

# Term Indexing for the LEO-II Prover

Frank Theiss and Christoph Benzmüller

FR Informatik
Universität des Saarlandes
Saarbrücken, Germany

Computer Laboratory
The University of Cambridge
Cambridge, UK



Motivation



- Motivation
- Some introductory conventions and remarks



- Motivation
- Some introductory conventions and remarks
- Closer look at some key features of the approach



- Motivation
- Some introductory conventions and remarks
- Closer look at some key features of the approach
- Preliminary Evaluation



- Motivation
- Some introductory conventions and remarks
- Closer look at some key features of the approach
- Preliminary Evaluation
- Conclusion



Implements extensional higher order resolution:



#### Implements extensional higher order resolution:

$$\forall B_{\alpha \to o}, C_{\alpha \to o}, D_{\alpha \to o} B \cup (C \cap D) = (B \cup C) \cap (B \cup D)$$

Negation and definition expansion with

$$\cup = \lambda A_{\alpha \to o}, B_{\alpha \to o}, X_{\alpha \bullet}(A X) \lor (B X) \qquad \cap = \lambda A_{\alpha \to o}, B_{\alpha \to o}, X_{\alpha \bullet}(A X) \land (B X)$$

leads to:

$$C_1: [\lambda X_{\alpha} (b X) \vee ((c X) \wedge (d X)) \neq^? \lambda X_{\alpha} ((b X) \vee (c X)) \wedge ((b X) \vee (d X)))]$$

Goal directed functional and Boolean extensionality treatment:

$$C_2: [(b\ x) \lor ((c\ x) \land (d\ x)) \Leftrightarrow ((b\ x) \lor (c\ x)) \land ((b\ x) \lor (d\ x)))]^F$$

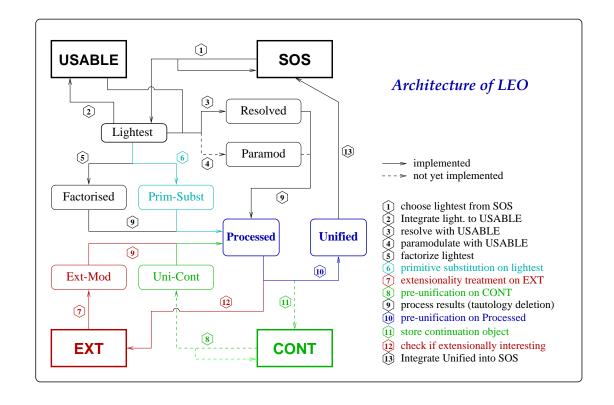
Clause normalization results then in a pure propositional, i.e. decidable, set of clauses. Only these clauses are still in the search space of  $\mathcal{LEO}$  (in total there are 33 clauses generated and  $\mathcal{LEO}$  finds the proof on a 2,5GHz PC in 820ms).



- Implements extensional higher order resolution:
- Based on extended set of support architecture



- Implements extensional higher order resolution:
- Based on extended set of support architecture





- Implements extensional higher order resolution:
- Based on extended set of support architecture
- Developed in LISP within the (old) OMEGA framework



- Implements extensional higher order resolution:
- Based on extended set of support architecture
- Developed in LISP within the (old) OMEGA framework
- Uses KEIM generic datastructures (how inefficient?)



- Implements extensional higher order resolution:
- Based on extended set of support architecture
- Developed in LISP within the (old) OMEGA framework
- Uses KEIM generic datastructures (how inefficient?)
- Can handle only a few thousand clauses in search space



- Implements extensional higher order resolution:
- Based on extended set of support architecture
- Developed in LISP within the (old) OMEGA framework
- Uses KEIM generic datastructures (how inefficient?)
- Can handle only a few thousand clauses in search space
- Often generates lots of first order or propositional clauses

