Section 2.3

Implementing Simple Theorem Provers

Subsection 2.3.1

Introduction

Overview

Theorem Prover

- theorem prover implements language and proof system
- used for proof checking and automated theorem proving

Goals of this subsection

- understanding the concepts and structure of a theorem prover
- illustrating how Isabelle/HOL can be used for system modeling

Subsection 2.3.2

A Very Simple Prover for Propositional Logic

Abstract syntax of propositional logic

```
datatype plformula  =
     Var   string
   | TTrue
   | FFalse
   | Neg  plformula
   | Imp  plformula  plformula      (infixr "~~>" 100)
```


Representation of variable assignments

A variable assignment is a mapping from string to bool.

```
definition mkVAssign :: "string set ⇒ (string ⇒ bool)"
where
  "mkVAssign strs s = (s ∈ strs)"


definition someva :: "(string \<Rightarrow>bool)"  where
   "someva = mkVAssign {''p'', ''q''}"
```


Semantics of propositional logic

```
primrec plfsem :: "(string ⇒ bool) ⇒ plformula ⇒ bool"
where
"plfsem va (Var s)    = va s"                    |
"plfsem va TTrue      = True"                    |
"plfsem va FFalse     = False"                   |
"plfsem va (Neg p)    = (¬ (plfsem va p))"       |
"plfsem va (Imp p q) = ((plfsem va p) ⟶ (plfsem va q))"
```


Sequents and their semantics

```
datatype plsequent =
     Sq "plformula list" plformula        (infix "⊢-" 50)


definition plf1 :: plformula where
"plf1 =  ((Var ''p'') ~~>  TTrue)"


definition pls2 :: plsequent where
"pls2 = ( [Var ''a'',plf1] ⊢- (Var ''p'') )"
```

Proof system

We build a simple proof system for backward proofs in natural deduction style. It consists of

- a proof state

```
type_synonym proofstate =  "plsequent list"
```

- a collection of rule-specific functions to manipulate the proof state according to the rules

Application of implication elimination and introduction

```
fun applImpE :: "proofstate ⇒ plformula ⇒ proofstate"
where
  "applImpE ((Sq asms q) # psl) plf =
       (Sq asms (plf ~~> q)) # (Sq asms plf)  # psl "
| "applImpE psl plf = psl"


fun applImpI :: "proofstate ⇒ proofstate" where
  "applImpI ((Sq asms (p~~>q)) # psl) = ((Sq (p#asms) q)
     # psl)"
| "applImpI psl =  psl"
```

Remark

The arguments of the application functions depend on the rule.

Application of negation elimination and introduction

```
fun applNegE :: "proofstate ⇒ plformula ⇒ proofstate"
where
  "applNegE ((Sq asms FFalse) # psl) plf =
      (Sq asms (Neg plf)) # (Sq asms plf) # psl "
| "applNegE psl plf = psl"


fun applNegI :: "proofstate ⇒ proofstate" where
  "applNegI ((Sq asms (Neg plf)) # psl) =
                          (Sq (plf#asms) FFalse) # psl "
| "applNegI psl =  psl"
```


Application of assumption axiom

```
primrec iselem :: "'a ⇒ 'a list ⇒ bool" where
  "iselem x [] = False"
| "iselem x (y#ys) = (if x=y then True else iselem x ys)"


fun assumpt :: "proofstate ⇒ proofstate" where
  "assumpt ((Sq asms q) # psl) =
     (if iselem q asms then psl else (Sq asms q) # psl )"
| "assumpt psl = psl"
```

Discussion

Properties of implementation

- Pattern matching mechanism of Isabelle/HOL is used to match the rules to the proof goal.
- For additional rules we have to add further functions to the prover.

Properties of logic

- Is the logic sound? I.e., is every derived formula a tautology?
- Is the logic complete? I.e., is every tautology derivable?
- Can these properties be formulated in Isabelle/HOL?

Subsection 2.3.3

Generic Application of Proof Rules

Extensible proof systems

Motivation

User should be able to prove rules and add them to the proof system.

