

# Introduction to formal protocol analysis

E-Voting Seminar  
18 November 2010

**Michael Schläpfer**  
(Some parts adapted from S. Mödersheim)



# Formal methods in information security

Formal methods are **techniques** and **tools** based on **mathematics** and **logic** that support the **specification**, **construction** and **analysis** of hardware and software systems.

## Some examples:

- Program logics (Hoare logic, dynamic logic)
- Temporal logics (LTL, CTL, TLA,  $\mu$ -calculus)
- Process algebras (CCS, CSP,  $\pi$ -calculus, Spi-calculus)
- Abstract data types (CASL, Z)
- Development tools (Rodin/Event-B, PVS, VSE)
- Theorem provers (Isabelle, Coq, HOL, Inka)
- Model checkers (Spin, SMV, Mur $\phi$ , OFMC, Scyther)

## Applying formal methods:

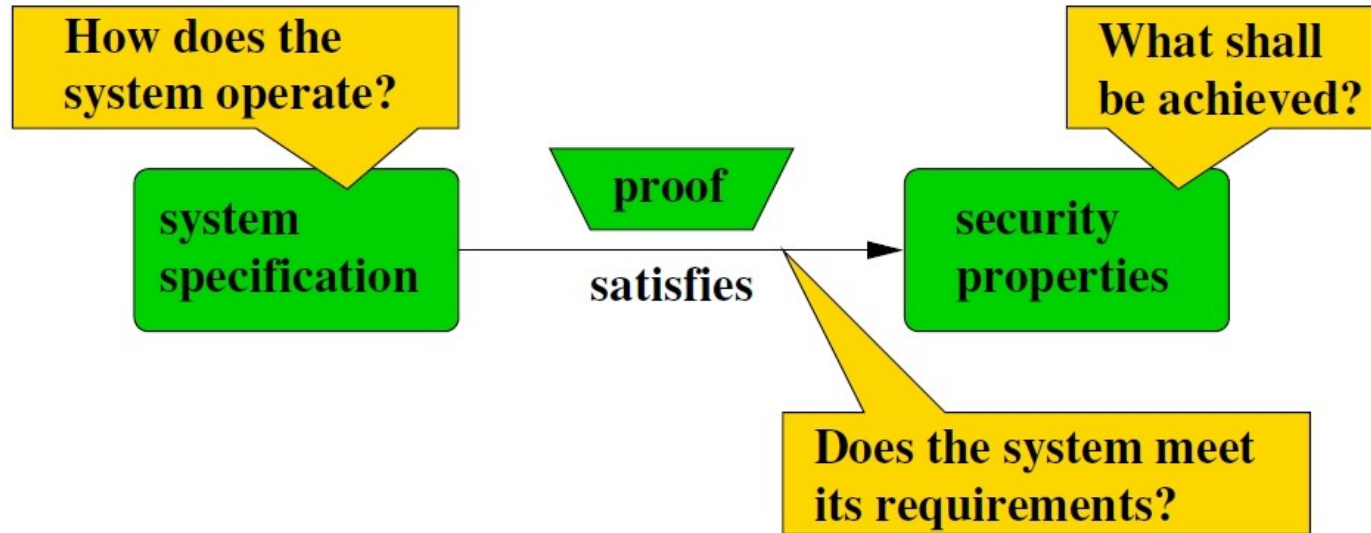
1. Formalize the system requirements as security properties
2. Construct a formal model of the system's behavior, an abstract specification or a concrete program
3. Verify that the system satisfies the properties at the level at which the system has been modeled

# Overview

## 1. Introduction

2. Formal languages for the specification of security protocols
3. The Dolev-Yao intruder model
4. Operational semantics of security protocols
5. Protocol goals and verification
6. Decidability of protocol security and deductive methods
7. Current research topics

# Formal security models



- Formal specification with formal languages
- Semantics of languages allow for verification and validation with mathematical methods

# Good crypto alone ...



# Security protocols

A **protocol** consists of a set of rules (conventions) that determine the exchange of messages between two or more principals. In short, a **distributed algorithm** with emphasis on communication.

**Security** (or **cryptographic**) protocols use cryptographic mechanisms to achieve security objectives.

## Some common security objectives:

- Entity or message authentication
- Key establishment
- Integrity
- Fair exchange
- Non-repudiation
- ...

