Program Schedule (January 21, 2026) - B10: Making Embodied AI Reliable with Testing and Formal Verification

Room: Garnet 215

08:30–08:45 — Opening Remarks

08:45–09:30 —Invited Talk

Speaker: Prof. Guy Van Den Broeck (UCLA)

Title: *Symbolic Reasoning in the Age of Large Language Models*

09:30–10:30 — Technical Session 1 - Embodied AI, Reinforcement Learning, World Models & Multi-Robot Safety (4 Papers)

Each paper: **10-minute presentation + 5-minute Q&A**

1. **Inferring Causal Graph Temporal Logic Formulas to Expedite Reinforcement Learning in Temporally Extended Tasks**

Hadi Partovi Aria and Zhe Xu

2. **Towards a Feedback-driven Adaptive Embodied Environment Generator for Training Robotic Agents**

Teresa Yeo, Dulaj Weerakoon, Dulanga Weerakoon and Archan Misra

3. **Explicit World Models for Reliable Human-Robot Collaboration**

Kenneth Kwok, Basura Fernando, Qianli Xu, Vigneshwaran Subbaraju, Dongkyu Choi and Boon Kiat Quek

4. **Robust Evacuation for Multi-Drone Failure in Drone Light Shows**

Minhyuk Park, Aloysius K. Mok and Tsz-Chiu Au

10:30–11:00 — Coffee Break

11:00–12:15 — Technical Session 2 - Verification, Certification, Program Analysis & LLM Safety (5 Papers)

Each paper: **10-minute presentation + 5-minute Q&A**

1. **Formal Safety Guarantees for Autonomous Vehicles using Barrier Certificates**

Oumaima Barhoumi, Mohamed H Zaki and Sofiène Tahar

2. **A New Strategy for Verifying Reach-Avoid Specifications in Neural Feedback Systems**

Sam Akinwande, Sydney Katz, Mykel Kochenderfer and Clark Barrett

3. **Toward Maturity-Based Certification of Embodied AI: Quantifying Trustworthiness Through Measurement Mechanisms**

Michael Darling, Alan Hesu, Michael Mardikes, Brian McGuigan and Reed Milewicz

4. ***BALI: Branch-Aware Loop Invariant Inference with Large Language Models***

Mingxiu Wang, Jiawei Wang and Xiao Cheng

5. **A Unified Framework for Jailbreak Attacks on Large Language Models**

Xiaobing Sun and Liangli Zhen

12:15- 12:45 Poster Session

(All accepted papers may optionally bring posters.)

12：45–14:00 — Lunch Break

14:00–14:45 — Panel Discussion

Moderator: *Xi Zheng*

Theme: *Certification Challenges for AI-Enabled Autonomy: What Can Neurosymbolic Methods Enable?*

Panellists:

- Guy Van Den Broeck (UCLA)
- Dinesh Manocha (UMCP)
- Taylor T. Johnson (Vanderbilt University)

14:45–15:30 — Breakout Discussions

Participants split into three focused groups:

1. **Scenario-Based Testing and LLM-Guided Specification Mining (led by Xi Zheng)**
2. **Compositional Verification and Modular Reasoning for Embodied AI (led by Guy Van Den Broeck)**
3. **Robustness Under Uncertainty, Sensing Noise, and Human–Robot Interaction (led by Archan Misra)**

Each group produces **2–3 actionable recommendations** for the roadmap.

15:30–16:00 — Coffee Break

16:00–16:30 — Breakout Report-Back Session

Each group presents their key insights and open research questions.

16:30–16:50 — Roadmap Consolidation Discussion

Collective shaping of a community **roadmap for reliable embodied AI**, covering:

- Neurosymbolic architectures
- Testing and scenario generation
- Specification mining
- Verification and certification
- Robustness and uncertainty management

16:50–17:00 — Closing Remarks

Plans for:

- Community follow-up
- Special issues / joint whitepaper
- Future AAAI Bridge events and co-located workshops