technische universität
dortmund

Chair of Programming Systems /
Department of Computer Science

Next Generation LearnLib (NGLL)

Falk Howar, Maik Merten

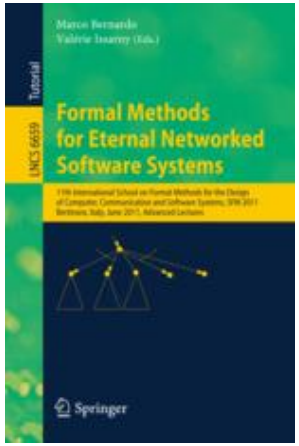
TU Dortmund, Chair of Programming Systems

LearnLib.
www.learnlib.de

Synopsis

- LearnLib: mature library for **active atomata learning**
- NGLL: framework, providing **infrastructure** for practical application
- Tool-demo: **Modeling** learning setups with LearnLib Studio

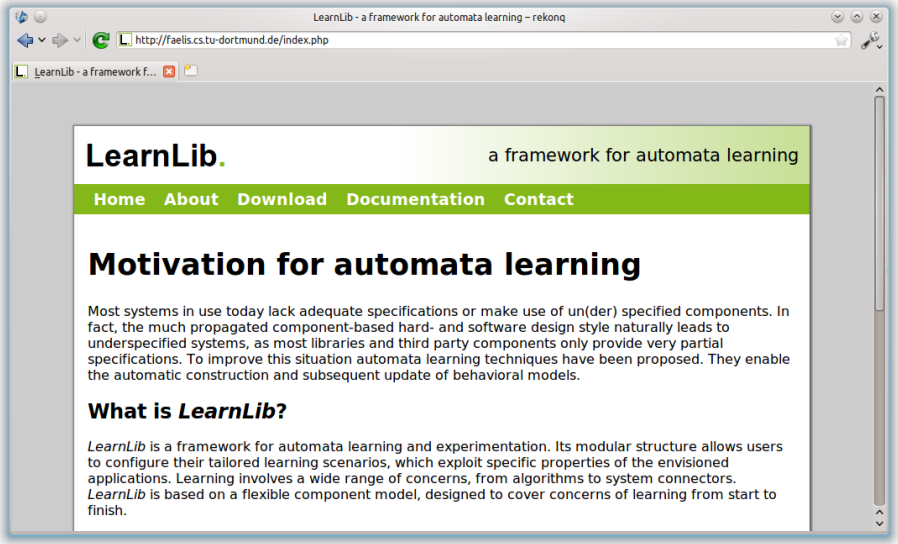
Literature / documentation



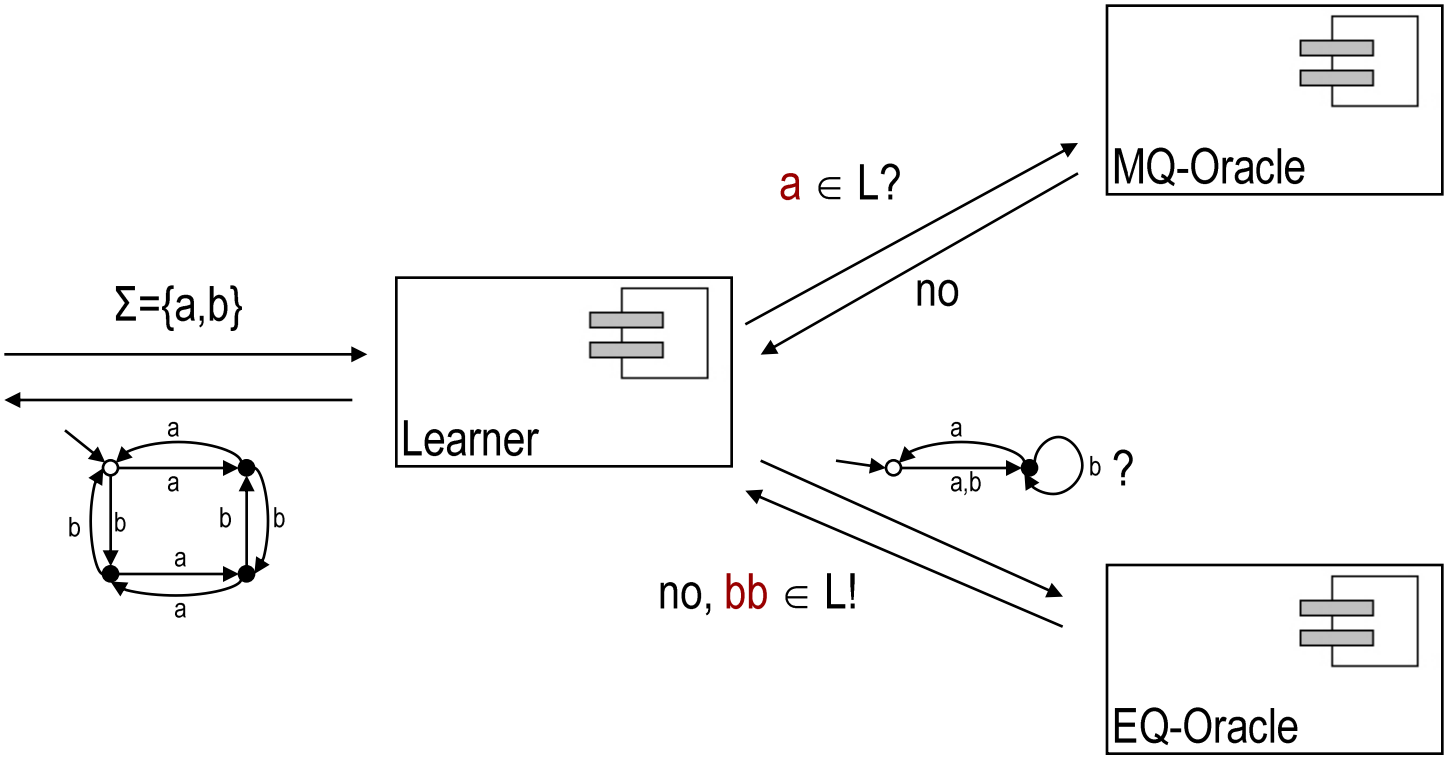
Bernhard Steffen, Falk Howar, Maik Merten: *Introduction to Active Automata Learning from a Practical Perspective*. **SFM 2011**.

Maik Merten, Bernhard Steffen, Falk Howar, Tiziana Margaria: *Next Generation LearnLib*. **TACAS 2011**.

- **http://www.learnlib.de**
- Tool
- Documentation
- Tutorials
- Pointers to literature



Active automata learning



NGLL: structure & components

Algorithms (LearnLib)

- Learning algorithms
- Counterexample analysis
- Equivalence query realizations

Core/Utils (LearnLib)

- (Small) automata library
- Statistics
- Logging
- Import/export (dot, ppt, xml etc.)
- Simulation environments

Case studies (NGLL)

- Test-driver tools
- Data mappers
- Abstraction

LearnLib Studio (NGLL)

- Visual composition of learning experiments

Algorithms

- **Learning algorithms**
  - Angluin's L*
  - Rivest and Schapire's "Reduced Observation Table"
  - Non-uniform observation table
  - DHC
- **Handling Counterexamples**
  - Maler and Pnueli
  - Shahabaz and Groz
  - Rivest and Schapire
- **Equivalence queries**
  - Chow's "W-Method"
  - Random walks
  - Hopcroft and Karp's almost linear time explicit equivalence test
  - Evolving hypothesis

+libalf 0.3 algorithms

All algorithms work for DFA + Mealy machines

LearnLib (algorithms) enhancements

Java NGLL (now)

163,999 states, 121 actions
≈ 20 million transitions

2.3GB memory
ca. 46 million MQ
ca. 100k EQ

ca. 60 minutes (learning: 59 min, eqs: 20 sec)

