

Six-level taxonomy of urban agent capabilities

6 levels · 24 categories

Capability level	Reasoning demand	Coverage categories	
		4 per level	
L1 Direct perception	View → Answer Interpret current evidence	01 Entity recognition & visual reference	02 Sign text & function reading
		03 Single-frame counting & grouping	04 Local relations & evidence judgments
L2 Active evidence gathering	Observe → Act → Observe Acquire missing evidence	01 Viewpoint selection & disocclusion	02 Function verification & intent grounding
		03 Cross-view association & deduplication	04 Local traversability & conditional counting
L3 Spatial reasoning	A — ○ — ○ — ○ — B Connect places & paths	01 Visible targets & legal connections	02 Multi-location conditional target selection
		03 Elevation & support-layer connectivity	04 Local road networks & relative position
L4 Long-term memory & reuse	Past -----> Now Retrieve prior experience	01 Delayed fact & quantity queries	02 Place relocation & historical reference
		03 New routes from known connections	04 Reuse of self-collected exploration records
L5 Constrained planning & revision	○ — ○ — ○ — ⊗ — ○ Coordinate & revise plans	01 Goal dependencies & conditional planning	02 Implicit composite-need decomposition
		03 Constraint conflicts & plan adjustment	04 Selective revision of local knowledge
L6 Integrated autonomous exploration	Explore → Act → Verify Organize an entire activity	01 Integrated tasks after open exploration	02 Long-horizon cross-region decisions
		03 Whole-task recovery after disruption	04 Autonomous verification & reliable termination

Categories specify capability coverage, not fixed task quotas. Schematics describe task requirements.

Current implementation priority: L1–L3. L4–L6 remain research designs.