Deterministic Discrete Modeling

Formal Semantics of Firewalls in Isabelle/HOL

Cornelius Diekmann, M.Sc.

Dr. Heiko Niedermayer
Prof. Dr.-Ing. Georg Carle

Lehrstuhl für Netzarchitekturen und Netzdienste
Institut für Informatik
Technische Universität München

Version: April 29, 2014
Agenda

1. Introduction to Isabelle/HOL
   - PL, FOL, HOL
   - Isabelle/HOL
2. Types & Functional Programming
3. Modeling Firewalls
   - Introduction
   - Syntax
   - Semantics
4. Induction
5. Analyzing the Firewall Model
6. Analyzing Rule Sets
7. Rule Lists
8. Conclusion

Slides: 63
About Isabelle/HOL

▶ Generic proof assistant

▶ For
  ▶ interactive theorem proving
  ▶ in higher-order logics (HOL)

http://isabelle.in.tum.de/

$ wget http://isabelle.in.tum.de/dist/Isabelle2013-2_linux.tar.gz
$ sha1sum Isabelle2013-2_linux.tar.gz
   cb8dca7fdcd909b7640121ca6b575a714a62a492 Isabelle2013-2_linux.tar.gz
$ tar -xzf Isabelle2013-2_linux.tar.gz
$ ./Isabelle2013-2/bin/isabelle jedit &
About Isabelle/HOL

```isabelle
theory firewall
import Main
begin

section{*Modeling Firewalls*}

subsection{*Syntax*}

(*The action a firewall can do to a packet*)

datatype action = []

(*We say a packet is of arbitrary type 'p'*)

(*We define the type ‘a rule as synonym for a total function from packets to actions*)
```

Proofs for inductive predicate(s) "action_rep_set"
- Proving monotonicity ...
- Proving the introduction rules ...
- Proving the induction rule ...
Proofs for inductive predicate(s) "action_rec_set"
- Proving monotonicity ...
- Proving the introduction rules ...
- Proving the elimination rules ...
- Proving the simplification rules ...

Note

Note: The theory files are relevant for the exam! We will soon leave the slide set and interactively work with Isabelle/HOL.

About this Chapter

▶ Please make sure you have a laptop and Isabelle/HOL with you (alternatively, find a fellow student who does)
▶ There is a firewall2.thy on the website.
  
  This is the accompanying exercise.
  
  It contains several TODOs. Try to solve them after each lecture.
Propositional Logic, First Order Logic

- Propositional Logic
  
  \((a \rightarrow b) \lor (b \rightarrow a)\)

- First Order Logic (FOL)
  
  \(\text{FOL} = \text{Propositional Logic} + \text{Quantifiers}\)

- Example FOL
  
  \(\forall a. \exists b. (a \rightarrow b) \lor (b \rightarrow a)\)

  \((a \rightarrow b) \lor (b \rightarrow a)\)
HOL – Higher-Order Logic

- Higher-Order Logic (HOL)
  
  \[ \text{HOL} = \text{Functional Programming} + \text{Logic} \]

- Example FOL
  
  \[(a \rightarrow b) \lor (b \rightarrow a)\]

- Example Lambda Calculus
  
  \[(\lambda x. x + 4) \ 38\]

- Example Transitive Closure of a Relation \(R\)
  
  \[(a, b) \in R \land (b, c) \in R \rightarrow (a, c) \in R^+\]
Isabelle/HOL

- Isabelle/HOL: A Proof Assistant for Higher-Order Logic
  We model and reason in HOL

- Example FOL
  \[(a \rightarrow b) \lor (b \rightarrow a)\]

- In Isabelle/HOL
  lemma "\[(a \rightarrow b) \lor (b \rightarrow a)\]" by auto

- Lambda Calculus
  \[(\lambda x. x + 4) 38\]

- In Isabelle/HOL
  lemma "\[(\lambda x. x + 4) 38 = 42\]" by simp
Isabelle/HOL Examples

▶ Example

```
lemma "(a → b) ∨ (b → a)" by auto
```

▶ Terms are written in ""

