

PERFORMANCE METRICS FOR INTELLIGENT SYSTEMS (PERMIS) WORKSHOP

National Institute of Standards and Technology, Gaithersburg, Maryland USA
September 21- 23, 2009



FOREWORD

Welcome to PerMIS'09!

The Performance Metrics for Intelligent Systems (PerMIS) workshop is dedicated to defining measures and methodologies of evaluating performance of intelligent systems. As the only workshop of its kind, PerMIS has proved to be an excellent forum for sharing lessons learned and discussions as well as fostering collaborations between researchers and practitioners from industry, academia and government agencies.

The main theme of the ninth iteration of the workshop, PerMIS'09, seeks to address the question: “**Does performance measurement accelerate the pace of advancement for intelligent systems?**” In addition to the main theme, as in previous years, the workshop will focus on applications of performance measures to practical problems in commercial, industrial, homeland security, and military applications.

The PerMIS'09 program consists of six plenary addresses and six general and special sessions. The topics that are to be discussed by the speakers cover a wide array of themes centered on many intricate facets of intelligent system research. The presentations will emphasize and showcase the interdisciplinary nature of intelligent systems research and why it is not straightforward to evaluate such interconnected system of systems. The three days of twelve sessions will span themes from manufacturing, mobile robotics, human-system interaction, theory of mind, testing and evaluation of unmanned systems, to name a few.

PerMIS'09 is sponsored by NIST, DARPA and NSF, with technical co-sponsorship of the IEEE Washington Section Robotics and Automation Society Chapter, and in-cooperation with the Association for Computing Machinery (ACM) Special Interest Group on Artificial Intelligence (SIGART). The Defense Advanced Research Projects Agency Information Processing Technology Office graciously provided funding to help support the workshop. Special thanks are due to the National Science Foundation for providing funding to allow undergraduate and graduate students to attend PerMIS this year. We also thank Professor Holly Yanco of the University of Massachusetts – Lowell for organizing the student support grants program. We gratefully acknowledge the support of our sponsors.

We thank the special session organizers for proposing interesting topics and assembling researchers related to their sessions. These focused sessions provide an opportunity to delve deeper into specialized topics and to hear from experts in the field. Our thanks are also due to the Program Committee members for publicizing the workshop and the reviewers for providing feedback to the authors, and for helping us to put together an exciting program.

The proceedings of PerMIS will be indexed by INSPEC, Compendex, ACM's Digital Library, and are released as a NIST Special Publication. Selected papers from last year's PerMIS have been published as an edited book volume by Springer Publishers entitled *Performance Evaluation and Benchmarking of Intelligent Systems* (Eds. Raj Madhavan, Edward Tunstel and Elena Messina). The book presents a detailed and coherent picture of state-of-the-art, recent developments, and further research areas in intelligent systems by drawing from the experiences and insights of experts gained both through theoretical development and practical implementa-

tion in a variety of diverse application domains. The book will be available for your perusal during the workshop.

It is our sincere hope that you enjoy the presentations, the social programs, renew old relationships, and forge new ones at PerMIS'09!

Raj Madhavan
Program Chair

Elena Messina
General Chair

SPONSORS

NIST
National Institute of Standards and Technology



PROGRAM COMMITTEE

General Chair:

[Elena Messina](#) (Intelligent Systems Division, NIST, USA)

Program Chair:

[Raj Madhavan](#) (Oak Ridge National Laboratory/NIST, USA)

Publicity Chair:

[Edward Tunstel](#) (Johns Hopkins University-APL, USA)

- S. Balakirsky (NIST USA)
- G. Berg-Cross (EM & I USA)
- F. Bonsignorio (Heron Robots Italy)
- R. Bostelman (NIST USA)
- S. Carpin (UC Merced USA)
- P. Courtney (PerkinElmer UK)
- G. Dimitoglou (Hood College USA)
- J. Evans (John Evans, LLC USA)
- P. Evans (SWRI USA)
- J. Gunderson (GammaTwo USA)
- L. Gunderson (GammaTwo USA)
- S. K. Gupta (UMD USA)
- J. Hallam (SDU Denmark)
- T. Kalmar-Nagy (Texas A & M USA)
- R. Lakaemper (Temple Univ USA)
- G. Moses (UMD USA)
- A. del Pobil (Univ. Jaume-I Spain)
- T. Potok (ORNL USA)
- E. Prassler (UApp Sci-BRS Germany)
- F. Proctor (NIST USA)
- D. Prokhorov (Toyota USA)
- L. Reeker (NIST USA)
- M. Reggiani (UPadua Italy)
- C. Schlenoff (NIST USA)
- M. Shneier (NIST USA)
- J. Steele (CSM USA)

