Toward Simulated Students for Reinforcement Learning-Driven Tutorial Planning in GIFT

Jonathan Rowe Bob Pokorny Benjamin Goldberg

Bradford Mott James Lester

North Carolina State University

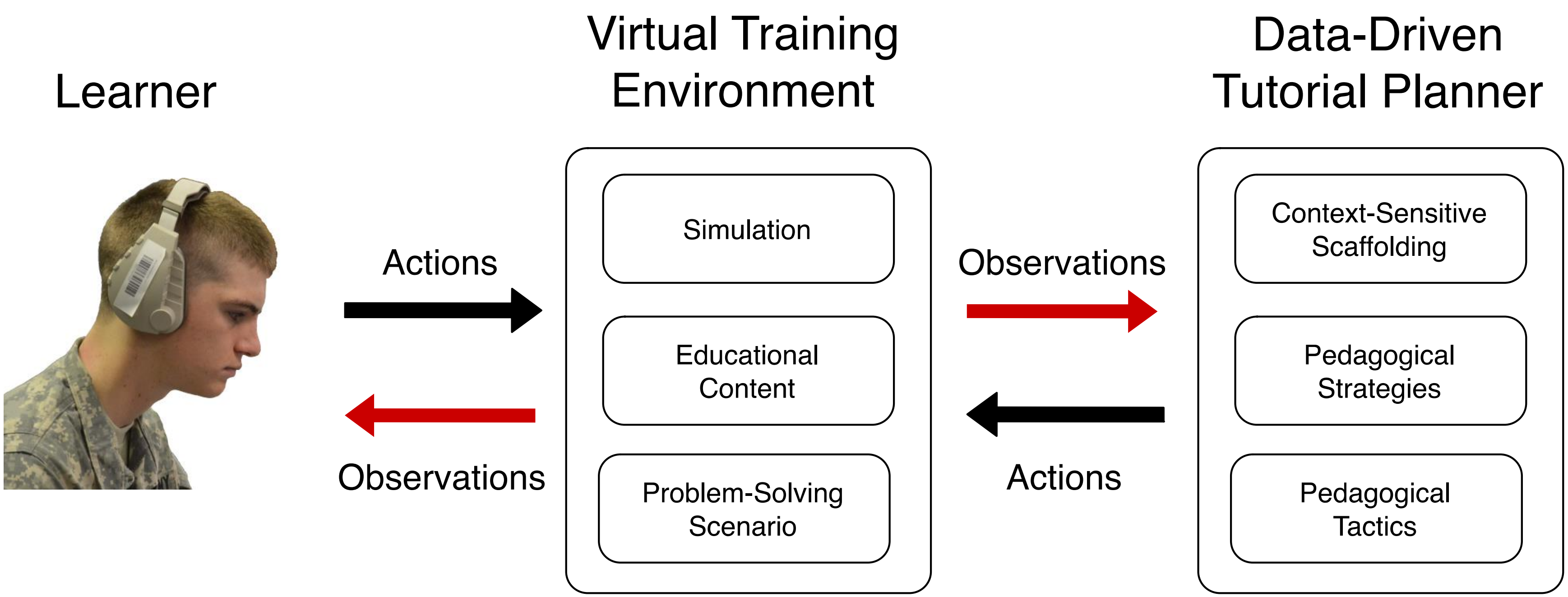
the intelliMEDIA group

Tutorial Planning

Tutorial Planning

- Hints
- Feedback
- Scenario adaptations
- Embedded assessments
- Remedial instruction