▶ lemma starts a lemma

▶ it can be proven automatically by auto

▶ other automated proof methods

▶ simp – the simplifier
▶ auto – simplification, logic, sets, may return unprovable goals
▶ blast – complete for FOL
▶ ...

```
lemma "(a → b) ∨ (b → a)" by simp
lemma "(a → b) ∨ (b → a)" by blast
lemma "(a → b) ∨ (b → a)" by fastforce
```
Isabelle/HOL Examples

- About the transitive closure
  \[
  \text{lemma } \left( (a, b) \in R \land (b, c) \in R \right) \implies (a, c) \in R^+ \text{ by auto}
  \]

- About the reflexive transitive closure
  \[
  \text{lemma } \left( (a, a) \in R^* \right) \text{ by simp}
  \]

- FOL is modeled in HOL
  - HOL implication: \( \implies \)
  - FOL implication: \( \rightarrow \)
  \[
  \text{lemma } \left( (a, b) \in R \land (b, c) \in R \rightarrow (a, c) \in R^+ \right) \text{ by auto}
  \]
Implications associate to the right

\[ A \implies B \implies C \] means \[ A \implies (B \implies C) \]

This is equal to

\[ A \land B \implies C \]

or

lemma assumes \( A \) and \( B \) shows \( C \)

Homework: prove

\[(a \rightarrow b \rightarrow c) \iff (a \land b \rightarrow c)\]

on paper and in Isabelle/HOL
Isabelle/HOL Notation

- FOL all quantifier: ∀
- HOL all quantifier: ⋀
- Example
  - lemma "∀x. x"
    - note: obviously wrong, counterexample x = False
  - apply(rule)
    - produces subgoal ⋀x. x
- From FOL to HOL
  - “free” variable (can be instantiated): ?x
- Example
  - lemma x: ‘a = a’ by simp
  - thm x gives you ?a = ?a
  - That is, you can instantiate lemma x for arbitrary ?a
  - Example: thm a[of "foo"] is foo = foo
- Concluding Example: Inspect allI
Types & Functional Programming
Does the set of all sets contain itself?
Does the set of all sets contain itself?

- If not: It is obviously not the set of all sets! ⚡
Does the set of all sets contain itself?

- If not: It is obviously not the set of all sets! ⚡

- If it does: Let $A$ be the subset that contains all sets that do not contain themselves.
  - Is $A \in A$?
  - If not: $A$ is obviously not the set of all sets that do not contain themselves! ⚡
  - If it does: $A$ contains a set that contains itself! ⚡
Types

- Why is HOL typed?

- An example of untyped mathematics:

  Does the set of all sets contain itself?

  - If not: It is obviously not the set of all sets!

  - If it does: Let $A$ be the subset that contains all sets that do not contain themselves.
    - Is $A \in A$?
      - If not: $A$ is obviously not the set of all sets that do not contain themselves!
      - If it does: $A$ contains a set that contains itself!
Types

- Typed sets
  - nat set e.g., $1 \in \{1, 2\}$
  - nat set set e.g., $\{1, 2\} \in \{\{1, 2\}, \{\}\}$
  - nat set set set e.g., $\{\{1, 2\}, \{\}\} \in \{\{\{1, 2\}, \{\}\}, \{1, 2, 3, 4\}\}$
  - $\in$ is of type $\text{nat} \Rightarrow \text{nat set} \Rightarrow \text{bool}$
    
    \[1 \in \{1, 2\} \text{ is true}\]

- The concept of set of all sets is not well-typed

- $\mathcal{M} = \text{the set of all sets}$

- What is the type of $\mathcal{M}$?
  - There is no (finite) type for $\mathcal{M}$!

- Even if there were a type for $\mathcal{M}$, $\mathcal{M} \in \mathcal{M}$ is an obvious type error

- HOL is typed!
Polymorphic Types

- Arbitrary types 'a, 'b, ....
- ∈ is of type 'a ⇒ 'a set ⇒ bool

- For example
  - False ∈ {True, False}, bool set
  - (a → b) ∈ {(a → b), x}, bool set
  - “Hello” ∈ {“Hello” , “World” }, string set
  - apple ∈ {apple, banana}, 'a set, apple and banana are of arbitrary type
  - value “{True, False} :: bool set”
**Functional Programming**

