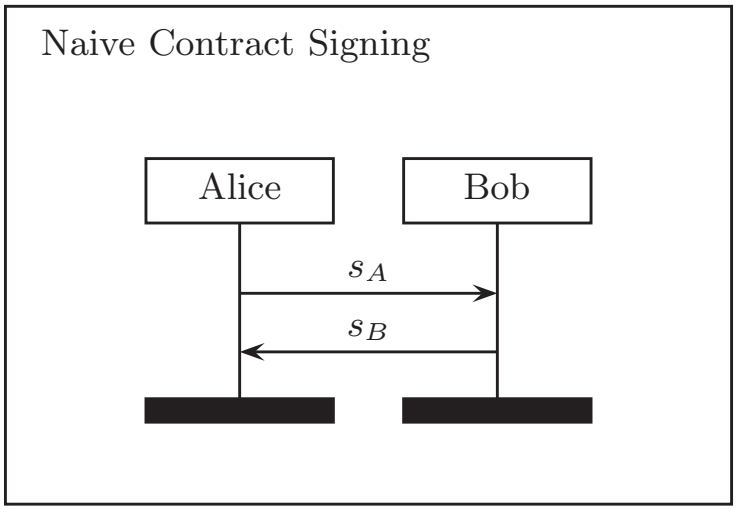
Security Protocols



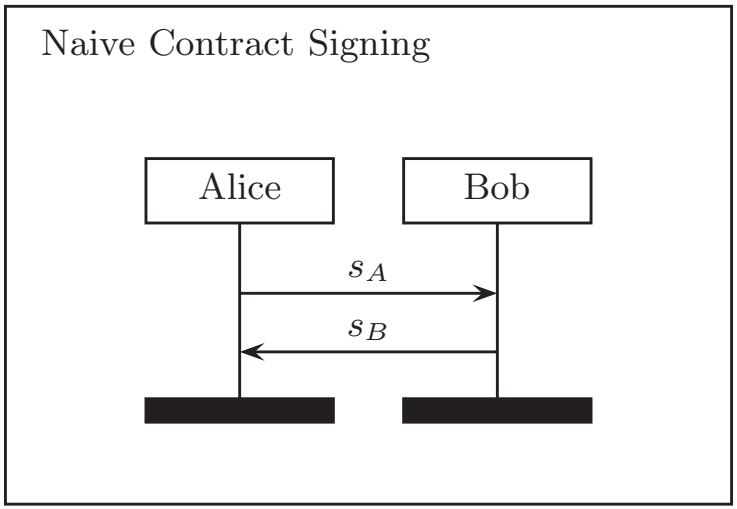
Security Protocols

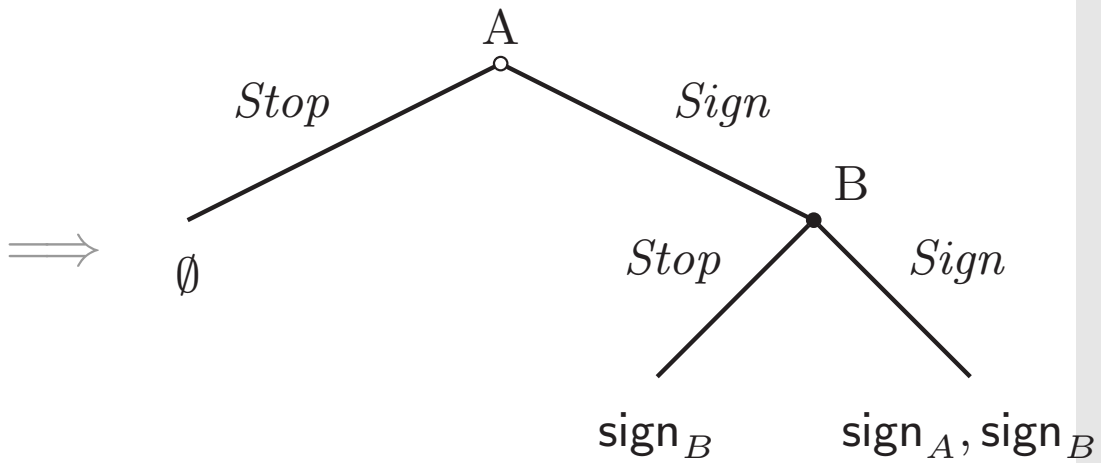


From Protocols to Games

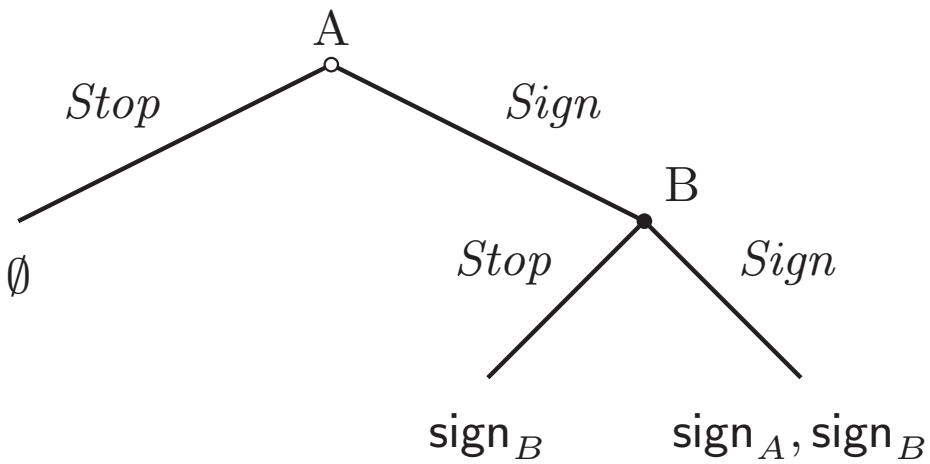


From Protocols to Games

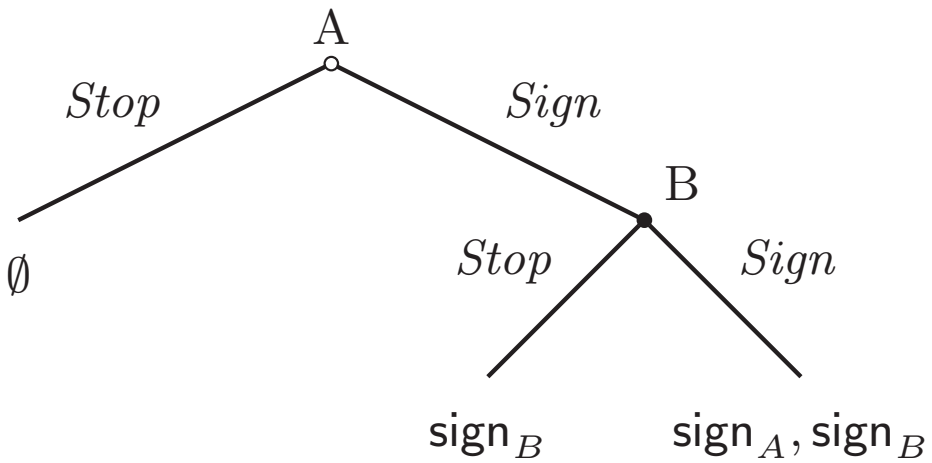




From Protocols to Games



From Protocols to Games



$\Longrightarrow$

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\emptyset$	$\emptyset$
$Sign$	$\{\mathsf{sign}_B\}$	$\{\mathsf{sign}_A, \mathsf{sign}_B\}$

Protocols as Games

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\emptyset$	$\emptyset$
$Sign$	$\{\mathsf{sign}_B\}$	$\{\mathsf{sign}_A, \mathsf{sign}_B\}$

Goal of the protocol $=$ a subset of behaviors

$=$ a subset of cells in the table

Protocols as Games

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\emptyset$	$\emptyset$
$Sign$	$\{\text{sign}_B\}$	$\{\text{sign}_A, \text{sign}_B\}$

Goal of the protocol $=$ a subset of behaviors

$=$ a subset of cells in the table

Protocols as Games

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\emptyset$	$\emptyset$
$Sign$	$\{\mathsf{sign}_B\}$	$\{\mathsf{sign}_A, \mathsf{sign}_B\}$

Incentives $=$ utility values

Protocols as Games

$A \backslash B$	$Stop$	$Sign$
$Stop$	$0, 0$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

Incentives $=$ utility values

Protocols as Games

$A \backslash B$	$Stop$	$Sign$
$Stop$	$0, 0$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

Rational play $=$ a subset of behaviors

$=$ a subset of cells in the table

Protocols as Games

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\mathbf{0, 0}$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

