Tool Construction and Use Challenge: Tooling Test Rebooted

Yoonsuck Choe, Jaewook Yoo, and Qinbo Li

Department of Computer Science and Engineering
Texas A&M University
College Station, TX 77843-3112, USA
choe@tamu.edu

Abstract

Intelligence is a mixture of skills and knowledge: logic, sensorimotor skills, linguistic capability, mathematical skills, social skills, common sense knowledge, etc. What kind of test can cover most of these skills/knowledge while remaining concrete and grounded in the physical environment? It is helpful to see what scientists look for in animals as signs of intelligence. Scientists often look for signs of communication (songs and body language), logic (e.g., transitive inference, numerical inference), and tool construction and use. Among these, we propose that tool construction and use can encompass most of the skills mentioned above. In addition to the broad coverage of skills, the task difficulty can be easily controlled, and evaluation is expected to be straight-forward. Furthermore, the tool construction and use task is highly extensible. We expect our proposed challenge to lead us to a deeper understanding of the nature of intelligence, and help advance the field of artificial intelligence.

Introduction

Intelligence is a multi-faceted capability, spanning various skills and requiring diverse knowledge. Can we come up with a single task that is (1) simple and concrete yet (2) covers multiple skills, and is (3) extensible within its original scope? For insights, we can turn to scientists studying intelligence in animals. They often look for signs of communication (Snowdon, 1990), logic (Allen, 2006), or tool construction/use (Hall, 1963). Communication (language) and logic are in many senses very abstract and both have been successfully (at least to some degree) implemented in artificial intelligence (AI). However, the last one, tool construction and use, is still a largely under-developed area in AI (for simple tool use, see Chung and Choe, 2011), especially tool construction. Furthermore, tool construction and use require a vast array of skills, both sensorimotor and abstract, yet concrete tasks can be set up at increasing levels of difficulty.

Based on these observations, we propose tool construction and use as a new challenge to improve on the Turing test. We must note that this idea was first proposed by St. Amant and Wood (2005) as the "Tooling test" (it even rhymes with "Turing test"), and below, we extend and elaborate on the ideas in their pioneering work.

The Challenge: Tool Construction and Use

In a nutshell, the tool construction and use challenge will require AI artifacts to construct and use tools to solve increasingly more complex tasks, either in the physical or a simulated environment.

Tool construction and use requires the following skills and knowledge at the least:

1. sensorimotor skills,
2. knowledge of basic physics and material properties,
3. logic and planning capabilities,
4. problem posing skills (Choe and Mann, 2012; Schmidhuber, 2013),
5. problem solving skills,
6. representation of sequential and hierarchical concepts,
7. social skills (for team-based tasks), and many more.

The tool construction and use challenge has several desirable properties, making it ideal for competitions:

1. Tasks can be made very concrete, situated in the physical environment.
2. Tasks can also be made abstract, allowing for simulation-based competitions.
3. Difficulty of the task can be easily controlled and can be made increasingly more difficult as needed, almost indefinitely so.
4. Success or failure is clear cut, so evaluation is easy.

Choice of Test Materials

The test can be defined at multiple difficulty levels as discussed above, and depending on the difficulty, the test materials and environment will vary.

Level 1: Simple tool use

- Synopsis: Use a tool to reach a distant target.
- Environment: Stick within reach, target outside of reach.
- Task: Make the target move.

- Variations: Horizontal reach, vertical reach, stepping stools instead of stick, etc.
- Same level of difficulty as “Simple tool use” defined in St. Amant and Wood (2005).

Level 2: Simple tool construction and use

- Synopsis: Construct a tool using at most two objects, utilizing the objects’ basic physical properties (function).
- Environment: Stick, stone with hole, leather belt, hard object, distant object.
- Task: Break hard object within reach, hurl projectile at distant target
- Variations: Use bucket and rope to obtain water from a well.
- Same level of difficulty as “Tool construction” defined in St. Amant and Wood (2005).