© Theiss and Benzmüller, 2006



- Implements extensional higher order resolution:
- Based on extended set of support architecture
- Developed in LISP within the (old) OMEGA framework
- Uses KEIM generic datastructures (how inefficient?)
- Can handle only a few thousand clauses in search space
- Often generates lots of first order or propositional clauses
- Has been successfully combined with first order ATPs



■ EPSRC project at University of Cambridge (with L. Paulson)

© Theiss and Benzmüller, 2006



- EPSRC project at University of Cambridge (with L. Paulson)
- Reimplentation of LEO in OCAML



- EPSRC project at University of Cambridge (with L. Paulson)
- Reimplentation of LEO in OCAML
- Shall combine the lessons of recent research with new ideas



- EPSRC project at University of Cambridge (with L. Paulson)
- Reimplentation of LEO in OCAML
- Shall combine the lessons of recent research with new ideas
- Shall be easy to integrate with proof assistants (ISABELLE/HOL, HOL, OMEGA, ...)

© Theiss and Benzmüller, 2006



- EPSRC project at University of Cambridge (with L. Paulson)
- Reimplentation of LEO in OCAML
- Shall combine the lessons of recent research with new ideas
- Shall be easy to integrate with proof assistants (ISABELLE/HOL, HOL, OMEGA, ...)
- HOTPTP as input syntax



- EPSRC project at University of Cambridge (with L. Paulson)
- Reimplentation of LEO in OCAML
- Shall combine the lessons of recent research with new ideas
- Shall be easy to integrate with proof assistants (ISABELLE/HOL, HOL, OMEGA, ...)
- HOTPTP as input syntax
- Shall cooperate with first order automated theorem provers



- EPSRC project at University of Cambridge (with L. Paulson)
- Reimplentation of LEO in OCAML
- Shall combine the lessons of recent research with new ideas
- Shall be easy to integrate with proof assistants (ISABELLE/HOL, HOL, OMEGA, ...)
- HOTPTP as input syntax
- Shall cooperate with first order automated theorem provers
- Shall be way more efficient than LEO (which was developed rather as an academic demonstrator than a prototype)

© Theiss and Benzmüller, 2006



Term indexing is standard in first order theorem proving



- Term indexing is standard in first order theorem proving
- Efficiency is now an important issue in LEO-II redevelopment



- Term indexing is standard in first order theorem proving
- Efficiency is now an important issue in LEO-II redevelopment
- Comparably few techniques studied for higher order logic



- Term indexing is standard in first order theorem proving
- Efficiency is now an important issue in LEO-II redevelopment
- Comparably few techniques studied for higher order logic
- One example: Brigitte Pientka's work [ICLP-03]
  - based on higher order substitution tree indexing
  - relies on unification of linear higher order patterns



- Term indexing is standard in first order theorem proving
- Efficiency is now an important issue in LEO-II redevelopment
- Comparably few techniques studied for higher order logic
- One example: Brigitte Pientka's work [ICLP-03]
  - based on higher order substitution tree indexing
  - relies on unification of linear higher order patterns
- Our approach
  - based on Stickel's coordinate and path indexing [SRI-Report-89]

low level operations (e.g., operations on hashtables)

# **HOL-Syntax: Simple Types**



(truth values)

Simple Types T:

(individuals)

 $(\alpha \rightarrow \beta)$  (functions from  $\alpha$  to  $\beta$ )

# **HOL-Syntax:** Simply Typed $\lambda$ -Terms



#### Typed Terms:

 $X_{\alpha}$  Variables  $(\mathcal{V})$ 

 $c_{\alpha}$  Constants & Parameters ( $\Sigma$ )

 $(\mathbf{F}_{\alpha \to \beta} \, \mathbf{B}_{\alpha})_{\beta}$  Application

 $(\lambda Y_{\alpha} \mathbf{A}_{\beta})_{\alpha \to \beta}$   $\lambda$ -abstraction

# **HOL-Syntax:** Simply Typed $\lambda$ -Terms



#### Typed Terms:

 $X_{\alpha}$  Variables ( $\mathcal{V}$ )

