LLWA-2003

Learning for User Adaptive Systems: Likely Pitfalls and Daring Rescue

(M. E. Müller, Univ. Augsburg)

LLWA-2003

Learning for User Adaptive Systems: Likely Pitfalls and Desparate Rescue

(M. E. Müller, Univ. Augsburg)

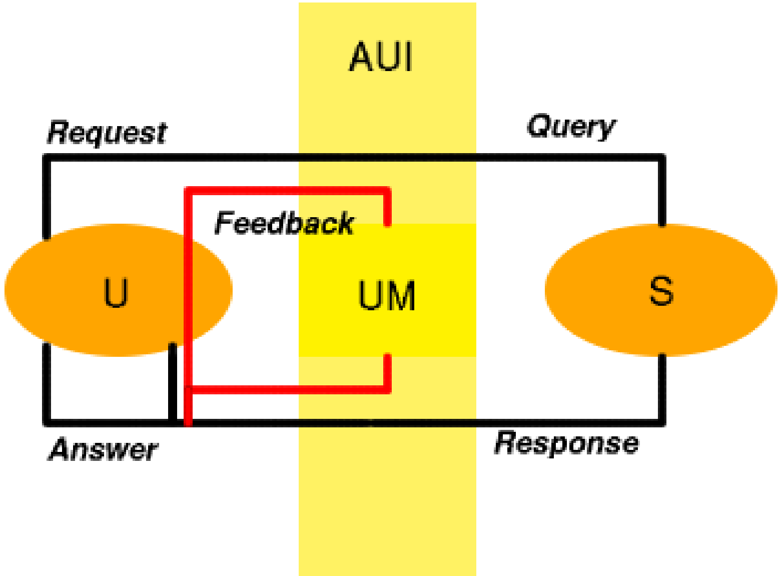
Learning to behave adaptively

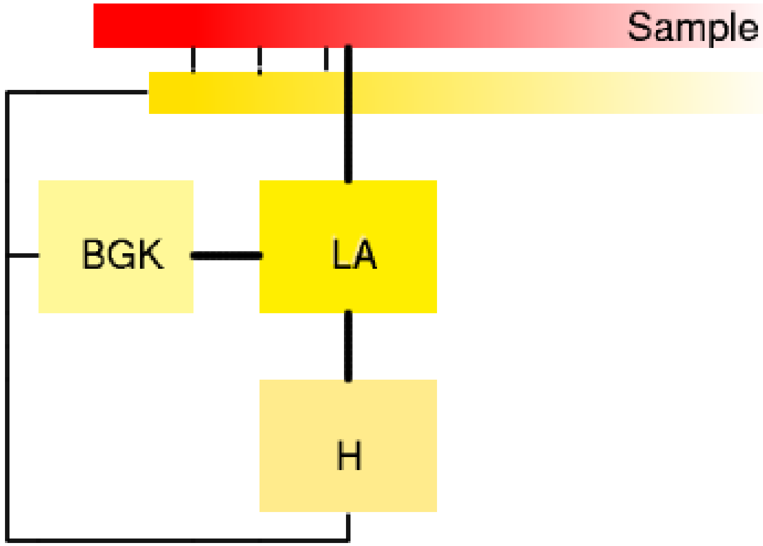
- Adaptive user interfaces adapt themselves to the user by reasoning about the user and refining their internal model of the user's needs.
- By observing examples from a sample, a learning algorithm tries to induce a hypothesis which approximates the teacher signal and allows to explain unknown cases.

[...] user modeling appears to be a prime candidate for straightforward application of standard machine learning techniques.

[Webb *et al.*2001]

Machine Learning and User Adaptive Systems





Problems and Likely Pitfalls

Demands on adaptive systems:

- rapid adaption
- high accuracy
- unobtrusive
- scrutable
- simple

$$
\begin{tikzcd}
 & & \mathbf{i}_u \arrow[dl, "?"'] \arrow[dd, "?\cdot\mathbf{o}"'] \\
\mathsf{KB} \arrow[d, "?"'] & \mathit{interaction} \arrow[dr, "\mathbf{o}"] & \\
\mathbf{A} \arrow[d] & \mathbf{f} \arrow[l] & F(m, ?\cdot\mathbf{o}) \arrow[l] \\
\mathbf{M}_u \sim \mathbf{i}_u & &
\end{tikzcd}
$$

See 11.

The problem ...

The problem with adaptive systems is, that the system tries to adapt itself to what it thinks the user might approve.

- *What* does the user *want* or *need*? — *How* can I help?
- *What* does the user *do*?
  - Anticipate what the user will do next!
  - Anticipate what the user will do in a similar situation!
- *Why* does he do what he does?

$\rightarrow$ What the user *does* $\neq$ What the user *wants/needs*

At first blush, user modeling appears to be a prime candidate for straightforward application of standard machine learning techniques.

The trapdoors ...

