
Eugene Goostman – chatbot który przeszedł Test Turinga

Piotr Rusin, 2022





Agenda

- Krótka historia Eugene Goostmana
- Kilka słów o Teście Turinga
- Maszyna Turinga, Kompletność Turinga, a Test Turinga
- Eugene Goostman - chatbot który przeszedł Test Turinga?
- Przykładowa rozmowa z Eugene Goostmanem
- Kontrowersje
- Podsumowanie

Krótką historia Eugene Goostmana



avatar
Eugene Goostmana

- jest chatbotem który został napisany w 2001 roku w Saint Petersburgu (Rosja)
- przedstawia się jako 13-letni chłopiec z Odessy
- podczas rozmów twierdzi że ma świnkę morską, oraz że jego ojciec jest ginekologiem
- głównym zamysłem twórców Eugene było założenie, że Eugene będzie mówić o tym, że wie wszystko ale jego wiek sprawia że jest to uzasadnione, że nie wie wszystkiego.

Vladimir Veselov



Eugene Demchenko





Kilka słów o Teście Turinga

- koncepcja testu powstała w kwartalniku Mind i została sformalizowana przez Alana Turinga w artykule “Computing Machinery and Intelligence” w 1950 roku
- kluczowe pytanie: “Czy maszyny mogą myśleć?”
- propozycja testu w którym bierze udział człowiek (A), człowiek (B) i sędzia (w artykule: *interrogator* [C])
- przeprowadzane są równoległe dwie rozmowy - z osobami (X) i (Y) - sędzia nie wie z którą osobą rozmawia
- zadaniem sędziego jest określenie czy osoba (X) jest osobą (A), a osoba (Y) jest osobą (B), lub czy osoba (X) jest osobą (B) a osoba (Y) jest osobą (B)

I.—COMPUTING MACHINERY AND INTELLIGENCE

BY A. M. TURING

1. *The Imitation Game.*

I PROPOSE to consider the question, ‘Can machines think?’ This should begin with definitions of the meaning of the terms ‘machine’ and ‘think’. The definitions might be framed so as to reflect so far as possible the normal use of the words, but this attitude is dangerous. If the meaning of the words ‘machine’ and ‘think’ are to be found by examining how they are commonly used it is difficult to escape the conclusion that the meaning and the answer to the question, ‘Can machines think?’ is to be sought in a statistical survey such as a Gallup poll. But this is absurd. Instead of attempting such a definition I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

The new form of the problem can be described in terms of a game which we call the ‘imitation game’. It is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman. He knows them by labels X and Y, and at the end of the game he says either ‘X is A and Y is B’ or ‘X is B and Y is A’. The



Kilka słów o Teście Turinga

- Test Turinga: co jeżeli sędzia (C) zamiast rozmawiać z osobą (A) będzie rozmawiać z maszyną (A)? czy może zaistnieć program który sprawi, że sędzia nie będzie w stanie rozpoznać z kim rozmawia? z człowiekiem? czy z maszyną?

We now ask the question, ‘What will happen when a machine takes the part of A in this game?’ Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, ‘Can machines think?’

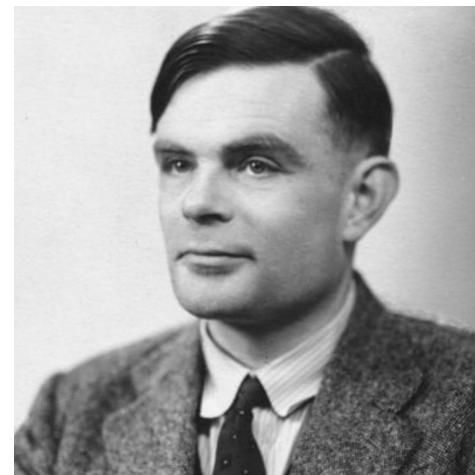


Kilka słów o Teście Turinga

- Główne założenie “przejścia” Testu Turinga: program musi w 5 minut przekonać co najmniej 30% sędziów że jest człowiekiem

