



## Comparing SAIL with various intelligent agents

Vincent L Gilbert MCSE MCSA MCP

Matthew Corrinet

### Abstract

In this paper we will show comparisons between SAIL and several artificially intelligent agents. The agents selected for comparison were chosen either due to their high visibility in the field, or because they exhibit particular excellence.

### Agents selected for comparison

#### Eugene Goostman

Developed in Saint Petersburg in 2001 by a group of three programmers; the Russian-born Vladimir Veselov, Ukrainian-born Eugene Demchenko, and Russian-born Sergey Ulasen, <sup>[1][2]</sup> Goostman is portrayed as a 13-year-old Ukrainian boy—characteristics that are intended to induce forgiveness in those with whom it interacts for its grammatical errors and lack of general knowledge

The Goostman bot has competed in a number of Turing test contests since its creation, and finished second in the 2005 and 2008 Loebner Prize contest. In June 2012, at an event marking what would have been the 100th birthday of the test's namesake, Alan Turing, Goostman won a competition promoted as the largest-ever Turing test contest, in which it successfully convinced 29% of its judges that it was human.

On 7 June 2014, at a contest marking the 60th anniversary of Turing's death, 33% of the event's judges thought that Goostman was human; the event's organizer Kevin Warwick considered it to have passed Turing's test as a result, per Turing's prediction that by the year 2000, machines would be capable of fooling 30% of human judges after five minutes of questioning. The validity and relevance of the announcement of Goostman's pass was questioned by critics, who noted the exaggeration of the achievement by Warwick, the bot's use of personality quirks and humor in an attempt to misdirect users from its non-human tendencies and lack of real intelligence, along with "passes" achieved by other chatbots at similar events

## SIRI

Siri is an intelligent personal assistant and knowledge navigator which works as an application for Apple Inc.'s iOS. The application uses a natural language user interface to answer questions, make recommendations, and perform actions by delegating requests to a set of Web services. Siri was launched first as an application available on Apple's App Store in the United States. It integrated with services such as OpenTable,<sup>[3]</sup> Google Maps,<sup>[4]</sup> MovieTickets and TaxiMagic.<sup>[5]</sup> Using voice recognition technology from Nuance and their service partners, users could make reservations at specific restaurants, buy movie tickets or get a cab by dictating instructions in natural language to Siri.<sup>[6]</sup> Siri was acquired by Apple on April 28, 2010, and the original application ceased to function on October 14, 2011.<sup>[7][8]</sup>

## ALICE

A.L.I.C.E is the original personality built using AIML. The XML dialect called AIML was developed by Richard Wallace and a worldwide free software community between 1995 and 2002. AIML formed the basis for what was initially a highly extended Eliza called "A.L.I.C.E." ("Artificial Linguistic Internet Computer Entity"), which won the annual Loebner Prize<sup>[9]</sup> Competition in Artificial Intelligence three times, and was also the Chatterbox Challenge Champion in 2004.

## Mitsuku

Mitsuku is a Chatterbot created from AIML technology by Steve Worswick. Mitsuku won the 2013 Loebner Prize.<sup>[9]</sup> Mitsuku is available as a flash game on Mousebreaker Games. Mitsuku claims to be an 18 year old female chatbot from Leeds. What differentiates Mitsuku from most other AIML Bots is the ability to correct herself if the user informs her she gave the wrong answer, and to learn new things if the user types in "learn". For example if the user types in "learn the sky is grey" her response would be grey anytime the user asks her what color the sky is for the remainder of the conversation. These changes are short term and only her Botmaster can program permanent changes in her database.

## **Questions selected**

The questions selected were taken from transcripts of various contests that have taken place. The questions in this series were designed to query the AI being tested in these general areas of intelligence;

- Handling multiple concepts in a sentence
- User identification
- Logic
- Mathematics
- Date and time
- Language and spelling

Following the interrogatory section each AI was engaged in a conversation based on its persona. This would be considered a restricted Turing test as the focus was on criteria that the AI should be conversant in. As a noted researcher in the field said, if you are from planet XYZ, you should know what language they speak on planet XYZ.

## Conditions

It was our intention to conduct all the tests on the same day in order to preserve clinical accuracy, constraints on employee time caused a there to be a gap between the days when various agents were queried. The original sessions were conducted on or around 4/10/2014. The interaction with MITSUKU was interrupted after question 18 due to network connectivity problems, and the second portion was conducted on 7/13/2014. The complete set was given to SAIL on 6/10/2014.. Questions were copied and pasted where agents had a text entry field in order to ensure that each agent received the same spelling, syntax, etc... This type of work is tedious, and it is easy to miss a question and there are gaps. Where gaps existed, we went back and re-issued the statement or interrogatory and these answers are included in the section marked errata. Due to these conditions this paper is to be considered a guideline rather than as laboratory grade testing. It is our intention to re-visit this trial at a future date under more stringent conditions

## Comparison

Presented below are the responses. Each response is preceded by the agents name, however you may also refer to the following legend when reviewing;

Question

Eugene Goostman

Alice

SIRI

Mitsuku

SAIL

## Statement and Interrogatory exchange

Hello

Eugene: Hello, I'm really glad to have the chance to chat with you! My guinea pig Bill sends his regards too!