 $c_{\alpha}$  Constants & Parameters ( $\Sigma$ )

 $(\mathbf{F}_{\alpha \to \beta} \, \mathbf{B}_{\alpha})_{\beta}$  Application

 $(\lambda Y_{\alpha} \mathbf{A}_{\beta})_{\alpha \to \beta}$   $\lambda$ -abstraction

### **Equality of Terms:**

 $\alpha$ -conversion Changing bound variables

 $\beta$ -reduction  $((\lambda Y_{\beta} \mathbf{A}_{\alpha}) \mathbf{B}_{\beta}) \stackrel{\beta}{\longrightarrow} [\mathbf{B}/Y] \mathbf{A}$ 

 $\eta$ -reduction  $(\lambda Y_{\alpha} (\mathbf{F}_{\alpha \to \beta} Y)) \xrightarrow{\eta} \mathbf{F}$   $(Y_{\beta} \notin \mathbf{Free}(\mathbf{F}))$ 

# **HOL-Syntax:** Simply Typed $\lambda$ -Terms



#### Typed Terms:

 $X_{\alpha}$  Variables ( $\mathcal{V}$ )

 $c_{\alpha}$  Constants & Parameters ( $\Sigma$ )

 $(\mathbf{F}_{\alpha \to \beta} \, \mathbf{B}_{\alpha})_{\beta}$  Application

 $(\lambda Y_{\alpha} \mathbf{A}_{\beta})_{\alpha \to \beta}$   $\lambda$ -abstraction

### **Equality of Terms:**

Every term has a unique  $\beta\eta$ -normal form (up to  $\alpha$ -conversion).



Instead of named bound variables we use de Bruijn indices

© Theiss and Benzmüller, 2006



- Instead of named bound variables we use de Bruijn indices
- Example:

$$\lambda a.\lambda b.(b = ((\lambda c.(cb))a))$$

translates to

$$\lambda \lambda . (x_0 = ((\lambda . (x_0 x_1)) x_1))$$



- Instead of named bound variables we use de Bruijn indices
- Example:

$$\lambda a.\lambda b.(b = ((\lambda c.(cb))a))$$

translates to

$$\lambda \lambda . (x_0 = ((\lambda . (x_0 x_1)) x_1))$$

Different occurrences of the same bound variable may have different de Bruijn indices:

b translates to both  $x_0$  and  $x_1$ 



- Instead of named bound variables we use de Bruijn indices
- Example:

$$\lambda a.\lambda b.(b = ((\lambda c.(cb))a))$$

translates to

$$\lambda \lambda . (x_0 = ((\lambda . (x_0 x_1)) x_1))$$

Different occurrences of the same bound variable may have different de Bruijn indices:

b translates to both  $x_0$  and  $x_1$ 

Different occurrences of the same de Bruijn index may refer to different λ-binders:

 $x_0$  relates to bound variable b and c

# **Some Important Remarks**



• Only terms in  $\beta\eta$  normal form ( $\eta$  short and  $\beta$  normal) are indexed



- Only terms in  $\beta\eta$  normal form ( $\eta$  short and  $\beta$  normal) are indexed
- **Example:**  $(\lambda.(\lambda.(x_1 x_0)))((\lambda.g x_0)c)$  normalises to (g c)



- Only terms in  $\beta\eta$  normal form ( $\eta$  short and  $\beta$  normal) are indexed
- $(\lambda.(\lambda.(x_1\,x_0))))((\lambda.g\,x_0)c)$  normalises to  $(g\,c)$ Example:
- Invariant for LEO-II: normalisation after each calculus step



- Only terms in  $\beta\eta$  normal form ( $\eta$  short and  $\beta$  normal) are indexed
- **Example:**  $(\lambda.(\lambda.(x_1 x_0)))((\lambda.g x_0)c)$  normalises to (g c)
- Invariant for LEO-II: normalisation after each calculus step
- Types
  - provide an additional criterion for distinction of terms (e.g., different occurrences of the same de Bruijn index may have different types)
  - no further impact on the indexing mechanism