# Overview

1. Introduction

2. Formal languages for the specification of security protocols

3. The Dolev-Yao intruder model

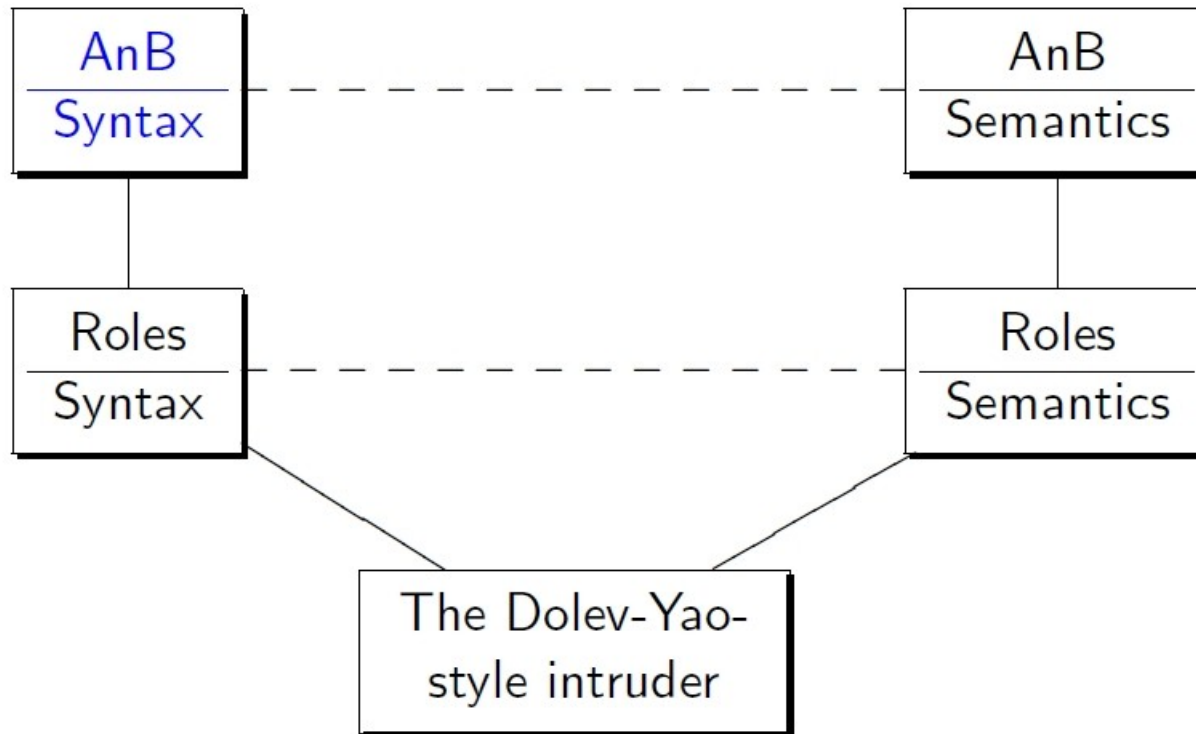
4. Operational semantics of security protocols

5. Protocol goals and verification

6. Decidability of protocol security and deductive methods

7. Current research topics

# Two formal languages





# Message notation

**Roles:**  $A, B$  or *Alice, Bob*

**Agents:**  $a, b, i$

**Symmetric Keys:**  $K, K_{AB}, \dots; \text{sk}(A, S)$

**Symmetric Encryption:**  $\{M\}_K$

**Public Keys:**  $K, \text{pk}(A)$

**Private Keys:**  $\text{inv}(K), \text{inv}(\text{pk}(A))$

**Asymmetric Encryption:**  $\{M\}_K$

**Signing:**  $\{M\}_{\text{inv}(K)}$

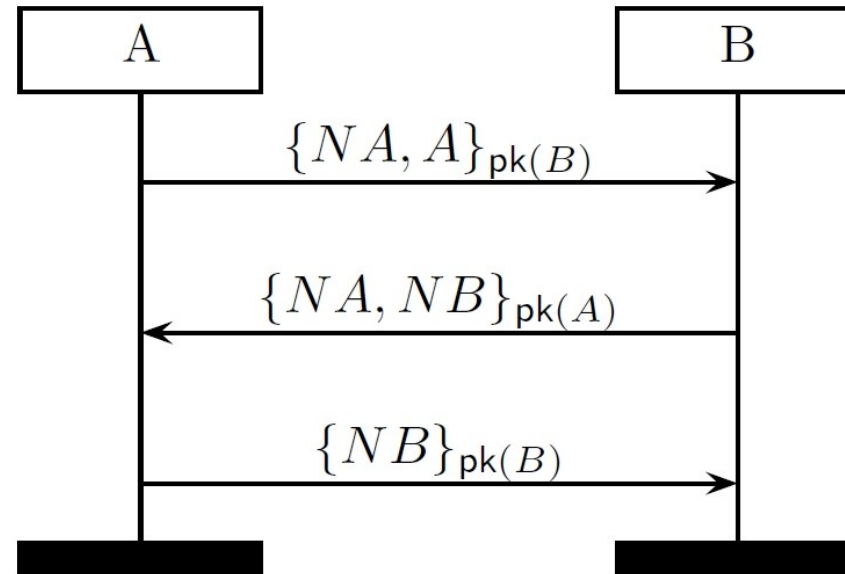
**Nonces:**  $N_A, N_1$  fresh data items used for challenge/response.

N.B.: sometimes subscripts are used, e.g.  $N_A$ , but it does not mean that principals can find out that  $N_A$  was generated by  $A$ .

**Timestamps:**  $T$ . Denote time, e.g. used for key expiration.

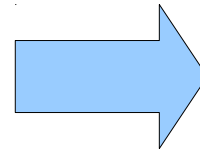
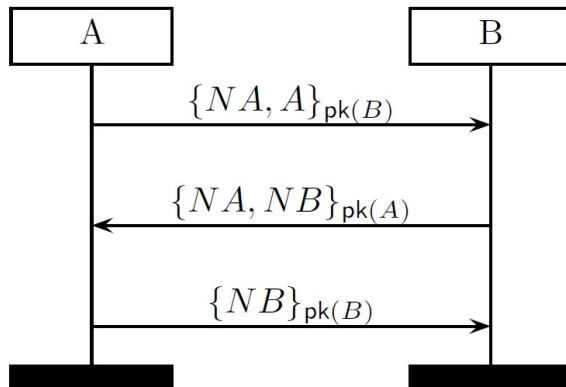
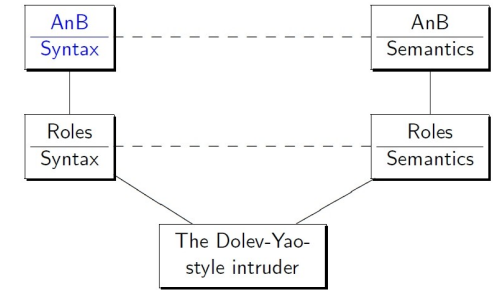
**Message concatenation:**  $M_1, M_2, M_3$