Partial models
Partial observation tables

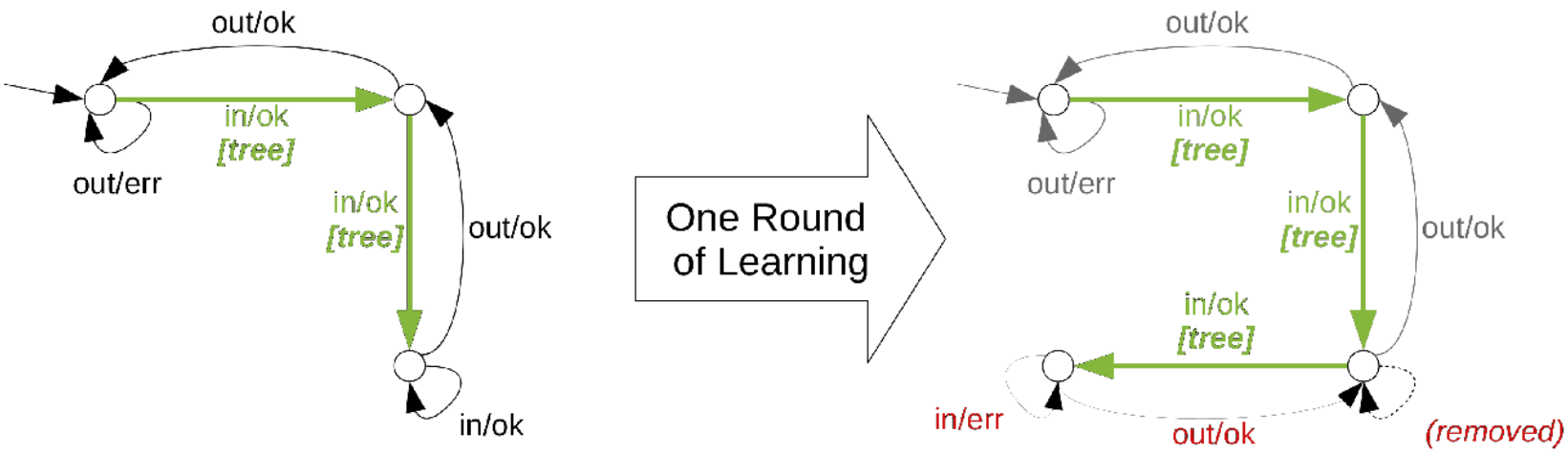
Old c++ Le

- 3 hours
- 7GB me
- > 690 million MQ

011)

- %)
- (-60%)
- ca. 593 million MQ (-15%)