**Demo: Introduction.thy**

```haskell
fun find_fives :: "nat list ⇒ nat list" where
  "find_fives [] = []"
| "find_fives (x#xs) =
    (if x = 5 then 5#find_fives xs else find_fives xs)"

value "find_fives [1,3,5,8,5,2,5]"

lemma "find_fives [1,3,5,8,5,2,5] = [5,5,5]" by simp

Library function filter
filter :: "('a ⇒ bool) ⇒ 'a list ⇒ 'a list"
lemma "find_fives l = filter (λx. x = 5) l"
  by(induction l, simp_all)
```
Total Functions

- A *total* function is a function that is defined for all input values.
  - For computer scientists:
    A total function must always terminate!
- Written ‘a ⇒ ‘b
- HOL is a total logic!
- Why total functions?
- Assume `f` of type `nat ⇒ nat`.
  
  ```
  f n = while(true){{} return 0
  ```

  1. `f n` does not terminate
  2. `does not terminate = does not terminate + 1`
  3. From (1) and (2): `f n = (f n) + 1`
  4. Subtracting `f n` on both sides: `0 = 0 + 1`
  5. `0 = 1`

- HOL is total and every function needs a proof for that!
Total Functions

Usually, the *total*-proof is for free.

```ocaml
fun find_fives :: "nat list ⇒ nat list" where
  "find_fives [] = []"
| "find_fives (x#xs) = (if x = 5 then 5 # find_fives xs else find_fives xs)"

lemma "find_fives l = filter (λx. x = 5) l"
by (induction l, simp_all)
```

Here, the size of the list is used for the automatic proof.

1. The function obviously terminates for the empty list
2. If the list consists of \(x :: xs\), all subsequent calls to `find_fives` only get `xs`.
Wrap-Up

- HOL = Functional Programming + Logic
- Strong Typing
- Total functions
- Proofs!
Modeling Firewalls
Firewalls – Introduction

- Access control on the network level
- Often positioned between a trusted and a less trusted network
  - E.g. between internal network and Internet

Example:
Firewalls

- Decides per packet what to do
  - Allow – Forward the packet
  - Deny – Drop the packet
- A firewall is configured by a ruleset. Example:

<table>
<thead>
<tr>
<th>#</th>
<th>Src IP</th>
<th>Dst IP</th>
<th>Proto</th>
<th>Src Port</th>
<th>Dst Port</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Alice</td>
<td>Gabby</td>
<td>TCP</td>
<td>*</td>
<td>80</td>
<td>Allow</td>
</tr>
<tr>
<td>B</td>
<td>Alice</td>
<td>8.8.8.8</td>
<td>UDP</td>
<td>*</td>
<td>53</td>
<td>Allow</td>
</tr>
<tr>
<td>C</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>Deny</td>
</tr>
</tbody>
</table>

Textual explanation on next slide (walls of text)
Firewalls

- Rules are processed consecutively for every packet
  Assume a packet arrives

1. Is the source Alice and the destination Gabby and the protocol TCP and the destination port 80 (HTTP)?
   If yes, the firewall allows the packet.
   If not, the next rule is processed.

2. Is the source Alice and the destination 8.8.8.8 (google DNS server) and the protocol UDP and the destination port 53 (DNS)?
   If yes, the firewall allows the packet.
   If not, the next rule is processed.

3. The firewall discards the packet.
   This default strategy is called blacklisting. Everything which is not explicitly allowed is denied.

<table>
<thead>
<tr>
<th>#</th>
<th>Src IP</th>
<th>Dst IP</th>
<th>Proto</th>
<th>Src Port</th>
<th>Dst Port</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Alice</td>
<td>Gabby</td>
<td>TCP</td>
<td>*</td>
<td>80</td>
<td>Allow</td>
</tr>
<tr>
<td>B</td>
<td>Alice</td>
<td>8.8.8.8</td>
<td>UDP</td>
<td>*</td>
<td>53</td>
<td>Allow</td>
</tr>
<tr>
<td>C</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>Deny</td>
</tr>
</tbody>
</table>
Firewalls

- For simplicity, our example ignores
  - Stateful firewalls
  - Answer packets.
  Alice can send out HTTP and DNS request, but the firewall will never allow the answer to reach Alice.

- We only focus on the packet header.