# A simple protocol description language



- Message sequence chart between roles A and B
- Represents the famous Needham-Schroeder Public Key protocol (NSPK, 1978)

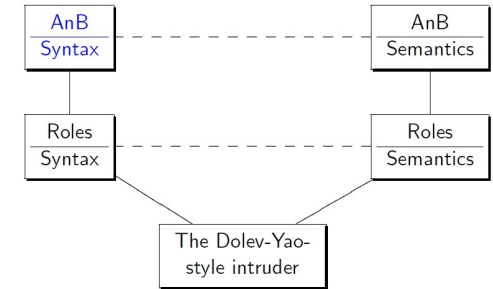
# AnB - Syntax



1.  $A \rightarrow B : \{NA, A\}_{pk(B)}$
2.  $B \rightarrow A : \{NA, NB\}_{pk(A)}$
3.  $A \rightarrow B : \{NB\}_{pk(B)}$

- Message sequence chart between roles A and B
- Represents the famous Needham-Schroeder Public Key protocol (NSPK, 1978)

# Informal correctness



1.  $A \rightarrow B : \{NA, A\}_{pk(B)}$
2.  $B \rightarrow A : \{NA, NB\}_{pk(A)}$
3.  $A \rightarrow B : \{NB\}_{pk(B)}$

“This is Alice and I have chosen a nonce  $NA$ .”

“Here is your nonce  $NA$ . Since I could read it, I must be Bob. I also have a challenge  $NB$  for you.”

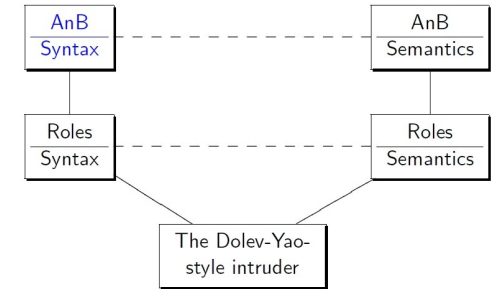
“You sent me  $NB$ . Since only Alice can read this and I sent it back, I must be Alice.”

▪ Secure?

# Lowe's attack

NSPK (1978)

$$\begin{aligned}
 A &\rightarrow B : \{NA, A\}_{pk(B)} \\
 B &\rightarrow A : \{NA, NB\}_{pk(A)} \\
 A &\rightarrow B : \{NB\}_{pk(B)}
 \end{aligned}$$



Attack (Lowe 1996):

$$1. \quad a \rightarrow i : \{na, a\}_{pk(i)}$$

$$1.' \quad i(a) \rightarrow b : \{na, a\}_{pk(b)}$$

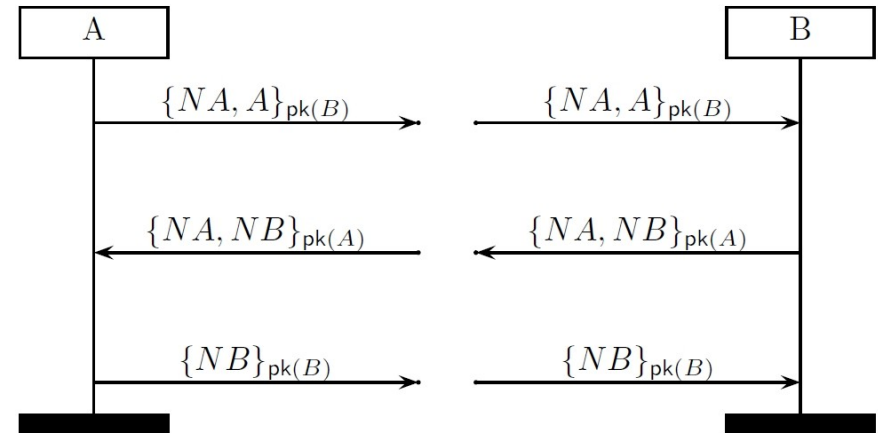
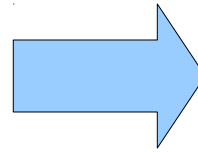
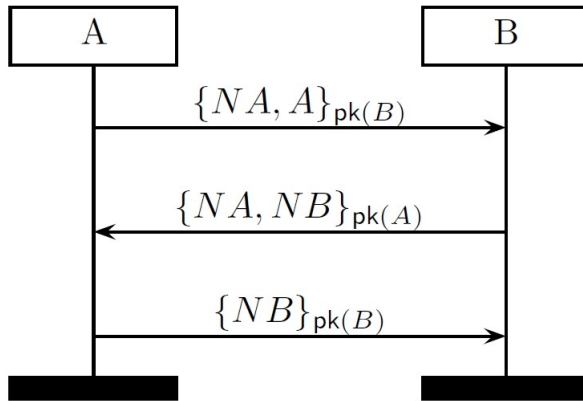
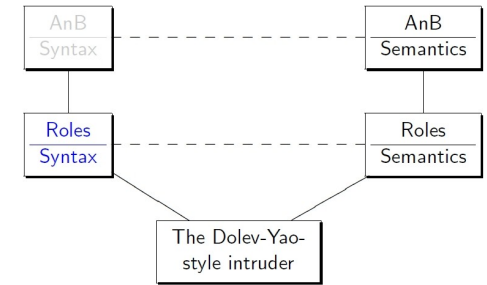
$$2.' \quad b \rightarrow i(a) : \{na, nb\}_{pk(a)}$$