Monday September 21		Tuesday September 22		Wednesday September 23	
8:00-8:30	Welcome/Overview	Overview	Overview	Overview	Overview
8:30-9:30	Plenary 1: Raffaello D'Andrea	Plenary 3: Ben Kuipers	Plenary 5: Lora Weiss		
9:30-10:00	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break
10:00 -12:30	MON-AM1: Model-based Performance Evaluation	TUE-AM1: The Role of Robotic Competitions in Advancing Intelligent Systems	TUE-AM2: Is an Agent Theory of Mind Valuable for Adaptive, Intelligent Systems?	WED-AM1: Performance Measures for Mobile Robots	WED-AM2: TRANSTAC: Performance Evaluation of Speech Translation Systems for Military Applications
12:30-14:00	Lunch Video Session	Food for Thought: Release of White Paper The Use of Reuse for Designing and Manufacturing Robots	Lunch	Lunch	Lunch
14:00 - 15:00	Plenary 2: David Bruemmer	Plenary 4: Paul Cohen		WED-PM1: Issues in Designing Intelligent Systems	WED-PM2: Performance Measurements Towards Improved Forklift Safety
15:00-15:30	Coffee Break	Coffee Break	Coffee Break	Coffee Break 16:00-16:30	Coffee Break 16:00-16:30
15:30 -17:30	MON-PM1: Performance Assessment and Reliability of Unmanned Systems	TUE-PM1: Ground Truth and Testbeds for Performance Testing	TUE-PM2: An Ontology for Robotics Science and Systems		
18:30 -		Banquet (19:00 –) Banquet Speaker: Tom Mitchell How does Brain Activity Represent Word Meanings?			Adjourn 16:30

PLENARY SPEAKER



**Prof. Raffaello
D'Andrea**

ETH, Zurich,
Switzerland

Towards a Ten Thousand Mobile Robot Warehouse

Mon. 8:30 am

ABSTRACT

Order fulfillment is a multi-billion dollar business. Existing solutions range from the highly automated, whose cost effectiveness is inversely related to their flexibility, to people pushing carts around in warehouses manually filling orders, which is very flexible but not very cost effective. In this talk I will describe a radical new approach to order fulfillment that is both flexible and cost effective. The key idea is to use hundreds of networked, autonomous mobile robots that carry inventory-storing pods to human operators. The result is a distribution facility that is dynamic, self-organizing, and adaptive.

Various challenges had to be overcome in order to make this an economically viable system, ranging from design of robust autonomous mobile robots, real-time wireless control of hundreds of moving agents, the coordination

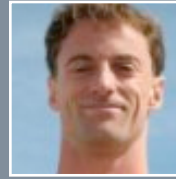
of these agents, and the design of various algorithms that allow the system to adapt and reconfigure itself based on the environment and operating conditions. I will discuss these challenges and how they scale to future warehouses with thousands—not just hundreds—of mobile robots.

BIOGRAPHY

Raffaello D'Andrea received the B.Sc. degree in Engineering Science from the University of Toronto in 1991, and the M.S. and Ph.D. degrees in Electrical Engineering from the California Institute of Technology in 1992 and 1997. He was an assistant, and then an associate, professor at Cornell University from 1997 to 2007. He is currently a full professor of automatic control at ETH Zurich. He is also a founder of, and chief scientific advisor for, Kiva Systems.

He is a co-recipient of the 2008 IEEE/IFR Invention and Entrepreneurship Award, a United States Presidential Early Career Award for Science and Engineering, and was the faculty advisor and system architect of the Cornell Robot Soccer Team, four-time world champions at the international RoboCup competition in Sweden, Australia, Italy, and Japan. He is a recipient of two best paper awards from the American Automatic Control Council and the IEEE, a National Science Foundation Career Award, and several teaching awards in the area of project-based learning. A creator of dynamic sculpture, his work has appeared at various international venues, including the National Gallery of Canada, the Venice Biennale, the Luminato Festival, Ars Electronica, and ideaCity.