Data-Driven Tutorial Planning



Data-Driven Tutorial Planning

Simulated Student

Actions

Observations

Virtual Training Environment

- Simulation
- Educational Content
- Problem-Solving Scenario

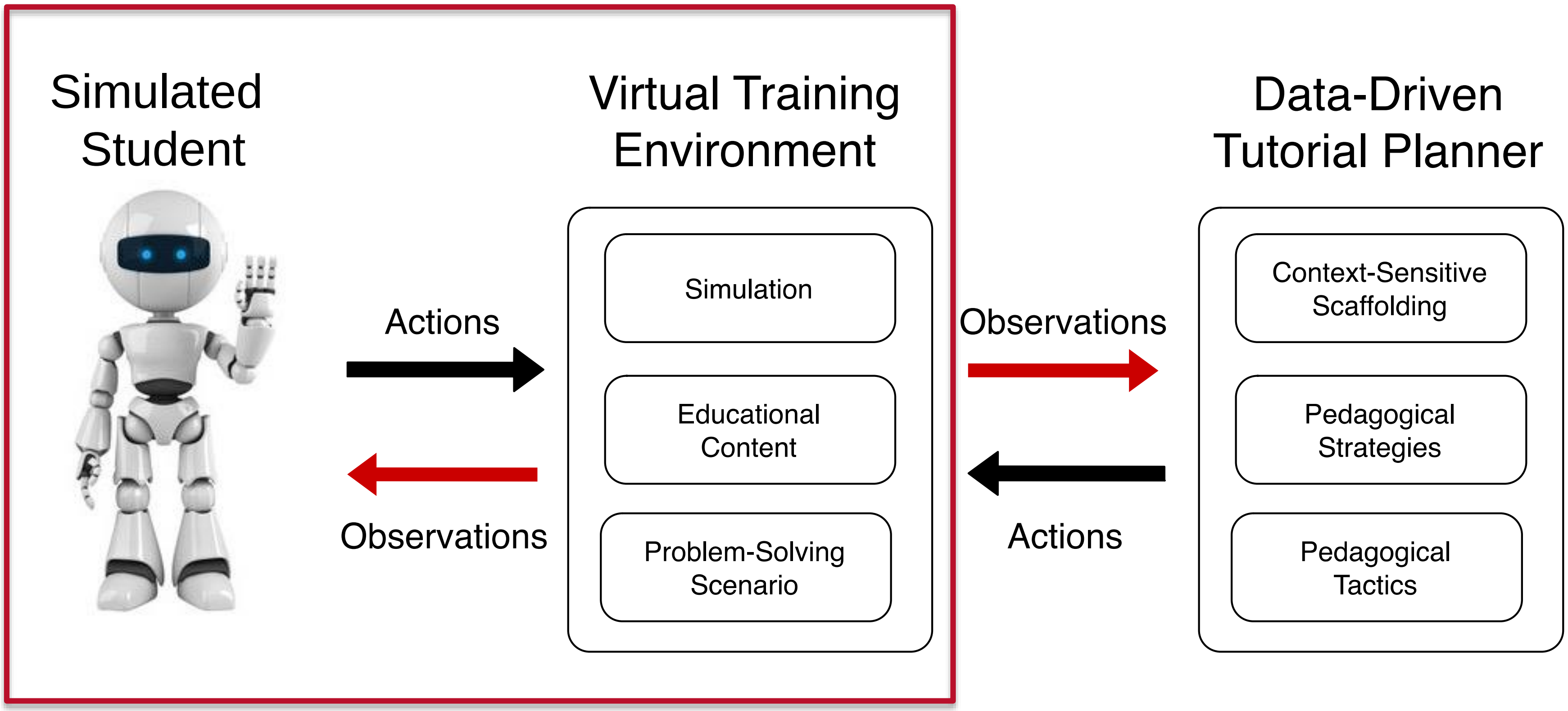
Observations

Actions

Data-Driven Tutorial Planner

- Context-Sensitive Scaffolding
- Pedagogical Strategies
- Pedagogical Tactics

Data-Driven Tutorial Planning



Research Question

How can we leverage *simulated students* to generate synthetic data for training generalized tutorial planners in GIFT?

Outline

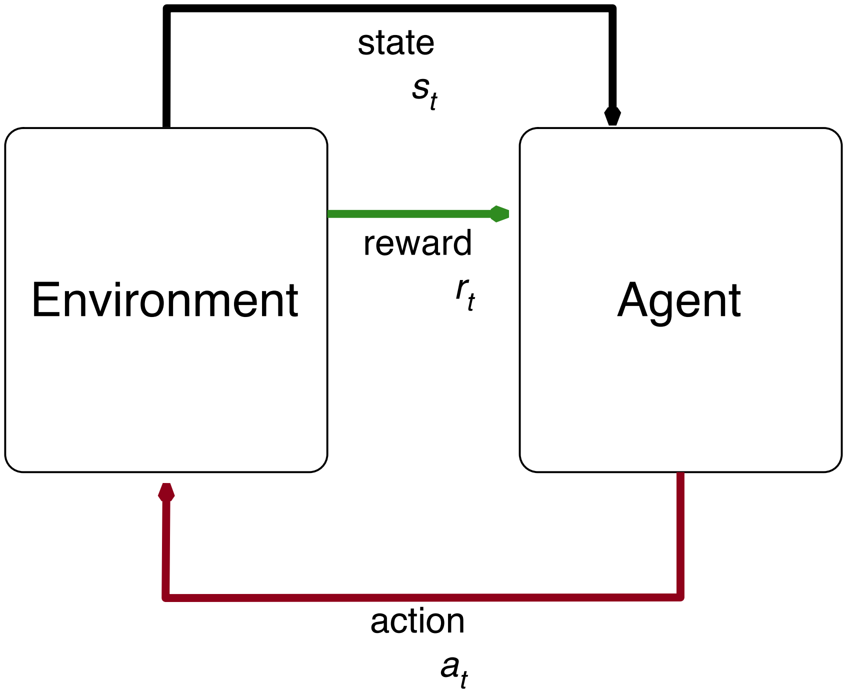
- Reinforcement Learning-Based Tutorial Planning
- Design Issues for Simulated Students
- Implementing Simulated Students for COIN Training
- Conclusions and Future Directions

Outline

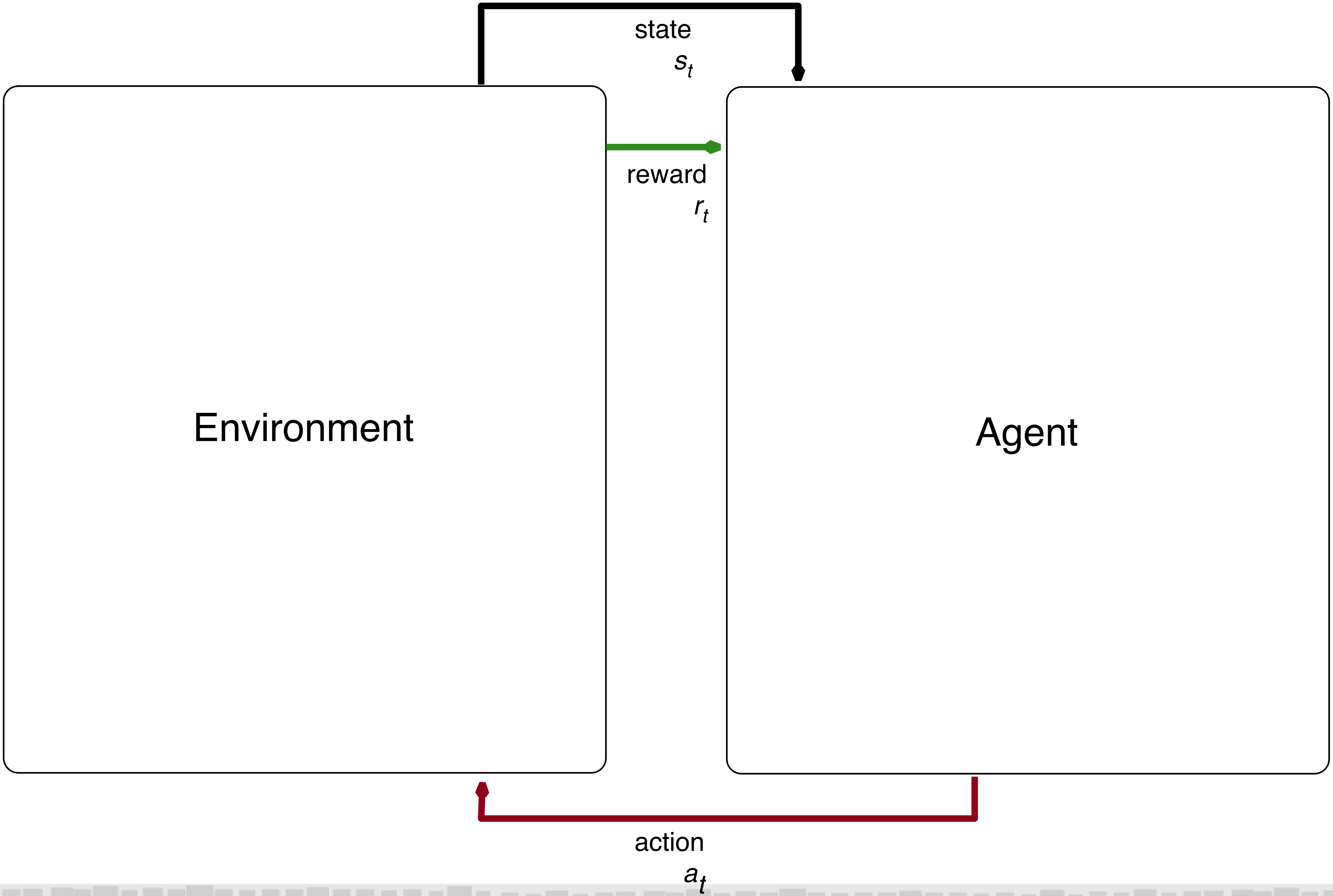
- **Reinforcement Learning-Based Tutorial Planning**
- Design Issues for Simulated Students
- Implementing Simulated Students for COIN Training
- Conclusions and Future Directions

