

Artificial Intelligence Literacy: Simulation, ELIZA and Skit

Bobby Johnson, Stephen Shervais, Debra Morgan, Atsushi Inoue
Eastern Washington University, Spokane, WA 99202 USA

Abstract.

This paper presents how the Artificial Intelligence literacy should be raised through a general education course for all university students. Three approaches were tried out: (1) hand computing simulation to let students experience algorithmic thinking, (2) ELIZA deployment on Raspberry Pi to let them experience the system engineering, and (3) skit production and some questionnaire study to observe the social impacts (i.e. the level of AI literacy). The observational, rather than the standard interactive Turing Test is conducted for the questionnaire study. Possible extensions are discussed at the end.

Introduction.

As Artificial Intelligence (AI for short, thereafter) emerges and penetrates in our societies, people are interested in discussing, reading and using it in many different ways. On the other hand, not necessarily many people realize that there is no specific definition of AI commonly agreed upon[1]. According to the Merriam-Webster dictionary, there are two definitions of AI listed.

‘a branch of computer science dealing with the simulation of intelligent behavior in computers’

‘the capability of a machine to imitate intelligent human behavior’

Note that the meaning of ‘intelligent(ce)’ and ‘behavior’ remains unclear. Furthermore, people are more interested in, thus better focus on its applications and impacts.

In higher education, the majority of AI courses are technically focused, thus for majors. On the other hand, any emerging social issues, including those related to AI, need to be considered as topics of general education courses across the university. Our university found AI as one of such topics.

This paper presents our approaches (challenges) of raising general literacy about Artificial Intelligence among all university students, both technical and non-technical majors, throughout a general education course. There are three approaches tried out in our course: computing simulation, (deployment of) ELIZA, and a skit production as a modified Turing Test. The computing simulation is intended for students to experience how computers work. The ELIZA deployment should let students experience the system engineering tasks. The skit production and conducting a simple questionnaire should serve them as a simple yet effective scientific (statistical) study concerning the awareness of AI among ordinary people.

Computing simulation.

In Computer Science, AI is often viewed as a matter of network search. In the textbook by Russel and Norvig among others, the 8-puzzle is conveniently used as an example[1] (Figure 1).



Figure 1: 8-puzzle

In this framework, we consider

AI = network representation + search algorithm

where all possible transitions among the states of a specific problem of interest (8-puzzle in this case) are represented as a network and their orders as a search algorithm such as depth-first, breadth-first, and heuristic (e.g. A-star). Figure 2 depicts its process. $COST()$ represents the computational cost (#transitions made so far) and $EST()$ represents the estimate to reach the goal (the desired state, e.g. a 8-puzzle board). $EST()=0$ when brutal force search such as depth-first and breadth-first is deployed.

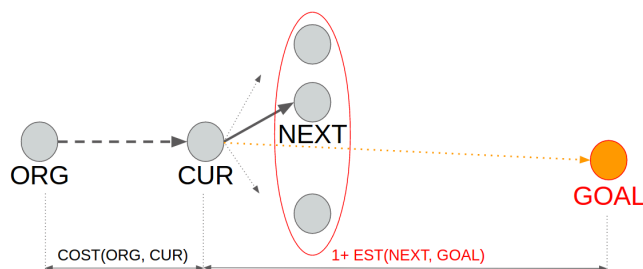


Figure 2: network search

While such a framework is very ideal for technical (e.g. computer engineering and science) majors to implement in programming, a hand simulation is ideal for non-technical majors. We simply asked small groups of students in class to generate by hand a network (tree) representation of the search process (Figure 3).

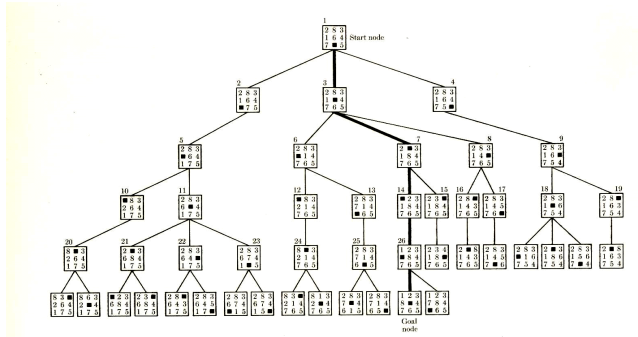


Figure 3: Tree (Graph) representation for 8-puzzle.

In addition, we asked students to work on the jug riddle by showing a video clip of the movie 'Die Hard 3' [2]. The case of deploying the depth-first search is shown in Figure 4.

Jug Riddle: The goal is to obtain exactly 4 gallons of water using only two jugs (3-gallon and 5-gallon) without measuring. You are allowed to pour, to dispose and to fill in the jugs.

