

Whole-Body Furniture Manipulation and Door Traversal with Steve

Supervisor: Rohit Menon

Why this project: Enable Steve to perceive and manipulate drawers and hinged furniture, then extend the same whole-body control principles to opening and traversing architectural doors.

Learning outcome: Integrate RGB-D affordance estimation, constrained whole-body control of an omnidirectional mobile manipulator, PSYONIC-hand grasping, and collision-aware loco-manipulation in one evaluated ROS2 system.

RGB-D perception → interaction frame and articulation
→ approach and grasp → constrained arm/base motion
→ open drawer or furniture door
→ unlatch architectural door → reposition and traverse



Figure 1: Steve coordinates its mobile base, arm, and PSYONIC hand while interacting with articulated furniture. Door traversal adds doorway and door-sweep constraints.

Motivation

Opening articulated objects requires the end effector to remain constrained by the drawer or hinge while the mobile base preserves reachability and avoids collisions. Whole-body mobile manipulation addresses these coupled arm–base motions rather than treating navigation and manipulation independently [1, 2]. Door traversal adds a distinct problem: after unlatching and opening the door, the robot must create clearance and pass its footprint through the frame while accounting for the moving door [3]. This project studies these problems on Steve using its omnidirectional base, manipulator, RGB-D sensing, and PSYONIC hand.

Goal and Task Boundary

Develop and evaluate an integrated system with three technical components:

1. **Perception and affordances:** estimate an actionable handle frame and articulation description from RGB-D observations, together with door-frame, door-sweep, and passage geometry for traversal.
2. **Whole-body constrained manipulation:** coordinate the base, arm, and hand while maintaining drawer or hinged-object constraints.
3. **Task coordination and door traversal:** sequence the skills and plan base motion around an architectural door and through its frame.

The project distinguishes three endpoints. **Furniture manipulation** ends when a drawer or cabinet door reaches a target opening. **Architectural-door manipulation** ends when a room door is unlatched and opened to a clearance angle. **Door traversal** additionally moves Steve through the doorway. All tasks must use the same WBC implementation and the same ROS2 launch. Drawers, hinged doors, sliding doors, and traversal are selected through runtime task references and constraint modes; task-specific controllers or launch files do not satisfy this requirement. The PSYONIC hand supplies a stable grasp; the arm and base generate the opening motion. A rated door-opening capability is not assumed and must be established experimentally.

O-STaR may be consulted as a technical reference, but it targets the Toyota HSR rather than Steve. Its software, sensor calibration, coordinate frames, and performance are therefore not assumed to transfer. Student A remains responsible for implementing and validating the required perception outputs on Steve within this project.

For initial integration, ArUco or other fiducial markers and measured furniture/door geometry may provide the interaction frame, hinge, door-frame, and clearance references. This allows the shared WBC and traversal logic to be tested before the full perception pipeline is ready. The input source must be selected explicitly at runtime within the same ROS2 launch. Marker-based input does not replace Student A’s minimum perception deliverable or the final integrated evaluation with RGB-D perception.

The project is designed for three students, each receiving 9 ECTS: approximately 270 hours per student and 810 student-hours in total.

System Representation

The perception output is

$$\mathcal{A} = \{\text{type}, \mathbf{T}_A^W, \mathbf{d}, \mathcal{M}\},$$

where $\mathbf{T}_A^W$ is the interaction-frame pose, $\mathbf{d}$ is a translation direction or rotation axis, and $\mathcal{M}$ denotes modes such as **PRISMATIC-PULL**, **GRASP-AND-PULL**, or **ROTATE-AND-PUSH**. For a drawer, the controller enforces motion along $\mathbf{d}_{\text{drawer}}$ while suppressing incompatible end-effector motion. For a hinged door, perception must distinguish left from right swing in a defined door-facing frame and provide the hinge axis, rotation direction, radius, and required opening arc.