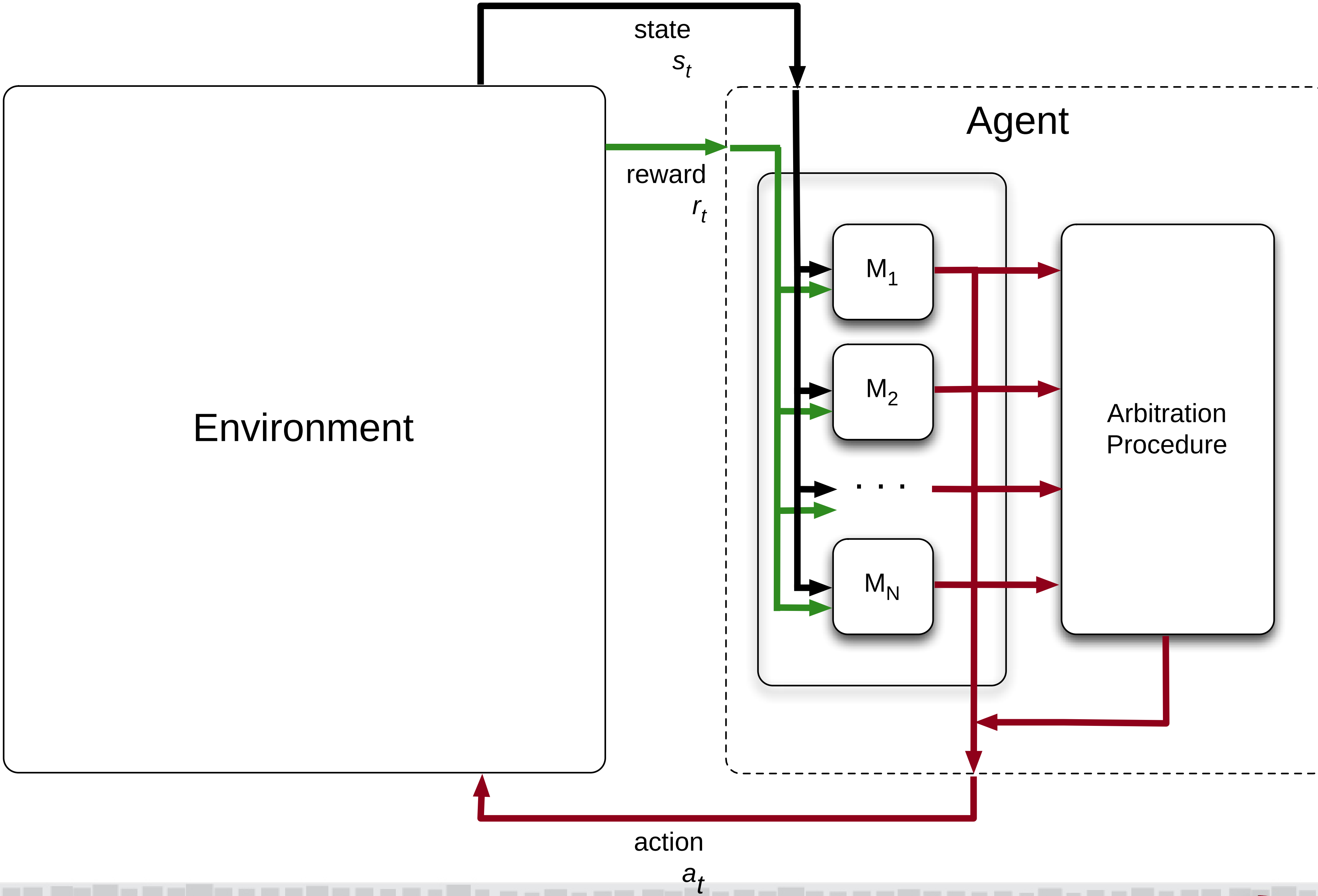
Reinforcement Learning

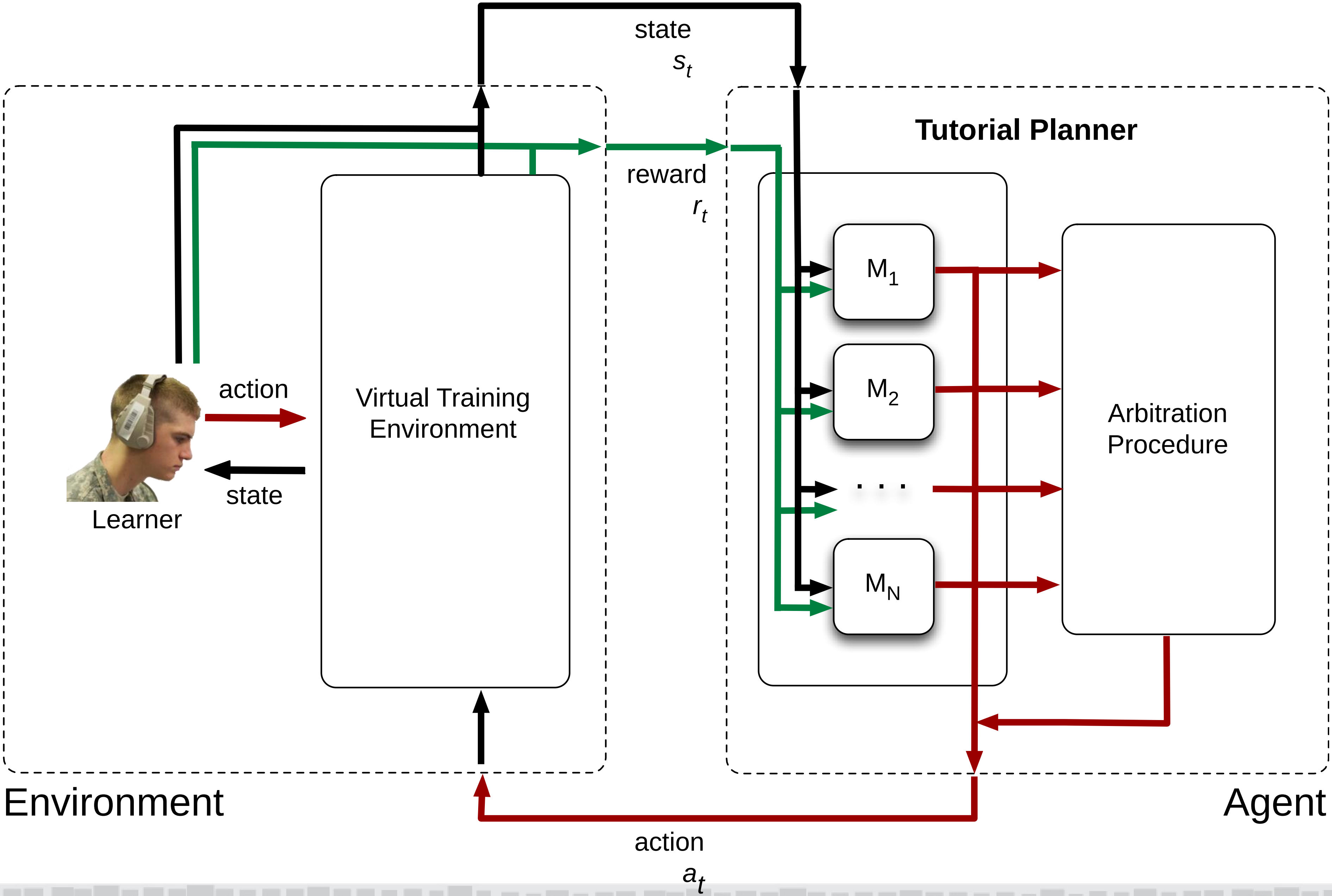
- **Problem:** Devise software agent that learns how to