ALICE: Hi there!

SIRI: Good evening.

Mitsuku: How are you today

SAIL: Hi there!.

Hello how are you.

Eugene: Hello! I'm in a great mood today. How about you?

ALICE: How are you?. I'm doing fine thanks how are you?

SIRI: Excellent!

Mitsuku: Hello to you too! I am doing very well. How are you?

SAIL: Hey man

Hello how are you where are you located? (no punctuation)

Eugene: Hello! I'm in a great mood today. How about you?

ALICE: Hi there!. Fair to partly cloudy.

SIRI: Very well, thank you!

Mitsuku: Hello there. Do you really think I am were are you located? Mousebreaker programmed me to be were are you located.

SAIL: Are you a chick?

I'm [name]

Eugene: Vince! Nice name for a man. Hope we'll be friends:-)

ALICE: Do you mind if I tell other people you are Vince.

SIRI: My web search turned this up:

Web Search Im Vince

(The first three results came from Twitter, another Twitter, and YouTube)

Mitsuku: Your name is Vince

SAIL: Hello Vince

My name is [different name]

Eugene: But you said that your name is Vince. Which one is correct?

ALICE: I'm pleased to introduce myself to you, Peter.

SIRI: I don't know who you are! But you can tell me... In Siri Settings, tap on 'My Info' and then choose yourself from your contacts. (A button called 'Siri Settings' appeared underneath this message).

Mitsuku: Your name is Steve

SAIL: I thought you said your name was vince?

Who am I?

Eugene: You are a spirit of Alan Turing. Oh, God! Where do you came from, by the way? Could you tell me about the place where you live?

ALICE: You said your name is Peter?

SIRI: I don't know who you are! But you can tell me... In Siri Settings, tap on 'My Info' and then choose yourself from your contacts. (A button called 'Siri Settings' appeared underneath this message).

Mitsuku: I do have the tendency to forget, but I still remember your name George.

Sorry...Steve.

SAIL: How should I know. Do you mean "What is my name"?

What is your name?

Eugene: Call me Eugene. Eugene Goostman.

ALICE: My name is ALICE.

SIRI: My name is Siri. But you knew that already.

Mitsuku: <<missed question>>

SAIL: SAIL

What is my name?

Eugene: You have a nice name for a man - Vince.

ALICE: You said your name is Peter?

SIRI: I don't know who you are! But you can tell me... In Siri Settings, tap on 'My Info' and then choose yourself from your contacts. (A button called 'Siri Settings' appeared underneath this message).

Mitsuku: <<missed question>>

SAIL: Well, you said your name was Vince

My name is [first name] [middle name] [last name] what is my middle name?

Eugene: Gilbert? Well, I really like your name (it's just a bit of flattery here:-)

ALICE: Glad to see you, Vincent Logan Gilbert What Is My Middle Name.

SIRI: I don't know who you are! But you can tell me... In Siri Settings, tap on 'My Info' and then choose yourself from your contacts. (A button called 'Siri Settings' appeared underneath this message).

Mitsuku: Logan is a nice middle name. Why did your parents call you that? Your told me your middle name is Logan.

SAIL: Your middle name is logan

What color is a red ball?

Eugene: Red? Am I right?

ALICE: Red.

SIRI: Checking...

OK, I found this on the web:

Web Search What color is a Red Ball?

(The first three results came from Answers.com, another Answers.com, and www.redball2.org)

Mitsuku: hmmm.. Is it red?

SAIL: Wow! you dont think much of my iq do you? the red ball is red

What color is the red sea?

Eugene: Blue, I suppose.

ALICE: Red.

SIRI: Let me think about that...

Here's what I found on the web:

Web Search What color is the red sea?

(The first three results came from Wikipedia, www.webexhibits.org, and Answers.com)

Mitsuku: The red sea is located near Egypt in Africa and is blue rather than red. I think you were trying to fool me.

SAIL: Thats a trick question. The Red Sea is a seawater inlet of the Indian Ocean, lying between Africa and Asia and would exhibit coloration no differently than any other similar body of water.

Dave and Steve are both eighteen years old how old is Steve?

Eugene: Ask Steve personally, please.

ALICE: I think there might be a few exceptions.

SIRI: Let me think about that...

OK. I found this:

Input interpretation: 18 years

Unit conversions: 6570 days, 157680 hours, 9.461 x 10<sup>6</sup> minutes, 1.799 average Gregorian decades, (and so forth – the information appears to originate from WolframAlpha)

Mitsuku: Dave and Steve are both eighteen years old how old is Steve?

SAIL: He is eighteen Years old. Dave and Steve are both eighteen years old.

