1 Abstract and Motivation

Robust robotic perception, manipulation, and interaction in domestic scenarios continues to present a hard problem: domestic environments tend to be unstructured, are constantly changing, and have to remain optimised for humans, not robots. Reliably perceiving and manipulating a wide range of objects in a person’s home therefore remains a key problem for future robotic applications.

With the Tidy Up My Room Challenge, we aim to highlight the existing research challenges in robotic perception and manipulation. We want to motivate the robotics, computer vision, and machine learning communities to solve these research challenges in a holistic fashion, requiring a tight integration of perception, reasoning, and actuation. By providing a standardized set of tasks and rules, we also foster reproducibility of experiments, which is often lacking in robotics research.

At the first annual “Tidy up my room” Challenge at ICRA 2018, participating teams will deploy their robotics technologies in a domestic living room setup and demonstrate capabilities in advanced perception and manipulation tasks. The Challenge also comprises a perception-only task to enable participants without a robot to demonstrate their capabilities in terms of scene understanding and advanced semantic and spatial reasoning. The Challenge will furthermore provide a forum for participants and the greater community to exchange ideas and discuss approaches as well as open research challenges.

We particularly encourage teams from previous Amazon Robotics Challenges, RoboCup@Home and European Robotics League challenges to participate, as the core skills required in these challenges go hand in hand.
2 Rules and Task Descriptions

2.1 Room Layout

For the Tidy Up My Room Challenge, the robots will operate in a controlled environment that resembles a living room. The size of this room will be 5x5 metres with fixed furniture objects such as bookshelves, small tables and chairs placed in a predetermined configuration within. The room will be enclosed by walls of around 1 metre height. The robot is not to navigate outside of this space during the course of the competition.

Each fixed furniture object will be at least 50 cm away from any and all other fixed objects to ensure adequate room for the robot, and mimic a real world scenario where humans, just as robots, need some space to navigate between objects. These fixed items may be include a number of chairs, bookshelves and tables along with larger furniture items that the robot may not move.

There will be an area of at least 1x1 metre in the room, on the floor that will be clear of any obstacles, allowing the robotic platform to be safely deployed. This area will be located near the center of the room, to provide good visibility of the room. The robot may navigate outside of this safe area, however there may be objects in its way.
Pictured below is an example of a furnished room with the 1x1 metre robot safety area in the center indicated by green color (for illustration only - area may or may not be highlighted in the real setup).

2.2 Objects and Open-Set Characteristics

We will place a selection of everyday objects in the mock living room. Many of these objects are of classes that are part of the popular COCO dataset, the YCB dataset, or the ACRV Picking Benchmark. To make our Challenge adhere to the open-set characteristics commonly encountered in domestic applications, the exact appearance, shape, nature, or types of these objects are not made known to the participants beforehand.

Participants best prepare for the Challenge by training their perception and manipulation systems to handle everyday items that can typically be found in a domestic environments. Among the broad categories that can be expected are diverse types of tableware, cutlery, toys, stationary, but also tools, and plants. There might be household objects that are not typically found in a living room (i.e. misplaced items that would belong in a different room).

The participants will not be allowed to remove objects from the room during test runs. Training specific object detectors for particular objects during the test / training days of the Challenge is discouraged. The organisers will introduce novel objects that were not seen during the training days on the final day of the competition.
2.3 Competition Task 1 - Scene Understanding

The task for the participants is to explore the room, automatically build a rich model of the environment, and answer questions such as “How many pieces of cutlery are in the room?”, “Where is the red bowl?”, or “Did you see any keys on the table?”.

The task begins with a dedicated exploration phase. The participants have 10 minutes to explore the room, and are not allowed to enter the room after this time has passed. The exploration can be done using a mobile or stationary robot, or by one team member who carries a sensor suite (including a laptop computer) through the room. If a robot is used, the robot can be teleoperated or do the exploration autonomously. There are no restrictions on the type and number of sensors used for this task.

After the exploration phase, the judges will hand a set of questions to the participating team, that will need to be automatically answered using the data and information gathered during the exploration, without human intervention or help. The questions will be in machine readable form (e.g. in a .json file, using WordNet synsets to identify object classes) and do not require natural language processing. Further details will be announced soon.

The questions can be broken down into several types, each with varying degree of difficulty:

a) Simple **inventory counting**. For example: How many spoons are in the room?

b) **Hierarchical inventory counting**. For example: How many pieces of cutlery are in the room?