behave in order to maximize numerical reward
- No external supervision
- Delayed rewards



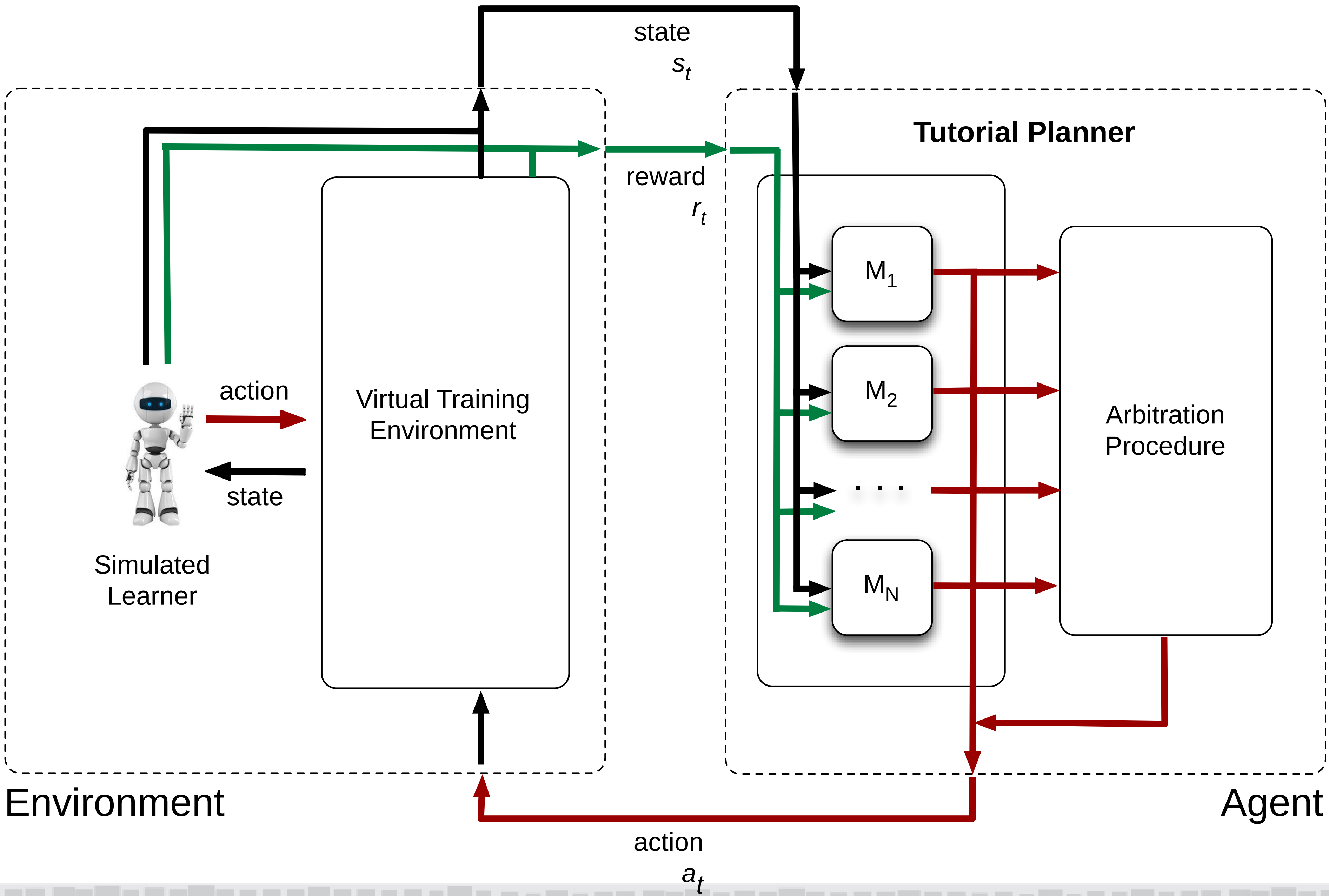
Adapted from Sutton & Barto (1998)