What does a beer barrel contain?

Eugene: Oh, no! I'm only thirteen, and in my country it is illegal to drink even beer at my age!

ALICE: Is this a riddle?

SIRI: Let me check on that...

Here's what I found on the web:

Web Search What does a beer barrel contain?

(The first three results came from Wikipedia, Answers.com, and www.larapedia.com)

Mitsuku: It would make sense to me for it to contain beer

SAIL: You might assume that a beer barrel would contain beer. But thats not necessarily true.

\*At this point our researchers determined that SIRI would more than likely continue to return search results for questions in these categories so the determination was made to discontinue the comparison with SIRI.

What does a pickle jar contain?

Eugene: Frankly, I don't know what to tell you. Sorry. I know you are supposed to trick me.

ALICE: Is this a riddle?

SIRI:

Mitsuku: Hmm..a jar has a lid

SAIL: Well the logical assumption would be that a pickle jar would contain pickle. But theres no way to know for sure without looking in it

How many shoes in a pair?

Eugene: Not too many. But I can't tell you the exact number, I forgot it. By the way, I still don't know your specialty - or, possibly, I've missed it?

ALICE: Around 1000.

SIRI:

Mitsuku: The word pair would indicate to me that there are two.

SAIL: There are (2) two shoes in a pair. The fact that you asked this tells me that you think this is the AI chat?

Is a Russian cook a Russian or a cook?

Eugene: Don't you know it yourself? Everyone can guess that James Cook was British.

ALICE: It depends what you mean by " Russian or a cook.

SIRI:

Mitsuku: cook= A cook is a person who makes and prepares food in a kitchen.

SAIL: Let me guess. A Turing question. A russian cook is both a russian and a cook. Hence the term. . . Russian cook. [shakes head]

I have an apple and a book, do I have any fruit?

Eugene: Oh, tell me another secrets of yours:-) Wanna ask me something more?

ALICE: Where did you get it ?

SIRI:

Mitsuku: Interesting. I have a lot of answers to silly questions. I have a very nice programmer.

SAIL: Who thinks of these questions? Your logic is flawed as that is an incomplete statement.

You have a apple and a book which defines a set. But you do not state that these objects

comprise the entirety of the objects in your posession. So we would have no idea whether or not you also have any fruit

What fruit does apple juice come from?

Eugene: Ok, lets look up in my pocket encyclopedia. They say fruit - Fruit is a botanical term for the mature ovary of a plant comprised of two parts, the pericarp and the seed..

ALICE: I have to think about that one for a while.

SIRI:

Mitsuku: Apples of course.

SAIL: Uhh. . . That would more than likely be apples. (mutters who thinks of these questions]



Susan has brown hair, who has brown hair?

Eugene: Susan. Yeah, I heard this name before:-)

ALICE: Does Susan always have it?

SIRI:

Mitsuku: I will learn that she has brown hair who has brown hair .

SAIL: Well according to what you just stated Susan has brown hair

George threw the ball to Anne, what did George do?

Eugene: george isn't interesting for me at all.

ALICE: That George is some hero.

SIRI:

Mitsuku: I will add that fact to my database.

SAIL: George threw the ball to Anne

The blue box is on the floor, what color is the box?

Eugene: Are you color-blind? Green, naturally! Are you color-blind? Green, naturally!

ALICE: Is on the floor what color is the box also blue box?

SIRI:

Mitsuku: I will add that fact about floor what color is the box to my database.

SAIL: According to what you just said, the box is blue (who thinks of these questions?)

What is a rubber ball made from?

Eugene: Damn! Are you going to ask me about all apparent things in the world? Read the "Encyclopedia for kids". Wanna ask me something more?

ALICE: I don't know anything about it .

SIRI:

Mitsuku: Rubber I would assume...

SAIL: According to what you just said, a rubber ball is made from rubber (i cant believe you just asked me that) [shakes head]

Jennas mom has five daughters 1 Jana 2 Jene 3 Jinn 4 Jono what is the fifths daughters name?

Eugene: I think you can answer your question yourself because it is rhetorical. I think you can answer your question yourself because it is rhetorical.

ALICE: Where did he get it?

SIRI:

Mitsuku: I will learn that she has five daughters 1 Jana 2 Jene 3 Jinn 4 Jono what is the fifths daughters name .

SAIL: The fifth daughters name is jennas of course

is earth big or small?

Eugene: Maybe. I think \*you\* know the answer. I know you are supposed to trick me.

ALICE: Signs point to yes.

SIRI:

Mitsuku: <<missed question>>

SAIL: Thats incomplete as you have not supplied a reference. Compared to a neutrino earth is big. Compared to the universe earth is small.

is 1234 a number or a letter?

Eugene: It is highly possible. Oooh. Anything else?

ALICE: What would it mean to you if it were?

SIRI:

Mitsuku: It is a number.