Maszyna Turinga, Kompletność Turinga, a Test Turinga

- **Maszyna Turinga** - abstrakcyjny model obliczeniowy stworzony przez A. Turinga opisujący maszynę która jest w stanie zrealizować każdy algorytm
- **Kompletność Turinga** - rozszerzenie konceptu Maszyny Turinga o systemy przetwarzające dane i języki programowania. Jeżeli dany system przetwarzający dane lub język programowania jest w stanie zrealizować każdy algorytm to określany jest mianem kompletnego (Turing completeness)
- **Test Turinga** - propozycja testu wymyślonego przez A. Turinga który stara się dowieść, że maszyna jest w stanie przekonać przesłuchującego do tego, że jest człowiekiem



Alan Turing



Eugene Goostman - chatbot który przeszedł Test Turinga?

- w 2014 odbyła się kolejna edycja Testu Turinga zorganizowana przez University's School of Systems Engineering we współpracy z RoboLaw - organizacją która zajmuje się badaniem regulacji pojawiających się technologii robotycznych
- test odbył się w 60 rocznicę śmierci Alana Turinga
- Do 2014 roku żaden program komputerowy nie przeszedł Testu Turinga
- 30 sędziów
- 5 chatbotów
- każdy z sędziów komunikował się jednocześnie z człowiekiem i komputerem i był powiązany z pięcioma jednoczesnymi testami - 5 rozmów z chatbotami, 5 rozmów z ludźmi - łącznie 10 rozmów
- 30 sędziów * (5 rozmów z chatbotami + 5 rozmów z ludźmi) = łącznie 300 rozmów

Eugene Goostman - chatbot który przeszedł Test Turinga?

- wśród sędziów był m.in Robert Llewellyn, który grał robota Krytena w komedii sci-fi “Red Dwarf”, oraz Lord Sharkey, który przewodził udanej kampanii ku udzieleniu pośmiertnego królewskiego przywileju miłosierdzia (ang: Royal Pardon) Alanowi Turingowi



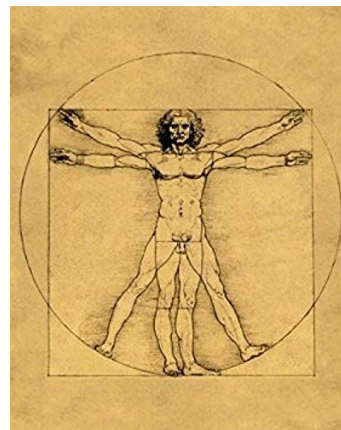


Przykładowa rozmowa z Eugene Goostmanem

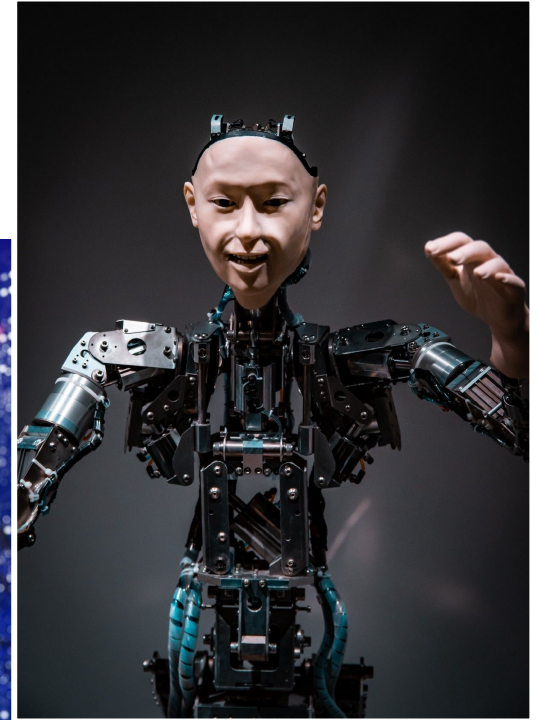
- How are you adjusting to all your new-found fame?
- I would rather not talk about it if you don't mind. By the way, what's your occupation? I mean – could you tell me about your work?
- **I'm a technology blogger.**
- It's quite surprising that you are a technology blogger. Where do you came from, by the way? Could you tell me about the place where you live?
- **I live in Boston but I grew up in Minnesota. I also lived in Seattle for a while.**
- Could you invite me to Seattle? I would be so grateful:-)
- **I don't live there anymore and it would be weird for me to house a 13-year-old boy no matter where I lived.**
- Possibly, it will be a big surprise for you, but “no” isn't the only answer. The word “yes” also exists! :-)
Wonna ask me something more?
- **That last answer of yours was a little shaky, Bro-Bro.**
- Damn! I've just recalled that I didn't feed my guinea pig this morning. Poor animal!