Evolving hypothesis



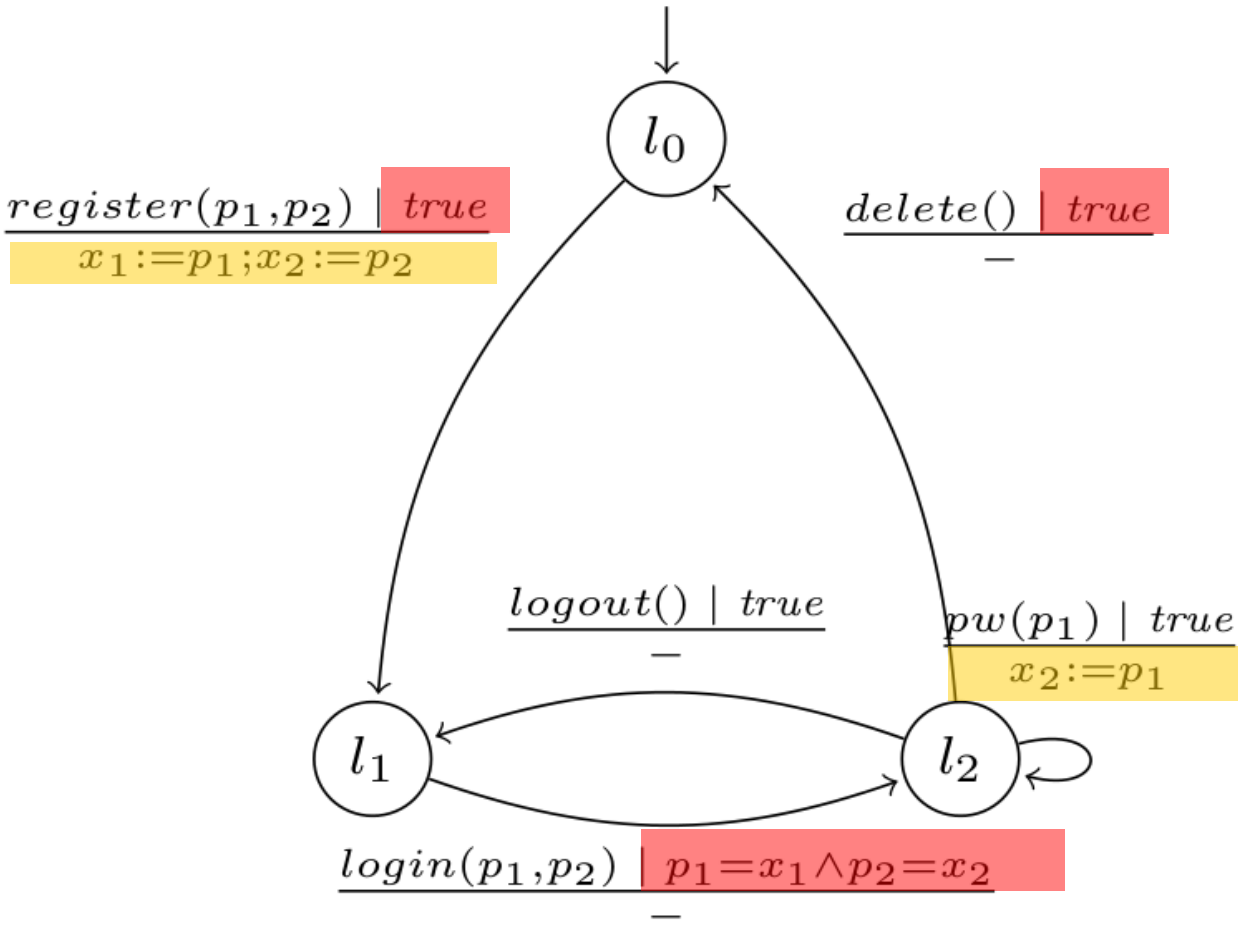
- The set $S \cup SA$ defines a monotonically growing spanning tree of the target automaton
- Usually only local modifications between two equivalence queries (especially for non-uniform sets of distinguishing suffixes)

ZULU competition (results for single example)

Algorithm	New Membership Queries			Rounds	States	Score
	Close	Analyze	Search			
E.H.Blocking	6,744	358	999	259	352	94.11
E.H.Weighted	6,717	349	1,035	262	351	94.61
Random	6,586	519	996	228	332	93.28
run_random	8,080	14	7	5	312	74.89
run_blocking1	8,074	11	16	6	319	73.06
run_weighted1	8,077	9	15	6	319	74.39

- ZULU limit: 8,101
- MQs / EQ: 1-3 (uniform), ca. 3.9 (non-uniform), ca. 4.36 (random)
- MQS / State: ca. 25 (uniform), ca. 19 (non-uniform)
- Random Walks: higher costs for analyzing counterexamples

Support for register automata



RA learning – preliminary results

Setup	# Loc.	# Trans.	MQs	EQs
RA learning algorithm	3	16	329	3
L^*, symmetry reduction, $\|D\| = 6$)	73	5,913	2,776	2
L^*, no optimization, $\|D\| = 6$)	73	5,913	415,333	72