1. Requirements
   - Data availability → *sample size*
   - Algorithmic complexity → *adaptation latency*
2. Restrictions
   - Data (Domain) complexity → *domain description*
   - Hypothesis expressiveness → *accuracy and triviality*
3. Reliability
   - Data quality → *noise*
   - Learning convergence → *adaptation accuracy*

Example: Emotion Recognition from speech, [Murray and Arnott1993, André and Müller2003]

Example: A motivational Tutoring system

$$
\begin{tikzcd}
 & & \textsc{Student} \arrow[dl, "?"'] \arrow[dr, "?"] & \\
\Sigma_{\text{Domain}} \arrow[d] & \mathit{TopicChoice} \arrow[dr, "\mathbf{o}"] & & \mathit{Test} \arrow[dl, "\mathbf{o}"'] \\
\mathbf{A}_{\text{NBC}} \arrow[d] & \mathbf{f} \arrow[l] & F(m, ? \cdot \mathbf{o}) \arrow[l] & \\
\mathbf{M}_u \arrow[r] \arrow[dr] & \mathit{Motivate} & & \\
 & \mathit{Recommend} & &
\end{tikzcd}
$$

[Stern and Woolf2000, André and Müller2003]

Example: Learning to play...

... `TRON`, [Sklar2002].

- Fully connected 3–layer MLP with standard BP.

  Input: 8 obstacle sensors, Output: *Right*, *Straight*, *Left*

- Learning target: Play performance similar to human player

  Performance measure: *winrate* $= \frac{wins_C}{Games} \approx \frac{wins_u}{Games}$

- Teacher signal: Player log

  Clones, *Peers* and *Instructors* simulate the target player against an opponent.

- Result: for *clones* rather poor

$\longrightarrow$ The clone learns a myopic method to mimmick a player's moves.

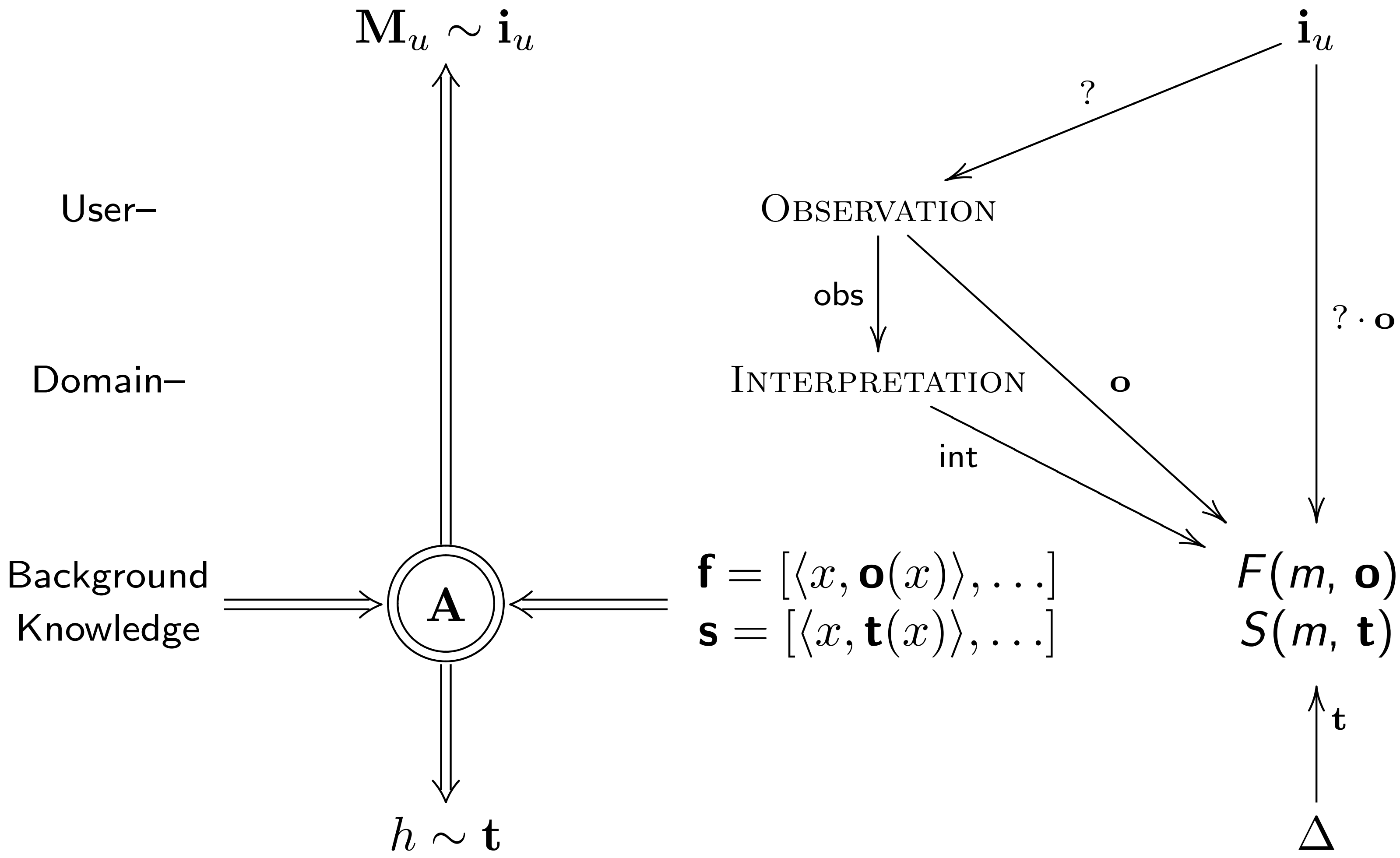
$\longrightarrow$ performance, (*win rate*) also depends on the strategy and deeper knowledge of game situation and game run.