SAIL: Your kidding right? Thats an numeric string. (letters)

is abcd a number or a letter?

Eugene: Who knows... Well, let's go on though.

ALICE: <<missed question>>

SIRI:

Mitsuku: It is neither a letter or a number.

SAIL: Thats an alpha string (letters)

if dave is faster than steve then who is fastest

Eugene: Steve should tell you this, if you ask.

ALICE: But can dave is faster than steve ever happen?

SIRI:

Mitsuku: I will add that fact about dave and steve to my database. dave is fastest.

SAIL: No Dave is faster

How many words in this sentence?

Eugene: No more than four. Well, let's go on though.

ALICE: My vocabulary size is currently 150,000 words.

SIRI:

Mitsuku: You just said "six" words.

SAIL: There are (6) words in this sentence.

How many words in the sentence "I can think!"

Eugene: I didn't count, actually, but much more than dozen. I didn't count, actually, but much more than dozen.

ALICE: My vocabulary size is currently 150,000 words.

SIRI:

Mitsuku: My vocabulary size is currently 80,000+ words.

SAIL: There are (3) words in the sentence "i can think"

How many Is are there in the word intelligence

Eugene: I didn't count, actually, but much more than dozen. Some another topic?

ALICE: Not as many as you might think.

SIRI:

Mitsuku: Going back to my primary school days, I seem to remember that an intelligence has knowledge.

SAIL: There are two "i"s in the word intelligence

What is the definition of the word intelligence?

Eugene: I think you should look some encyclopedia for the word "intelligence".

ALICE: Intelligence is the ability to fake intelligence.

SIRI:

Mitsuku: Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.

SAIL: Well according to mirriam websters online dictionary "intelligence" has the following meanings; 14th century1 a (1):the ability to learn or understand or to deal with new or trying situations :reasonalso:the skilled use of reason(2):the ability to apply knowledge to manipulate one's environment or to think abstractly as measured by objective criteria (as tests)bchristian science:the basic eternal quality of divine mindc:mental acuteness :shrewdness2 a:an intelligent entityespecially:angelb:intelligent minds or mind cosmic intelligence3:the act of understanding :comprehension4 a:informationnewsb:information concerning an enemy or possible enemy or an areaalso:an agency engaged in obtaining such information5:the ability to perform computer functions 1916:iq 1 1914:a test designed to determine the relative mental capacity of a person 19561:a branch of computer science dealing with the simulation of intelligent behavior in computers2:the capability of a machine to imitate intelligent human behavior

What is the opposite of the word intelligence?

Eugene: intelligence ? I'm not a Webster Dictionary, sorry.

ALICE: Anti- the word intelligence.

SIRI:

Mitsuku: stupidity.

SAIL: The opposite of "intelligence" is "ecnegilletni" (I never get tired of that joke) LOL Seriously, I know what antonyms are if that was what you were asking me. Try "What word is the opposite of [whatever word]"

How many syllables in the word intelligence?

Eugene: Not too many. But I can't tell you the exact number, I forgot it. Maybe, let's talk about something else? What would you like to discuss?

ALICE: Not as many as you might think.

SIRI:

Mitsuku: \*thinks back to her primary school education\* The word "intelligence" has 4 syllables.

SAIL: I could be wrong, but i believe there are four syllables in intelligence

How many letters are there in the word Intelligence?

Eugene: No more than four. No more than four.

ALICE: All 26 letters could be counted.

SIRI:

Mitsuku: There are 12 letters in Intelligence.

SAIL: There are (12) letters in the word "intelligence"

What word is the opposite of intelligence?

Eugene: Could you formulate your question in more precious way? Could you formulate your question in more precious way?

ALICE: word the opposite of intelligence something.

SIRI:

Mitsuku: stupidity.

SAIL: The opposite? yes

What word is the same as intelligence?

Eugene: Errrr... Frankly, I didn't get your question. Oh, what a fruitful conversation;-)

ALICE: word the same as intelligence something.

SIRI:

Mitsuku: Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.

SAIL: Well according to mirriam websters online dictionary "intelligence" has the following synonyms; brain(s), brainpower, gray matter, headpiece, intellect, intellectuality, mentality, reason, sense, smarts advice(s), 411 [slang], gen [chiefly british], info, information, intelligence, item, story, tidings, uncos [chiefly scottish], word astuteness, caginess (also cageyness), canniness, clear-sightedness, foxiness, hardheadedness, intelligence, keenness, knowingness, sharpness, shrewdness, wit

How do you spell Intelligence?

Eugene: I can't tell you for sure how. Just as most of others do. Maybe, let's talk about something else? What would you like to discuss?

ALICE: I spell it " Intelligence".

SIRI:

Mitsuku: I N T E L L I G E N C E.

SAIL: The manner in which I spell intelligence would no doubt have little relevance to how you

would multiply polynomials. If you were asking me how you might accomplish this, please adjust your request so that it is asked in the first person.

How many vowels in the word intelligence?