PLENARY SPEAKER



**Mr. David
Bruemmer**

5D Robotics,
Inc., USA

Measuring the Benefits of Intelligent Behavior for Robotic Threat Detection

Mon. 2:00 pm

ABSTRACT

For robotic applications in hazardous, critical environments, the intelligence needed to provide functional value (i.e. reduced time, increased probability of detection, increased hazard source localization accuracy) cannot be derived from a single behavior (such as obstacle avoidance, mapping, or mine detection). Rarely do we find an integrated suite of capabilities that is capable of accomplishing an end-to-end mission. Intelligence requires not simply behavior, but also the ability to use behaviors effectively towards a highly complex set of real-world, mission-level requirements. If the level of robot initiative and autonomy used in real-world missions is to increase, the underlying mechanisms for behavior composition and human interaction must also change.

Many approaches to creating behaviors as well as orchestrating them have been offered by the community

including a variety of machine learning based techniques. These methods and algorithms are often highly elegant, formalized methods intended to streamline the development and testing methodologies. Unfortunately, these all too often fail to provide truly intelligent systems that provide value in the real world. Why is this?

One clue may be found if we consider biology. Is there anywhere in biology where we can find an elegant, formalized, understandable method for behavior composition? Functional intelligence may be, in part, derived from many interwoven heuristics for sequencing and interleaving behavior. In the brain these heuristics are learned over time through experience and perhaps not in an elegant fashion. Artificial Neural Networks (ANNs) are intended to model the behavior derivations we find in biology, but although ANNs allow us to effectively capture particular perceptual and action pairings, we are still left with the fundamental problem of how to sequence and compose behaviors to get a real job done. Without this behavior composition, we may have capability, but enjoy meager intelligence.

Although this talk will not submit a solution to this fundamental challenge, I would like to share a variety of experiments which, over the past few years, have allowed us to metric various components of intelligence for mobile robots used in a variety of real world missions. These missions include chemical plume localization, radiological characterization, urban search and rescue, mine detection and defeat of improvised explosive devices. To accomplish end-to-end missions in the hands of operators with no or little experience with robots requires a means to fuse components of robot intelligence while hiding the behavioral complexity from the user.

The Robot Intelligence Kernel (RIK) is being used to coalesce software components for perception, communication, behavior, world modeling, and human interaction into a single behavior architecture that can be easily transferred for use with a wide variety of robots and sensor-suited, low-level proprietary controls. This talk will discuss implementation strategies employed to integrate these components into a functional system that provides high-performance utility for various real-world tasks. Of particular interest is the cognitive glue, a fuzzy logic rule base, used to sequence and blend these behaviors into mission-level capabilities, such as minesweeping or radiological characterization. Lastly, the paper discusses agents within the interface that fuse various forms of robot and world representation. The interface agents also filter and interpret human input in order to incorporate it seamlessly into the behavioral intelligence of the robotic system. Our strategy is to hide sensor and behavior complexity while providing a means to integrate human intelligence at an appropriate level. In reviewing the benefits and limitations of the RIK approach, the talk will provide system-oriented results from recent hazard detection experiments. In particular, the talk will detail a number of measurements focused on the complete (i.e. human + robot + software + interface) system metrics as well as various component measurements.

BIOGRAPHY

Mr. David J. Bruemmer is Vice President for Research and Development at 5D Robotics, Inc. where he is also a founder and board member. Prior to joining 5D Mr. Bruemmer was Technical Director for Unmanned Vehicles at the Idaho National Laboratory (INL.) For more than 14 years Mr. Bruemmer has enjoyed finding ways to fuse emerging science and engineering into innovative technologies that can change the way robots interact with humans and their environment. He has authored over 50 peer reviewed journal articles, book chapters and conference papers in the area of intelligent robotics. Mr. Bruemmer has been recognized by the President's Office of Science and Technology Policy for his work to forge effective interagency research collaborations across the Federal government (e.g. NASA, Dept. of Energy, Dept. of Defense, Dept. of Commerce, Dept. of Homeland Defense). He is a winner of the R & D 100 Award, the Stoel Reeves Idaho Innovation Award and the Federal Lab Consortium Award for Excellence in Technology Transfer.

The Robot Intelligence Kernel (RIK), developed by Mr. Bruemmer and his team, is being used as a framework for integrating robot software into a standardized, interoperable architecture. Mr. Bruemmer has developed robot behaviors used for a wide variety of robots for applications including remote characterization of high radiation environments, mine sweeping operations, military reconnaissance, IED defeat, chemical plume tracing and search and rescue operations. These efforts have yielded 11 Patents (Issued and Pending) and 10 copyrighted software inventions. His research in the area of countermining operations has demonstrated a four fold decrease in time necessary to find landmines and an improvement of over 20% in probability of detection when compared with the current military baseline. Before working at the INL, Mr. Bruemmer served as a consultant to the Defense Advanced Research Projects Agency, where he worked to

coordinate development of autonomous robotics technologies across several offices and programs.