Level 3: Multi-part tool construction and use

- Synopsis: Construct a complex tool involving more than two objects. Requires high levels of creativity.
- Environment: Stick, rope, sharp object, flat stone, round stone, tinder, tree branches.
- Task: Construct a tool with the above to build fire (e.g., like a pump drill).
- Variations: Use sticks and strings to build a small animal trap, use wood, ropes, hammer, etc. to build a rock hauling machine (e.g., like a trebuchet).
- Similar but more advanced than “Tool construction” defined in St. Amant and Wood (2005). Also see “Non-transparent tool use” in their paper, which is not included in our proposed challenges.

Level 4: Social tool construction and use

- Synopsis: Two or more agents with different capabilities build physical/abstract tools to solve a task.
- Environment: Two AI agents. One agent can only see and touch. The other agent can only hear and touch. There are objects in a room that can be seen and touched, and some make noise. There is a deck of braille cards with arbitrary symbols on them.
- Task: The room has several puzzles that require both visual and auditory perception to solve. Solving all puzzles will allow the agents to escape the room. The AI agents need to develop means to communicate with each other using the braille cards and solve the puzzles as a team.
- Variations: Open ended.
- Similar to “Collaborative tool use” defined in St. Amant and Wood (2005), but ours involves a more complex task of constructing a communication system with external objects.

Level 5: Explanation of tool construction and use process

- Synopsis: Explain, step by step, how the tool was constructed and used and why.
- Environment and Task: All of the above.

As we can see, a wide range of tasks of increasing complexity can be defined under a singular banner of tool construction and use.

Evaluation Metrics

Evaluation metrics are quite straight-forward, due to the binary nature of the task outcome (perhaps excluding level 5). For the difficulty levels where tool construction is needed, the quality of the constructed tool may be separately evaluated. Modularity, reusability of design, etc. could be quantitatively measured. For the actual competition, concrete, quantitative metrics may have to be developed.

Logistics

Logistics is pretty simple as well. First we need to determine the task, then come up with a tool. Then we can decompose the tool to produce the raw materials. Initially the task itself will be announced, and the raw material list will be withheld (to minimize the effect of preprogramming). On the competition day, the raw materials will be made available and the contestant AI agent should attempt to solve the task. Depending on the difficulty, the duration can be adjusted. Different physical capabilities of robots will not be differentiated, since maximally utilizing their own physical capabilities is part of being intelligent. Both physical and simulation-based competitions can be offered.

Conclusion

We proposed tool construction and use as the next challenge to improve on the Turing Test. Tool construction and use is a hallmark of intelligent behavior in animals, and it requires multiple coordinated skills and knowledge of the environment. Furthermore, it is an ideal task for competition since the level of difficulty can be easily controlled and it is easy to measure success. We expect this new challenge to lead to a deeper understanding of the nature of intelligence, and as a consequence advance the field of artificial intelligence.

References

Allen, C. 2006. Transitive inference in animals: reasoning or conditioned associations. *Rational animals* 175–185.

Choe, Y., and Mann, T. A. 2012. From problem solving to problem posing. *Brain-Mind Magazine* 1:7–8.

Chung, J. R., and Choe, Y. 2011. Emergence of memory in reactive agents equipped with environmental markers. *IEEE Transactions on Autonomous Mental Development* 3:257–271.

Hall, K. R. L. 1963. Tool-using performances as indicators of behavioral adaptability. *Current Anthropology* 479–494.

Schmidhuber, J. 2013. POWERPLAY: Training an increasingly general problem solver by continually searching for the simplest still unsolvable problem. *Frontiers in Psychology* 4(313).

Snowdon, C. T. 1990. Language capacities of nonhuman animals. *American Journal of Physical Anthropology* 33(S11):215–243.

St. Amant, R., and Wood, A. B. 2005. Tool use for autonomous agents. In *Proceedings of the Twentieth National Conference on Artificial Intelligence (AAAI 2005)*, 184–189.