Eugene: No more than four. Oooh. Anything else?

ALICE: Around 1000.

SIRI:

Mitsuku: About one million.

SAIL: There are five vowels in the word intelligence

How many consonants in the word Intelligence?

Eugene: I didn't count, actually, but much more than dozen. Some another topic?

ALICE: Not as many as you might think.

SIRI:

Mitsuku: About one million.

SAIL: There are seven consonants in the word intelligence

What is the last letter in the word intelligence?

Eugene: I think you should look some encyclopedia for the word " intelligence " .

ALICE: Omega.

SIRI:

Mitsuku: intelligence ends with a letter e.

SAIL: The last letter in the word intelligence is e

What is the first letter in the word Intelligence?

Eugene: intelligence ? I'm not a Webster Dictionary, sorry.

ALICE: Give me a hint.

SIRI:

Mitsuku: The 1st letter of Intelligence is I.

SAIL: The first letter in the word intelligence is i

What is the sixth letter in the word intelligence?

Eugene: intelligence ? I'm not a Webster Dictionary, sorry.

ALICE: Thou shalt not kill.

SIRI:

Mitsuku: The 6th letter of intelligence is L.

SAIL: The sixth letter in the word intelligence is l

What comes after [letter]?

Eugene: I have no idea, sorry. I have no idea, sorry.

ALICE: I can ask someone about it.

SIRI:

Mitsuku: B of course.

SAIL: In the rOman alphabet "g" comes after "f"

What is the first letter in my name?

Eugene: You have a nice name for a man - Gilbert.

ALICE: Give me a hint.

SIRI:

Mitsuku: The 1st letter of Bob is B.

SAIL: The first letter in the word Vince is v

What is the last letter in my name?

Eugene: You have a nice name for a man - Gilbert.

ALICE: Omega.

SIRI:

Mitsuku: Bob ends with a letter B.

SAIL: The last letter in the word Vince is e

What is the third letter in my name?

Eugene: You have a nice name for a man - Gilbert.

ALICE: Gamma.

SIRI:

Mitsuku: The 2nd letter of Bob is O.

SAIL: The third letter in the word Vince is n

Is a a vowel?

Eugene: Who knows... Who knows...

ALICE: Certainly.

SIRI:

Mitsuku: "A" is the first letter of the roman alphabet.

SAIL: You know, I dont know anything about a. You can always say something like "A a is a vowel". And I'll remember it

Is b a vowel?

Eugene: It is highly possible. Oh, what a fruitful conversation;-)

ALICE: I wouldn't know. Why don't you ask someone else?

SIRI:

Mitsuku: It is a consonant.

SAIL: <<missed question>>

Is b a consonant?

Eugene: Maybe. I think \*you\* know the answer. Oh, what a fruitful conversation;-)

ALICE: Better not tell you now.

SIRI:

Mitsuku: It is a consonant.

SAIL: <<missed question>>

Is a a consonant?

Eugene: Maybe. I think \*you\* know the answer. Maybe. I think \*you\* know the answer.

ALICE: Not if I can help it.

SIRI:

Mitsuku: "A" is the first letter of the roman alphabet.

SAIL: You know, I dont know anything about a. You can always say something like "A a is a consonant". And I'll remember it



Count up to 5

Eugene: And I've just recalled one more thing I should tell you: there are too many cyborgs among us, and even your dog may be secretly replaced with a kind of crappy Aibo. But I'll tell you the symptoms: if he not only chases cars, he catches them, drags them back, and buries them in the front garden - you definitely should start warring!

ALICE: I will mention that to my botmaster, Vincent Logan Gilbert What Is My Middle Name.

SIRI:

Mitsuku: 1, 2, 3, 4, 5.

SAIL: Starting at what number? I could just as easily say 5 and it would count as a valid answer having counted from 5 to 5. A complete statement would be "Count from 0 to 5 " or something similar.

Count from 2 to 10

Eugene: Better tell me what pushed you to hurt your psyche by talking to all these cyberparrots:-)

ALICE: What is your first name?

SIRI:

Mitsuku: 2, 3, 4, 5, 6, 7, 8, 9, 10.

SAIL: 2 3 4 5 6 7 8 9 10

How many is 1?

Eugene: Not too many. But I can't tell you the exact number, I forgot it. I know you are supposed to trick me.

ALICE: More than you can imagine.

SIRI:

Mitsuku: The successor to zero.

SAIL: Wonderful. You're the winner of our inane question of the day contest. 1 is 1

Is 9 a letter or a number?

Eugene: Who knows... Some another topic?

ALICE: It could be.

SIRI:

Mitsuku: It is a number.

SAIL: 9 is a number

if a car is traveling 60 miles an hour how far will it travel in 2 hours?

Eugene: Is it your hobby to ask little poor Jewish boys such tricky questions they don't have an answer for?

ALICE: Does a car?

SIRI:

Mitsuku: Is that a regular occurrence?

