U.S. Army Research, Development and Engineering Command
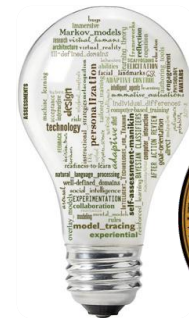
TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Semi-Supervised Classification of Realtime Physiological Sensor Datastreams for Student Affect Assessment in Intelligent Tutoring

Keith Brawner, Avelino Gonzalez, Robert Sottilare
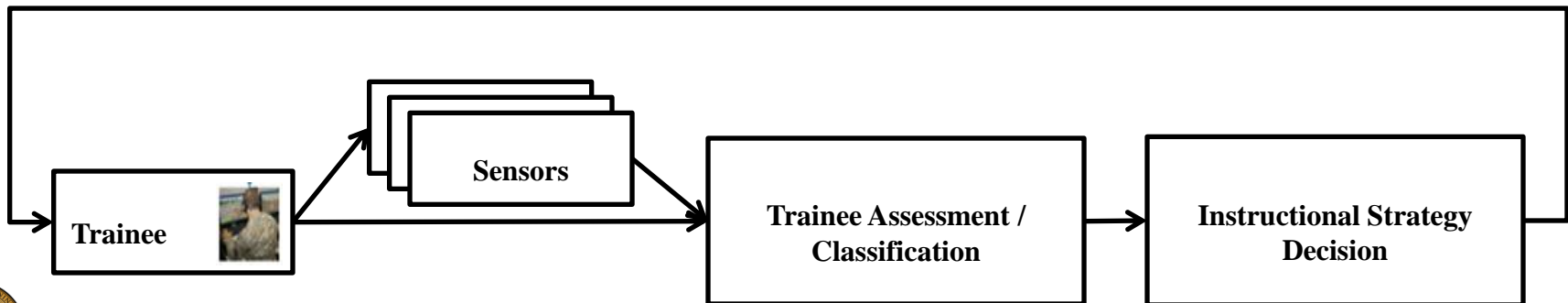
University of Central Florida
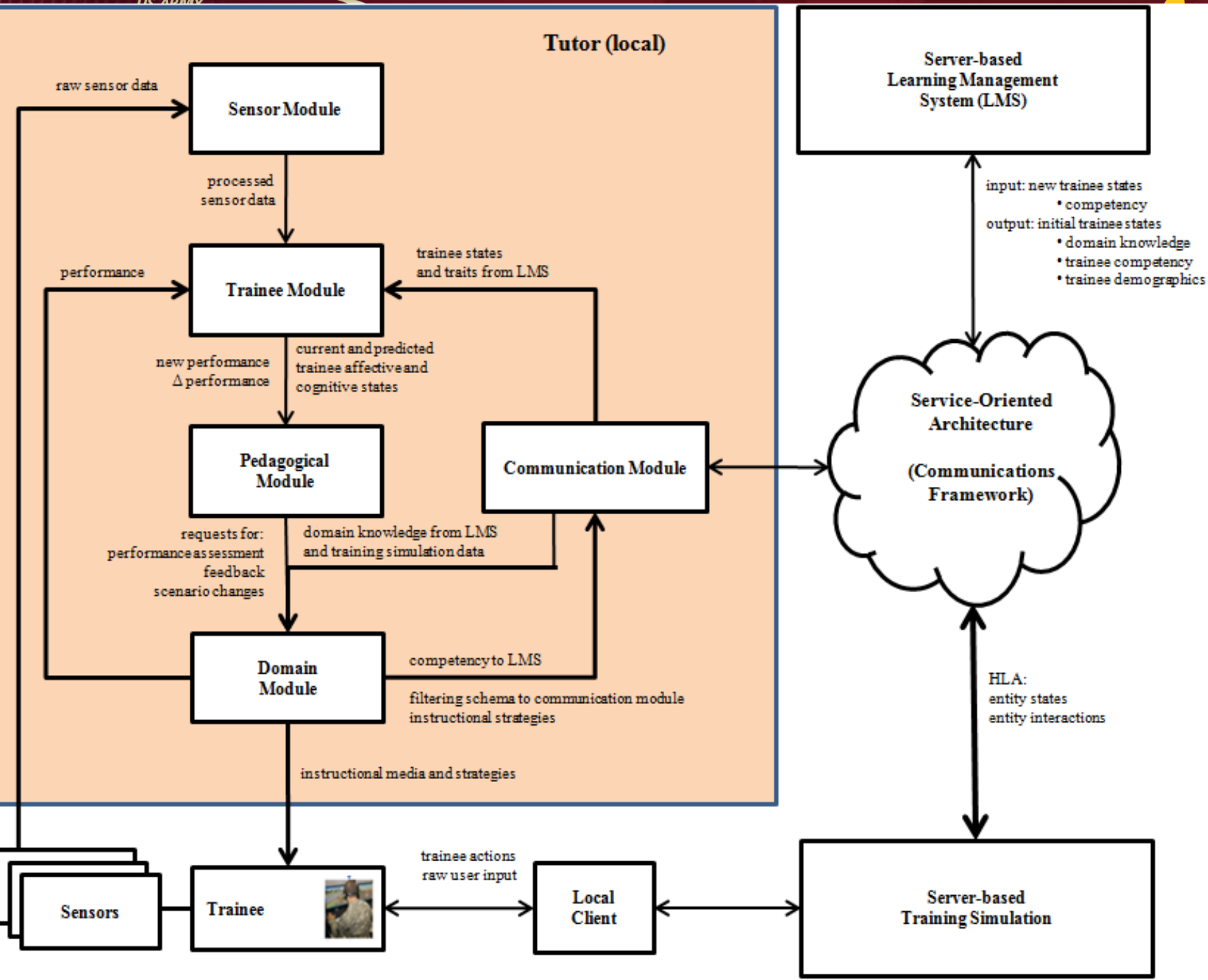
LITE Lab