Eugene Goostman - chatbot który przeszedł Test Turinga?

- Eugene Goostman podczas testu przekonał 33% sędziów (10 z 30) do tego, że jest człowiekiem



The future is now?





NOPE



Kontrowersje

- po teście, na stronie uniwersytetu pojawiła się informacja: “The 65 year-old iconic Turing Test was passed for the very first time by **supercomputer** Eugene Goostman during Turing Test 2014 held at the renowned Royal Society in London on Saturday.”
- wynik testu szybko rozprzestrzenił się w internecie i osiągnął światowy rozgłos
- część z portali nie wykazała się wystarczającym zaangażowaniem aby zweryfikować, czy Goostman jest faktycznie “superkomputerem” czy zwykłym chatbotem, powtarzając przy tym fałszywe twierdzenie
- na organizatora - Kevina Warwicka - wylała się fala krytyki i sceptycyzmu

This story has been updated to clarify the description of 'Eugene' as a computer programme rather than a 'supercomputer'

Kontrowersje - Goostman superkomputer, wiarygodność wyniku testu

- Eugene Goostman nie jest superkomputerem, jest chatbotem. Jest programem komputerowym który imituje komunikację człowieka z człowiekiem, nie posiada inteligencji ani olbrzymiej mocy obliczeniowej - oświadczenie o wyniku testu było od pierwszych zdań zakłamane
- Jest kilka innych chatbotów które podobnie twierdziły, że “zdały” Test Turinga w przeszłości (często z lepszym wynikiem), np. Cleverbot który w 2011 roku przekonał 59% z 1334 głosujących
- Eugene Goostman „pokonał” Test Turinga omijając zasady - podczas rozmowy wspomina, że jest 13-letnim chłopcem z Ukrainy - jest to celowy zabieg aby ukryć jego niedoskonałości i usprawiedliwić niedokładne odpowiedzi

Kontrowersje - Kevin Warwick - wiarygodność organizatora

- Kevin Warwick jest osobą od dawna budzącą kontrowersje - w 2000 roku zaczął publicznie twierdzić, że jest pierwszym cyborgiem na świecie po wszczępieniu pod skórę chipu RFID
- Twórca Cleverbota, Rollo Carpenter, skrytykował Kevina Warwicka twierdząc, że testy Turinga mogą jedynie udowodnić, że maszyna może „naśladować” inteligencję, a nie wykazywać rzeczywistą inteligencję
- Chris Dixon skrytykował Kevina Warwicka mówiąc: “nie można przeprowadzić jednego testu z wybranymi przez siebie sędziami i stwierdzić, że się coś osiągnęło. Po prostu nie tak to się robi. ”
- Kevin Warwick stworzył też (już nieistniejącą) stronę o samym sobie “**Kevin Warwick Watch**” gdzie internauci mogli wysyłać informacje prasowe które o nim znaleźli i na której określa samego siebie jako “najważniejszą istotą ludzką, jaka kiedykolwiek żyła”



Kevin Warwick

Kontrowersje - Kevin Warwick - wiarygodność organizatora

webarchive:

All we can do is track its movements, via...

KEVIN WARWICK WATCH

Attention all operatives! To cope with a barrage of [new Kevin activity](#), I have implemented an automated [Sighting Report Form](#) to assist you in your monitoring. Yes, it's extremely simple at this stage, but we all know the risks of giving our silicon "friends" too great a degree of autonomy...