Rational play $=$ a subset of behaviors

$=$ a subset of cells in the table

Outline

Incentive-Based Correctness

Definition

A protocol P with utility profile u is **correct** with respect to goal γ under solution concept SC iff:

$$\begin{cases} SC(P,u) \subseteq \gamma & \text{if } SC(P,u) \neq \emptyset \\ \gamma = \textit{all outcomes} & \text{otherwise.} \end{cases}$$

Incentive-Based Correctness

$A \backslash B$	$Stop$	$Sign$
$Stop$	$0, 0$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

Incentive-Based Correctness

$A \backslash B$	$Stop$	$Sign$
$Stop$	$0, 0$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

Incentive-Based Correctness

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\mathbf{0, 0}$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

Incentive-Based Correctness

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\mathbf{0, 0}$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

For this distribution of incentives, and under Nash equilibrium, the naive protocol is correct!

Incentive-Based Correctness

$A \backslash B$	$Stop$	$Sign$
$Stop$	$\mathbf{0, 0}$	$0, 0$
$Sign$	$-1, 2$	$1, 1$

For this distribution of incentives, and under Nash equilibrium, the naive protocol is correct!

What if the actual incentives are **unknown**?

Outline

Defendability of Protocols

Definition

A group of agents D **supports** goal γ in game (P, u) iff, for all $i \in D$, if $s \in \gamma$ and $s' \in \overline{\gamma}$ then $u_i(s) > u_i(s')$.

Defendability of Protocols

Definition

A group of agents D **supports** goal γ in game (P, u) iff, for all $i \in D$, if $s \in \gamma$ and $s' \in \overline{\gamma}$ then $u_i(s) > u_i(s')$.

Protocol P is **defended by agents** D with respect to goal γ and solution concept SC iff (P, u) is correct wrt γ, SC for all utility profiles u such that D support γ in (P, u).

Example: Fairness in Contract Signing

Is fairness defendable by Alice under Nash equilibrium?

$A \backslash B$	$Stop$	$Sign$
$Stop$		
$Sign$		

Example: Fairness in Contract Signing

Is fairness defendable by Alice under Nash equilibrium?

$A \backslash B$	$Stop$	$Sign$
$Stop$	$1, 0$	$0, 1$
$Sign$	$-1, 1$	$1, 0$

Example: Fairness in Contract Signing

Is fairness defendable by Alice under Nash equilibrium?

$A \backslash B$	$Stop$	$Sign$
$Stop$	$1, 0$	$0, 1$
$Sign$	$-1, 1$	$1, 0$

Fairness in contract signing is not defendable by Alice under NE

Example: Fairness in Contract Signing

Is fairness defendable by the grand coalition (Alice and Bob together)?

$A \backslash B$	$Stop$	$Sign$
$Stop$	$1, 0$	$0, 1$
$Sign$	$-1, 1$	$1, 0$

Example: Fairness in Contract Signing

Is fairness defendable by the grand coalition (Alice and Bob together)?

$A \backslash B$	$Stop$	$Sign$
$Stop$	$1, 0$	$0, 1$
$Sign$	$-1, -1$	$1, 0$

Example: Fairness in Contract Signing

Is fairness defendable by the grand coalition (Alice and Bob together)?

$A \backslash B$	$Stop$	$Sign$
$Stop$	$1, 0$	$0, 1$
$Sign$	$-1, -1$	$1, 0$

Seems like it should be...

Defendability under Nash Equilibrium

Theorem

Let P be a finite protocol, and γ a nontrivial objective in P.

Then, γ **is defendable in P by the grand coalition under Nash equilibrium** *iff*

1. *the deviation closure of γ covers the whole payoff table, and*
2. *there is a strategy profile in γ that belongs to no strategic knots in γ.*

Defendability under Nash Equilibrium

Deviation closure:

$A \backslash B$	$Stop$	$Sign$
$Stop$		
$Sign$		

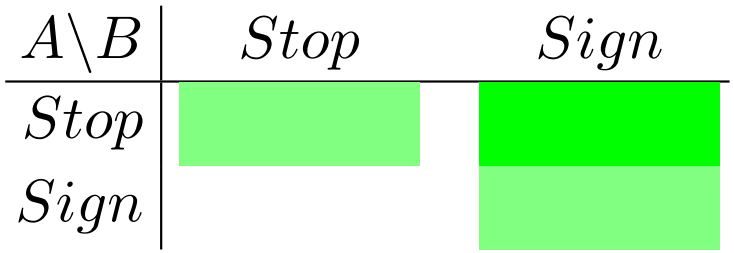
Defendability under Nash Equilibrium

Deviation closure:

$A \backslash B$	$Stop$	$Sign$
$Stop$		
$Sign$		

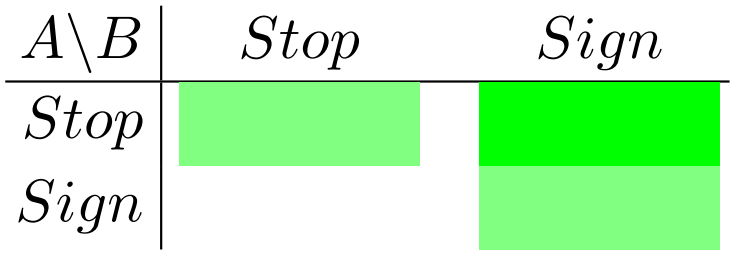
Defendability under Nash Equilibrium

Deviation closure:

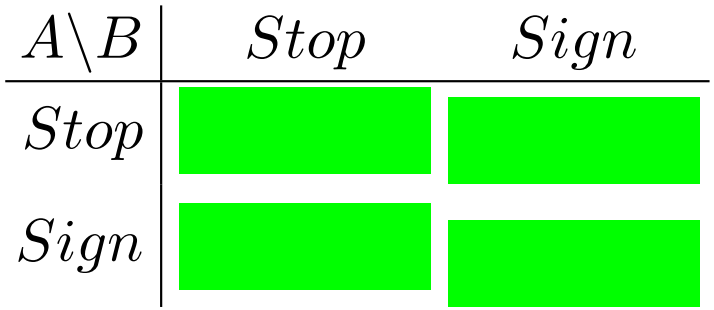


Defendability under Nash Equilibrium

Deviation closure:

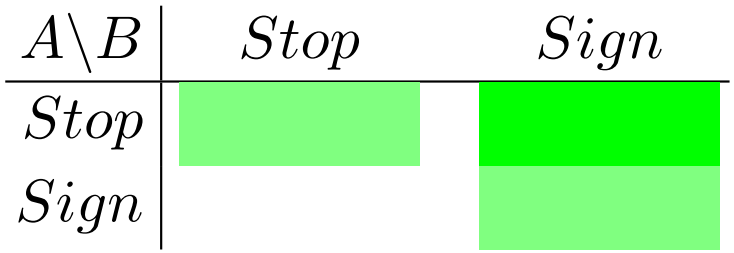


Strategic knot:

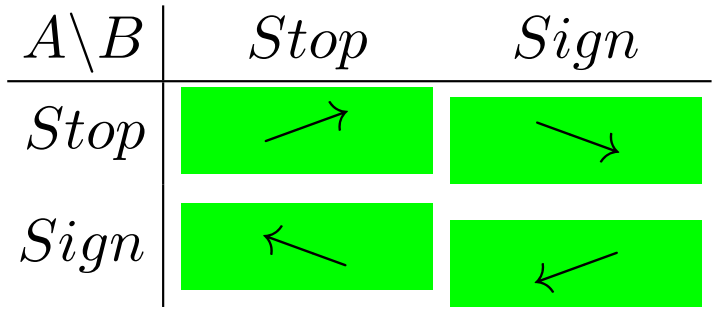


Defendability under Nash Equilibrium

Deviation closure:



Strategic knot:



Defendability of Fairness in Contract Signing

$A \backslash B$	$Stop$	$Sign$
$Stop$		
$Sign$		

Fairness in contract signing **is** defendable by Alice and Bob under Nash Equlibrium

Other Results

What else is in the paper:

- Characterization of defendability under Pareto-optimal Nash equilibrium
- Definition and characterization of defendability for mixed strategies
- Characterizations of defendability for non-injective games
- Complexity results for related decision problems
  Complexity between NP/co-NP and PSPACE
  $\rightsquigarrow$ Characterization results decrease the complexity!
- Analysis of the ASW contract signing protocol

Wojciech Jamroga, Matthijs Melissen and Henning Schnoor (2013), *Defendable Security in Interaction Protocols*. Proceedings of the 16th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA 2013). LNCS vol. 8291, Springer.

Outline

Conclusions

- A framework for analyzing interaction protocols, based on incentives and rationality of agents.
- Novel notion of **defendability**: the protocol is correct as long as a given subset of the participants (the "defenders") is in favor of the security property
- Formal results: characterizations of defendability, computational complexity

Thank you for your attention!