Consequences

- Rules have to be derivable and become syntactical objects.
- We have to distinguish between formulas and formula schemas:
  → introduce schema variables
- Generic *methods* are needed that can apply different rules, in particular rules derived by the user.


```
datatype plformula  =
      Var  string
    | SVar string    (* schema variables *)
    | TTrue
    | FFalse
    | Neg  plformula
    | Imp  plformula  plformula     (infixr "~~>" 100)
```

Example (Formula with schema variables)

```
(SVar ''A'') ~~> (SVar ''B'')
```

Can be used to express rules (see below)

Substitutions and matching

Definition (Substitution)

A *substitution* σ is a partial function from (schema) variables to formulas (or terms).

Definition (Matching)

Let f_1, f_2 be formulas (or terms).
f_1 *matches* f_2 if f_2 can be obtained by substituting the (schema) variables in f_1 by suitable formulas/terms:

$$\text{matches}(f_1, f_2) \;\Leftrightarrow\; \exists\, \sigma.\ (\text{applSubst}\ \sigma\ f_1) \;=\; f_2$$

Unification

Definition (Unifier, unifiability)

A substitution σ is a *unifier* of two formulas/terms if it makes them equal:

$$\text{unifier}(\sigma, f_1, f_2) \;\Leftrightarrow\; (\text{applSubst}\ \sigma\ f_1 \;=\; \text{applSubst}\ \sigma\ f_2)$$

Two formulas/terms are *unifiable* if they have a unifier:

$$\text{unifiable}(f_1, f_2) \;\Leftrightarrow\; \exists\, \sigma.\ \text{unifier}(\sigma, f_1, f_2)$$

Remarks

- Two formulas/terms might have several unifiers. Usually, one is interested in the most general unifier (MGU).
- For many logics, it is a non-trivial task to compute the MGU.

Rules (→I) and (→E) as data

```
datatype plrule = Rule "plsequent list" plsequent

definition impI :: plrule where
"impI =  (Rule
  [ [(SVar ''A'')] ⊢- (SVar ''B'') ]
  ( [] ⊢- ((SVar ''A'') ~~> (SVar ''B'')) )
 )"

definition impE :: plrule where
"impE = (Rule
  [ []⊢-((SVar ''A'')~~>(SVar ''B'')), []⊢-(SVar ''A'') ]
  ( []  ⊢- (SVar ''B'') )
 )"
```

Remark

Rule representation leaves Gamma implicit.

Rules (¬I) and (¬E) as data

```
definition negI :: plrule where
"negI = (Rule
    [ [(SVar ''A'')] ⊢- FFalse ]
    ( [] ⊢- (Neg (SVar ''A'')) )
 )"

definition negE :: plrule where
"negE =  (Rule
    [ [] ⊢- (Neg (SVar ''A'')), [] ⊢- (SVar ''A'') ]
    ( [] ⊢- FFalse )
 )"
```

Rule application and matching

Rule application

- Goal: function that applies rules to sequents.
- Central aspect: formula matching

Algorithm for matching formulas

- traverse two formulas f_x, f_y where f_x might contain schema variables
- recursively descent into f_x, f_y and build up a substitution σ
- when a schema variable s is encountered in f_x, and sf_y is the corresponding subformula in f_y, check whether the current σ already has a binding for s:
  - if yes and $\sigma(s) = sf_y$ then continue with σ
  - if yes and $\sigma(s) \neq sf_y$ then matching fails
  - if no, add a binding $(s \mapsto sf_y)$ to σ


```
type_synonym substitution = "string ⇀ plformula"

fun matchf :: "plformula ⇒ plformula ⇒ substitution
               ⇒ (substitution × bool)"
where
"matchf (Var s1) (Var s2) σ =
  (if s1=s2 then (σ,True) else (empty,False))"          |
"matchf (SVar s) fy  σ = (case (σ s) of
  None   ⇒ (σ(s↦fy), True) |
  Some f ⇒ (if f=fy then (σ,True) else (empty,False)))"|
"matchf TTrue    TTrue    σ = (σ, True)"               |
"matchf FFalse   FFalse   σ = (σ, True)"               |
"matchf (Neg px) (Neg py) σ = matchf px py σ"          |
"matchf (Imp px1 px2) (Imp py1 py2) σ =
   (let  (σ1,success1) = matchf px1 py1 σ in
     if success1 then matchf px2 py2 σ1
                 else (empty, False)        )"        |
"matchf _        _        _     =  (empty, False)"
```