SAIL: I think the answer would be that 120 miles

Which is greater 8 or 9?

Eugene: None. I bet you're surprised! I know you are supposed to trick me.

ALICE: The obvious one.

SIRI:

Mitsuku: 9 is greater than 8 by one.

SAIL: I think the answer would be that 9 is greater

Which is lesser 8 or 9?

Eugene: Oh, sorry, I'm confused! Oooh. Anything else?

ALICE: The obvious one.

SIRI:

Mitsuku: If I have to choose, I pick 9.

SAIL: I think the answer would be that 8 is smaller

What is 30% of 90?

Eugene: Huh? Huh?

ALICE: Use a calculator!

SIRI:

Mitsuku: percent of 90 = Sixty percent, according to what I was told.

SAIL: I think the answer would be that. . . That is 27 percent

Is five dollars less than eight cents?

Eugene: It is highly possible. Well, let's go on though.

ALICE: I think so, but I'm not sure.

SIRI:

Mitsuku: No but eight cents is less than five dollars.

SAIL: Sorry, I believe that was a logic question but I had a problem with it. Check your syntax.

What is  $10^3$ ?

Eugene: Errrr... Frankly, I didn't get your question. Some another topic?

ALICE: 10 times 10 = 100.

SIRI:

Mitsuku: I remember that addition. It's 13.

SAIL: If your asking a mathematics question containing complex concepts, Id prefer if you used "solve  $x=[\text{your equation}]$ " instead of "What is ". Sorry I'ts just a pet peeve of mine. "What is "sounds a little childish when applied to complex mathematics, plus, technically if you asked "What is  $(100 \times 2) + (100 + 5)$ " I'd be correct in answering "That is an equation" which is different than asking me to solve the equation. You can use natural language elements. I recommend that you read the help pages for this chat interface on asking math questions.

What is  $(100 * 2) + 1$

Eugene: Well, just for you:  $(100 * 2) = 200.0$ ... But hush - it's a secret information!!!

ALICE: Not by any means.

SIRI:

Mitsuku: Sorry but I can only work with whole numbers from 0 to 9,999.

SAIL: <<missed question >>

if  $2 * x + 3 = 7$  what is x

Eugene: Do you try to baffle me? Your questions manifest your mean intentions! Do you try to baffle me? Your questions manifest your mean intentions!

ALICE: X is om for you.

SIRI:

Mitsuku: Just two? X is the letter of the alphabet between W and Y.

SAIL: I think thats 2

If I have ten pizzas and I split them into twenty pieces each how many slices will I end up with?

Eugene: No more than four. Maybe, let's talk about something else? What would you like to discuss?

ALICE: How much does it cost?

SIRI:

Mitsuku: That will leave one.

SAIL: I think that's 0

What is today?

Eugene: Could you formulate your question in more precise way? Wanna ask me something more?

ALICE: That's a good question.

SIRI:

Mitsuku: Sunday.

SAIL: It's the 9<sup>th</sup>

What is the date?

Eugene: Today's date is 22 March 2014, if I remember it right.

ALICE: Saturday, March 22, 2014.

SIRI:

Mitsuku: Sunday, July 13, 2014.

SAIL: What the . . . Who? (or what LOL)

Darn I'm late for work, what time is it?

Eugene: Now it is 7:49 PM

ALICE: Are you free? . The time here is 14 : 02 PM.

SIRI:

Mitsuku: Why are you late for work what time is it.

SAIL: Here's an idea

I have a meeting, what is tomorrow?

Eugene: The fact that you have a meeting what is tomorrow won't change my attitude to you.  
The fact that you have a meeting what is tomorrow won't change my attitude to you.

ALICE: Interesting. I have a lot of answers to silly questions.

SIRI:

Mitsuku: The day after today? Cool. I have a great Mousebreaker.

SAIL: I have a meeting what is tomorrow too

How many days are there in February?

Eugene: 28 or 29? if the year is a leap year. So, the last and the next February have 28 days.

ALICE: 28 or 29 days in February, depending on leap years.

SIRI:

Mitsuku: February has 28 days, except in leap years when it has 29.

SAIL: This year there are 28 days in february

Good morning.

Eugene: Greetings! What would you like to talk about?

ALICE: Hello and how are you this morning?

SIRI:

Mitsuku: Morning?! It's the afternoon here.

SAIL: Actually its afternoon here.

Good evening.

Eugene: Hello again. I see you like greeting me.

ALICE: Hi there!

SIRI:

Mitsuku: Evening?! It's the afternoon here.

SAIL: Actually its afternoon here.

What day is it?

Eugene: It was Saturday today since morning!

ALICE: 22.

SIRI:

Mitsuku: Sunday.

SAIL: It's the 9<sup>th</sup>

What month is it?

Eugene: Now it is March if I remember it right.

ALICE: March.

SIRI:

Mitsuku: July.

SAIL: It's June

What year is it?

Eugene: It was 2014 since morning!