- Only terms in  $\beta\eta$  normal form ( $\eta$  short and  $\beta$  normal) are indexed
- **Example:**  $(\lambda.(\lambda.(x_1 x_0)))((\lambda.g x_0)c)$  normalises to (g c)
- Invariant for LEO-II: normalisation after each calculus step
- Types
  - provide an additional criterion for distinction of terms (e.g., different occurrences of the same de Bruijn index may have different types)
  - no further impact on the indexing mechanism
- Due to Currying all our applications have just one argument





- 1. Shared representation of terms
- 2. Use of partial syntax trees to speedup logical computations



- 1. Shared representation of terms
- 2. Use of partial syntax trees to speedup logical computations
- 3. Indexing of subterm occurrences



- 1. Shared representation of terms
- 2. Use of partial syntax trees to speedup logical computations
- 3. Indexing of subterm occurrences
- 4. Indexing of bound variable occurrences (support for explicit substitutions)



Terms are represented as sets of term nodes



- Terms are represented as sets of term nodes
  - ightharpoonup symbol  $s \in \Sigma$
- $\longrightarrow$  a term node n : symbol(s) is created
- bound variable  $x_d$  (d is de Bruijn index)
  - $\longrightarrow$  a term node m: bound(type, d) is created
- ▶ application (s t) (s, t) already represented by nodes i, j)  $\longrightarrow$  a term node l: application(i, j) is created
- ightharpoonup abstraction  $\lambda t$  (t is already represented by i)
  - $\longrightarrow$  a term node k: abstraction(type, i) is created



- Terms are represented as sets of term nodes
- Example:



- Terms are represented as sets of term nodes
- Example: Original problem terms

$$\emptyset := \lambda x_{\iota}.\bot$$

$$\in := \lambda y_{\iota}.\lambda s_{\iota \to o}.(s\ y)$$

$$\neg (A_{\iota} \in \emptyset)$$

$$(A_{\iota} \in \emptyset) \Rightarrow \bot$$

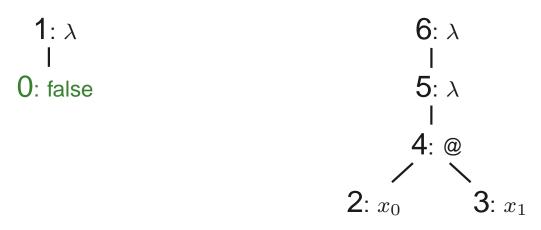


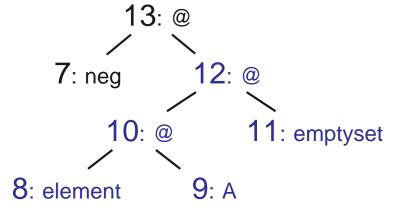
- Terms are represented as sets of term nodes
- Example: HOTPTP encoding

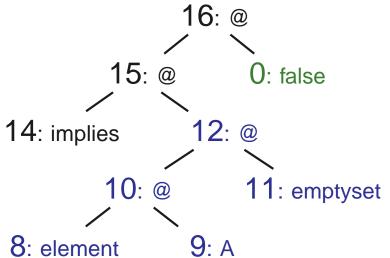
```
thf(emptyset, definition,
        (emptyset :=
                 (^ [Z : $i] : $false))).
thf(element, definition,
        (element :=
                 (^{(Y:\$i, S:(\$i > \$o)]} : (S @ Y)))).
thf(theorem1,conjecture,
        (~ ((element @ A) @ emptyset))).
thf(theorem1alt,conjecture,
        (((element @ A) @ emptyset) => $false)).
```



- Terms are represented as sets of term nodes
- Example: Graph representation of terms









