MANIPULATION BENCHMARK PROPOSAL

Reference No / Version	RAL-SI-2020-B19-0823-V1.0 (for the latest versions of the benchmark, please refer to http:/www.ycbbenchmarks.com/protocols-and-benchmarks)
Authors	Fabrizio Bottarel, Giulia Vezzani
Institution	Istituto Italiano di Tecnologia
Contact information	fabrizio.bottarel@iit.it, giulia.vezzani@iit.it
Adopted Protocol	RAL-SI-2020-P19-0823-V1.0
Scoring	Different scores must be obtained before combining them into the final score for each layout. For the details of each score, please refer to the paper. A hands-on tutorial and code to automatically compute every score is available at https://github.com/robotology/GRASPA-benchmark According to whether the benchmarked grasping pipeline works in clutter or not, users need to place all the objects at once in the layout (cluttered mode) or one at a time (in isolation). **Reachability score S0.** This score makes use of the poses logged during the reachability check. S0 is computed (for each of the 6 regions the layout area is subdivided) as the fraction of the target poses inside a region that are reachable within a position and orientation error thresholds. **Camera calibration score S1.** To be computed only if a vision system is used as input. Makes use of the poses logged during the camera calibration check. S1 is computed (for each of the 6 regions the layout area is subdivided) as the fraction of the target poses inside a region that feature a camera calibration error within a position and orientation thresholds. S1 accounts for whether the vision system is well calibrated with respect to the workspace. **Graspability score S2.** Defines whether each *k-th* object of the protocol can be grasped by the robot, according to gripper size and arm payload. $S2_k$=1 if the *k-th* object can be grasped, $S2_k$=0 otherwise. **Grasp quality score S3.** Requires the end effector kinematic and collision model. Given a layout object and a grasp pose planned by the benchmarked pipeline, computes the Grasp Wrench Space and Object Wrench Space in simulation. Please refer to the paper for the details about this metric. $\overline{S3}_k^L$ is computed for each object in each layout by averaging the grasp quality over a number of planned grasps. The minimum number of planned grasps per object is 5. **Grasp execution score S4.** Computing this score requires execution of the planned grasps in the physical setup. For each object *k* in each layout *L* and grasp *t*, assign $S4_{k,t}^L$=1 if the robot can successfully grasp the object and lift it by 0.15 m, holding it for 5 seconds. $\overline{S4}_k^L$ averages over all grasps for the same object in the same layout. **Grasp stability score S5.** Once the object has been grasped and lifted, for each planned grasp perform a trajectory with 5 waypoints. The waypoints are rotations around the approach axis in the end effector reference frame (+45 degrees, 0 degrees, -45 degrees, 0 degrees) followed by a rotation in

	the vertical plane containing such axis (30 degrees towards the table surface). The movement between one waypoint to the next lasts for 2 seconds. Assign a score of 0.2 to $S5_{k,t}^{L}$ for every waypoint reached without the object falling. $\overline{S5}_{k}^{L}$ averages over all grasps for the same object *k* in the same layout. **Obstacle avoidance score S6.** To be computed only if the grasp and movement planner account for obstacle avoidance (cluttered mode). For each planned grasp *t*, $$\bar{S6}_k^L = \frac{1}{T}\sum_{t=1}^{T}\left(1 - \frac{N_{hit,t}}{N_L^{obj}}\right) \in [0,1]$$ according to how many objects were hit during the execution of the task. $\overline{S6}_{k}^{L}$ averages over all grasps for the same object *k* in the same layout. **Final score S_L.** The composite score is computed for each object that is graspable (S2) and reachable with accurate calibration (S0, S1). $$\bar{S}_L = \frac{1}{M_L^{obj}}\sum_{m=1}^{M_L^{obj}} \bar{S}_m^L$$ where $$\bar{S}_m^L = \frac{1}{T}\sum_{t=1}^{T}(S3_{m,t}^L + S5_{m,t}^L)\cdot S4_{m,t}^L \in [0,2]$$ or $$\bar{S}_m^L = \frac{1}{T}\sum_{t=1}^{T}(S3_{m,t}^L + S5_{m,t}^L + S6_{m,t}^L)\cdot S4_{m,t}^L \in [0,3]$$ if the algorithm is benchmarked in cluttered mode.
Details of Setup	The GRASPA benchmark is designed to be deployed on any robot arm platform, regardless of the gripper or vision system. This benchmark is not compatible with grasping pipelines that make use of environmental constraints (since it does not provide any supporting surface beyond the tabletop). Since the scope of the benchmark is assessing performance of grasp planning pipelines, users can employ any grasping strategy (that does not use environmental constraints) and any grasp planner they wish to evaluate.
Results to Submit	Users need to fill the attached table with their computed scores. A brief description of the robot setup (arm, hand/gripper, vision system) and the grasping pipeline being tested is also needed. Additionally, researchers can report observations about failure cases or edge cases that are not taken into account by the scoring procedure (e.g. the robot vision system cannot perceive the whole layout area).

Layout	$\bar{S}^{L}$	Object	$S0_k^L$	$S1_k^L$	$S2_k^L$	$\overline{S3}_k^L$	$\overline{S4}_k^L$	$\overline{S5}_k^L$	$\overline{S6}_k^L$	$\bar{S}_k^L$
Layout 0		Banana								

		Foam brick								
		Gelatin box								
		Mustard bottle								
		Potted meat can								
Layout 1		Banana								
		Hammer								
		Chips can								
		Tennis ball								
		Cracker box								
		Mustard bottle								
		Potted meat can								
Layout 2		Pear								
		Scissors								
		Chips can								
		Strawberry								
		Tennis ball								
		Power drill								
		Mustard bottle								
		Medium clamp								
		Master chef can								
		Potted meat can								
		Tomato soup can								