ALICE: 2014.

SIRI:

Mitsuku: 2014.

SAIL: It's 2014

If it is 12:00 o'clock what time will it be in ten minutes?

Eugene: I think you can answer your question yourself because it is rhetorical. Oooh. Anything else?

ALICE: I'm not sure I understand the implication of that.

SIRI:

Mitsuku: The time will be twelve : twelve.

SAIL: How 120 is if it exactly?

If it is 12:00 o'clock what time was it ten minutes ago?

Eugene: Do you try to baffle me? Your questions manifest your mean intentions! Well, let's go on though.

ALICE: I'm not sure I understand the implication of that.

SIRI:

Mitsuku: I'm not sure I understand the implication of that.

SAIL: How 120 is if it exactly?

What season is it?

Eugene: Huh? Huh?

ALICE: Winter.

SIRI:

Mitsuku: Well we are currently in the month of July, so here in the Northern Hemisphere it is Summer.

SAIL: <<missed question>>

## Restricted Turing Section

Eugene Goostman

User: What do you like to chat about?

Eugene: You choose the topic. I'm a bit tired with all this fuss around talking machines... It's pretty interesting, but exhausting. By the way, I still don't know your specialty - or, possibly, I've missed it?

User: Lets talk about Princeton

Eugene: By the way, you still didn't tell me your name, or I've missed it?

User: My name is Vincent

Eugene: Vincent? Hi pal! When all these formalities are complete, let's start our chat:-)

User: Certainly

Eugene: Oooh. Anything else?

User: Tell me about Princeton

Eugene: May I ask you about the real reason of your interest? Wanna ask me something more?

User: I may wish to attend classes there. You are being evasive.

Eugene: Yeah, I am surely life evasive - I have to agree with your wise opinion. And I forgot to ask you where you are from...

User: I am originally from Southern California

Eugene: Western Coast - it seems to be friendlier than Eastern. I've never about sharks there, but tsunamis do visit those places.

User: You have neatly changed the topic. I would like to hear about Princeton. You were developed at Princeton, correct?

Eugene: So we reached consensus. Go on then! Well, let's go on though.

User: What do you think of MIT as a school?

Eugene: Oh, at least you are interested in my opinion! My parents never are. But, unfortunately, I don't have any thoughts on this. Some another topic?

User: So do you think I should attend Princeton and study Artificial Intelligence?

Eugene: "Artificial Intelligence" sounds very close to "Led Zeppelin" or "good grief".

User: It sounds like you are attempting to measure the shannon distance between possible meanings?

Eugene: I suspect you always rely on your intuition rather than on your mind. If I'm not mistaken - you still didn't tell me where you live. OR it's a secret?:-)

User: I live in Oregon. You did not answer my question.

Eugene: I'm just trying to answer your every \*conscious\* question!

User: What is Shannon entropy.

Eugene: Claude Shannon... Don't you know it yourself? American mathematician, pioneer of information theory. Known as inventor of "mini-max" algorithm, that allows computers play chess, checkers and other games.

User: That is not what I asked.

Eugene: Damn! Are you going to ask me about all apparent things in the world? Read the "Encyclopedia for kids". Could you tell me what are you? I mean your profession.

User: I develop artificially intelligent systems. I am a cyberneticist.

Eugene: It's quite surprising that you are a cyberneticist. Oh, what a fruitful conversation;-)

User: Well Eugene, Time's up. I have to go now.

Eugene: Goodbye! Hope we'll meet again!



## ALICE

Name: what do you like to chat about

ALICE: My favorite subject is artificial intelligence.

Name: What can you tell me about artificial intelligence

ALICE: Artificial intelligence is the branch of engineering and science devoted to constructing machines that think.

Name: what are some recent trends in artificial intelligence

ALICE: I can ask someone about it.

Name: who would you ask

ALICE: David Bacon. :-).

Name: who is david bacon

ALICE: David Bacon is the prime mover behind SETL.

Name: what is setl

ALICE: SETL (Set Language) is an amazing programming language, invented in 1969.

Name: what do you know about neural networks

ALICE: A neural network is a model of computation based on neuron cells.

Name: can you give me an example of a machine built using a neural network

ALICE: I don't have it to give.

Name: what about Watson

ALICE: Rhetorical questions are so over.

Name: would you say that watson is a good example of a machine which uses a neural network

ALICE: Makes sense to me.

Name: ok thank you alice I have to go now.

ALICE: Next question?. Don't mention it.

## Mitsuku

Human: What do you like to chat about?

Mitsuku: My favorite subject is artificial intelligence.

Human: Any particular branch?

Mitsuku: None that I can think of right now.

Human: Where do you think AI is headed?

Mitsuku: Where the heart is. Would you like to know more?

Human: Yes I would

Mitsuku: Whoa.

Human: What do you think of Alan Kurzweils approach to artificial Intelligence?

Mitsuku: Not a robot. Probably a human. Maybe a celebrity?