Work Packages

WP	Student A: Perception	Student B: Whole-Body Control	Student C: Traversal
WP1 – Interfaces	Handle/knob/lever class, world-frame interaction pose, articulation axis/direction, confidence, and failure.	One WBC interface for all tasks: base, arm, and hand state; runtime task references and constraint mode; grasp state; limits; collision distances; and controller outcome.	Task phase, robot footprint, door/frame geometry, base path, grasp transition, and success/failure.
WP2 – Core	Detect and localize drawer handles and door hardware; distinguish prismatic drawers from hinged doors; estimate 3D grasp pose, left/right door swing, hinge axis, opening arc, door-frame pose, doorway width, and passage clearance.	End-effector tracking, grasp maintenance, joint/workspace limits, preferred configuration, collision avoidance, mobile-base motion, and prismatic/revolute constraints.	Approach the affordance; invoke manipulation; monitor execution; consume perceived door/frame geometry; reposition the omnidirectional base; and coordinate release.
WP3 – Extension	Estimate push/pull semantics and update the affordance, moving-door pose, swept region, and traversable clearance during motion; add sliding-door perception.	Generalize the controller across drawers and left- and right-swing furniture doors of different sizes; extend it to sliding furniture doors and add compliant or adaptive motion where sensing permits.	Plan around the moving door and frame using Student A's geometry; maintain arm reachability; execute push-open, push-and-enter, and pull-and-enter.
WP4 – Evaluation	Pose/orientation error, valid-grasp rate, direction/axis error, and failures.	Tracking error, constraint violation, slip, force/current proxy, collisions, joint margin, and task success.	Clearance, path length/time, interventions, collisions, phase failures, and traversal success.
Shared	ROS2 integration on Steve, simulation/fixture tests, calibrated benchmark furniture and doors, matched repeated trials, run logs, demonstration, and final report.		

Scope by Student

Student A – Perception and Affordance Estimation. The minimum output must be implemented and validated on Steve and must be sufficient for control: a world-frame 3D grasp pose, classification as prismatic drawer or hinged door, left/right swing in the door-facing frame, and the hinge axis, rotation direction, radius, and opening arc for a hinged door. For architectural doors, the minimum also provides the door-frame pose, doorway width, door pose/opening angle, and available passage clearance needed by the traversal module. O-STaR may be used as an HSR-specific reference, but its implementation, calibration, and performance are not assumed to transfer. Expected work generalizes perception across multiple drawers and left- and right-swing furniture doors of different sizes, handle types, and viewpoints; it adds handle orientation, door plane, push/pull mode, and the predicted door-swept region. Desired work recognizes sliding furniture doors, estimates their sliding axis and travel direction, and updates affordance, moving-door, sweep, and traversable-clearance geometry during manipulation with a confidence or uncertainty estimate. Training a new affordance network is optional; existing segmentation/detection with RGB-D geometry is sufficient.

Student B – Whole-Body Constrained Manipulation. The minimum demonstrations are one drawer and one hinged furniture door using the PSYONIC hand and coordinated arm/base motion. Expected work generalizes the same controller to multiple drawers and conventional hinged furniture doors, explicitly covering both left- and right-swing doors and furniture of different sizes. Desired work adds a sliding furniture/cabinet door with lateral prismatic motion, compliance, adaptation to imperfect articulation parameters, and stronger base/arm coordination near limits. Every level must retain one WBC implementation and one ROS2 launch; only runtime task references, objectives, and constraint modes may change.

Student C – Task Coordination and Door Traversal. Minimum work coordinates unlatching and pushing an architectural door open to a defined clearance angle without requiring traversal. Expected work continues the push-opening sequence with collision-aware base motion through the doorway. Desired work completes pull-open-and-traverse, moving Steve out of the incoming door sweep before repositioning and crossing the frame.

PSYONIC-Hand Feasibility Check

The official interface exposes motor current/torque control and touch sensing, but does not specify a rated whole-hand grip force, handle torque, or allowable external wrist load [4]. Before architectural-door trials, students therefore measure: (i) sustained tangential load before slip for selected power/hook grasps, (ii) lever/knob actuation torque, (iii) target-door opening force, and (iv) grasp retention during arm/base motion. Current and thermal limits remain at documented settings. A high-friction pad or simple task-specific grasp aid is permitted and must be reported. Failure to meet the measured load is an explicit hardware limitation, not a controller success.

Workload

Per student: literature review and Steve familiarization (25 h), design/formulation (35 h), individual implementation (120 h), component experiments (30 h), joint integration (40 h), and evaluation/documentation/presentation (20 h), totaling 270 h. The 40 integration hours per student are reserved throughout the semester rather than deferred to the end.

Project Schedule

Period	Planned Milestone
November	Steve/ROS2 familiarization; benchmark furniture and door selection; interfaces, safety limits, and logging; PSY-ONIC load/slip and target-door force measurements.
December	Initial ArUco/measured-geometry integration, RGB-D affordance pipeline, drawer/furniture-door constraints, task coordinator, and component tests in simulation or a fixed fixture.
January	Integrate perception, grasp, and whole-body drawer plus hinged-furniture manipulation.
End of January	Lab midterm: measured PSYONIC feasibility; one perceived drawer task and one hinged-furniture task with explicit component outcomes.
February	Architectural lever/knob door, task switching, base repositioning, and opening-to-clearance trials.
March	Push/pull traversal attempts, repeated evaluation, final demonstration, presentation, and report.