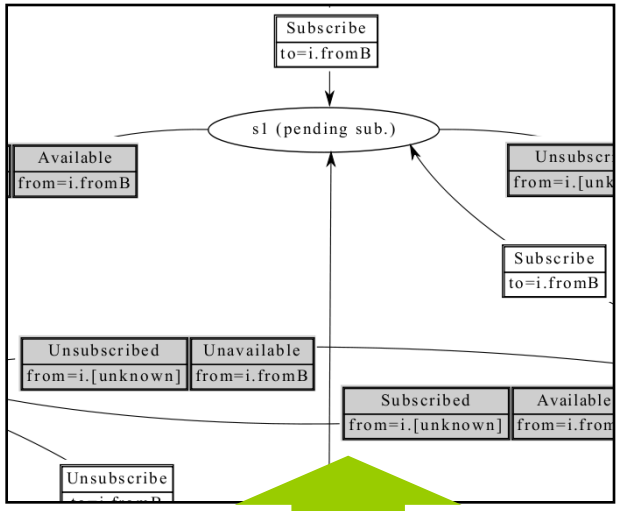
Practical challenges

```
<xs:complexType name="Login">
 <xs:sequence>
  <xs:element minOccurs="0" name="
  <xs:element minOccurs="0" name="
 </xs:sequence>
</xs:complexType>
<xs:complexType name="BeginTransa
 <xs:sequence>
  <xs:element minOccurs="0" name="
  <xs:element minOccurs="0" name="
  <xs:element minOccurs="0" name="
```

Interface description

etc.