$$2. \quad i \rightarrow a : \{na, nb\}_{pk(a)}$$

$$3. \quad a \rightarrow i : \{nb\}_{pk(i)}$$

$$3.' \quad i(a) \rightarrow b : \{nb\}_{pk(b)}$$

# Role scripts for A and B



## Textual:

$$\text{NSPK}(A) := \text{snd}(\{NA, A\}_{pk(B)}) \cdot \text{rcv}(\{NA, NB\}_{pk(A)}) \cdot \text{snd}(\{NB\}_{pk(B)})$$

# Overview

1. Introduction
2. Formal languages for the specification of security protocols
3. The Dolev-Yao intruder model
4. Operational semantics of security protocols
5. Protocol goals and verification
6. Decidability of protocol security and deductive methods
7. Current research topics

# Modelling the attacker

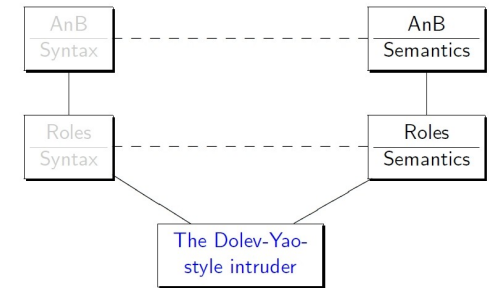
## On the Security of Public Key Protocols (IEEE Trans. Inf. Th. 1983):

- Danny Dolev
- Andrew C. Yao

### The Dolev-Yao Intruder:

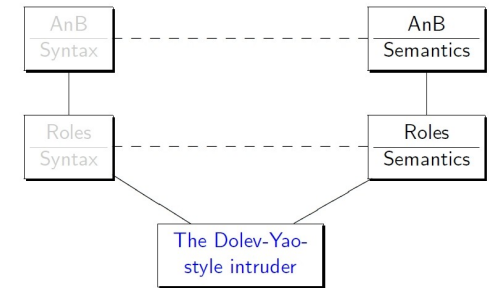
- Controls the network (read, intercept, send)
- Is a legitimate user
- Can apply every publicly available information or function
- Can apply his private information and functions
- Cannot break cryptography

### Following, a semi-formal overview!





# Modelling the attacker



## Definition

Given a set of terms  $M$  we define  $\mathcal{DY}(M)$  as the **least closure of  $M$**  under the following rules:

$$\frac{}{m \in \mathcal{DY}(M)} \text{Axiom } (m \in M) \quad \frac{s \in \mathcal{DY}(M)}{t \in \mathcal{DY}(M)} \text{Algebra } (s \approx t)$$

$$\frac{t_1 \in \mathcal{DY}(M) \quad \dots \quad t_n \in \mathcal{DY}(M)}{f(t_1, \dots, t_n) \in \mathcal{DY}(M)} \text{Composition } (f \in \Sigma_p)$$

$$\frac{\langle m_1, m_2 \rangle \in \mathcal{DY}(M)}{m_i \in \mathcal{DY}(M)} \text{Proj}_i \quad \frac{\{m\}_k \in \mathcal{DY}(M) \quad k \in \mathcal{DY}(M)}{m \in \mathcal{DY}(M)} \text{DecSym}$$

$$\frac{\{m\}_k \in \mathcal{DY}(M) \quad \text{inv}(k) \in \mathcal{DY}(M)}{m \in \mathcal{DY}(M)} \text{DecAsym} \quad \frac{\{m\}_{\text{inv}(k)} \in \mathcal{DY}(M)}{m \in \mathcal{DY}(M)} \text{OpenSig}$$

# A simple example

## Example

$$M = \{ x, \{ \{ b, \text{exp}(g, y) \} \}_k, k, m \}$$

$$\{ \{ m \} \}_{\text{exp}(\text{exp}(g, x), y)} \stackrel{?}{\in} \mathcal{DY}(M)$$

$$\frac{\{ \{ b, \text{exp}(g, y) \} \}_k \in \mathcal{DY}(M) \quad k \in \mathcal{DY}(M)}{\langle b, \text{exp}(g, y) \rangle \in \mathcal{DY}(M)}$$

$$\frac{\langle b, \text{exp}(g, y) \rangle \in \mathcal{DY}(M)}{\text{exp}(g, y) \in \mathcal{DY}(M)}$$

$$\text{exp}(g, y) \in \mathcal{DY}(M)$$

$$\frac{x \in \mathcal{DY}(M)}{\text{exp}(\text{exp}(g, y), x) \in \mathcal{DY}(M)}$$