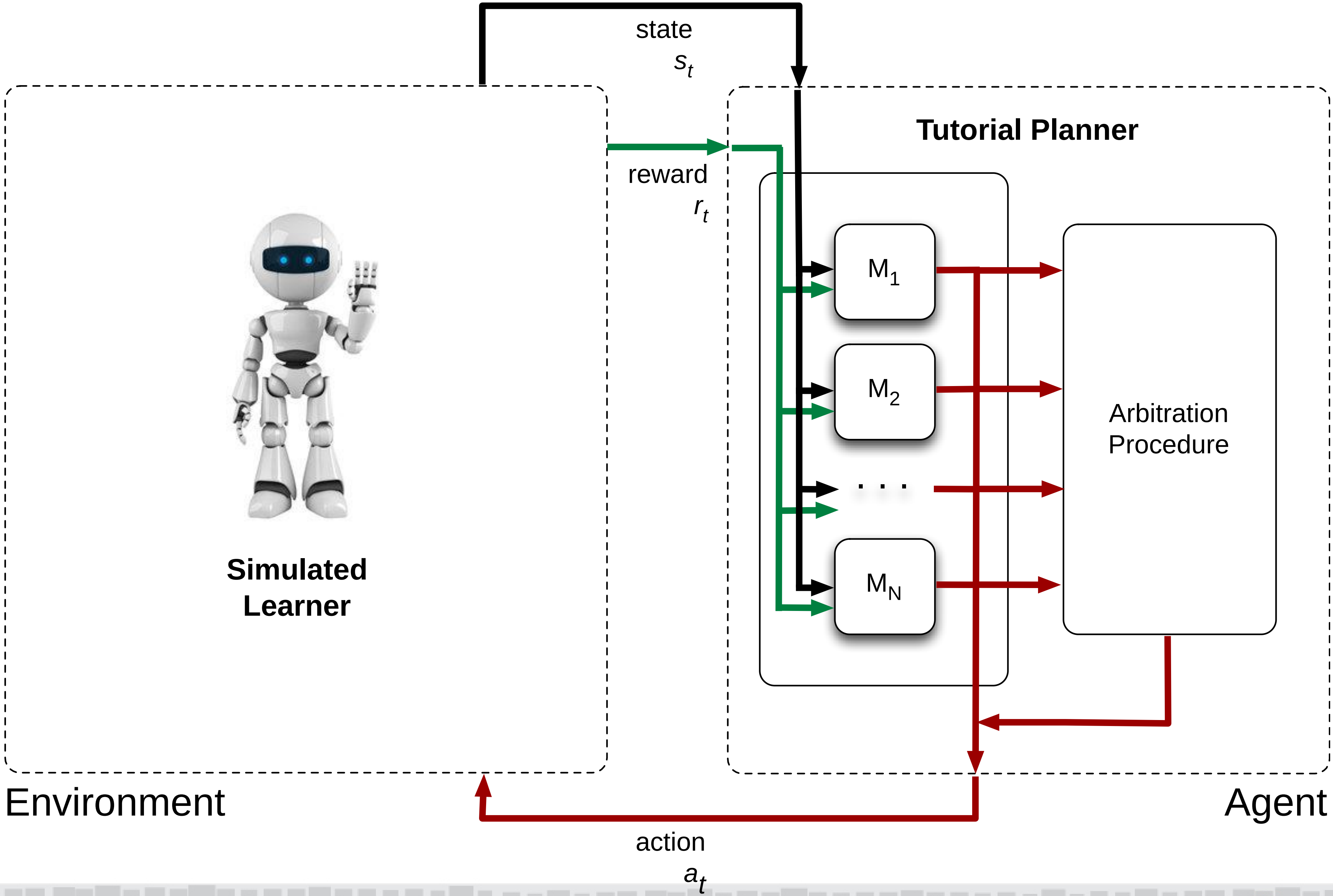

state
s_t
reward
r_t
Environment
Agent
action
a_t


state
s_t
reward
r_t
Agent
Environment
M_1
M_2
. . .
M_N
Arbitration Procedure
action
a_t



state
s_t
Tutorial Planner
reward
r_t
action
Virtual Training
Environment
state
Learner
M_1
M_2
. . .
M_N
Arbitration
Procedure
Environment
Agent
action
a_t


state
s_t
Tutorial Planner
reward
r_t
action
state
Virtual Training Environment
Simulated Learner
M_1
M_2
. . .
M_N
Arbitration Procedure
Environment
Agent
action
a_t


state
s_t
reward
r_t
Tutorial Planner
M_1
M_2
. . .
M_N
Arbitration Procedure
Simulated Learner
Environment
Agent
action
a_t

Markov Decision Processes

- Reinforcement learning problems are often modeled as Markov decision processes (MDP)

- Defined by a tuple $\mathcal{M} = (\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$
  - Environment state set $\mathcal{S}$
  - Action set $\mathcal{A}$
  - State transition model $\mathcal{P}: \{\mathcal{S} \times \mathcal{A} \times \mathcal{S}\} \rightarrow [0, 1]$
  - Reward model $\mathcal{R}: \{\mathcal{S} \times \mathcal{A} \times \mathcal{S}\} \rightarrow \Re$

- Solution is optimal policy $\pi^*: \{\mathcal{S}\} \rightarrow \mathcal{A}$

Policy Learning

- Online learning
  - Interleave data collection and model operation
  - Temporal-difference methods
  - Works well with simulation-generated training data
- Offline learning
  - Separate data collection and model operation
  - Certainty equivalent learning (Kaelbling, Littman & Moore 1996)
  - Approximate state-transition model and reward model using collected corpus

Outline

- Reinforcement Learning-Based Tutorial Planning
- Design Issues for Simulated Students
- Implementing Simulated Students for COIN Training
- Conclusions and Future Directions

Outline

- Reinforcement Learning-Based Tutorial Planning
- **Design Issues for Simulated Students**
- Implementing Simulated Students for COIN Training
- Conclusions and Future Directions

Simulated Students Overview

Definition:

"… computer systems that simulate human students as they learn educationally significant subject matter … "

(VanLehn, Ohlsson, & Nason, 1993)

Applications

1. Provide teachers with practice opportunities
2. Serve as co-learners for human students
3. Conduct formative evaluations of learning materials

Simulated Students Overview

Definition:

"… computer systems that simulate human students as they learn educationally significant subject matter … "

(VanLehn, Ohlsson, & Nason, 1993)