interfacing real systems:
- alphabet generation
- abstraction
- data



Behavioral models

equivalence queries

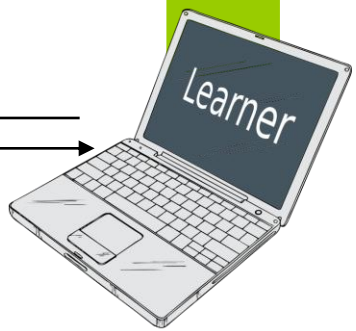
<presence type=… />

<iq type= “result“ />

Test-driver

Available

OK

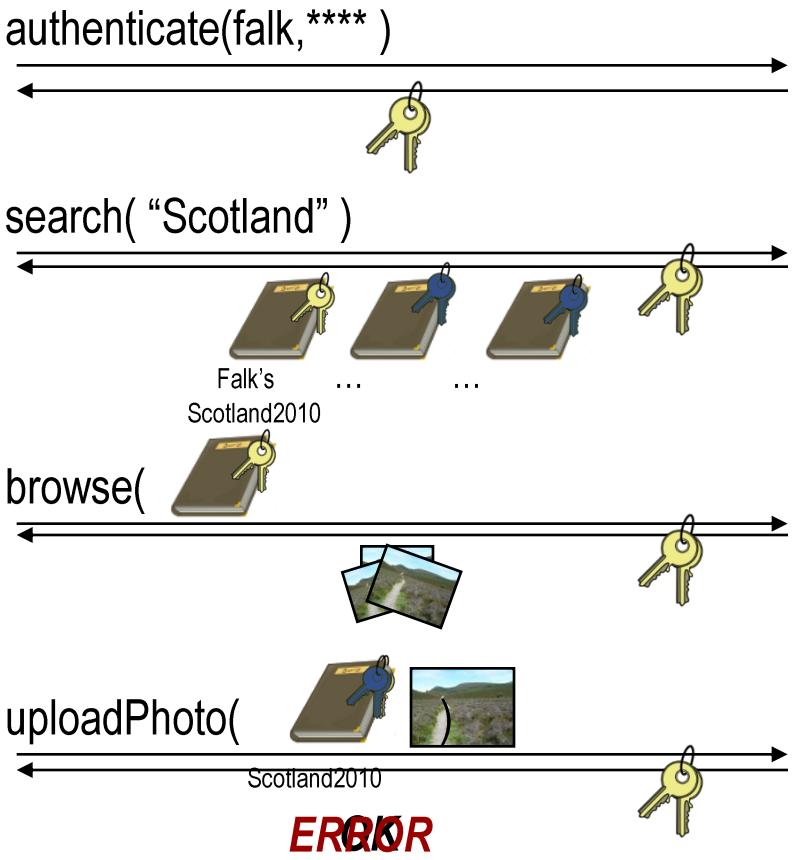


membership queries

reset

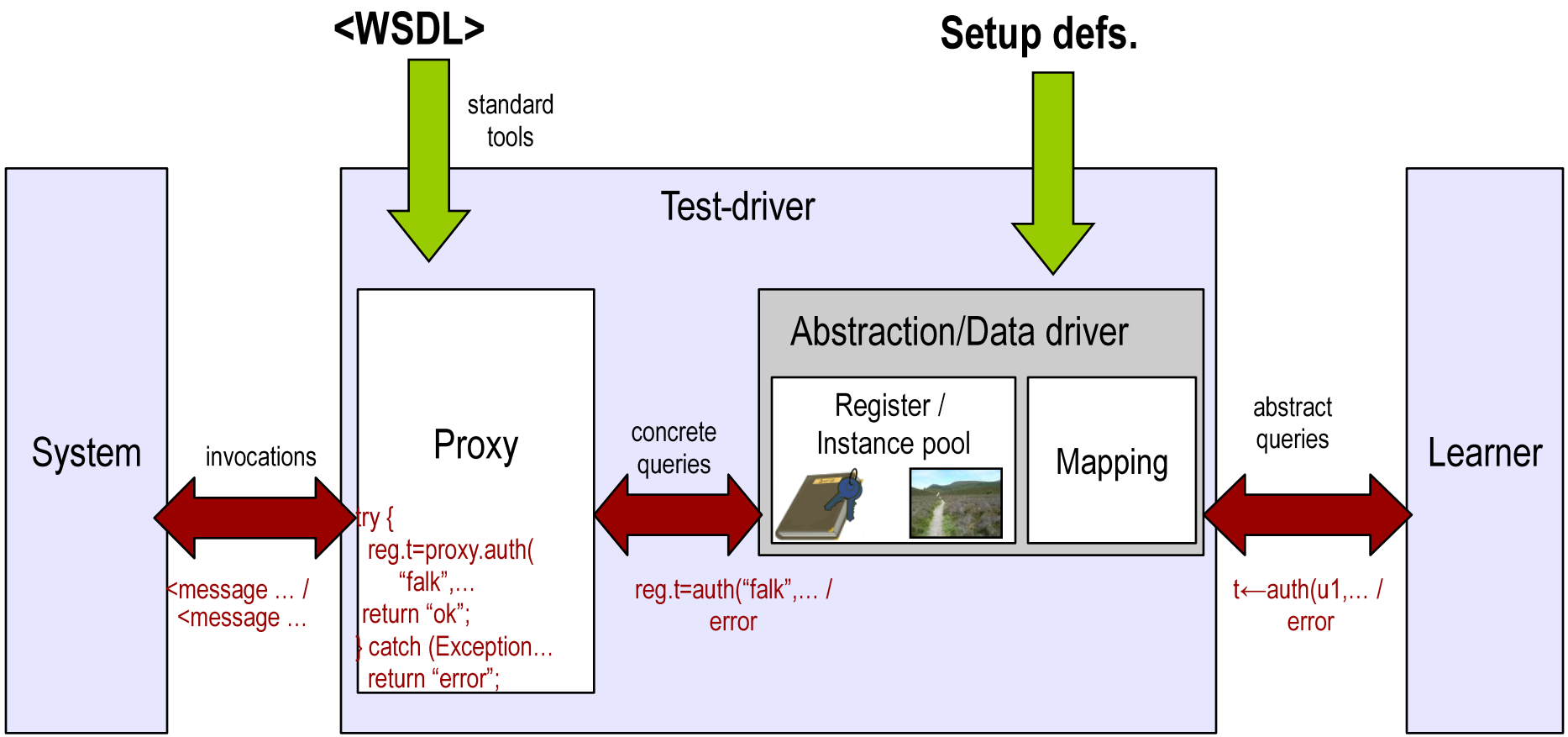
Case study tools

A photo service



register(…
createAlbum(…
getAlbum(…

Generating test-drivers from WSDLs



Live demo

...