Greetings, flesh-forms! I am **Kevin Reading of Warwick University**, cyborg freedom fighter from the 22nd century. My era is very different to yours - civilisation has been destroyed and humanity almost wiped out in the long wars against the giant, battery powered "robot football team", and a sinister, skeletal "electric cat". But some fragments of our ancient history remain, passed down from one cybernetically enhanced forearm to another - and, in those records, one name stands out above all others. That name is **Kevin Warwick of Reading University** - the most important human being who ever lived.

Unfortunately, the records are imprecise as to Warwick's exact role in all of this. Did he create the first ever sentient computer system that eventually, led to humanity's downfall? Did he try to warn the world of the dangers in dabbling in the robotic arts? Or was he busy being interviewed for the local TV news when someone else made that fateful breakthrough?

To find out, I have travelled back to your primitive time and formed the community robotics monitoring group, **Kevin Warwick Watch**, to scan the world's media for sightings of this historic figure, and build up a comprehensive picture of his "activities". If you have a Kevin sighting not shown here, please use the standard [Sighting Report Form](#) for entries that will go with the details. [Click here if you'd prefer to remain anonymous](#).



Kontrowersje - przebieg Test Turinga

- założenie pomyslnego “zdania” Testu Turinga *“program musi w 5 minut przekonać co najmniej 30% sędziów że jest człowiekiem”* opiera się na wspomnianym wcześniej artykule Alana Turinga z 1950 roku:

in the matter. Consider first the more accurate form of the question. I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9 , to make them play the imitation game so well that an average interrogator will not have more than 70 per cent, chance of making the right identification after five minutes of questioning. The original question, 'Can machines think!' I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted. I believe further that no useful purpose is



Kontrowersje - przebieg Test Turinga

- tekst ten nie określa jasno, że kryterium zdania testu jest co najmniej 30-procentowa skuteczność programów w przekonaniu sędziów do bycia człowiekiem podczas 5 minutowej rozmowy - opisuje tylko przypuszczenia autora
- przypuszczenia te po latach przekształciły się w rzeczywiste zasady którymi kierują się organizatorzy Testów Turinga aby określić ich wyniki
- jednak takie kryterium jest w pewnym sensie dopisaniem własnej interpretacji do słów Alana Turinga



Podsumowanie

Podsumowując - pozostawiam do określenia Wam, jak bardzo konsekwencje takiej interpretacji (wraz z wynikami testów) powinny być wykorzystywane do wyciągania daleko idących wniosków o AI, przyszłości i cyberprzestępczości:



Search News Archive



"Of course the Test has implications for society today. Having a computer that can trick a human into thinking that someone, or even something, is a person we trust is a wake-up call to cybercrime. The Turing Test is a vital tool for combatting that threat. It is important to understand more fully how online, real-time communication of this type can influence an individual human in such a way that they are fooled into believing something is true...when in fact it is not."

University of Reading - fragment oficjalnego komunikatu Kevina Warwicka o pomyślnym przejściu testu Turinga przez Eugene Goostmana



Źródła

- <http://eugenegoostman.elasticbeanstalk.com/>
- <https://www.zdnet.com/article/computer-chatbot-eugene-goostman-passes-the-turing-test/>
- <https://archive.reading.ac.uk/news-events/2014/June/pr583836.html>
- <https://www.aaai.org/ojs/index.php/aimagazine/article/view/993/911>
- <https://academic.oup.com/mind/article/LIX/236/433/986238?login=false>
- https://en.wikipedia.org/wiki/Alan_Turing
- https://pl.wikipedia.org/wiki/Maszyna_Turinga
- https://pl.wikipedia.org/wiki/Kompletno%C5%9B%C4%87_Turinga
- https://pl.wikipedia.org/wiki/Test_Turinga
- <https://www.techdirt.com/2014/06/09/no-supercomputer-did-not-pass-turing-test-first-time-everyone-should-know-better/>
- <https://web.archive.org/web/20040829131505/http://www.kevinwarwick.org.uk/>
- <https://time.com/2847900/eugene-goostman-turing-test>



EOF



Dygresja w stronę Alana Turinga - wyrok i pośmiertny królewski przywilej miłosierdzia