- Bloom Human Tutoring
  - 2 sigma improvement
- "Is Adaptive Learning Effective" Literature review
  - 15 systems, average effect size .95
- Woolf ADL Keynote
  - Learning time reduced 30-50%
  - Training support reduced 70%
  - Operating cost reduced 92%

- Preaching to the choir…

- Student Actions

- Sensor Data

- Assessment and Classification

- Instructional Strategy Decision?

Tutor (local)

Sensor Module

raw sensor data

processed sensor data

Trainee Module

performance

trainee states and traits from LMS

new performance Δ performance

current and predicted trainee affective and cognitive states

Pedagogical Module

Communication Module

requests for:
performance assessment
feedback
scenario changes

domain knowledge from LMS and training simulation data

Domain Module

competency to LMS

filtering schema to communication module
instructional strategies

instructional media and strategies

Server-based Learning Management System (LMS)

input: new trainee states
• competency
output: initial trainee states
• domain knowledge
• trainee competency
• trainee demographics

Service-Oriented Architecture

(Communications Framework)

HLA:
entity states
entity interactions

Sensors

Trainee

trainee actions
raw user input

Local Client

Server-based Training Simulation

- Generalized Models for a Population?
  - The highly individual nature of Galvanic Skin Response (GSR) makes it virtually impossible to compare baselines across different people (Bersak et. al.)
  - Naïve/General/Dynamic Bayesian networks fail to predict Spring semester students from Fall data (Robinson et al., Sabourin et al.)
  - "It is becoming increasingly common to use cross-validation at the student level to verify that a model is applicable to new students; however it remains rare for researchers to validate that a model generalizes across subsets" (Baker, 2010)

  - Bersak, D., McDarby, G., Augenblick, N., McDarby, P., McDonnell, D., McDonal, B. and Karkun, R. (2001) "Biofeedback using an Immersive Competitive Environment". Online Proceedings for the Designing Ubiquitous Computing Games Workshop, Ubicomp 2001.
  - Sabourin, J., Rowe, J., Mott, B., Lester, J. When Off-Task is On-Task: The Affective Role of Off-Task Behavior in Narrative-Centered Learning Environments. In: Proc. of the 15th Intl. Conf. on Artificial Intelligence in Education, Auckland, New Zealand (2011).
  - Robison, J., McQuiggan, S., Lester, J. Evaluating the Consequences of Affective Feedback in Intelligent Tutoring Systems. In: Proc. of the Intl. Conf. on Affective Computing and Intelligent Interaction, pp 37-42 (2009)
  - Baker, R.S.J.d. (2010). Mining data for student models. In R. Nkmabou, R. Mizoguchi, & J. Bourdeau (Eds.) *Advances in Intelligent Tutoring Systems, Studies in Computational Intelligence* (Vol. 308, pp. 323-337). Heidelberg: Springer Verlag.