- Terms are represented as sets of term nodes
- Example: Representation as term sets

```
0: symbol false
                             10: appl(8,9)
1: abstr(i,0)
                             11: symbol emptyset
2: bound(i -> 0,0)
                             12: appl(10,11)
3: bound(i,1)
                             13: appl(7,12)
                             14: symbol implies
4: appl(2,3)
5: abstr(i -> o,4)
                             15: appl(14,12)
6: abstr(i,5)
                             16: appl(15,0)
7: symbol neg
8: symbol element
9: symbol A
```

© Theiss and Benzmüller, 2006



- Terms are represented as sets of term nodes
- Example: Parsing returns pointers to this term set / index

```
emptyset: lambda [Z]: false
->index: 1

element: lambda [Y]: lambda [S]: S Y
->index: 6

theorem1: neg ((element A) emptyset)
->index: 13

theorem1alt: (implies ((element A) emptyset)) false
->index: 16
```



- Terms are represented as sets of term nodes
- Example: Term set representation supported via hashtables

- ht abstr\_with\_scope :  $I\!\!N \to I\!\!N$ : lookup abstractions with a given scope i
- ht appl\_with\_func :  $I\!\!N \to I\!\!N \to I\!\!N$ : lookup application with a given function i and argument j
- ht appl\_with\_arg :  $I\!\!N \to I\!\!N \to I\!\!N$ : lookup application with a given argument j and function i



 Represention of the paths to particular symbol or subterm occurrences within a term



- Represention of the paths to particular symbol or subterm occurrences within a term
- Called partial because they only represent relevant parts of a term



- Represention of the paths to particular symbol or subterm occurrences within a term
- Called partial because they only represent relevant parts of a term
- Allow for early detection of branches in a term's syntax tree with no occurrences of a specific symbol/subterm, since these branches are not represented



- Represention of the paths to particular symbol or subterm occurrences within a term
- Called partial because they only represent relevant parts of a term
- Allow for early detection of branches in a term's syntax tree with no occurrences of a specific symbol/subterm, since these branches are not represented
- Need to define position before we can give an example PST

#### **Positions**



• Consider term  $(\lambda.x_0)@(f@a)$ :

```
(\lambda . x_0) @ (f@a) : []
```

 $\lambda . x_0$  : [func]

 $x_0$ : [func; arg]

f@a : [arg]

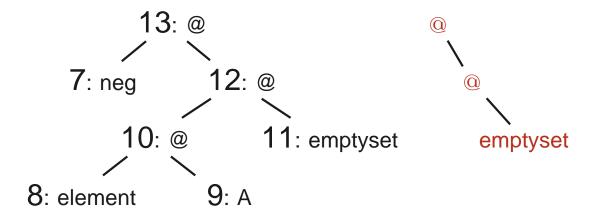
f: [arg; func]

a : [arg; arg]



#### Example

PST for occurrences of symbol emptyset in theorem1

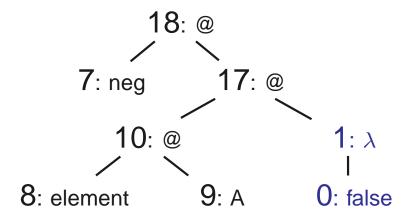


```
\begin{array}{ll} \text{positiontable:} & p_{\text{emptyset-1}} = pst(\_,\_,p_{\text{emptyset-2}}) \\ \text{[arg; arg]: emptyset} & p_{\text{emptyset-2}} = pst(\_,\_,p_{\text{emptyset-3}}) \\ \text{end.} & p_{\text{emptyset-3}} = pst(\_,\_,\_) & -\text{emptyset-3} \end{array}
```