$$\text{exp}(\text{exp}(g, y), x) \in \mathcal{DY}(M)$$

$$\text{exp}(\text{exp}(g, x), y) \in \mathcal{DY}(M)$$

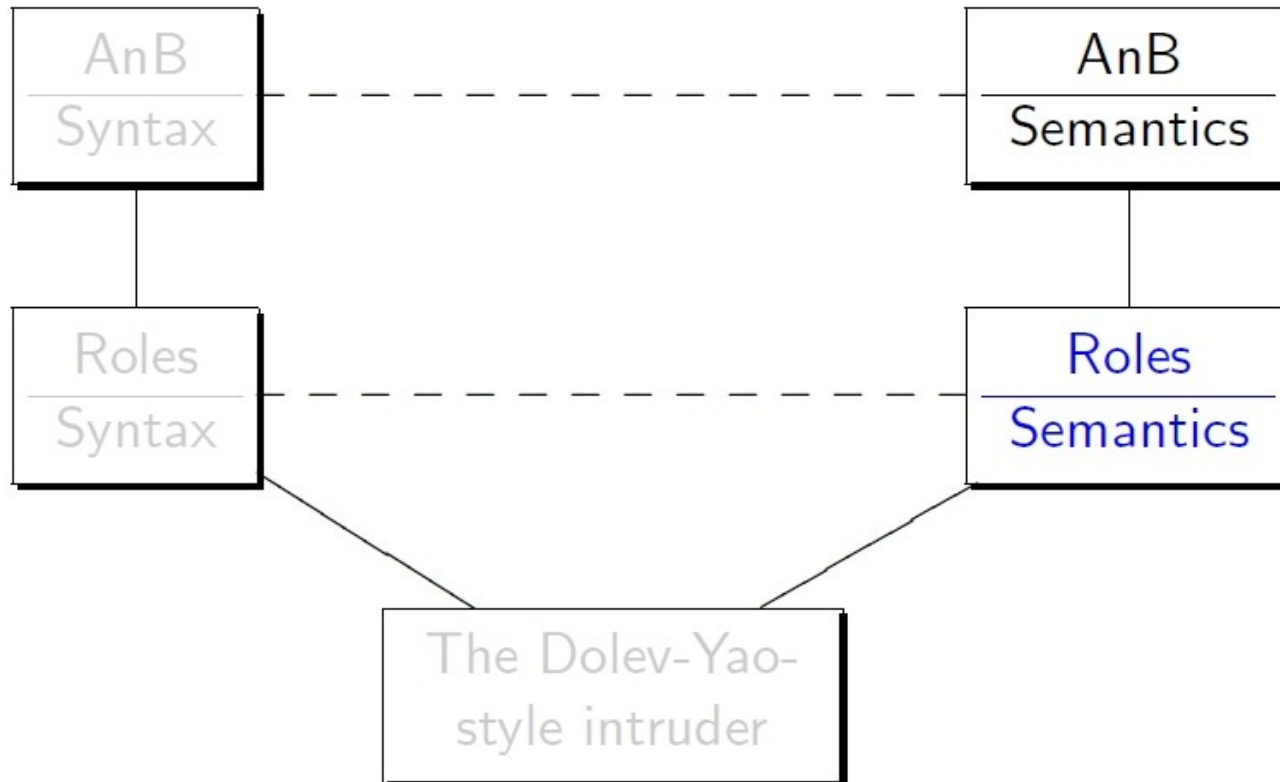
$$\frac{m \in \mathcal{DY}(M)}{\{ \{ m \} \}_{\text{exp}(\text{exp}(g, x), y)} \in \mathcal{DY}(M)}$$

$$\{ \{ m \} \}_{\text{exp}(\text{exp}(g, x), y)} \in \mathcal{DY}(M)$$

# Overview

1. Introduction
2. Formal languages for the specification of security protocols
3. The Dolev-Yao intruder model
4. Operational semantics of security protocols
5. Protocol goals and verification
6. Decidability of protocol security and deductive methods
7. Current research topics

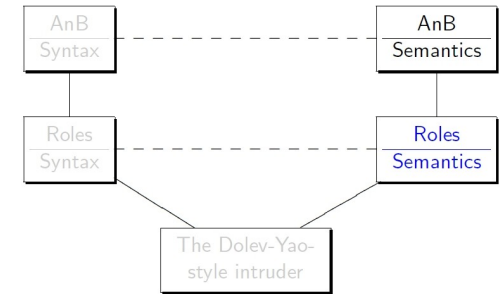
# Formal semantics of the languages



# Operational semantics

- Defined by a transition system

$$TS(P, IK_0, th_0) = (State, \rightarrow, ([], IK_0, th_0))$$



## Definition (State)

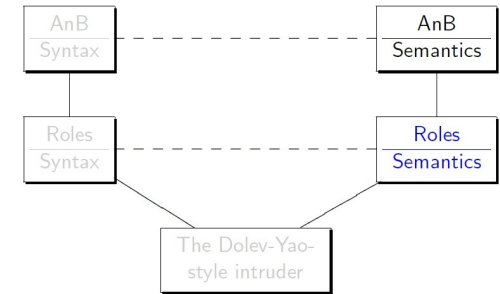
- $State = Trace \times IntruderKnowledge \times Threads.$
- $Trace = (TID \times Event)^*$
- $IntruderKnowledge = \mathcal{P}(Term)$
- $Threads = TID \multimap Role$

where the trace and the intruder knowledge are ground and the threads are closed.

