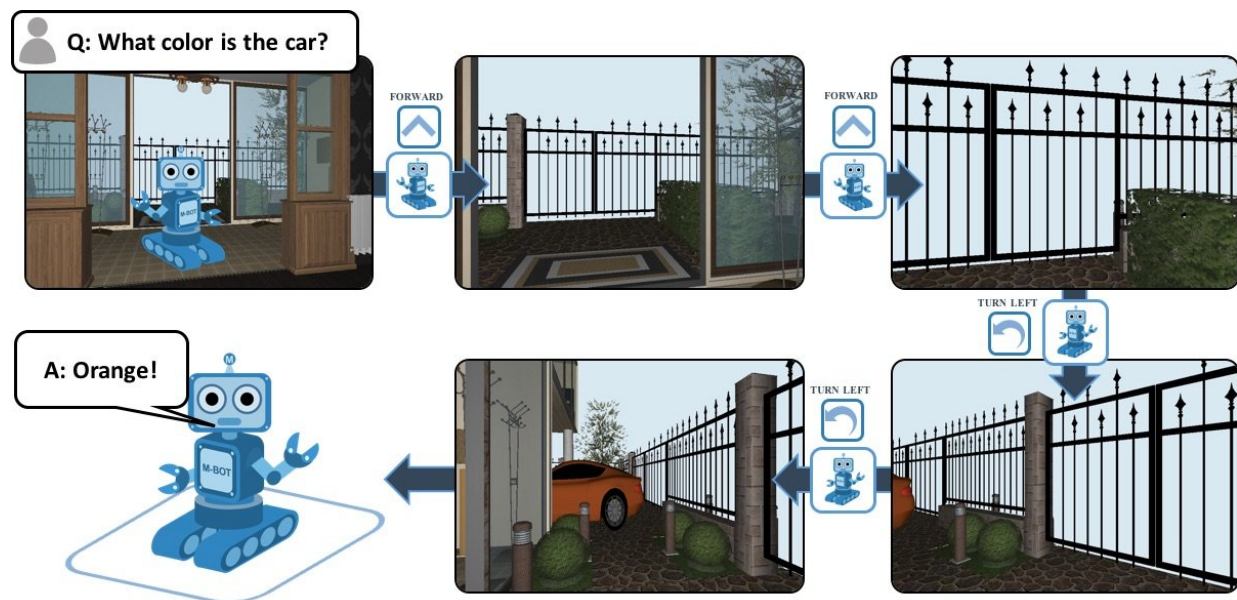# When little robot will go through your rooms to find the orange purse

May 4 2018, by Nancy Owano

Credit: embodiedqa.org

Hmm, once upon a time, we were impressed that this search phenomenon called Google could instantly answer questions and that is by just typing in words into a space bar. Mirabile dictu if you asked where is Miani Google would fire back, Did you mean Miami?

The question and answer scene has grown leaps ahead and now scientists are working on another level where intelligent systems see, plan, and

reason out the answer.

Embodied Question Answering is the name of a project and the title of a [paper on arXiv](). The six authors, with Georgia Institute of Technology and Facebook AI Research affiliations, describe their work encompassing a range of AI skills.

EmbodiedQA, as it is called, tasks agents with navigating rich 3-D environments in order to answer [questions](). Will Knight, *MIT Technology Review*, referred to this "scavenger-hunt challenge."

These agents must jointly learn language understanding, visual reasoning, and goal-driven navigation to succeed.

What it's all about: An agent is spawned at a random location in a 3-D environment. The agent is asked a question ("What color is the car?"). To get the answer, the agent must navigate to explore the environment, gather information through "first-person (egocentric) vision," and then answer.

The team developed a dataset of questions and answers in House3D environments. (You can find out more about [House3D a virtual 3-D environment, on GitHub]()).

Their paper goes into further detail on the question types and templates in the EQA dataset. location: What room? What color is the [object]()? What is above, below, next to, the object? Existence: Is there an object in the room? How many? Is Object 1 closer to Object 2 than Object 3?

The questions test abilities: object detection, scene recognition, counting, spatial reasoning, color recognition and logic.

Also, the authors said that "EQA is easily extensible to include new

elementary operations, question types, and templates as needed to increase the difficulty of the task to match the development."

The authors stressed that EQA is not a static dataset. Rather, it is a test for "a curriculum of capabilities that we would like to achieve in embodied communicating agents."

Why this matters: *Fast Company* made note that this Facebook and Georgia Tech project is actually training artificial intelligence systems to parse natural language questions and find specific objects.

Why this matters, to Will Knight in *MIT Technology Review*: "Imagine asking a Roomba to go vacuum the bedroom. Even if the machine could understand your voice and see its surroundings, it has no idea what a bedroom is, or where one might be found. But future home robots might use AI software that has learned such simple facts about ordinary homes by exploring lots of virtual homes first."

How did the researchers do it? Daniel Terdiman in *Fast Company* wrote that the team "utilized numerous types of machine learning to train the bots to answer questions about the virtual home."

"Learning" is an important part of what the team accomplished. The agent learned what Knight called "a rudimentary form of common sense." With trial and error, it figured out the best places to look for the object in question. Maybe, for example, the agent learns that cars are usually found in the garage. It may figure out that the garages are out the front or back door.

  More information: — embodiedqa.org/

— Embodied Question Answering, arXiv:1711.11543 [cs.CV] arxiv.org/abs/1711.11543

Abstract

We present a new AI task—Embodied Question Answering (EmbodiedQA)—where an agent is spawned at a random location in a 3D environment and asked a question ("What color is the car?"). In order to answer, the agent must first intelligently navigate to explore the environment, gather information through first-person (egocentric) vision, and then answer the question ("orange").

This challenging task requires a range of AI skills—active perception, language understanding, goal-driven navigation, commonsense reasoning, and grounding of language into actions. In this work, we develop the environments, end-to-end-trained reinforcement learning agents, and evaluation protocols for EmbodiedQA.

Citation: When little robot will go through your rooms to find the orange purse (2018, May 4) retrieved 24 April 2024 from https://techxplore.com/news/2018-05-robot-rooms-orange-purse.html