PLENARY SPEAKER



Prof. Ben Kuipers

University of
Michigan, USA

Evaluating the Robot Cognitive Mapper

Tues. 8:30 am

ABSTRACT

A robot observes the space within range of its sensors. In this “small-scale” space, it detects hazards and makes local motion plans. As it explores its global environment, it knits local spatial models together to build a cognitive map—a representation of the global structure of “large-scale” space that extends beyond the sensory horizon of the robot at any given time.

We have developed the Hybrid Spatial Semantic Hierarchy (HSSH), a model of the cognitive map that covers both large-scale and small-scale space, as experienced by the exploring robot. The key idea behind the HSSH is to combine the strengths of multiple different representations (ontologies) for space, each relatively simple: the Local Metrical,

Local Topological, Global Topological, and Global Metrical maps.

This hierarchy of representations supports a relatively simple and robust way for the robot to construct a useful cognitive map from exploration experience. It also supports robust and efficient planning of routes from one place to another, as well as multiple ontologies for communication between a robot and a human directing it in how to reach a desired destination.

The structure of the HSSH allows us to factor the evaluation task into simpler elements. Each level of the hierarchy can be evaluated according to its ability to meet the needs of the other levels, and the hierarchy as a whole is evaluated according to the different ways it can meet the needs of the robot agent, and how well each of those ways is accomplished. As a result of this factoring, each component is easier to evaluate, and has a lower bar for successful performance.

BIOGRAPHY

Benjamin Kuipers joined the University of Michigan in January 2009 as Professor of Computer Science and Engineering. Prior to that, he held an endowed Professorship in Computer Sciences at the University of Texas at Austin. He received his B.A. from Swarthmore College, and his Ph.D. from MIT. He investigates the representation of commonsense and expert knowledge, with particular emphasis on the effective use of incomplete knowledge. His research accomplishments include developing the TOUR model of spatial knowledge in the cognitive map, the QSIM algorithm for qualitative simulation, the Algernon system for knowledge representation, and the Spatial Semantic Hierarchy model of knowledge for robot exploration and mapping. He has served as Department Chair at UT Austin, and is a Fellow of AAAI and IEEE.

PLENARY SPEAKER



Prof. Paul Cohen

University of
Arizona, USA

Against Sophistication: Why Worry About Performance Assessment

Tues. 2:00 pm

ABSTRACT

The theme of the 2009 PerMIS is, "Does performance measurement accelerate the pace of advancement for intelligent systems?" Surely, performance measurement is necessary but not sufficient for the advancement of intelligent systems, and no measurement can compensate for badly designed performance tasks or for performance becoming an end in itself. AI is drunk on performing hard tasks at high levels. Given a choice between power and generality, most of us choose power. Our programs depend on designed exploits, or on designed search spaces in which programs can learn exploits. Divide-and-conquer, specific function, power over generality, and exploits are valuable engineering methods in many disciplines. They are apt to build machines that do one thing well. Human intelligence isn't that kind

of machine.

Fixing the current situation will require a disciplined stand against sophistication. It will require investments in general, child-like intelligence, and the investors might not see a return—high performance from cognitive systems—for some time. I think this is a deal worth making, both because it is likely to succeed and because the pursuit of high performance returns low dividends.

BIOGRAPHY

Paul Cohen is Professor and Head of Computer Science at the University of Arizona. Before that he worked at UMass Amherst and the USC Information Sciences Institute. His research is on planning, learning, cognitive development and language. He wrote a textbook on empirical methods for computer science and has worked on the evaluations of several DARPA programs, most recently PAL, Coordinators and Machine Reading.

