

CS 472 HW 1

Aaron Havens

September 1, 2017

1 The Turing Test and Progress since the Loebner Prize

The Turing Test The *Turing Test* was developed by Alan Turing in 1950, introduced in his paper “Computing Machinery and Intelligence” [1]. The purpose of the test is to measure a machine’s ability to behave intelligently in way that is equivalent or indistinguishable from a human. The test involves a human interrogator and two other players, one human and one machine. The task of the interrogator is to observe digital text responses to a written question and to discern the machine player from the human.

The Loebner Prize Several decades following the conception of Alan’s Turing Test an annual competition, the *Loebner Prize*, commenced which sought to award the most human-like AI subject to the Turing Test [2]. The Loebner Prize is somewhat controversial among the research community due to the subjective nature of the competition and varying success criteria. Many also argue that the test is not a means of pursuing intelligence, only human-like intelligence, which may just be an arbitrary form. Over the years, many bots have fooled human interrogators, but none have achieved the necessary criteria to pass the Turing Test. These competitions mostly facilitated advancement in human error detection, in order to interface with uncertain information input. Over the years, the interaction time increased from 5 minutes to 25 minutes, with often restricted conversation.

How will AI Stack Up in Ten Years? Although most academic researchers aren’t necessarily working towards human-like intelligence, I believe

that given rapid advancements in deep learning and access to enterprise-size data, AI that satisfies and exceeds the criteria of the Turing test is impending. Primarily because large tech companies have so much to benefit from a human AI as a service.

2 Problem 2.3, Russel

For each of the following assertions, say whether it is true or false and support your answer with examples or counterexamples where appropriate.

a.) **An agent that senses only partial information about the state cannot be perfectly rational.** False: If your performance criteria is only concerned with the state you can observe, then your agent is able to act perfectly rationally. For example, the vacuum cleaner always make a rational decision without knowing the other grid states.

b.) **There exist task environments in which no pure reflex agent can behave rationally.** True: If past state information about the environment is required for the agent to act rationally, a pure reflex agent can not be perfectly rational. An example could be a card matching game, were the agent's performance depends on it's past observations of cards.

c.) **There exists a task environment in which every agent is rational.** True: An environment where every possible action (or no action) will give equal utility makes an action from any agent achieve equal and maximum utility. All agents act rationally.

d.) **The input to an agent program is the same as the input to the agent function.** False: The agent program only takes in current observable state information about the environment. However, the agent program can store this information and provide a history of states as input to the agent function.

e.) **Every agent function is implementable by some program/machine combination.** False: Suppose there was an agent function where every preceding state is needed to make a rational decision. Eventually, the agent will run out of memory and the function will fail.

f.) **Suppose an agent selects its action uniformly at random from the set of possible actions. There exists a deterministic task environment in which this agent is rational. True:** This may be a similar task environment as discussed in c.), where every action gives equal reward and maximizes utility.

g.) **It is possible for a given agent to be perfectly rational in two distinct task environments. True:** It is possible for two unique task environments to both be spanned by an agent's rational action set.

h.) **Every agent is rational in an unobservable environment. False:** A priori information about the task environment can allow an agent to make rational decisions, even when no observations can be made. In contrast to this agent, an agent with no task knowledge would not act rationally. One example could be a racetrack with a constant turning radius where the goal is for a car agent to stay on the road. If the car was designed with an a priori to constantly turn for all of time, it could be rational, but without a prior, it would not be.

i.) **A perfectly rational poker-playing agent never loses. False:** A game of poker has stochastic elements. Even if two perfectly rational agents were pitted against each other, one may lose due to an unfavorable outcome of an expectation.

References

- [1] A. Turing ,*Computing Machinery and Intelligence*, Mind 49: 433-460, 1950.
- [2] *Homepage of the Loebner Prize in Artificial Intelligence, The first "Turing Test"*, loebner.net, 2015.