Learning Revised Models For Planning In Adaptive Systems

Daniel Sykes, Domenico Corapi, Jeff Magee, Jeff Kramer, Alessandra Russo, and Katsumi Inoue

Imperial College, London

ICSE, 22nd May 2013
Adaptive systems

System

Domain model

Environment

action

reaction
Adaptive systems

**System** should adapt to **System’**, but domain model out of date
Adaptive systems

Behavioural model revision through **probabilistic rule learning**
Factory floor reactive plan

1. pickup
2. move(3)
3. move(3)
4. move(5)
5. putdown
6. Goal

Sykes et al., SEAMS 2008
Factory floor at runtime

$P(\text{success}) = r$
Factory floor at runtime

Sensors report location 4

P(success) = p

move(3)

P(success) = r

move(5)

move(5)

putdown

1

2

3

4

5

6

pickup
Factory floor at runtime

Probability of overall path success = p.q.r

P(success) = r
P(success) = q
P(success) = p
Model revision

• Model does not reflect real environment
  – Unmodelled states or transitions
  – Original model not probabilistic
  – Difficult to estimate probabilities without testing

Task: update model according to observed environment
Probabilistic rule learning

NoMProL

1. Rule learning
2. Probab. estimation

Planning

Traces

Plan

Plan execution

Kramer & Magee, FOSE 2007
Corapi et al., CLIMA 2011
Probabilistic rule learning

1. Rule learning
2. Probabilistic estimation

Domain model

NoMProL

Planning

Goal management layer

Change & component layers

Plan execution

Traces

Plan

Corapi et al., CLIMA 2011

Kramer & Magee, FOSE 2007

daniel.sykes@imperial.ac.uk
Inductive logic programming

\[ B \cup H \models E \]

Background knowledge, Hypothesis (rules), Observations (traces)

Many possible hypotheses, some very specific, some more general

Muggleton 1995
Mode declarations

\[
\text{mode}(h, 2, \text{succeeds}(act, +time))
\]
\[
\text{mode}(b, 2, \text{holdsAt}(\text{cond}, +time))
\]
\[
\text{mode}(b, 2, \text{not holdsAt}(\text{cond}, +time))
\]

Head or body

Maximum occurrences in a rule

Want to learn action success under holdsAt conditions

Mode declaration for each action act and condition cond
possible(pickup, T) :-
    not holdsAt(holdingObject, T),
    holdsAt(at(loc1), T).
possible(putdown, T) :-
    holdsAt(holdingObject, T),
    holdsAt(at(loc5), T).
possible(move(L1, L2), T) :-
    holdsAt(at(L1), T),
    connected(L1, L2).

... initiates(pickup, holdingObject, T).
terminates(putdown, holdingObject, T).
initiates(move(L1, L2), at(L2), T).
terminates(move(L1, L2), at(L1), T).
Step 1: Rule learning

**Observations (traces)**

holdsAt(at(loc1), 0).
do(pickup, 0).

holdsAt(at(loc1), 1).
holdsAt(holdingObj, 1).
do(move(loc1, loc3), 1).

holdsAt(at(loc3), 2).
holdsAt(holdingObj, 2).
do(move(loc3, loc5), 2).

holdsAt(at(loc5), 3).
holdsAt(holdingObj, 3).
do(putdown, 3).

**Explanatory rules (hypothesis)**

succ(move(loc3, loc5), T) :-
    holdsAt(at(loc3), T),
    holdsAt(holdingObj, T).

**Learned rules result in new transitions in the domain model**
Many traces, many hypotheses

- Traces may exhibit inconsistent behaviour

Maximum likelihood hypotheses has greatest probability of explaining observations

```prolog
holdsAt(at(loc1), 0).
do(pickup, 0).
holdsAt(at(loc1), 1).
holdsAt(holdingObject, 1).
do(move(loc1, loc3), 1).
holdsAt(at(loc3), 2).
holdsAt(holdingObject, 2).
do(move(loc3, loc5), 2).
holdsAt(at(loc5), 3).
holdsAt(holdingObject, 3).
do(putdown, 3).
succeeds(move(loc3, loc5), T) :-
    holdsAt(at(loc3), T),
    holdsAt(holdingObject, T).
```
### Step 2: Probability estimation

**Probability of a hypothesis h**

\[ P_0^\theta(h) = \prod_{a \in h} \theta_a \prod_{a \in A \setminus h} (1 - \theta_a) \]

<table>
<thead>
<tr>
<th>Rule 1</th>
<th>Rule 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explains trace – increase probability</td>
<td>Does not explain trace – decrease probability</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>holdsAt(holdingObject, T)</th>
<th>011</th>
<th>021</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>holdsAt(at(loc3), T)</td>
<td>012</td>
<td>022</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Step 2: Probability estimation

Minimise (by gradient descent)

\[ \text{MSE of current estimates } \theta \]

\[ \text{MSE}(\bar{\theta}) = \frac{1}{|X|} \sum_{i} (1 - P^\theta(x_i|\Delta \cup \chi_i))^2 \]

i.e. maximise prob. of hyp. predicting observations

Predictive ratio for observation \( x \)

\[ P^\theta(x_i|\Delta \cup \chi_i) = \frac{\sum_{\{h \in \Delta, h \cup \chi_i = x_i\}} P^\theta_0(h)}{\sum_{\{h \in \Delta\}} P^\theta_0(h)} \]
Applying learned rules

r1: 0.7 : succeeds(pickup, T).
r2: 0.9 : succeeds(move(L1, L2), T) :-
      holdsAt(at(L1), T),
      connected(L1, L2),
      L2 != loc3.
r3: 0.9 : succeeds(putdown, T) :-
      not happened(move(loc2, loc3), T-2).
r4: 0.1 : succeeds(putdown, T) :-
      happened(move(loc2, loc3), T-2).

Rule probabilities calculated from condition probabilities

Learned rules result in new states and transitions in the domain model – with probabilities
Updated factory floor model

1. pickup 0.7
2. move(4) 0.9
3. move(4) 0.9
4. move(5) 0.9
5. putdown 0.1
6. putdown 0.9
New factory floor plan

1. pickup
2. move(4)
3. move(4)
4. move(5)
5a. putdown
5b. putdown
6.
Experience

Robot navigation: global failure rate reduced from 30% to 10%

Human-readable explanations

\[
\text{succeeds} \left( \text{move}(L_1, L_2), T \right) : - \\
\text{holdsAt}(\text{at}(L_1), T), \\
\text{connected}(L_1, L_2), \\
L_2 \neq \text{loc3}.
\]
Challenges

- High complexity of ILP
  - Limit rule length with mode declarations
- Improve tool support (GD, integration)
- Scope for adaptation is limited by set of actions and sensed conditions
  - Cannot learn rules based on conditions not present in traces
- Opportunity for starting from a minimal model
  - Exploration of environment
Summary

• Behavioural model revision using ILP
  – Traces gathered from plan execution
  – Missing states/transitions
  – Estimated probabilities, find maximum likelihood hypothesis
  – Mitigate inaccuracy and incompleteness (uncertainty) in model

• Revised model remains human-readable
Thanks, questions?