

---

# IGLU : Databases and software

Simon Brodeur\*<sup>1</sup>, Pablo Azagra Millan\*<sup>2</sup>, Luca Celotti\*<sup>1</sup>, and Florian Golemo\*<sup>3</sup>

<sup>1</sup>Univ. Sherbrooke – Canada

<sup>2</sup>UNIZAR – Espagne

<sup>3</sup>INRIA – bordeaux – France

## Résumé

IGLU : Databases and software

- Multimodal object learning from Human-Robot Interaction, Pablo Azagra Millan, UNIZAR, Spain

A dataset focuses on interactions where the user teaches new objects to the robot in various ways. It contains synchronized recordings of visual (3 cameras) and audio data which provide a challenging evaluation framework for different tasks..An end-to-end system that learns object models using object patches extracted from the recorded natural interactions.

- HoME: a Household Multimodal Environment, Simon Brodeur, Univ. de Sherbrooke, Canada

HoME is an open-source and extensible platform for artificial agents to learn at large-scale from vision, audio, semantics, physics, and interaction with objects and other agents, all within a realistic context of thousands of simulated 3D household environments. In this short tutorial, we will explain how HoME can be used as a learning and benchmarking platform for human language understanding (HLU) tasks. For instance, visual question answering, visual object discovery through multi-modal dialogue, and applications of grounding not only in vision but also in acoustics and physics.

- GuessWhat?! Visual object discovery through multi-modal dialogue, Luca Celotti, Univ. de Sherbrooke, Canada

A two-player guessing game as a testbed for research on the interplay of computer vision and dialogue systems. The goal of the game is to locate an unknown object in a rich image scene by asking a sequence of questions. Higher-level image understanding, like spatial reasoning and language grounding, is required to solve the proposed task. Our key contribution is the collection of a large-scale dataset consisting of 150K human-played games with a total of 800K visual question-answer pairs on 66K images. We explain our design decisions in collecting the dataset and introduce the oracle and questioner tasks that are associated with the two players of the game. We prototyped deep learning models to establish initial baselines of the introduced tasks.

- Robotics Simulators Overview, Florian Golemo, INRIA

Overview of the different robots simulators with some examples in pybullet & mujoco.

---

\*Intervenant