#### Example

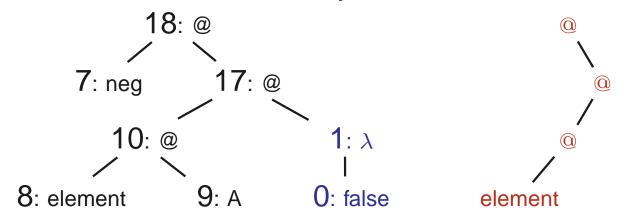
Modified term after replacement





#### Example

PST for occurrences of symbol element in term

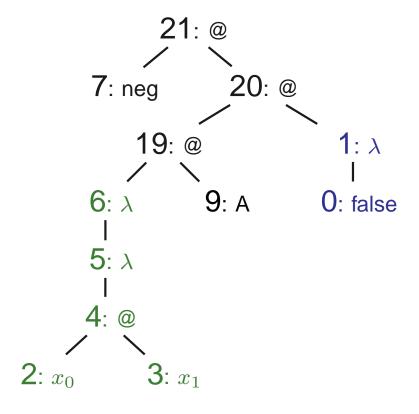


```
\begin{array}{ll} p_{\text{emptyset-1}} = pst(\_,\_,p_{\text{emptyset-2}}) \\ positiontable: & p_{\text{emptyset-2}} = pst(\_,p_{\text{emptyset-3}},\_,) \\ [arg; func; func]: & \text{element} \\ p_{\text{emptyset-3}} = pst(\_,p_{\text{emptyset-4}},p_{\text{emptyset-4}},\_,) \\ p_{\text{emptyset-4}} = pst(\_,\_,\_) & -\text{element} \end{array}
```



#### Example

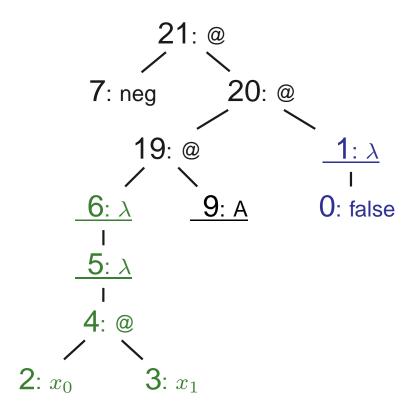
Modified term after replacement





#### Example

Use available information for normalisation





#### Example

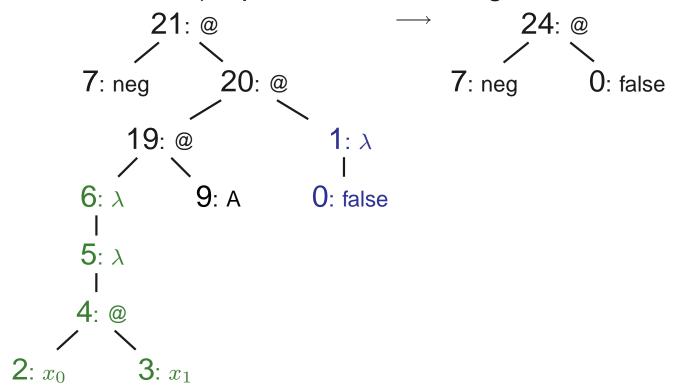
We also index occurrences of bound variables





#### Example

Normalisation (required before term goes to index again)

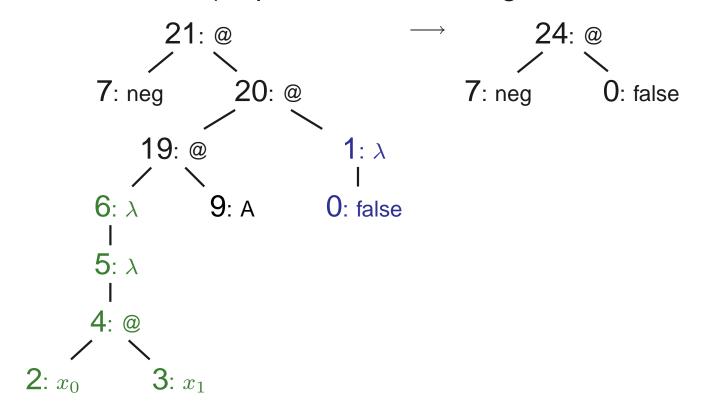


(LEO-II will later immediately say: Proof found!)