Matching sequents

Simplification

For the application of rules, we need to match sequents.
In particular, we have to match Γ to the list of assumptions.
To keep things simple here, we

- assume that sequent schemas in consequences only have the schema variable Γ as assumption list.
- handle the binding for Γ as the first component of the result.

```
fun match :: "plsequent ⇒ plsequent
              ⇒ (plformula list × substitution × bool)"
where
"match (gamma ⊢- fx) (fly ⊢- fy) =
                          (fly, (matchf fx fy empty))"
```


Application of substitutions to formulas

```
fun applySubst::"substitution ⇒ plformula ⇒ plformula"
where
"applySubst sigma (Var s)   =   (Var s)"       |
"applySubst sigma (SVar s)  =
     (case (sigma s) of None    ⇒ (SVar s)
                      | Some f  ⇒ f      )"     |
"applySubst sigma TTrue     =   TTrue"         |
"applySubst sigma FFalse    =   FFalse"        |
"applySubst sigma (Neg plf) =
   (Neg (applySubst sigma plf))"                |
"applySubst sigma (Imp plf1 plf2) =
   (Imp (applySubst sigma plf1) (applySubst sigma plf2))"
```

Application of substitutions to sequents

```
fun applySubstSq ::
"plformula list ⇒ substitution ⇒ plsequent ⇒ plsequent"
where
"applySubstSq gammasubst sigma (Sq asms concl) =
   ( gammasubst@(map (applySubst sigma) asms)
    ⊢- (applySubst sigma concl)
   )"
```


Application of rules

Application “methods”

Application of a rule to a proof state might take different arguments.
In our setting, applying

- the introduction rules (→I) and (¬I) need only the rule as argument (handled by `appl`).
- the elimination rules (→E) and (¬E) need an additional formula as argument to instantiate the schema variable in the premisses that does not occur in the consequent (handled by `applE`).

Definition of function *appl*

```
fun appl :: "plrule ⇒ proofstate ⇒ proofstate"
where
"appl pr [] = []"  |
"appl (Rule prems conseq) (currentgoal#psl) =
  (let (gsubst,sigma,success) = match conseq currentgoal
    in if success
       then (map (applySubstSq gsubst sigma) prems) @ psl
       else (currentgoal # psl)
  )"
```


Definition of function *applE*

```
fun applE ::
"plrule ⇒ proofstate ⇒ substitution ⇒ proofstate"
where
"applE pr [] sigma1 = []"  |
"applE (Rule premisses conseq) (currentgoal#psl) sigma1 =
  (let (gsubst,sigma2,success) = match conseq currentgoal
    in if success
       then (map (applySubstSq gsubst (sigma1++sigma2))
                 premisses) @ psl
       else psl
  )"
```

Concluding remarks

- With *appI* and *applE* rules of the natural deduction calculus can be applied to a proof state.
- To prove a formula, start with a single goal and apply rules until no subgoal is left.
- The simple prover illustrates many concepts of Isabelle/HOL's proof system like schema variable, rule application, different arguments depending on the used method/tactic.
- Having described the prover in Isabelle/HOL allows to specify and verify its correctness!

Chapter summary

- Introduction to functional programming and modeling in Isabelle/HOL
- Illustration of basic concepts underlying interactive proof systems
- Mentioning the idea that we could develop a theorem prover (or some other software system) in Isabelle/HOL and prove it correct

Questions

1. What is the relationship between the data type construct and the case expression? Illustrate the relationship by an example.
2. Why are there different forms of function definitions in Isabelle/HOL, but only one in ML?
3. Why is there a distinction between types with equality and types without equality in ML, but not in Isabelle/HOL? I.e., why can we compare functions by equality in Isabelle/HOL, but not in ML?
4. What is the distinction between matching and unification?
5. Why did we develop two functions for rule application (*appI*, *applE*) in our prover?