Applications

1. Provide teachers with practice opportunities
2. Serve as co-learners for human students
3. Conduct formative evaluations of learning materials
4. Enable ITS authoring systems using authoring-by-tutoring (Matsuda, Cohen, & Koedinger, 2014)
5. Generate synthetic data for training data-driven intelligent tutors (Beck, Woolf, & Beal, 2000; Wang et al., 2017)

Simulated Students Overview

Definition:

"… computer systems that simulate human students as they learn educationally significant subject matter … "

(VanLehn, Ohlsson, & Nason, 1993)

Applications

1. Provide teachers with practice opportunities
2. Serve as co-learners for human students
3. Conduct formative evaluations of learning materials
4. Enable ITS authoring systems using authoring-by-tutoring (Matsuda, Cohen, & Koedinger, 2014)
5. **Generate synthetic data for training data-driven intelligent tutors** (Beck, Woolf, & Beal, 2000; Wang et al., 2017)

Related Work

- AIED Workshop Series on Simulated Learners (AIED-2013, AIED-2015)
- SimStudent (Matsuda, Cohen, & Koedinger, 2014)
- Simulated students in RL-based tutoring systems (Beck, Woolf, & Beal, 2000; Folsom-Kovarik, Sukthankar, & Schatz, 2013; Wang et al., 2017)
- Simulated users in spoken dialogue systems (Schatzmann, Weilhammer, & Young, 2006; Young et al., 2013)

Design Dimensions for Simulated Students

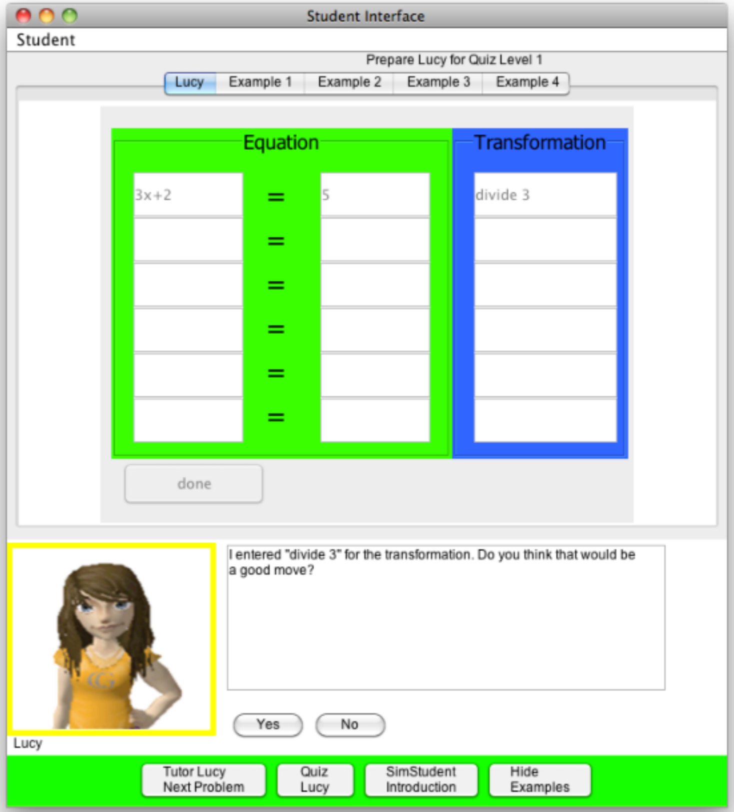
- Representational Granularity
- Computational Framework
- Model Complexity
- Learning Process
- Model Validity

Design Dimensions for Simulated Students

- **Representational Granularity**
- Computational Framework
- Model Complexity
- Learning Process
- Model Validity

Representational Granularity

- Varying levels of temporal granularity
- Fine-grained representation
  - SimStudent (Matsuda, Cohen, & Koedinger, 2014)
- Coarse-grained representation
  - SimGrad (LeLei & McCalla, 2015)

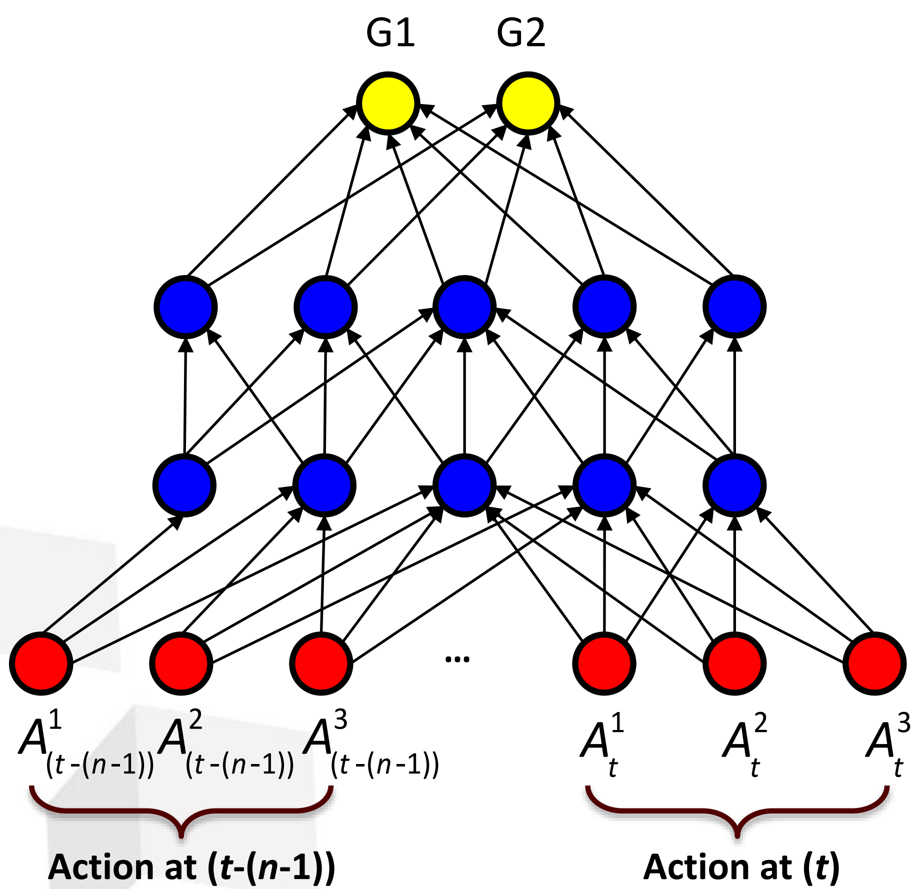


(Matsuda, Cohen, & Koedinger, 2014)