#### Example

Normalisation (required before term goes to index again)

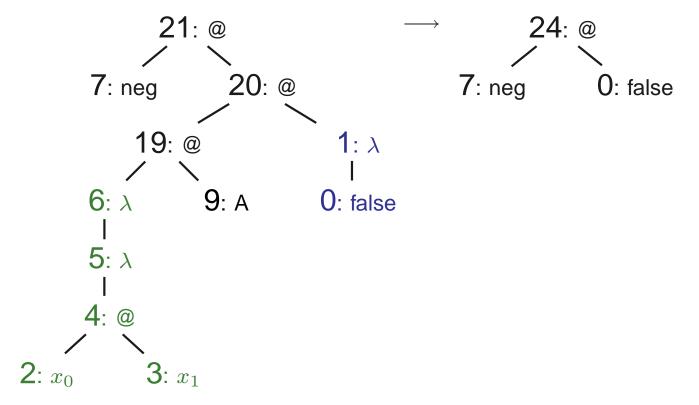


Hence, we guide replacement & normalization with PSTs



#### Example

Normalisation (required before term goes to index again)



- Hence, we guide replacement & normalization with PSTs
- Provide also PSTs for bound variable occurrences



 Index records whether and at which positions a subterm occurs in a term (symbols and nonprimitive subterms can be handled – tradeoff?)



- Index records whether and at which positions a subterm occurs in a term (symbols and nonprimitive subterms can be handled – tradeoff?)
- For each new (normalised) term to be indexed the term set is computed and termset hashtables are updated



- Index records whether and at which positions a subterm occurs in a term (symbols and nonprimitive subterms can be handled – tradeoff?)
- For each new (normalised) term to be indexed the term set is computed and termset hashtables are updated
- At the same time term is analysed for contained subterms



- Index records whether and at which positions a subterm occurs in a term (symbols and nonprimitive subterms can be handled – tradeoff?)
- For each new (normalised) term to be indexed the term set is computed and termset hashtables are updated
- At the same time term is analysed for contained subterms
- Result is a PST for each subterm of a given term to be indexed



- Index records whether and at which positions a subterm occurs in a term (symbols and nonprimitive subterms can be handled – tradeoff?)
- For each new (normalised) term to be indexed the term set is computed and termset hashtables are updated
- At the same time term is analysed for contained subterms
- Result is a PST for each subterm of a given term to be indexed
- PSTs are added to hashtable occurrences :  $I\!N \to I\!N \to PST$

-p.15(c) Theiss and Benzmüller, 2006

## **Building and Using the Index**



- Index records whether and at which positions a subterm occurs in a term (symbols and nonprimitive subterms can be handled – tradeoff?)
- For each new (normalised) term to be indexed the term set is computed and termset hashtables are updated
- At the same time term is analysed for contained subterms
- Result is a PST for each subterm of a given term to be indexed
- PSTs are added to hashtable occurrences :  $I\!\!N \to I\!\!N \to PST$
- lacksquare Second hashtable occurs\_in :  $I\!\!N o I\!\!N^*$

## **Building and Using the Index**



- Index records whether and at which positions a subterm occurs in a term (symbols and nonprimitive subterms can be handled – tradeoff?)
- For each new (normalised) term to be indexed the term set is computed and termset hashtables are updated
- At the same time term is analysed for contained subterms
- Result is a PST for each subterm of a given term to be indexed
- PSTs are added to hashtable occurrences :  $I\!\!N \to I\!\!N \to PST$
- lacksquare Second hashtable occurs\_in :  $I\!\!N o I\!\!N^*$
- lacksquare Third hashtable occurs\_at :  $pos o I\!\!N o I\!\!N^*$



