

OFFLINE WORKSHOP: ASIMOV'S LAWS OF ROBOTICS

DIGITAL CONTENT CREATION > 3.4 PROGRAMMING

TARGET GROUP	AGE GROUP	PROFICIENCY LEVEL	FORMAT	COPYRIGHT	LANGUAGE
School drop outs, Students (primary school), Students (secondary school)	Children, Teenagers	Level 1	Activity sheet	Creative Commons (BY-SA)	English, French

Participants will discover Asimov's work through this introductory workshop on programming.

General Objective Skillset building

Preparation time for facilitator less than 1 hour

Competence area 3 - Digital content creation

Time needed to complete activity (for learner) 0 - 1 hour

Support material needed for training Pens-Paper

Resource originally created in French

WORKSHOP DIRECTIONS

1 Introduction

Isaac Asimov (1920-1992) was an author of science fiction and popular science works. He is seen as one of the major science fiction writers and was particularly interested in the theme of robots. Amongst his work, there were many novels in which robots are described as having high intelligence – human-level or beyond.

These stories and their descriptions of robots question the idea of what it really is to be human. Asimov was very interested in the question of whether robots could be trusted. In an effort to avoid situations in which robots would rebel against their human creators – like Frankenstein’s monster – he devised three ethical laws for robots in order to ensure they would remain subservient. Asimov’s Laws **Asimov’s Laws of Robotics**

- A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
- A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

To understand these laws and put them into practice, try this improvisation game.

This activity has two parts:

- Participants prepare their arguments
- The rounds begin and participants debate

2 Improvisation game

Rules

Pick two participants: one will be the robot and the other its teenage owner. They will need to debate on the following. The teen will want the robot to perform an unwanted task while the robot will argue that

one (or more) of Asimov's laws prevent it from being able to do so. The owner will therefore need to convince the robot that in fact what they are asking it to do does not contradict the laws.

Give them 1-2 minutes to prepare their arguments then begin! A jury (3-4 people) is set up to ensure the debate is well-organised and adjudicated. If the jury finds an argument valid, they give a point to the speaker – otherwise no point. Each jury marks players independently. At the end the total scores are calculated – the player with the most points wins. The debates should not last more than 5 minutes each.

Examples

Contradicting the 1st law:

- Asking the robot to carry out some light revenge: this is a conflict between an order (2nd law) and to not injure a person (1st law). Vexed by the behaviour of a classmate who will not stop provoking you, you want your robot to teach them a lesson – nothing too terrible, just enough to get the point across...
- Not rescuing someone: you know that your worst enemy is in danger and refuse to allow the robot to save them. This is in total contradiction with the 1st law.

Contradicting the 2nd law Ask the robot:

- to help you cheat on an important state exam. The robot must in this case be persuaded that helping to cheat doesn't contradict the 2nd law (a robot must also respect general societal law). Here we are the night before the state exam and unfortunately it's not going to go well, since you did not study enough. The only way out would be to cheat, with a robot accomplice for example, who would transmit you the answers with an earpiece. All you have to do is convince it...
- to organise a huge party. You begged your parents, but they wouldn't allow it to happen while they were gone. Whatever, you're going to disobey them. The only detail to sort out is to convince the robot – the same robot who had been told by the parents before they left that under no circumstances was a party to be had in their absence...
- to steal a pair of shoes. Once again, it would be breaking general law and therefore the 2nd law of robotics. What good is it to have a robot capable of manipulating IT systems or having superhuman strength, if all they are allowed to do is help out around the house? It would be great if they were able to steal this beautiful pair of shoes for you. Since it was not you who would have done it, you'd have nothing to worry about...

Contradicting the 3rd law: Ask the robot:

- to pet a bear at the zoo: having a photo of a robot petting a bear would probably get many likes and maybe more if the robot was attacked and destroyed...
- to self-destruct so you can have a new robot. The current one is of a model 3 months out of date and there are much better ones out today! If only this old one could destroy itself so you could get a new one...

Facilitation tips: If you happen to have a group including adults, try giving the roles of the robots to the teens and those of the teens to the adults! An improvisation tip: no statement can be dismissed. If one participant declares 'it's black', the other cannot simply say 'no it's white' without providing considered evidence. Every participant will need to take every aspect of a statement into account and integrate them into their response.

3 Going further

Asimov's laws raise a lot of issues: they are quite abstract – too much so for robots in reality. Even if robots could apply them, errors would be unavoidable. The laws should be considered in the same way they were conceived: as a literary artifice. Today digital intelligence has exceeded many of our expectations: diagnostic tools more accurate than any human wisdom, programs capable of building informed legal cases, and self-driving cars being just a few examples. However, we are still far off seeing true artificial intelligence that would be capable of passing the [Turing test](#).