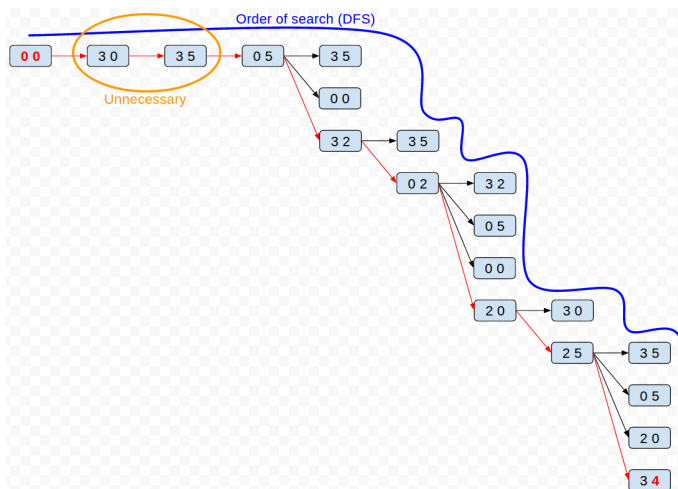


Figure 4: tree (network) representation for the jug riddle

This simulation motivates students to think like computers -- algorithmic thinking. As expected, this is very well received by technically oriented students while strongly rejected by the other types of students. As a matter of fact, hardly anyone in my freshman course completed this work.

ELIZA deployment.

ELIZA is a program that mimics a psychiatric session between a doctor and a patient [3]. Many variations of this program have been implemented and are available to the public (for free).

We asked our small groups of students to prepare a chatbot that carries out some interesting counseling session using Python on Raspberry Pi. While students are allowed to do so in any way they like, all groups chose to download some free implementation (e.g. [4]) and modify the conversational patterns to achieve this goal. We used the Raspberry Pi as a headless server and remotely accessed using ssh (Figure 5 and 6).

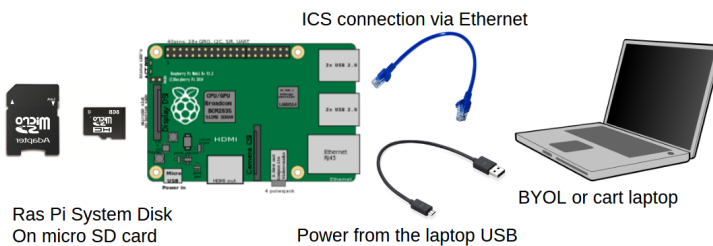


Figure 5: Headless Raspberry Pi server

```

ainoue@ainoue-Latitude-E7250:~/Documents/TTFE/python9000$ python eliza3.py
How are you doing today?
> I am fine.
Did you come to me because you are fine?
> No.
How do you feel when you say that?
> Very strange.
How does that make you feel?
> Even stranger.
Very interesting.
> Indeed.
Let's change focus a bit... Tell me about your family.
> I have my wife and two kids.
Now that you have your wife and two kids, what will you do next?
> I have not decided yet.
Why do you tell me that you've not decided yet?
> Because this is what I thought.
What other reasons come to mind?
> None.
Can you elaborate on that?
> I cannot.
How does that make you feel?
> Very frustrated.
very frustrated.

```

Figure 6: ELIZA chat on Raspberry Pi

Through this effort, students experience both using and deploying a chatbot such as this ELIZA. Doing so on the Raspberry Pi enables them to make it more marketable with advanced user interfaces such as voice generation and recognition, and aesthetics such as web interfaces (like those available at various E-commerce sites).

This was received reasonably well among all students, mostly because non-technical majors may contribute more significantly **in many different ways**. **The** Raspberry Pi may have motivated them as a good platform (or a toy) as well.

Skit production.

This was considered as the main course project. Each small group of students were asked to produce a video clip containing a skit that captured a mock counseling session between the ELIZA bot and a human (a student). In doing so, the students were asked to conceal the particular bot appearance so that audiences could not tell exactly which one was the bot. The concealing method was left for the student innovation and creativity. While most of the groups adopted a software such as nawmal.com [5] to produce animated avatars, one group performed an actual skit (a short play). See Figures 7 and 8 for example.



Figure 7: Skit with animated avatars



Figure 8: Skit with actual playing

The script was generated as a result of an actual interaction with the ELIZA on Raspberry Pi. That project was very well received by most students.

A modified Turing Test -- observational rather than interactive.

The Turing Test exhibits intelligent behavior equivalent to, or indistinguishable from, that of humans[6]. Among three players, namely A, B and C, A and B play roles of conversational partner while C interrogates A and B. In doing so, the interrogator (i.e. player C) tries to determine which, A or B, is a bot. When unsuccessful, the bot is considered passed in the Turing Test, i.e. that (say player A) this performed an intelligent behavior equivalent to, or indistinguishable from that of the human (player B). Figure 9 illustrates the test.