# Operational semantics

$$TS(P, IK_0, th_0) = (State, \rightarrow, ([], IK_0, th_0))$$

- We start from an initial state
- Roles are instantiated (free variables set)



## Example (An initial state that is sufficient to find the attack against NSPK)

$$tr_0 = []$$

$$IK_0 = \{ a, b, i, pk(a), pk(b), pk(i), inv(pk(i)) \}$$

$$th_0(0) = NSPK(A)[A \mapsto a, B \mapsto i, NA \mapsto na_0]$$

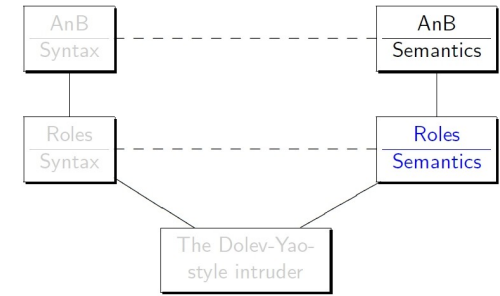
$$th_0(1) = NSPK(B)[B \mapsto B, NB \mapsto nb_1]$$

$$NSPK(A) := snd(\{NA, A\}_{pk(B)}) \cdot rcv(\{NA, NB\}_{pk(A)}) \cdot snd(\{NB\}_{pk(B)})$$

# Operational semantics

$$TS(P, IK_0, th_0) = (State, \rightarrow, ([], IK_0, th_0))$$

- Transition relation defined by a set of deduction rules
- Signals sig will be explained later



## Rules

$$\frac{th(tid) = \text{snd}(t) \cdot tl}{(tr, IK, th) \rightarrow (tr \cdot (tid, \text{snd}(t)), IK \cup \{t\}, th[tid \mapsto tl])} \text{snd}$$

$$\frac{th(tid) = \text{rcv}(t) \cdot tl \quad \text{dom}(\sigma) = \text{var}(t) \quad t\sigma \in \mathcal{DY}(IK)}{(tr, IK, th) \rightarrow (tr \cdot (tid, \text{rcv}(t\sigma)), IK, th[tid \mapsto tl\sigma])} \text{rcv}$$

$$\frac{th(tid) = \text{sig}(\text{sig}, t) \cdot tl}{(tr, IK, th) \rightarrow (tr \cdot (tid, \text{sig}(\text{sig}, t)), IK, th[tid \mapsto tl])} \text{sig}$$

# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$	$\text{snd}(\{na_0, i\}_{pk(i)})$	$\text{rcv}(\{NA, A\}_{pk(b)})$
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$	$\text{rcv}(\{na_0, NB\}_{pk(a)})$	$\text{snd}(\{NA, nb_1\}_{pk(A)})$
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$	$\text{snd}(\{NB\}_{pk(i)})$	$\text{rcv}(\{nb_1\}_{pk(b)})$
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		



# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$	$\text{snd}(\{na_0, i\}_{pk(i)})$	$\text{rcv}(\{NA, A\}_{pk(b)})$
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$	$\text{rcv}(\{na_0, NB\}_{pk(a)})$	$\text{snd}(\{NA, nb_1\}_{pk(A)})$
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$	$\text{snd}(\{NB\}_{pk(i)})$	$\text{rcv}(\{nb_1\}_{pk(b)})$
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$		$\text{rcv}(\{NA, A\}_{pk(b)})$
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$	$\text{rcv}(\{na_0, NB\}_{pk(a)})$	$\text{snd}(\{NA, nb_1\}_{pk(A)})$
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$	$\text{snd}(\{NB\}_{pk(i)})$	$\text{rcv}(\{nb_1\}_{pk(b)})$
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$		
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$	$\text{rcv}(\{na_0, NB\}_{pk(a)})$	$\text{snd}(\{na_0, nb_1\}_{pk(a)})$
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$	$\text{snd}(\{NB\}_{pk(i)})$	$\text{rcv}(\{nb_1\}_{pk(b)})$
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$		
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$	$\text{rcv}(\{na_0, NB\}_{pk(a)})$	
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$	$\text{snd}(\{NB\}_{pk(i)})$	$\text{rcv}(\{nb_1\}_{pk(b)})$
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$		
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$		
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$	$\text{snd}(\{nb_1\}_{pk(i)})$	$\text{rcv}(\{nb_1\}_{pk(b)})$
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$		
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$		
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$		$\text{rcv}(\{nb_1\}_{pk(b)})$
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

# Attack on NSPK

## Example (NSPK Attack)

<i>Trace</i>	<i>th(0)</i>	<i>th(1)</i>
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$		
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$		
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

**Attack trace!**

# Overview

1. Introduction
2. Formal languages for the specification of security protocols
3. The Dolev-Yao intruder model
4. Operational semantics of security protocols
5. Protocol goals and verification
6. Decidability of protocol security and deductive methods
7. Current research topics



# Protocol goals

**Goals**, the protocol should achieve:

- **Authenticate** messages, binding them to their originator
- Guarantee **secrecy** of certain items (e.g. keys)
- Sender invariance
- Anonymity
- Non-repudiation
- ...

**Most common goals:**

- Secrecy
- Authentication (many different forms)

# Protocol properties

## Properties:

- Semantics of a security protocol  $P$  is a set of traces  $\|P\| = \text{traces}(P)$
- Security goal / property  $\phi$  also denotes a set of traces  $\|\phi\|$

## Correctness:

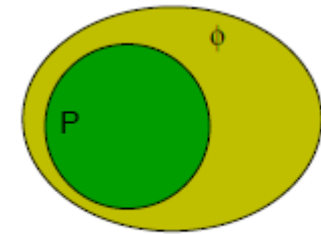
- Protocol  $P$  satisfies property  $\phi$ , written  $P \models \phi$ , iff

$$\|P\| \subseteq \|\phi\|$$

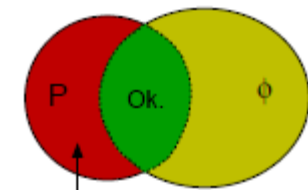
- Attack traces are those in

$$\|P\| - \|\phi\|$$

- Every correctness statement is either true or false



Ok, no attacks.