PLENARY SPEAKER



Dr. Lora Weiss

Georgia Institute of
Technology,
USA

Assessing Autonomous Systems As They Evolve

Wed. 8:30 am

ABSTRACT

Today, unmanned systems are operating in-theater with untested collaborative capabilities. The vehicles are heterogeneous, in that they are developed by different contractors, they have different levels of autonomy, they have different sensors and capabilities, and they are physically disparate. Unmanned air vehicles built by one contractor have never autonomously collaborated with unmanned sea surface vehicles built by another contractor, and no one knows how they would perform if deployed together today. Their integrated use, however, is rapidly growing in the military. As improvements in autonomy, sensing, and reasoning advance, collaborating, multi-vendor unmanned systems will be increasingly employed to support challenging, tactical operations. The anticipated increase in sophistication drives the need for an ability to robustly test,

measure, and evaluate heterogeneous unmanned vehicles for full spectrum dominance and joint operations. We need to consider assessment methods to evaluate force-on-force and mission level the effectiveness of disparate unmanned systems collaborating in theater-wide scenarios. A key requirement for assessing autonomous unmanned systems is the realization that unmanned vehicles pose new challenges that are distinct from traditional approaches to assessing systems. These challenges stem from the upcoming capabilities of unmanned systems being able to autonomously collect and process data, turn it into valued information and knowledge, and then intelligently act upon it with little to no operator involvement. Autonomy at the individual vehicle level involves transitioning cognition into decisions that drive actions. Based on the mission or operational environment, these unmanned systems may execute behaviors that cannot be precisely predicted. Assessments need to support evaluation of autonomous vehicle actions and judge whether the actions are reasonable and acceptable, without having precisely quantifiable metrics. Evaluating these systems will focus more on capabilities and missions rather than mechanics. New approaches to measuring their effectiveness will be adopted to support advances in autonomy and cognition, where the metrics and methods evolve and adapt, just as the systems do.

BIOGRAPHY

Dr. Lora G. Weiss is a lab Chief Scientist at the Georgia Tech Research Institute, where she conducts research on the design, development, and implementation of autonomy and control for manned and unmanned systems. She has supported intelligent autonomy for unmanned underwater vehicles, unmanned air vehicles, and unmanned ground vehicles, and is currently engaged in research in exploring all aspects of the behavior of

these systems. Dr. Weiss has chaired sessions at IEEE conferences, ASA conferences, and Navy Symposia and currently chairs the ASTM Standards Development Subcommittee F41.01, on Unmanned Maritime Vehicle Autonomy and Control. Dr. Weiss is on the Board of Directors for AUVSI, the world's largest non-profit unmanned systems organization. She has developed a video for IEEE Educational Services and has received several publication awards. Dr. Weiss has been Principal Investigator on numerous DoD programs sponsored by offices such as DARPA, the Office of Naval Research, and various Navy Program Executive Offices. She has provided over 150 technical briefs to high-ranking DoD officers and DoD technology offices.

VIDEO SESSION

Dr. Gary Berg-Cross

Knowledge Strategies, USA

**Developmental Robotics in Theory and Action:
a new way to Understand Cognition and Build
Robots with Adaptive Abilities?**

Mon. 12:45 - 1:15 pm

This video session serves as an introduction to the topic of developmental robotics (DR). It also serves to discuss some topics in the broader field of cognitive development, which can be explored by the DR research program. DR is a newly emerging interdisciplinary field that builds on 2 of the best tools we have to study cognition –robots and computer modeling. DR studies how autonomous robots can acquire/construct skills, processes & knowledge on their own, strictly through their interactions with the surrounding environment. A core idea is that intelligence is not solely explained by innate mechanisms that modularly organize the human brain. Instead the hypothesis is that much of intelligence/cognition results from a much dynamic process constructing cognitive ability through a long personal development involving “embodied interactions” in rich environments.

FOOD FOR THOUGHT: RELEASE OF WHITE PAPER

Prof. Erwin Prassler

UAppSci. Bonn-Rhein-Sieg, Germany

**The Use of Reuse for Designing and
Manufacturing Robots**

Tues. 12:45 - 1:30 pm

BANQUET SPEAKER



Prof. Tom Mitchell

Carnegie
Mellon
University, USA

**How does
Brain Activity
Represent
Word
Meanings?**

Tues. 7:00 pm

Gaithersburg
Hilton

ABSTRACT

How does the human brain represent meanings of words and pictures in terms of the underlying neural activity? This talk will present our research using machine learning methods together with fMRI brain imaging to study this question. One line of our research has involved training classifiers that identify which word a person is thinking about, based on their neural activity observed using fMRI. A more recent line involves developing a computational model that predicts the neural activity associated with arbitrary English words, including words for which we do not yet have brain image data. Once trained, the model predicts fMRI activation for any other concrete noun appearing in the text corpus, with highly significant accuracies over the 100 nouns for which we currently have fMRI data. Professor Mitchell's research was recently featured on a CBS 60 Minutes story “Reading your

Mind.”