Here we expect teams to reason that spoons, knives, and forks are all pieces of cutlery, and therefore the correct answer is the sum of occurrences for spoons, knives, and forks combined. We will use the WordNet hierarchy to determine the hypernyms used for this task.

c) **Inventory counting with simple attributes**. For example: How many blue bowls are in the room? Possible attributes for this task are color...

d) **Object detection with size attributes**. For example: Show me the big plate. Here we expect teams to show an image where the big plate is the dominant object in the image. For questions like this, there will be multiple instance of the same class, but different sizes in the room.

e) **Object detection with location attributes**. For example: Is there a wine glass on the shelf?

We expect teams to positively answer this question by showing an image where the wine glass in the shelf is the dominant object, if the glass is there. For this task there can be multiple instance of the same class in the room, but in different locations. Example locations might be on the shelf, on the table, on the chair. The simplest spatial relations we will use are “in” or “on / on top of”. More complex relations can comprise directions (left of, right of, next to, under), or distance (close to, near).

Points will be awarded for providing the correct answers. Further details will be announced soon.
2.4 Competition Task 2 - Autonomous Scene Understanding

The motivation of this task is to highlight the difficulties of autonomous navigation and scene understanding. This task is very similar to Task 1. The key differences are:

- The exploration needs to be done autonomously on a robot without any human intervention.
- The questions will be handed to the team at the beginning of the exploration phase. This way, the robot can use the questions to guide the exploration.

2.5 Competition Tasks 3 & 4 - Object Manipulation

The task for the participants is to move objects that are ‘out of place” to their designated spots. Building on previous tasks the robot should be performing these operations automatically. No human intervention is allowed after starting the task.

The participants have a specified time to detect out of place objects and move them to a dedicated space. There is no restrictions on the manipulators used for the task. For the first Tidy Up My Room Challenge we will provide a list of appropriate places for objects as a simple “object : furniture” relation model, as in Competition Task 1. Points are awarded for placing objects in their dedicated space - destroying an object incurs negative points. Extra points are awarded for placing objects in their canonical poses, e.g. a wine glass placed upright, like a human would do. Additionally time remaining after finishing the task will provide additional points. Further details on scoring will be announced soon.

To appeal to a wider audience there are two tasks, either a stationary task or mobile version.

2.6 Competition Task 3 - Manipulation with a Stationary Robot

This tasks aims at demonstrate simple robotic manipulation fused with reasoning about the objects currently on a table. The robot can be placed, by the competing team, in front of the specified fixed furniture item that needs to be tidied for best reach or perception of the scene. The task is to be performed automatically within a maximum of 15 minutes.

Example Scenario: “Tidy up the coffee table”

- Robot shall be placed in front of the table by the team
- A variety of objects are then placed (according to predefined, but unknown to the participants schematic) and the team are given a task description file
- The robot will need to clean up the coffee table by grabbing or pushing objects to their desired positions on the table, or off the table into a box if they do not belong onto the table
- This will need to be performed autonomously
This task may be performed in front of any of the stationary objects as mentioned above.

### 2.7 Competition Task 4 - Manipulation with a Mobile Robot

For this task, a robot has to autonomously tidy up the room, i.e. it is required to move all objects from the top of the table and the floor into the shelves. The robot is to be placed within the safety area by the competing team and has then 30 minutes of autonomous operation to tidy up the room.

The robot should explore, map and navigate the space to find the objects that are out of place, e.g. a book on the floor. Each object should be placed in their desired space as provided by the task description. Correctly placed objects at the end of the run are awarded with points.

Extra points are awarded for placing objects in their canonical position, e.g. a wine glass should be placed upright, like a human would do it.

Further extra points are rewarded for tidying up systematically and grouping objects: The robot should move a glass from the table into a bin in the shelf that already contains other glasses, or move a plate onto an existing pile of plates.

### 2.8 Robot Specifications

The teams can field a robot of their choosing. There are no limitations to sensors, compute, budget or volume. Participants can for example use a combination of stereo camera, depth camera, and laser scanner.

There is no restrictions on the manipulators used for the task. Participants can for example use a single arm, vacuum suction or grippers on a mobile or stationary base.

For the stationary setups power can be extended to the teams chosen fixed location.

### 3 Disclaimer

The organisers of the challenge encourage teams to explore a large range of possibilities within the scope of the rules described herein, as long as they are within the spirit of the challenge. Final decisions on rules, scoring and rankings are at the sole discretion of the organisers.