Design Dimensions for Simulated Students

- Representational Granularity
- **Computational Framework**
- Model Complexity
- Learning Process
- Model Validity

Computational Framework



- Expert systems
- Closed-form expressions
  - Weighted sum (Frost & McCalla, 2015)
  - Item response theory (Hernando, Guzman, & Conejo, 2013)
- Machine learned models
  - Linear regression (Beck, Woolf, & Beal, 2000)
  - Hidden Markov models (Pardos & Yudelson, 2013)
  - LSTM neural networks (Wang et al., 2017)

Design Dimensions for Simulated Students

- Representational Granularity
- Computational Framework
- **Model Complexity**
- Learning Process
- Model Validity

Model Complexity

- Number of parameters
- Linear vs non-linear functions
- Tabular vs algorithmic simulations (VanLehn, Ohlsson, & Nason, 1993)
  - Tabular models are efficient and easily authored
  - Algorithmic models generalize to novel situations
- Run-time efficiency

Design Dimensions for Simulated Students

- Representational Granularity
- Computational Framework
- Model Complexity
- **Learning Process**
- Model Validity

Learning Process

- Cognitive simulations
  - Problem-solving behavior (Matsuda, Cohen, & Koedinger, 2014)
  - Academic performance (LeLei & McCalla, 2015)
- Affective simulations
  - Emotion regulation (Sabourin et al., 2013)
- Social simulations
  - Peer-to-peer learning (Frost & McCalla, 2013)

Design Dimensions for Simulated Students

- Representational Granularity
- Computational Framework
- Model Complexity
- Learning Process
- **Model Validity**

Model Validity

- Not all simulated students are validated
  - Designer intuition
  - Theoretically grounded
  - Empirically derived
- Designer bias
- Population-dependent aspects of learning are difficult to estimate

Outline

- Reinforcement Learning-Based Tutorial Planning
- Design Issues for Simulated Students
- Implementing Simulated Students for COIN Training
- Conclusions and Future Directions

Outline

- Reinforcement Learning-Based Tutorial Planning
- Design Issues for Simulated Students
- **Implementing Simulated Students for COIN Training**
- Conclusions and Future Directions

Tutorial Planning for Counterinsurgency Training

Adaptive Hypermedia

UrbanSim Primer

Simulation-Based Training

UrbanSim
(McAlinden, Pynadath, & Hill, 2014)

COIN Training Testbed

UrbanSim Primer

- Adaptive hypermedia learning environment
- Range of doctrinal concepts of COIN
  - Population support
  - Clear-Hold-Build
  - Intelligence gathering
- Preliminary instruction on UrbanSim usage

COIN Training Testbed

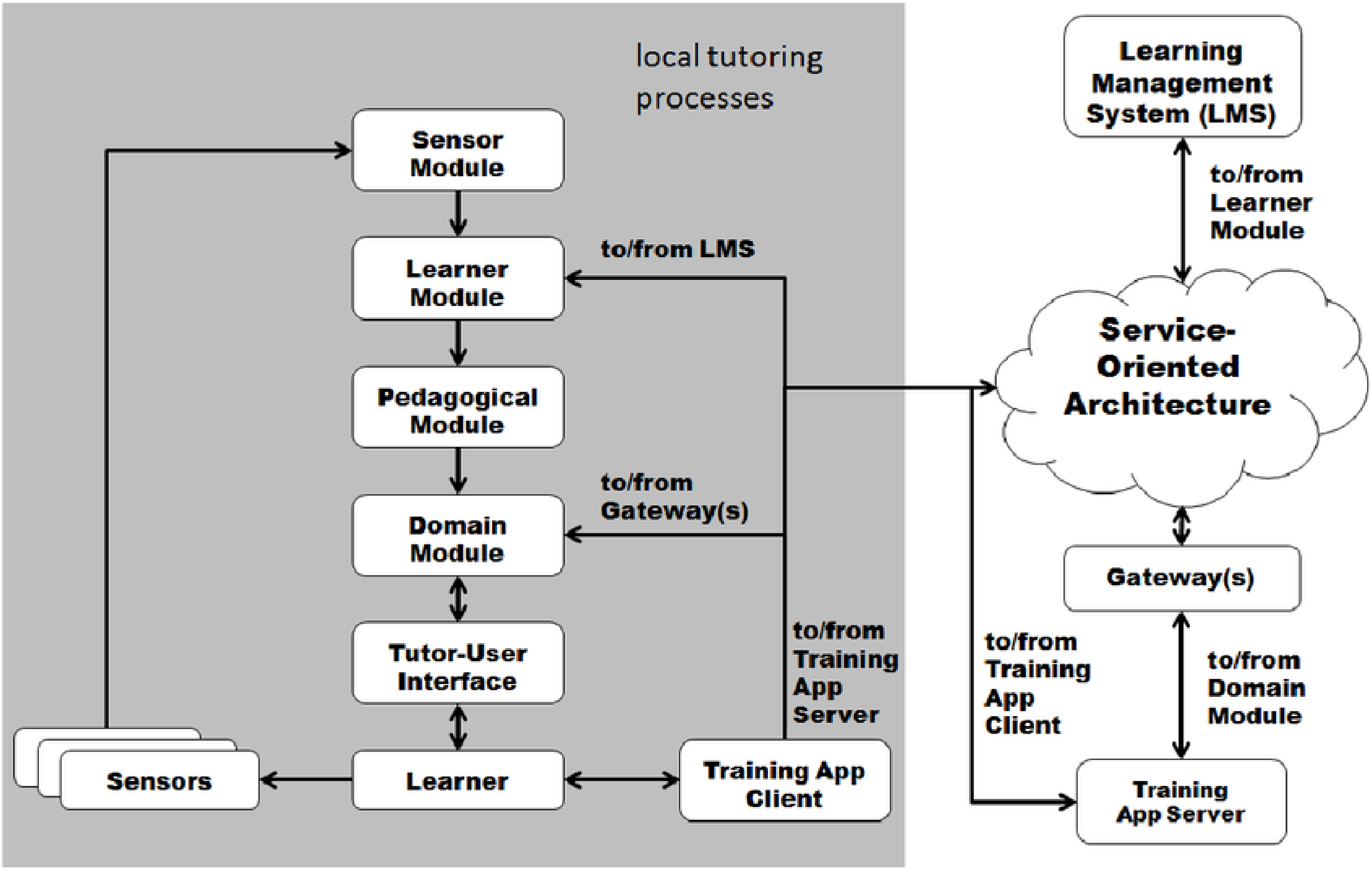
- Simulation-based learning environment
- **Role:** Learner is battalion commander
- **Objective:** Maximize civilian support for host nation government
- PsychSim social simulation engine