BIOGRAPHY

Tom M. Mitchell is the E. Fredkin University Professor and head of the Machine Learning Department at Carnegie Mellon University. Mitchell is a past President of the American Association of Artificial Intelligence (AAAI), and a Fellow of the AAAS and of the AAAI. His general research interests lie in machine learning, artificial intelligence, and cognitive neuroscience. Mitchell's web home page is www.cs.cmu.edu/~tom.



WONDAY

08:00	Welcome & Overview
08:30	Plenary Presentation: Raffaello D'Andrea Towards a Ten Thousand Mobile Robot Warehouse
09:30	Coffee Break
10:00	MON-AM1 Model-based Performance Evaluation <i>Chairs: M. Ani Hsieh & Paul Evans</i> <ul style="list-style-type: none"> Validating Extended Neglect Tolerance Model with Humanoid Soccer Robots across Varying Task Complexities [Rajesh Elara, Wijerupage Sardha Wijesoma, Carlos Acosta Calderon, Changjiu Zhou] Modeling Multiple Human Operators in the Supervisory Control of Heterogeneous Unmanned Vehicles [Brian Mekdeci, Mary Cummings] Internal Model Generation for Evolutionary Acceleration of Automated Robotic Assembly Optimization [Jeremy Marvel, Wyatt Newman] Development of Top-Down Analysis of Distributed Assembly Tasks [Anthony Cowley, M. Ani Hsieh, C.J. Taylor] Context-Based Object Recognition [Shaun Edwards, Meredith Wright, Ben Abbott]
12:30	Lunch 12:45 - 13:15 Video Session (Gary Berg-Cross)
14:00	Plenary Presentation: David Bruemmer Measuring the Benefits of Intelligent Behavior for Robotic Threat Detection
15:00	Coffee Break
15:30	MON-PM1 Performance Assessment and Reliability of Unmanned Systems <i>Chairs: Hui-Min Huang & Coire Maranzano</i> <ul style="list-style-type: none"> A Mission Taxonomy-Based Approach to Planetary Rover Cost-Reliability Tradeoffs [David Asikin, John Dolan] Towards a Systematic Assessment of the Functions of Unmanned Autonomous Systems [Robin Jaulmes, Eric Moline, Laurent Vielle] Performance Measures Framework for Unmanned Systems (PerMFUS) [Hui-Min Huang] Optimum Combination of Full System and Subsystem Tests for Estimating the Reliability of a System [Coire Maranzano, James Spall]



08:00	Welcome & Overview
08:30	Plenary Presentation: Raffaello D'Andrea Towards a Ten Thousand Mobile Robot Warehouse
09:30	Coffee Break
10:00	MON-AM2 Special Session I: Performance Metrics for Sustainable Manufacturing <i>Organizers: Kevin Lyons, Mahesh Mani & Ram Sriram</i> <ul style="list-style-type: none"> • Manufacturing Unit Process Life Cycle Inventories (Uplci) [Michael Overcash, Janey Twomey, Jacqueline Isaacs] • Conceptual Foundations of Energy Aware Manufacturing [Soundar Kumara] • Discrete Event Simulation as Requirements Specification for Sustainable Design of Manufacturing Systems [Björn Johansson, Anders Skoogh, Mahesh Mani, Swee Leong] • Towards A New Geometric Metric for Sustainability Assessment [Gaurav Ameta] • Panel Discussion
12:30	Lunch 12:45 - 13:15 Video Session (Gary Berg-Cross)
14:00	Plenary Presentation: David Bruemmer Measuring the Benefits of Intelligent Behavior for Robotic Threat Detection
15:00	Coffee Break
15:30	MON-PM2 Special Session II: Test and Evaluation of Unmanned and Autonomous Systems <i>Organizers: Mauricio Castillo-Effen & Nikita Visnevski</i> <ul style="list-style-type: none"> • Unmanned and Autonomous Systems Mission Based Test and Evaluation [Philipp Djang, Frank Lopez] • Modeling and Simulation for Unmanned and Autonomous System Test and Evaluation [Mauricio Castillo-Effen, Nikita Visnevski, Raj Subbu] • Evolutionary Framework for Test of Autonomous Systems [Raj Subbu, Nikita Visnevski, Philipp Djang] • Metrics for Co-evolving Autonomous Systems [Jack Ring]