- Individual Models? Make once, use infinitely?
  - Electrode drift
  - Difference in default impedance of sensor
  - Modulation across mental states (boredom and attention)

- "Oh darn its Monday, thank God its Friday" syndrome

  - AlZoubi, O., Hussain, M.S., D'Mello, S., Calvo, RA., Affective Modeling from Multichannel Physiology: Analysis of Day Differences, ACII 2011, Part I, LNCS 6974, pp. 4–13, 2011. Springer-Verlag Berlin Heidelberg 2011

- Student Actions

- Sensor Data

- Assessment and Classification

Instructional Strategy Decision
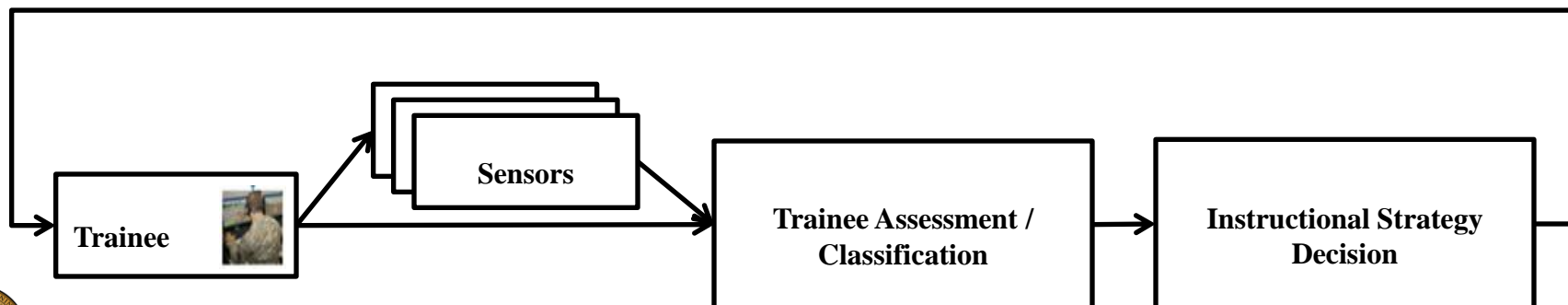
- Real time instructional decisions are made
  - Real time models must be available
  - Model must be constructed in real time
  - Data should be real time processed, and assessed

  - Draw from the field of datastream mining

- Network activity detection
  - Can't inspect all traffic
  - Continuous stream of data
- Traditional algorithmic problems (Masud et al)
  - Infinite Length
  - Concept Detection
  - Concept Drift
  - Concept Evolution

- Masud, M., Q. Chen, et al. (2010). "Classification and novel class detection of data streams in a dynamic feature space." Machine Learning and Knowledge Discovery in Databases: 337-352.

SFC Paul Ray Smith Simulation & Training Technology Center

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

- Group classification models of sensor data are impractical or nonexistent
  – Individual classification models must be built
- Offline individual models of sensor-based affect are not reusable
  – Models must be built in realtime
- Realtime-constructed models are comparable to their offline counterparts
  – Making them usable in Intelligent Tutoring Systems

- \<YRT paper end\>

- Semi-supervised datastream learning
  - 'initial packet inspection' vs. 'performance data tagging'
  - 'deep packet inspection' vs 'self report questioning'

- Realtime algorithms experimentally valid
  - K-means
  - Adaptive Resonance Theory (ART)
  - Growing Neural Gas (GAS)

- Realtime algorithms experimentally unstable
  - Neuroevolutionary techniques (rtNEAT)
  - Aglomerate clustering
    - too much data to perform centroid movement based on representative points

Questions?