Attacks.

# Formalizing security properties

## Direct formulation:

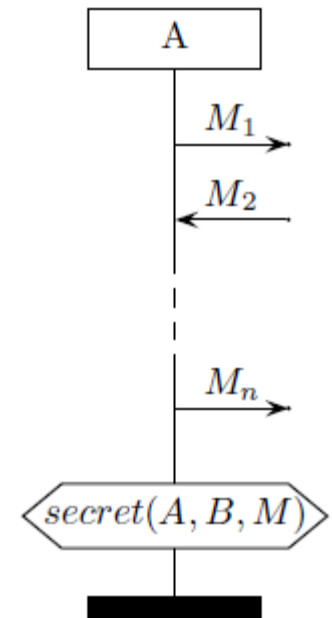
- Formulate property  $\phi$  directly in terms of send and receive events occurring in protocol traces, i.e., as a set of (or predicate on) traces
- Drawback: Standard properties like secrecy and authentication become highly protocol-dependent, since they need to refer to the concrete protocol messages.

## Protocol instrumentation

- Insert special signal events into the protocol roles
- Possible to express properties independently of protocol
- Example:

$\text{sig}(\text{secret}, A, B, M)$

claims that  $M$  is a secret shared by roles  $A$  and  $B$



# Signal events

## Remember:

### Signal rule

$$\frac{th(tid) = sig(sig, t) \cdot tl}{(tr, IK, th) \rightarrow (tr \cdot (tid, sig(sig, t)), IK, th[tid \mapsto tl])} sig$$

## Properties of signal events:

- Used to record facts of claims in the protocol trace
- Since they are artificially inserted into the protocol, the intruder cannot observe or modify or generate them
- Properties formulated from the point of view of a given role, thus yielding security guarantees for that specific role

# Formalizing secrecy

## Example (NSPK Attack)

Trace	$th(0)$	$th(1)$
$(0, \text{snd}(\{na_0, i\}_{pk(i)}))$		
$(1, \text{rcv}(\{na_0, a\}_{pk(b)}))$		
$(1, \text{snd}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{rcv}(\{na_0, nb_1\}_{pk(a)}))$		
$(0, \text{snd}(\{nb_1\}_{pk(i)}))$		
$(1, \text{rcv}(\{nb_1\}_{pk(b)}))$		

## Definition (Secrecy)

The property  $\text{Secret}(A, B, M)$  consists of all traces  $tr$  satisfying

$$\forall tid. (tid, \text{sig}(\text{secret}, A, B, M)) \in \text{set}(tr) \wedge B \neq i \Rightarrow M \notin \mathcal{DY}(IK(tr))$$

$$IK(tr) = \{m \mid \exists tid. (tid, \text{snd}(m)) \in \text{set}(tr)\}$$

# Formalizing authentication

## Two new signals:

- *running*
- *commit*

## Different definitions:

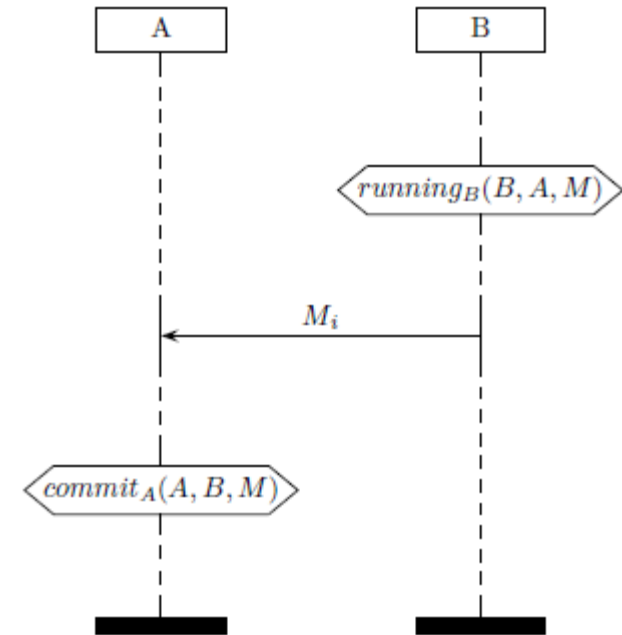
- Aliveness
- Weak agreement
- Non-injective agreement
- Injective agreement
- ...