22
YEARS
TUESDAY

08:15	Overview
08:30	Plenary Presentation: Ben Kuipers Evaluating the Robot Cognitive Mapper
09:30	Coffee Break
10:00	TUE-AM1 The Role of Robotics Competitions in Advancing Intelligent Systems <i>Chairs: Stephen Balakirsky & Jason Gorman</i> <ul style="list-style-type: none"> • The Role of Robotics Competitions in Advancing Intelligent Systems: A Practitioner's Perspective • History and Evolution of Robot Rescue Competitions [Adam Jacoff] • Evaluating The RoboCup 2009 Virtual Robot Rescue Competition [Stephen Balakirsky, Stefano Carpin, Arnoud Visser] • RoboCupRescue Interleague Challenge 2009: Bridging the Gap between Simulation and Reality [Alexander Kleiner, Chris Scrapper, Adam Jacoff] • Mobile Microrobot Characterization through Performance-Based Competitions [Jason Gorman, Craig McGray, Richard Allen]
12:30	Lunch 12:45 - 13:30 Food for Thought: Release of White Paper <i>The Use of Reuse for Designing and Manufacturing Robots (Erwin Prassler)</i>
14:00	Plenary Presentation: Paul Cohen Against Sophistication: Why Worry About Performance Assessment
15:00	Coffee Break
15:30	TUE-PM1 Ground Truth and Testbeds for Performance Testing <i>Chairs: Tsai Hong & Barry Bodt</i> <ul style="list-style-type: none"> • Data Collection Test-bed for the Evaluation of Range Imaging Sensors for ASME/ITSDF B56.5 Safety Standard for Guided Industrial Vehicles [William Shackelford, Roger Bostelman] • Ground Truth Data Using 3D Imaging for Urban Search and Rescue Robots [Nick Scott, Alan Lytle] • Performance Measurements of Evaluating Static and Dynamic Multiple Human Detection and Tracking Systems in Unstructured Environments [Barry Bodt, Richard Camden, Harry Scott, Adam Jacoff, Tsai Hong, Tommy Chang, Rick Norcross, Anthony Downs, Ann Virts] • Mathematical Metrology for Evaluating a 6DOF Visual Servoing System [Mili Shah, Tommy Chang, Tsai Hong, Roger Eastman]
18:30	Banquet 19:00 - Banquet Speech <i>How does Brain Activity Represent Word Meanings? (Tom Mitchell)</i>

27
 YVCSER
 TUESDAY

08:15	Overview
08:30	Plenary Presentation: Ben Kuipers Evaluating the Robot Cognitive Mapper
09:30	Coffee Break
10:00	TUE-AM2 Special Session III: Is an Agent Theory of Mind Valuable for Adaptive, Intelligent Systems? <i>Organizer: Gary Berg-Cross</i> <ul style="list-style-type: none"> • Is an Agent Theory of Mind (ToM) Valuable for Adaptive, Intelligent Systems? [Gary Berg-Cross] • Robotic Theory of Mind [Kyung-Joong Kim, Hod Lipson] • Resilient Behavior through Controller Self-Diagnosis, Adaptation and Recovery [Juan Cristobal Zagal, Hod Lipson] • Neurodynamics of Cognition and Consciousness [Robert Korzma, Walter Freeman] • Theory of Mind, Computational Tractability, and Mind Shaping [Tad Zawidzki]
12:30	Lunch 12:45 - 13:30 Food for Thought: Release of White Paper <i>The Use of Reuse for Designing and Manufacturing Robots (Erwin Prassler)</i>
14:00	Plenary Presentation: Paul Cohen Against Sophistication: Why Worry About Performance Assessment
15:00	Coffee Break
15:30	TUE-PM2 Special Session IV: An Ontology for Robotics Science and Systems <i>Organizers: Erwin Prassler & Herman Bruyninckx</i> <ul style="list-style-type: none"> • An Ontology of Robotics Science [Herman Bruyninckx] • Ontology Formalisms: What is Appropriate for Different Applications? [Craig Schlenoff] • Universal Core Semantic Layer: A Roadmap to Semantic Interoperability [Lowell Vizenor, Barry Smith] • Mobile Robot Map-building in the Hybrid Spatial Semantic Hierarchy [Ben Kuipers]
18:30	Banquet 19:00 - Banquet Speech <i>How does Brain Activity Represent Word Meanings? (Tom Mitchell)</i>