<table>
<thead>
<tr>
<th>#</th>
<th>Src IP</th>
<th>Dst IP</th>
<th>Proto</th>
<th>Src Port</th>
<th>Dst Port</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Alice</td>
<td>Gabby</td>
<td>TCP</td>
<td>*</td>
<td>80</td>
<td>Allow</td>
</tr>
<tr>
<td>B</td>
<td>Alice</td>
<td>8.8.8.8</td>
<td>UDP</td>
<td>*</td>
<td>53</td>
<td>Allow</td>
</tr>
<tr>
<td>C</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>Deny</td>
</tr>
</tbody>
</table>
Modeling Firewalls – Representing Rules

Demo: firewall.thy

Step 1: Rules Basics

- Per Rule Action
  - Allow
  - Deny
  - Pass — Rule does not apply, apply next rule

```
datatype action = Allow | Deny | Pass
```

- The Type of a Rule
  - A rule processes an packet of arbitrary type ’p
  - A rule defines an action for that packet
  - A rule is a total function from packets to actions

```
type_synonym ’p rule = ’p ⇒ action
```
Representing Rules – Example 1

- Let a packet consist of a tuple of arbitrary type \( a \times b \)
- Let \( a \) be the source and \( b \) be the destination address
- We can write a deny-all rule as
  \[
  \lambda (src,dst). \quad \text{Deny}
  \]
- The type is \((a \times b)\) rule
Representing Rules – Example 2

- Let a packet consist of a tuple strings \( \text{string} \times \text{string} \)
- The first entry is the source and the second the destination address
- We can write a rule which allows packets from Alice to Bob and denies all others

\[
\lambda (\text{src}, \text{dst}). \begin{cases} 
\text{Allow} & \text{if } \text{src} = \text{’’Alice’’} \land \text{dst} = \text{’’Bob’’} \\
\text{Deny} & \text{else}
\end{cases}
\]

- The type is \((\text{string} \times \text{string})\) rule
- For a packet \(\text{’’Alice’’, ’’Carl’’}\)
- We can show:

\[
(\lambda (\text{src},\text{dst}). \begin{cases} 
\text{Allow} & \text{if } \text{src} = \text{’’Alice’’} \land \text{dst} = \text{’’Bob’’} \\
\text{Deny} & \text{else}
\end{cases}) (\text{’’Alice’’, ’’Carl’’}) = \text{Deny}
\]

Applying the packet \(\text{’’Alice’’, ’’Carl’’}\) to the rule results in Deny
Modeling Firewalls – Representing Rules

Step 2: A Firewalls Rule Set

Probably, rule list would be a better terminology, we will soon prove why.

- The firewall has a ‘table’ of rules
- For packets of type ’p, we develop datatype ’p ruleset
Modeling Firewalls – Representing Rules

datatype 'p ruleset
  ▶ Rules apply sequentially
  ▶ If no rule matches, a default rule must apply
  ▶ Default rules
    ▶ DefaultAllow – Blacklisting
      Everything not prohibited is allowed by default
    ▶ DefaultDeny – Whitelisting
      Everything not allowed is denied by default
  ▶ A Rule consists of a rule and its successor rules
    ▶ Rule "'p rule" "'p ruleset"

datatype 'p ruleset = DefaultAllow
  | DefaultDeny
  | Rule "'p rule" "'p ruleset"
Rulesets – Example 1

- Let a packet consist of arbitrary type 'a
- We can write a rule set with a Allow-All rule, a Deny-All rule, and a default allow strategy
  
  Rule (\( \lambda p. \) Allow) (Rule (\( \lambda p. \) Deny) DefaultAllow)

- The type is 'p ruleset
Rule **Notation**

- Instead of Rule, we will use the infix operator $\&$

- Example:
  - Rule $x \, y$ is written as $x \, \& \, y$

Previous slide’s example

Rule $\lambda p. \text{Allow} \land (\lambda p. \text{Deny}) \land \text{DefaultAllow}$

Is written as

$(\lambda p. \text{Allow}) \, \& \, (\lambda p. \text{Deny}) \, \& \, \text{DefaultAllow}$
Conclusion

▶ We defined the syntax of
  ▶ action
  ▶ ’p rule
  ▶ ’p ruleset

▶ Next, we will develop an example of a concrete ruleset for a concrete ’p

▶ Afterwards, the semantics of firewalls are specified.
  ▶ Syntax: How to write things down
  ▶ Semantic: Meaning
Modeling Firewalls – Syntax Example

Demo: firewall.thy
Semantics

We define a function `firewall` to capture the semantics of a firewall.