Outcome Levels

Minimum Outcome

Student A: Implement and validate perception on Steve that detects drawer handles and door knobs; distinguishes a prismatic drawer from a hinged door; estimates a world-frame 3D grasp pose; classifies left/right door swing in the door-facing frame; and provides the hinge axis, rotation direction, radius, and opening arc required by the controller. For traversal, provide the architectural door-frame pose, doorway width, current door pose/opening angle, and available passage clearance. O-STaR may be consulted as an HSR-specific reference, but cannot be assumed as a transferable working component. Perception failure must be reported explicitly. **Student B:** Open one drawer and one hinged furniture door with the PSYONIC hand and coordinated arm/base motion. **Student C:** Unlatch and push an architectural door open to a defined clearance angle, with explicit phase monitoring and outcomes; traversal is not required at this level. **Overall project:** Demonstrate perceive → grasp → open for one drawer and one hinged furniture door, and push an architectural door open to the defined clearance angle using the integrated system on Steve, or on the agreed simulation and hardware fixture if robot access is delayed.

Expected Outcome

Student A: Generalize perception across multiple drawers and left- and right-swing furniture doors of different sizes, handle types, and viewpoints; add handle orientation, door plane, plausible interaction direction, and the predicted architectural-door swept region. **Student B:** Generalize the same WBC to multiple drawers and hinged furniture doors, covering both left- and right-swing doors and furniture of different sizes while maintaining articulation, grasp, joint, workspace, and collision constraints. **Student C:** Use Student A's door/frame and clearance estimates to continue the push-opening sequence with base repositioning and collision-aware traversal through the doorway. **Overall project:** Using one WBC implementation and one ROS2 launch, demonstrate automatic perception and whole-body manipulation across the expanded drawer and hinged-furniture set, then push an architectural door open and traverse the doorway.

Desired Outcome

Student A: Recognize sliding furniture doors, estimate their sliding axis and travel direction, update affordance, moving-door, swept-region, and traversable-clearance geometry during motion, and quantify uncertainty. **Student B:** Extend the same WBC to a sliding furniture door using a lateral prismatic constraint and add compliance or online articulation adaptation. **Student C:** Complete pull-open-and-traverse using the updated perception, including clearance from the incoming door sweep and an explicit maintain/release-grasp decision. **Overall project:** Without changing the WBC implementation or ROS2 launch, demonstrate the sliding furniture door and both push-open and pull-open doorway traversal, plus one recovery from slip, lost perception, or an inaccurate axis estimate.

Evaluation

Furniture manipulation and traversal are evaluated separately. Furniture trials record perception error, grasp acquisition, slip, articulation tracking error, opening distance/angle, base and arm motion, joint margin, interventions, outcome, and failure reason. Architectural-door trials additionally record door/frame pose error, opening-angle and passage-clearance error, unlatching success, maximum clearance, door/frame distance, doorway crossing, and the phase at which failure occurs. Minimum reporting includes repeated trials for each selected object and a component-level failure breakdown; negative results and hardware force limits remain part of the report.

Risk and Contingency

The main risks are insufficient grasp retention, uncertain articulation geometry, door forces beyond the hand/arm limits, collision near the frame, and limited access to Steve. Tests progress from a force gauge and fixed handle fixture to drawers/furniture doors, then to a selected low-force architectural door. Motions use bounded velocity, workspace/joint limits, collision monitoring, timeout, and emergency stop. If the selected room door exceeds measured capability, the students report the limitation and evaluate traversal with a validated low-force door or instrumented mock-up; the furniture demonstrations remain required.

Prerequisites and Deliverables

Prerequisites are Python or C++, ROS2, coordinate frames, linear algebra, and basic robot kinematics/control. Deliverables are three reusable ROS2 components, one shared WBC implementation and ROS2 launch for all tasks, configuration and calibration files, benchmark geometry and force measurements, simulation/fixture assets, run logs and evaluation tables, a video-ready Steve demonstration, setup documentation, final presentation, and final report.

References

- [1] M. Mittal, D. Hoeller, F. Farshidian, M. Hutter, and A. Garg, "Articulated Object Interaction in Unknown Scenes with Whole-Body Mobile Manipulation," *arXiv:2103.10534*, 2021.
- [2] M. V. Minniti, F. Farshidian, R. Grandia, and M. Hutter, "Whole-Body MPC for a Dynamically Stable Mobile Manipulator," *IEEE Robotics and Automation Letters*, 2019.
- [3] G. Kang, H. Seong, D. Lee, and D. H. Shim, "A Versatile Door Opening System with Mobile Manipulator through Adaptive Position-Force Control and Reinforcement Learning," *Robotics and Autonomous Systems*, 2024, doi:10.1016/j.robot.2024.104760.
- [4] PSYONIC, "Ability Hand API and Interface Control Document," software and technical documentation, accessed October 2026.