23

WEDNESDAY

08:15	Overview
08:30	Plenary Presentation: Lora Weiss <i>Assessing Autonomous Systems As They Evolve</i>
09:30	Coffee Break
10:00	WED-AM1 Performance Measures for Mobile Robots <i>Chairs: Alan Bowling & Rolf Lakaemper</i> <ul style="list-style-type: none"> • Performance Measures of Agility for Mobile Robots [Alan Bowling, Shih-Chien Teng] • Measuring Rover Performance in Real-time for NASA Robotic Recon Operations [Debra Schreckenghost, Terrence Fong, Tod Milam, Hans Utz] • A Biologically Inspired Sensory Driven Method for Tracking Wind-Borne Odors [Brian Taylor, Brandon Rutter, Roger Quinn] • A Confidence Measure for Segment Based Maps [Rolf Lakaemper] • Evaluation of Robocup Maps [Ben Balaguer, Stefano Carpin, Stephen Balakirsky, Arnoud Visser]
12:30	Lunch
14:00	WED-PM1 Issues in Designing Intelligent Systems <i>Chairs: Danil Prokhorov & Satyandra Gupta</i> <ul style="list-style-type: none"> • Performance Measurement and Its Role in Advancement for Intelligent Systems: Discussion Points [Danil Prokhorov, Yasuo Uehara] • Collective Intelligence: Toward Classifying Systems of Systems [Alan Ramsbotham] • A Decision-Theoretic Formalism for Belief-Optimal Reasoning [Kris Hauser] • Evaluation of Automatically Generated Reactive Planning Logic for Unmanned Surface Vehicles [Max Schwartz, Petr Svec, Atul Thakur, Satyandra Gupta]
16:00	Coffee Break
16:30	Adjourn

23

WEDNESDAY

08:15	Overview
08:30	Plenary Presentation: Lora Weiss Assessing Autonomous Systems As They Evolve
09:30	Coffee Break
10:00	WED-AM2 Special Session V: TRANSTAC: Performance Evaluation of Speech Translation Systems for Military Applications <i>Organizers: Craig Schlenoff & Brian Weiss</i> <ul style="list-style-type: none"> • Evaluating Speech Translation Systems: Applying SCORE to TRANSTAC Technologies [Craig Schlenoff, Brian Weiss, Michelle Potts, Greg Sanders, Frederick Proctor, Ann Virts] • Development and Internal Evaluation of Speech-to-Speech Translation Technology at BBN [David Stallard, Rohit Prasad, Prem Natarajan] • The Impact of Scenario Development on the Performance of Speech Translation Systems Prescribed by the SCORE Framework [Brian Weiss, Craig Schlenoff] • Probability of Successful Transfer of Low-Level Concepts via Machine Translation: A Meta-Evaluation [Greg Sanders, Sherri Condon] • Automated Metrics for Speech Translation [Sherri Condon, Mark Arehart, Christy Doran, Dan Parvaz, John Aberdeen, Karine Megerdoomian, Beatrice Oshika, Greg Sanders] • Utility Assessment in TRANSTAC: Using a Complementary Set of Methods [Michelle Steves, Emile Morse]
12:30	Lunch
14:00	WED-PM2 Special Session VI: Performance Measurements Towards Improved Forklift Safety <i>Organizer: Roger Bostelman</i> <ul style="list-style-type: none"> • Fork Lift Awareness [Mark Austin] • AGV Forklifts - Current and Future Safety Systems [Benny Forsman] • Where AGV's and Forklifts Roam: Preserving Operational Safety in a Shared Workspace [Richard Ungerbuehler] • Performance Measurements Towards Improved Manufacturing Vehicle Safety [Roger Bostelman, Will Shackelford] • Recommendations for Next Generation Forklifts to Become Safer – Group Discussion [Roger Bostelman]
16:00	Coffee Break
16:30	Adjourn

ACKNOWLEDGMENTS

These people provided essential support to make this event happen. Their ideas and efforts are very much appreciated.

Website and Proceedings

Debbie Russell

Local Arrangements

Jeanenne Salvermoser

Jennifer Peyton

Conference and Registration

Mary Lou Norris

Angela Ellis

Teresa Vicente

Kathy Kilmer



Thank you
PerMIS
attendees!

Intelligent Systems Division
Manufacturing Engineering Laboratory
National Institute of Standards and Technology
100 Bureau Drive, MS 8230
Gaithersburg, MD 20899-8230
<http://www.nist.gov/mel/isd>