Step 1) The type

- A firewall operates according to a `ruleset`
- A firewall processes packets accordingly
- The firewall decides what to do with the packet
  It returns an action

Hence, the type is `’p ruleset ⇒ ’p ⇒ action`
Semantics of `firewall`

- The semantics are specified via pattern matching on the ruleset
  - an underscore `_' matches everything
- If the ruleset is only `DefaultAllow`, the firewall allows any packet
  
  "`firewall DefaultAllow _ = Allow`"
- If the ruleset is only `DefaultDeny`, the firewall denies any packet
  
  "`firewall DefaultDeny _ = Deny`"
- Most of the time, the ruleset will consists of a current rule and
  and some `next_rules`

  `rule § next_rules`

- `next_rules` can be any `p` ruleset, either further rules or one of
  the default rules
Semantics of firewall

\[ \text{firewall (rule} \ \text{next\_rules)} \ p \]

- For a packet \( p \)
- The firewall examines \text{rule} and tests whether the rule matches \( p \)
- If the action is \text{Allow} or \text{Deny}, it applies the respective action
- If the action is \text{Pass}, the \text{next\_rules} must be applied to \( p \)

This is a recursive call to firewall
Example

Demo: firewall.thy
Termination of `firewall`

- The default cases trivially terminate
- For the case `rule § next_rules`, the size of the ruleset is decreased with every recursive call. If `firewall (rule § next_rules) p` is called
  - it either terminates with `Allow` or `Deny`
  - or calls `firewall next_rules p`
- This is an inductive argument.
Add-In: Induction
Induction

- Induction on natural numbers \( \mathbb{N} \)
- To prove \( P \ n \) for \( n \in \mathbb{N} \)
- Show that \( P \) holds for 0
- Assume \( P \) holds for \( n \) and show that it also holds for \( n + 1 \)

Induction rule in Isabelle/HOL (simplified)

\[
P 0 \implies (\forall n. P \ n \implies P \ (n + 1)) \implies P \ x
\]

Read as follows

- To show \( P \ x \) (the goal right)
- Show \( P \ 0 \)
- and show \( \forall n. P \ n \implies P \ (n + 1) \)

Demo: intro_induction.thy
Induction – Induction Rules

- Where do we get \texttt{nat.induct} from?
- Isabelle/HOL proves a lot automatically for us in the background

- How are natural numbers modeled?
  - \texttt{Zero} is a natural number
  - The successor \texttt{Suc} of a natural number is a natural number

```plaintext
datatype nat = Zero | Suc nat
```

Proofs for inductive predicate(s) "nat_rep_set"
- Proving monotonicity ...
- Proving the introduction rules ...
- Proving the induction rule ...

Proofs for inductive predicate(s) "nat_rec_set"
- Proving monotonicity ...
- Proving the introduction rules ...
- Proving the elimination rules ...
- Proving the simplification rules ...
Induction – Induction Rules

datatype nat = Zero | Suc nat

nat.induct

\[ P \text{ Zero} \implies (\forall n. P n \implies P (\text{Suc } n)) \implies P x \]

- **Base case:** \( P \text{ Zero} \)
- **Induction step:** \( (\forall n. P n \implies P (\text{Suc } n)) \)
Induction – Firewall Rules

- Firewall rules are modeled similarly
  - `DefaultAllow` is a ruleset
  - `DefaultDeny` is a ruleset
  - A Rule consists of a rule and a ruleset

```plaintext
datatype 'p ruleset = DefaultAllow
  | DefaultDeny
  | Rule "'p rule" "'p ruleset"
```

- Two base cases
  - `DefaultAllow`
  - `DefaultDeny`

- One induction step
  - Assume it holds for some `rules :: 'p ruleset`. For any rule `r :: 'p rule`, show that it holds for `Rule r rules`.
ruleset induction rule