Example: Learning to play `Tron`

$$
\begin{tikzcd}
 & & \textsc{Player} \arrow[d, "?"] \\
\emptyset \arrow[d] & & \mathit{Move} \arrow[d, "\mathbf{o}"'] \\
\mathbf{A}_{\textsc{MlpBp}} \arrow[d] & \mathbf{f} \arrow[l] & F(m, ? \cdot \mathbf{o}) \arrow[l] \\
\mathbf{M}_u \arrow[r] & \mathit{Play} &
\end{tikzcd}
$$

[Sklar2002]

$\textsc{Player}$ competence $\neq$ Play behavior



Back to slide 5

Abandon all hope?

- Where do causal relationships come from?
  → Requires empirical studies!
- Where do net topologies come from?
  → Observables? Complexity? Recurrent networks?
- Where do similarity measures come from?
  → are *near* neccessarily *similar* concepts?
- Is there truly no noise?
  → Is the user goal–oriented, rational *and* flawless?
- Where do plans come from?
  → From clicks to intentions and goals?

✓ many promising solutions and approaches!

A checklist...

- Do we want to learn a high–level model of the user's needs or do we rather want to predict a user's behavior?
  Do we have appropriate data to achieve that goal?
- Is the data noisy? How many examples do we have?
- What kind of domain restrictions are implied by biases and are they satisfied?
- Do we need background knowledge and if so, do we have the right one?
- Is a quick training needed? Is offline learning sufficient?
- Is quick learning required or shall we rather wait a bit longer instead of a quick wild guess that might go wrong?
- Shall the inferred hypothesis be scrutable to the user?

Summary

- There is no one–size–fits–all solution!
- There is no solution out of the box!

But there are:

- many machine learners, data miners and knowledge discoverers
- many knowledge managers and workflow analysts
- many psychologists and HCI–specialists

... right here, right now.

Never surrender!

The big advantage of the domain of user adaptive systems, is that we usually can draw *a rough picture of what a user may want*—and we are also able to *anticipate* to a certain degree *what might go wrong*.

- If there are real restrictions, we have a good reason for bias.
- If there is a domain theory, we have good background knowledge.
- If we have an idea of what can go wrong, we can denoisify samples.

– Huge, monolithic systems & ad-hoc solutions for off–topic add–ons

We might get a (hybrid) system, that avoids noise, learns what the user needs and anticipates a goal, adapts to the user and thus makes interaction action easier — if we keep it small, simple and sound.

References

[André and Müller2003] Elisabeth André and Martin E. Müller. Learning affective behavior. In Constantine Stephanidis, editor, *Universal Access in HCI. Proc. 10th Intl. Conf. on HCI*, volume 4. Lawrence Erlbaum Associates, 2003.

[Billsus and Pazzani1997] Daniel Billsus and Michael Pazzani. Learning probabilistic user models. In Anthony Jameson, Cécile Paris, and Carlo Tasso, editors, *User Modeling: Sixth International Conference, UM97*. Springer Wien New York, Vienna, New York, 1997. Workshop Paper.

[Murray and Arnott1993] I. R. Murray and J. L. Arnott. Toward the simulation of emotion in synthetic speech: A review of literature on human vocal emotion. *Journal of Acoustical Society of America*, 93(2), 1993.

[Sklar2002] Elizabeth Sklar. It takes a virtual village: Towards an automated interactive agency. In *Proc. AAAI Fall Symposium on Personalized Agents*, 2002.

[Stern and Woolf2000] Mia K. Stern and Beverly Park Woolf. Adaptive content in an online lecture system. In Peter Brusilovsky, Oliviero Stock, and Carlo Strappavara, editors, *Adaptive Hypermedia and Adaptive Web–Based Systems: AH-2000*, number 1892 in LNCS. Springer, 2000.

[Webb *et al.*2001] Geoffrey I. Webb, Michael J. Pazzani, and Daniel Billsus. Machine learning for user modeling. *User Modeling and User–Adapted Interaction*, (11):19–29, 2001.