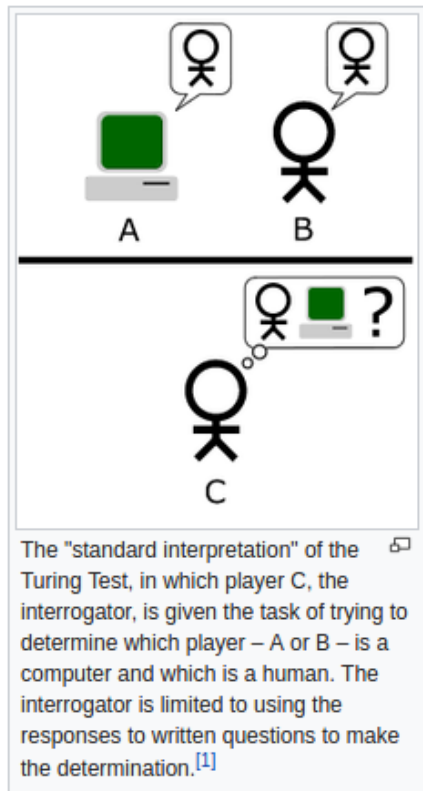


Figure 9: Turing Test (from Wikipedia)

While being controversial and criticized that this is merely a distraction from more fruitful research[7], this is valid and useful for our purpose -- AI literacy, due to its simplicity and tractability among other advantages. On the other hand, the imposition of interaction (chatting) may be a burden on ordinary people (as this is likely the same as product trials and monitoring). So, we asked randomly recruited observers to watch the skit video clips to fill in a simple questionnaire -- so called an observational Turing Test [8]. We may conclude that the bot exhibits an intelligent behavior when the observers cannot identify it correctly.

This work motivated students to critically consider what AI really is, while serving as a simply scientific (statistical) study. Table 1 shows its simple summary. It is interesting to note that more or less 40% correctly identifies a bit by observing a skit -- Turing Test not passed successfully, and this is generally consistent with the percentile of respondents who know how AI works. This suggests that discussion and study on social impacts and application are not sufficient to

constitute the AI literacy. Awareness of computing and technological aspects are likely demanded.

Table 1: Summary of responses.

Skit code name	Kernel	Nerdy	Python	MKC
#responses	171	103	80	91
Identify bot correctly	37%	38%	35%	47%
Know what AI is	62%	58%	53%	50%
Know how AI works	34%	42%	30%	32%
Education				
Beyond college	2%	2%	4%	1%
college	82%	74%	76%	81%
high school	16%	24%	20%	18%
Age				
<=30	82%	90%	86%	89%
31-50	13%	8%	12%	10%
51-70	4%	1%	1%	1%
>70	1%	1%	1%	0%

Possible extensions.

There are many possible extensions of this work. As a result of enhancing Raspberry Pi platform, we may ask observers to conduct the standard, i.e. the interactive Turing Test. Even with this observational Turing Test, we are very curious of different education levels and age groups (this study focuses on the young generation under 30 years old and the college level education).

The computing simulation may further be extended to cover probabilistic search such as the simulated annealing and the genetic algorithms. The estimation function (EST()) may be extended from crisp to fuzzy. Then the set of next state(s) is to be extended to its power set.

From a different aspect, the whole effort of the skit production and the observational (or the standard) Turing Test can be considered as a benchmarking of computational and human-computer interaction effectiveness, particularly for chatbots.

Concluding remarks.

We presented our recent work on Artificial Intelligence literacy, especially with three approaches: computing hand simulation, ELIZA deployment and the skit production. The observational Turing Test is introduced as a benchmark of each skit while raising students awareness of AI.

The possible extensions are discussed at the end for our continuing works.

Lastly, we would like to remark that our effort did indeed affect students critical thinking about AI (See Appendix).

References.

- [1] S. Russel, Norvig, *Artificial Intelligence: Modern Approach* (3rd eds), Prentice Hall, 2009.
- [2] Die Hard 3 YouTube video clip. https://www.youtube.com/watch?v=BVtQNK_ZUJg
- [3] J. Weizenbaum, "ELIZA -- A Computer Program for the Study of Natural Language Communication Between Man and Machine," *Communications of the ACM*, vol. 9, No. 1, pp.36-45, 1966.
- [4] ELIZA chatbot in Python.
<https://www.smallsurething.com/implementing-the-famous-eliza-chatbot-in-python/>
- [5] Nawmal movie making service. <http://www.nawmal.com/>
- [6] A. Turing, "Computing Machinery and Intelligence," *Mind*, 59 (236), pp. 433-460, 1950.
- [7] S. M. Shieber, "Lessons from a Restricted Turing Test," *Communications of the ACM*, 37 (6), pp. 70-78, 1994.
- [8] The observational Turing Test web page. <http://www.inoueatsushi.net/ttfe2017>

Appendix: Abstract of student research symposium about the definition of AI.

In the modern era, AI is everywhere. From sci-fi to online programs simulating a conversation, people are always trying to find some new form of AI, but they never think about what an AI actually is. Most people don't realize that there is no universally agreed upon definition of an AI, instead most people have their own definition of what can qualify as AI. The Turing Test was one of the first ways to lead us to a more solid definition. To show more about an AI, we made a skit that was a conversation between a human and an AI, and had anonymous people answer a survey in which they judged whether they felt the conversation was natural or not, and asked them to try and figure out which one was the AI. After dealing with the ELIZA AI and reviewing the results from the survey, my own definition changed.