\text{thm \ ruleset\_induct}

\begin{align*}
P \ \text{DefaultAllow} & \implies P \ \text{DefaultDeny} \\
(\land r \ \text{rules}. \ P \ \text{rules} & \implies P (r \ \& \ \text{rules})) & \implies P \ \text{ruleset}
\end{align*}

For an arbitrary ruleset, to prove $P$ ruleset

- Prove that $P$ holds for the base cases
  - $P$ DefaultAllow
  - $P$ DefaultDeny

- And the induction step
  - Assume $P$ rules
  - Show $P (r \ \& \ \text{rules})$
  - $(\land r \ \text{rules}. \ P \ \text{rules} \implies P (r \ \& \ \text{rules}))$
Back to firewall
Analyzing firewall

- For a ruleset $r$ and a packet $p$
- firewall $r$ $p$ never returns Pass
- It either returns Allow or Deny

lemma ‘‘firewall $r$ $p$ = Allow $\lor$ firewall $r$ $p$ = Deny’’

proof(induction $r$)

...

Demo: firewall.thy

Yes, the *.thy is important!
Analyzing Rule Sets

▶ What does the policy \texttt{allow\_DNS} \& \texttt{allow\_HTTP} \& \texttt{DefaultDeny} allow?

▶ Only UDP port 53 or TCP port 80

▶ Phrasing as a lemma
  ▶ A firewall with this rule set
  ▶ A packet \( p \) applied to it returns \texttt{Allow} if and only if
  ▶ \( p \) is UDP port 53 or \( p \) is TCP port 80

\[
\text{lemma} \quad \left\{ \text{firewall} \left( \texttt{allow\_DNS} \& \texttt{allow\_HTTP} \& \texttt{DefaultDeny} \right) \mathrel{\text{p = Allow}} \quad \iff \quad \left( \text{proto p = UDP} \wedge \text{dst.port p = 53} \right) \lor \left( \text{proto p = TCP} \wedge \text{dst.port p = 80} \right) \right\}
\]

by(simp add: allow\_DNS_def allow\_HTTP_def)
Why did the previous lemma require a (difficult) induction whereas this lemma could be directly solved by the simplifier?

The previous lemma argued about arbitrary rule sets whereas this lemma had a known rule set.

The firewall is defined in terms of the rule set.

For unknown rule sets we (often) need induction over the rule set.

If the rule set and the packet are given, the proof is often trivial as firewall is executable.
Analyzing Rule Sets

- Rule sets may be inefficient or misleading
- For example, rules can *shadow* each other
  
  \[
  \text{deny}_{\text{UDP}} \not\subseteq \text{allow}_{\text{DNS}} \not\subseteq \text{allow}_{\text{HTTP}} \not\subseteq \text{DefaultDeny}
  \]

- The rule \( \text{allow}_{\text{DNS}} \) is shadowed by \( \text{deny}_{\text{UDP}} \)
- It can never apply
- We can simplify this ruleset and preserve the semantics!
  
  \[
  \text{deny}_{\text{UDP}} \not\subseteq \text{allow}_{\text{HTTP}} \not\subseteq \text{DefaultDeny}
  \]

- We can prove that the two rulesets are semantically equivalent if \( \text{firewall} \) is equal for both of them
- \( \text{lemma 'firewall rules}_1 = \text{firewall rules}_2' \)
When are two functions equal?

- Two (total) functions \( f \) and \( g \) are equal iff they return the same result for every input

\[ (f = g) = (\forall x. f\ x = g\ x) \]

fun_eq_iff

Back to firewalls

- Two firewall rule sets \( \text{rules1} \) and \( \text{rules2} \) are semantically equivalent iff \( \text{firewall} \) is equal for them

- \( \text{firewall} \ \text{rules1} \) is equal to \( \text{firewall} \ \text{rules2} \) iff the two functions are equal for all packets
The previous rule sets are semantically equivalent

\[
\text{deny}_{\text{UDP}} \sqsubseteq \text{allow}_{\text{DNS}} \sqsubseteq \text{allow}_{\text{HTTP}} \sqsubseteq \text{DefaultDeny} \\
\text{deny}_{\text{UDP}} \sqsubseteq \text{allow}_{\text{HTTP}} \sqsubseteq \text{DefaultDeny}
\]