977 random terms from a HOTPTP version of Jutting's
 Automath encoding of Landau's book Grundlagen der Analysis



 977 random terms from a HOTPTP version of Jutting's Automath encoding of Landau's book Grundlagen der Analysis

Number of indexed terms	977
Number of created term nodes	11618
Average term size	54
Number of nodes with no parent nodes	904
Number of nodes with one parent node	9633
Number of nodes with two more more parent nodes	1083
Maximum number of parent nodes	$2778$ (symbol $\forall$ )
Average number of parent nodes	1.68
Average number of terms a node occurs in	33.5
-"-(for symbols)	493.9
-"-(for nonprimitive term nodes)	24
Average PST/term size for symbol occurrences	0.21
Average PST/term size for bound variable occurrences	0.33
Average PST/term size for all term nodes	0.12



- 977 random terms from a HOTPTP version of Jutting's
   Automath encoding of Landau's book Grundlagen der Analysis
- Average of 1.68 parent nodes seems low / relativised by average of 33.5 terms a node occurs in



- 977 random terms from a HOTPTP version of Jutting's
   Automath encoding of Landau's book Grundlagen der Analysis
- Average of 1.68 parent nodes seems low / relativised by average of 33.5 terms a node occurs in
- One indicator for term retrieval performance: 99,7% of candidate nodes can be excluded (average of occurrences 33.5 versus 11618 terms)



- 977 random terms from a HOTPTP version of Jutting's
   Automath encoding of Landau's book Grundlagen der Analysis
- Average of 1.68 parent nodes seems low / relativised by average of 33.5 terms a node occurs in
- One indicator for term retrieval performance: 99,7% of candidate nodes can be excluded (average of occurrences 33.5 versus 11618 terms)
- Occurs check for symbols, bound variables, and non primitive subterms in constant time is expected to have strong impact on crucial operations (e.g., replacement, substitution, global unfolding of definitions)



- 977 random terms from a HOTPTP version of Jutting's
   Automath encoding of Landau's book Grundlagen der Analysis
- Average of 1.68 parent nodes seems low / relativised by average of 33.5 terms a node occurs in
- One indicator for term retrieval performance: 99,7% of candidate nodes can be excluded (average of occurrences 33.5 versus 11618 terms)
- Occurs check for symbols, bound variables, and non primitive subterms in constant time is expected to have strong impact on crucial operations (e.g., replacement, substitution, global unfolding of definitions)
- One indicator for improvement for replacement operations is PST/term size rate: 0.21 (= speedup factor 5) for symbols and 0.33 (speedup factor 3) for bound variables



Presented a term indexing approach that



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,
  - in a higher order context



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,
  - in a higher order context
- Orthogonal/complementary to Pientka's approach combination ?



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,
  - in a higher order context
- Orthogonal/complementary to Pientka's approach combination ?
- Future work:



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,
  - in a higher order context
- Orthogonal/complementary to Pientka's approach combination?
- Future work:
  - Study alternative term representation techniques (suspension calculus, spine representation, explicit substitutions) in our framework



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,
  - in a higher order context
- Orthogonal/complementary to Pientka's approach combination ?
- Future work:
  - Study alternative term representation techniques (suspension calculus, spine representation, explicit substitutions) in our framework
  - Proper evaluation in LEO-II



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,
  - in a higher order context
- Orthogonal/complementary to Pientka's approach combination ?
- Future work:
  - Study alternative term representation techniques (suspension calculus, spine representation, explicit substitutions) in our framework
  - Proper evaluation in LEO-II



- Presented a term indexing approach that
  - combines ideas from first order coordinate and path indexing [Stickel-89]
  - with novel ideas, especially partial syntax trees,
  - in a higher order context
- Orthogonal/complementary to Pientka's approach combination?
- Future work:
  - Study alternative term representation techniques (suspension calculus, spine representation, explicit substitutions) in our framework
  - Proper evaluation in LEO-II