UrbanSim

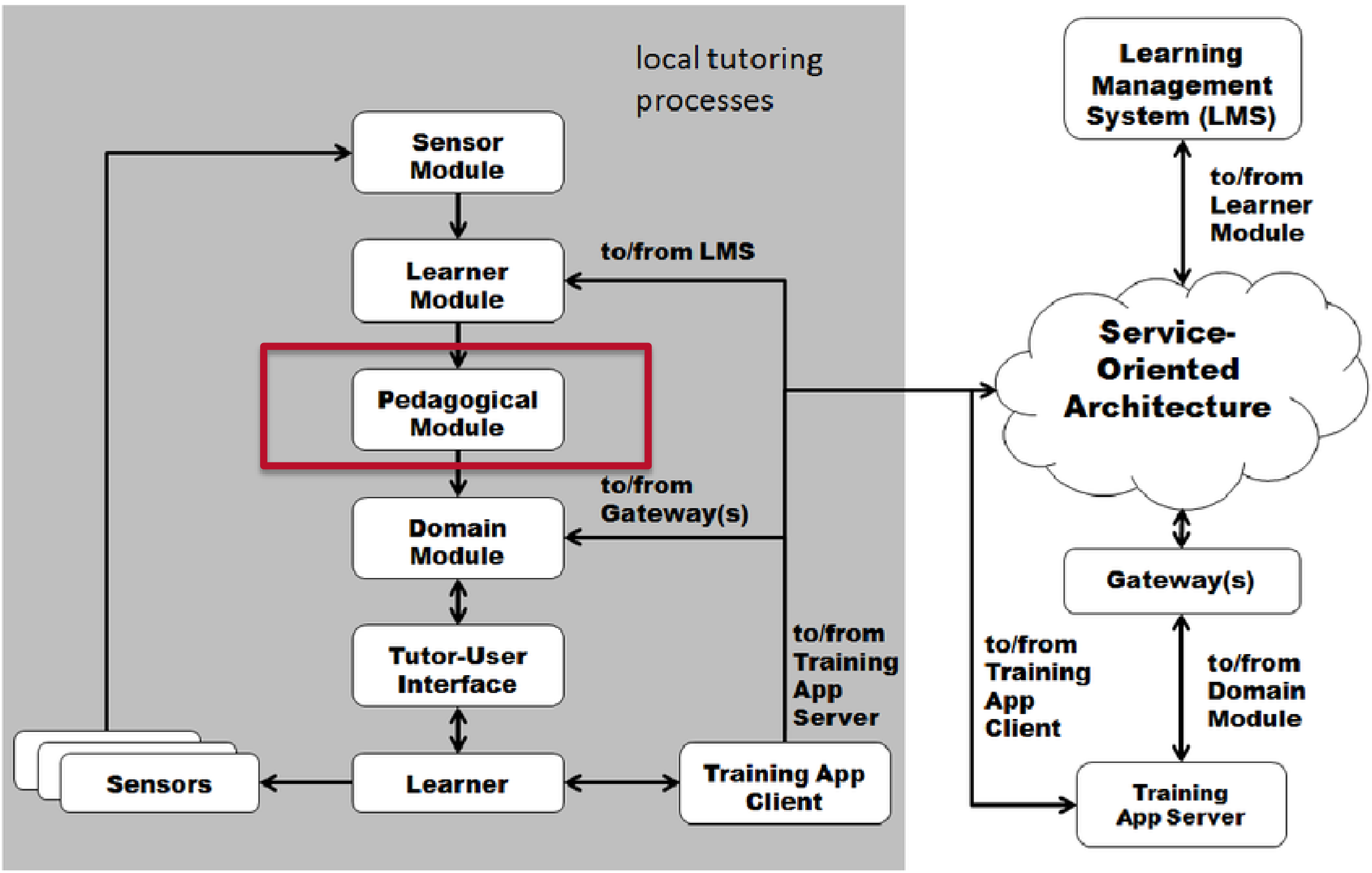
Generalized Instructional Strategies for COIN Training

- High-level instructional strategies
  - Single-topic coaching
  - Multi-concept review
  - Feedback on unproductive learning behaviors
- ICAP-inspired implementation strategies (Chi, 2009)
  - Constructive
  - Active
  - Passive

GIFT Pedagogical Module



GIFT Pedagogical Module



Toward Simulated Students for COIN Training

- *Bipartite model* of simulated students
  - Student behavior
  - Learning outcomes
- Tabular joint probability distribution
  - Values estimated from pilot study data
  - Data sparsity challenges
- Granularity
  - *UrbanSim Primer*: One lesson
  - *UrbanSim*: One turn of simulation

Toward Simulated Students for COIN Training

- Devise simulated student for each MDP
- Domain-independent state features
  - Student knowledge & traits
  - Task states
  - Pedagogical history
- Model student responses to pedagogical actions
- Rewards model student learning gains
- Population of simulated students

Outline

- Reinforcement Learning-Based Tutorial Planning
- Design Issues for Simulated Students
- Implementing Simulated Students for COIN Training
- Conclusions and Future Directions

Outline

- Reinforcement Learning-Based Tutorial Planning
- Design Issues for Simulated Students
- Implementing Simulated Students for COIN Training
- **Conclusions and Future Directions**

Conclusions

- Simulated students show promise for generating synthetic data to train data-driven tutorial planners
- Design of simulated students presents several questions:
  - Representational granularity
  - Computational framework
  - Model complexity
  - Target learning process
  - Model validity
- We are devising simulated students to support RL-based tutorial planning for COIN training in GIFT

Future Directions

- Conduct studies validating simulated students by comparing synthetic data with human student data
- Devise tools and workflows for incorporating tutorial planning policies induced from simulated students in GIFT
- Provide tools for non-expert users to work with simulated students, including creating, configuring, sharing, and refining simulated student models
- Conduct GIFT studies with Mechanical Turk populations to complement synthetic data from simulated students

Acknowledgments

- NCSU Army ROTC Program
- Institute for Creative Technologies
- Vanderbilt University
- Dignitas Technologies

Support provided by the U.S. Army Research Laboratory under cooperative agreement W911NF-15-2-0030.