## Example:

### Definition (Non-injective agreement)

We define  $tr \in \text{Agreement}_{NI}(A, B, M)$  for a trace  $tr$  by

$$\forall tid. (tid, \text{sig}(\text{commit}_A, A, B, M)) \in \text{set}(tr) \wedge B \neq i \\ \Rightarrow \exists tid'. (tid', \text{sig}(\text{running}_B, B, A, M)) \in \text{set}(tr)$$



# Formalizing authentication

## Two new signals:

- *running*
- *commit*

## Different definitions:

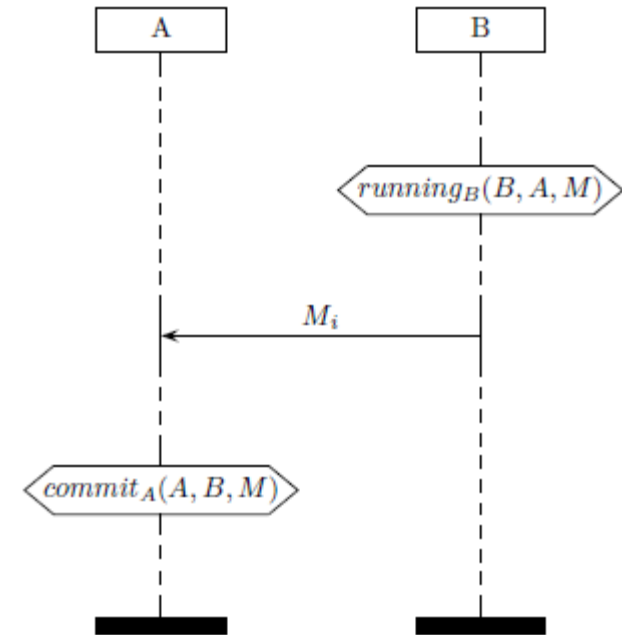
- Aliveness
- Weak agreement
- Non-injective agreement
- **Injective agreement**
- ...

## Example:

### Definition (Injective agreement)

We define  $tr \in \text{Agreement}(A, B, M)$  for a trace  $tr$  iff there is an **injective function**  $g : TID \rightarrow TID$  such that

$$\forall tid. (tid, \text{sig}(\text{commit}_A, A, B, M)) \in \text{set}(tr) \wedge B \neq i \\ \Rightarrow (g(tid), \text{sig}(\text{running}_B, B, A, M)) \in \text{set}(tr)$$



# Overview

1. Introduction
2. Formal languages for the specification of security protocols
3. The Dolev-Yao intruder model
4. Operational semantics of security protocols
5. Protocol goals and verification
6. Decidability of protocol security and deductive methods
7. Current research topics



# Decidability of protocol security

## Algorithmic analysis:

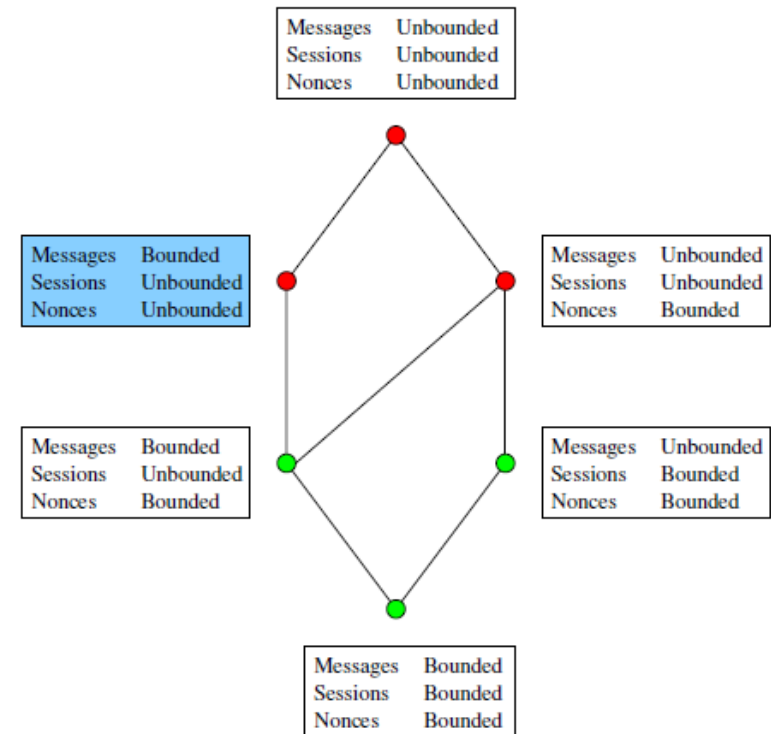
- Fully automatic analysis
- Correctness in general undecidable

## Sources of infinity:

- Messages
- Sessions
- Nonces

## Solutions:

- Various kinds of abstractions  
(Not covered here!)



# Deductive methods (Maybe next talk?)

## Generality:

- Deductive methods can handle all infinite state spaces
- No need for finiteness bounds (e.g. on messages, nonces, sessions)
- Properties are defined over reachable states and proven by induction

## Expressiveness:

- Flexible platform for experimentation
- Possibility to prove meta-results about a model

## Insights:

- Modeling and proving process yields insights into the problem
- Insights may lead to simplifications of model and/or properties
- Simplifications often foster an increased proof automation

## Drawback:

- Loss of automation, proofs generally require user interaction and profound knowledge of both, the used tools and protocol itself

# Overview

1. Introduction
2. Formal languages for the specification of security protocols
3. The Dolev-Yao intruder model
4. Operational semantics of security protocols
5. Protocol goals and verification
6. Decidability of protocol security and deductive methods
7. Current research topics

# Current research topics

## **Modeling more complex protocols:**

- Modeling complex protocols is non-trivial
- Some work in progress at ETH (e.g. KERBEROS)
- Electronic voting protocols, an interesting application?

## **Formalizing electronic voting specific properties and goals:**

- Receipt-freeness?
- Coercion-resistance?
- ...?

## **Open Issues:**

- Secure Platform Problem (some work in progress at ETH)
- Adaptive Corruption (some work in progress at ETH)
- Side-channel attacks

# Questions

