

ONR MURI 2015-2018 Three-Year Review Meeting on

Understanding Scenes and Events through Joint Parsing, Cognitive Reasoning and Lifelong Learning

White Mountain, NH

September 25-27, 2018



Location: The Mountain Club on Loon
90 Loon Mountain Road
White mountain, New Hampshire
1-800-229-7829

Arrival: Monday, September 24

Reception: 6:30-7:30PM, at the Seasons on Loon

Dinner: 7:30-8:30 PM, at the Seasons on Loon

Meeting room: Hancock Function Room

Meals: Buffet breakfast, lunch and dinners are included in your hotel rate (+ Gondola tickets).

Day 1: Tuesday, September 25 Formal Review

7:30-8:30 *Breakfast and coffee*

Session 1: Project Overview

08:30-08:55 Project Overview: Themes, Progress and Perspective
--- Song-Chun Zhu, UCLA

08:55-09:00 Comments from ONR
--- Behzad Kamghar-Parsi

09:00-09:10 Comments from DSTL
--- Bob Madahar

Session 2: Visual Commonsense: Tasks, Physics, Causality, Animacy and Intentionality

09:10-09:40 Task-Oriented Representations: Human Mechanisms
--- Philippe Schyns, Glasgow

09:40-10:10 Human Hierarchy of Tasks and Active Spatial Perception
--- Andrew Glennerster, Reading

10:10-10:25 *Coffee break*

10:25-10:55 Material Property and Depth Field Prediction and Completion
--- David Forsyth

10:55-11:25 Visual Commonsense Reasoning with Intuitive Physics
--- Josh Tenenbaum, MIT

11:25-11:55 Perceiving Animacy, Causality and Intentionality
--- Brian Scholl, Yale

11:55-12:25: Panel discussion on
Human mechanisms on visual commonsense reasoning

12:25-14:00: *Lunch break*

Session 2: Technology Integration: Vision meets Learning, Language and Robotics

14:00-14:30 Compositional Models of Vision and Language for Interpretation and Synthesis
--- Derek Hoiem, UIUC

14:30-15:00 Learning with Less Supervision through Data Hallucination and Geometric Reasoning
--- Martial Hebert, CMU

15:00-15:30 Understanding Complex Tabletop Scenes Using Physics in Embodied Setting
--- Ales Leonardis by Hector Basevi, Birmingham

15:30-16:00 *Coffee break*

- 16:00-16:30 Real-time 3D Scene Relocalization
--- Phil Torr Oxford
- 16:30-17:00 Acquiring Interpretable Commonsense through Physical Interactions
--- Abhinav Gupta, CMU
- 17:00-17:30 Grounding Visual relations in Instructional Videos
--- Feifei Li by Juan Carlos, Stanford
- 17:30-18:00: Panel discussion on
The Era of Integration
- 18:00-19:00 *Dinner, and discussion may continue after dinner*
-

Day 2: Wednesday, September 26 In-depth Technical Discussion and Collaborations

7:30-8:30 *Breakfast and coffee*

Topic 1: Physical Commonsense: Causal and Reinforcement Learning

- 08:30-08:55 Fish Gills to Faces: Human Hierarchies of Tasks and Lessons for Reinforcement Learning
--- Luis Gootjes-Dreesbach, Reading
- 08:55-09:20 Understanding the Dynamics and Interactions of Objects
--- Jiajun Wu, MIT
- 09:20-09:45 Scene parsing system incorporating stability-based learning
--- Hector Basevi and Paulo Ferreira, Birmingham
- 09:45-10:10 Task-dependent representations in deep RL with multi-task & multi-environment objectives
--- Arnab Ghosh and Viveka Kulharia, Oxford
- 10:10-10:30 *Coffee break*
- 10:30-10:55 Human Causal Transfer: Challenges for Deep Reinforcement Learning
--- Mark Edmonds, UCLA
- 11:00-12:00 **Group breakout for collaborations on physical commonsense**
(Led by Josh Tenenbaum, Ales Leonardis, Phil Torr...)
- 12:00-13:00: *Lunch break*
- 13:00-15:00: *Team Activity: Hiking in the White Mountain*

Topic 2: Social Commonsense: Multi-agent Learning and Theory-of-Mind

15:00 - 15:25 Perceiving Animacy with Causal Constraints
--- Tao Gao, UCLA

15:25-15:50 Social Perception of Heider-Simmel Animations and Interactive Agent Modeling
--- Tianmin Shu, UCLA

15:50-16:20 The Five Minds: Multi-agent interactions with Theory-of-Mind
--- Siyuan Qi, UCLA + Michael Walton, SPAWAR

16:20-16:30 *Coffee break*

16:30-16:55 TBD
--- Sarah Schwettman, MIT

17:00- 18:00 **Group breakout for collaborations on social commonsense**
(Tao Gao, Brian Scholl, Song Chun Zhu, ...)

18:00-19:00 *Dinner*

Day 3: Thursday, September 27 In-depth Technical Discussion and Collaborations

7:30-8:30 Breakfast and coffee

Topic 3: Commonsense Acquisition in 3D Physical and Virtual Environment

08:30-08:55 3D Scene Parsing and Reconstruction from a Single Image
--- Chuhan Zou, UIUC

08:55-09:20 Object and Scene Understanding: From Passive Observation to Active Interaction
--- Yixin Zhu, UCLA

09:20 -09:45 Improving confidence of predictions for novel examples
--- Zhizhon Li, UIUC

09:45-10:00 *Coffee break*

10:00 -10:25 Using psychophysical methods to understand info processing in deep networks
--- Tian Xu, Glasgow

10:25 -10:50 TBD
--- Siddarth N, Oxford

10:50 -11:15 Algorithms and Architectures for Robust Multi-Class Generations
--- Arnab Ghosh and Viveka Kulharia, Oxford

11:15-11:30 Wrap-up summary

11:30-12:00 *Lunch and Departure*
Lunch Box will be provided by 11:30.