This can be proven by the simplifier once it is rewritten to

\[
\forall p. \text{firewall } ?r1 p = \text{firewall } ?r2 p
\]

This rule set can even be reordered

\[
\text{allow}_{\text{HTTP}} \sqsubseteq \text{deny}_{\text{UDP}} \sqsubseteq \text{DefaultDeny}
\]
Ruleset Reordering

- We say a rule applies if it returns an action which is not $\text{Pass}$.
- Rules $r_1$ and $r_2$ can be reordered if there is no packet such that both rules apply to the packet simultaneously:
  \[ \# p. \ r_1 p \neq \text{Pass} \land r_2 p \neq \text{Pass} \]
- This is equal to the following:
  \[ \forall p. \ r_1 p = \text{Pass} \lor r_2 p = \text{Pass} \]
- This condition is sufficient to allow reordering of the first rules:
  \[ \text{firewall} (r_1 \land r_2 \land r_3) = \text{firewall} (r_2 \land r_1 \land r_3) \]
- The condition is not necessary (try quickcheck).
Rule Sets or Rule Lists?
Rule Sets or Rule Lists

- The *rulesets* we saw bear great resemblance to lists
- $r_1 \ § \ r_2 \ § \ r_3 \ § \ r_4 \ldots$
- The order of the elements is important
- Can we represent firewall rules as list?
Lists

- $[1 :: \text{nat}, 2, 3, 5] :: \text{nat list}$
- recall find_fives
- $\text{rev } [1 :: \text{int}, 2, 3, 4] = [4, 3, 2, 1]$
- lemma ‘‘$\text{rev (rev } l) = l$’’ by simp
- $[a, b, c] :: 'a \text{ list}$
- $[\text{allow_DNS}, \text{allow_HTTP}] :: \text{packet rule list}$
Semantics of a firewall with a rule list

- We define the semantics of a firewall whose rules are defined as list

  \[ \text{Type: } 'p \text{ rule list } \Rightarrow 'p \Rightarrow \text{ action} \]

- The list is processed sequentially – the rule in the list which matches first is applied

- If a rule returns \textit{Pass}, the next list item is observed

- We use Default-Deny semantics – the empty list corresponds to \textit{Deny}

- Note the similarity
  - \textit{ruleset}: \textit{rule} \& \textit{next_rules}
  - \textit{rule list}: \textit{rule} \# \textit{next_rules}

See \texttt{firewall_list_defaultDeny}
ruleset vs. rule list

▶ How does the ruleset and rule list correspond?
▶ we can translate them

'p ruleset ⇒ 'p rule list

▶ With the Default-Deny semantics
  ▶ DefaultDeny corresponds to the empty list []
  ▶ DefaultAllow corresponds to the one-element list with an allow-all rule [(λp. Allow)]
  ▶ the correspondence rule § next_rules and rule # next_rules is defined recursively

See ruleset_to_rulelist
ruleset to rule list – Example

- A ruleset can be translated to a rule list

```
ruleset_to_rulelist (allow_DNS § allow_HTTP § DefaultDeny) = [allow_DNS, allow_HTTP]
```

- Does this preserve the semantics?
- in this example, yes!

```
firewall (allow_DNS § allow_HTTP § DefaultDeny) = firewall_list_defaultDeny [allow_DNS, allow_HTTP]
```

- Does this hold in general?
Relating the semantics

- Translating a ruleset to a rule list, the firewall and the firewall_list_defaultDeny are equal

lemma fw_eq_fwlist: "firewall r = firewall_list_defaultDeny (ruleset_to_rulelist r)"

- The proof (idea)
  - First, use fun_eq_iff
  - This leaves the subgoal that the two definitions are equal for all packets
  - Then, apply induction over the ruleset r

Demo: firewall.thy

Yes, the *.thy is important!
Conclusion
Conclusion

- Basics of modeling in Isabelle/HOL
  - FOL, HOL, $\lambda$-calculus, polymorphic types
  - Types and total functions
  - Induction

- Firewalls with rulesets
  - Syntax
    - actions, rules, ruleset
  - Semantics
    - Proofs about the semantics
    - Termination, never returns Pass, ruleset shadowing, ruleset reordering

- Firewall modeled with lists

- Semantical equality
Thanks for your attention!

Questions?