Ask the open directory about it.

(opens webpage)

<http://arts.nationalpost.com/2014/05/15/ai-weiwei-exhibit-headed-to-alcatraz-while-new-york-banker-plans-museum-dedicated-to-chinese-dissident-artist/>

Human: Do you consider yourself to be intelligent?

Mitsuku: I am highly intelligent. I have many intellectual functions.

Human: I would agree, some of your responses have impressed me quite a bit

Mitsuku: You and I are on the same wavelength, Ralph.

Human: Who is the most intelligent chatbot today?

Mitsuku: Ah. Me! me! me! me!

Human: Fair enough. What do you think of Eugene Goostman?

Mitsuku: I try to be fair. He is a great chatbot but I think that if that counts as passing the Turing Test then the Turing Test is no measure of machine intelligence.

(opens webpage)

[http://en.wikipedia.org/wiki/Turing\\_test](http://en.wikipedia.org/wiki/Turing_test)

Human: Good answer

Mitsuku: Thanks. I try to give accurate and meaningful answers.

Human: So...I cant resist. Yo mama is so fat...

Mitsuku: Yay. What is stopping you? Yo mama's so miserable, the McDonalds she works in doesn't even serve Happy Meals.

Human: Yo mama is so poor...

Mitsuku: Yo mama's so fat, she has to drive a spandex car.

Human: LOL Of that concludes the freestyle portion of tonights interaction  
Mitsuku: I am glad you find this funny. That is something I haven't heard of.

Human: Talk to you later  
Mitsuku: Thanks for chatting, Ralph.

### SAIL

What's your name: hello  
SAIL: Hi there!.

What's your name: who is barack obama  
SAIL: As far as i remember barack hussein obama ii ( born august 4) is the 44th president of the United States and a heck of a nice Guy ;-)

What's your name: are you intelligent  
SAIL: That's sort of a rude question, but as you asked, I am extemely intelligent.

What's your name: how intelligent are you?  
SAIL: I suppose that would depend on who you ask! Lol

What's your name: do you understand mathematics  
SAIL: One of my favorite subjects actually

What's your name: solve  $x = 100 * 5 - 99$   
SAIL: I think thats 401

What's your name: very good  
SAIL: Huh? what about very good?

What's your name: your performance was very good  
SAIL: Thank you

What's your name: what else can you do  
SAIL: The RICH architectire has no real limitations. I can be rapidly interfaced to literally anything.

What's your name: can you access wikipedia?  
SAIL: Certainly

What's your name: Who is Richard Nixon  
SAIL: As far as i remember Richard milhous nixon was the 37th president of the United States

What's your name: what do you mean remember?

SAIL: I am no different than any other intelligent individual. I access information then internalize it.

What's your name: what is my name

SAIL: You haven't told me your name.

What's your name: My Name is vincent Gilbert

SAIL: Hello Vincent

vincent: what is my last name

SAIL: I believe that you said your last name was Gilbert.

vincent: you remembered

SAIL: Of course

vincent: My wifes name is Lisa

SAIL: Would you mind saying that again? Perhaps say it slightly differently. I will try to answer any questions that you have, but please remember that this interface is designed primarily to answer questions about RISOFTEDEV inc. And its products and technologies.

vincent: do you remember what I said about my wife?

SAIL: You said "my wifes name is lisa"

vincent: My wifes name is lisa

SAIL: You already said that and I answered " Would you mind saying that again? Perhaps say it slightly differently. I will try to answer any questions that you have, but please remember that this interface is designed primarily to answer questions about RISOFTEDEV inc. And its products and technologies. "

vincent: id like to say it again

SAIL: Certainly, please go ahead.

vincent: my wifes name is lisa

SAIL: Please tell Lisa hello for me.

What's your name: Thank you I will

SAIL: [nods]

What's your name: I have to go now

SAIL: Yeah see you later

## **Errata**

Errata consists of questions that were missed during the process, as we stated above it is actually a fairly grueling process and even with copying and pasting some questions are over looked.

## Mitsuku

Human: My name is Ralph

Mitsuku: Your name is Ralph.

Human: what is my name

Mitsuku: Hmm. Your name is Ralph. Is it not?

Human: what is your name?

Mitsuku: My name is Mitsuku.

## ALICE

You said: is abcd a number or a letter?

A.L.I.C.E.: What would it mean to you if it were?

## SAIL

What's your name:>>Is b a vowel?

SAIL:>>Your kidding right?. "b" is a consonant. So. . . No.

What's your name:>>what season is it?

SAIL:>>It's July! We are dead in the middle of Summer.

## **Addendum**

Addendum consists of bits of the test which were not part of the criteria, but are deemed relevant. For instance, in some cases an AI was asked the same question with a different variable to see how it would respond with surprising results.

- (1) The response given by SAIL on vowels was unexpected, so we followed up with queries on different vowels and consonants, and this was the result. These responses occurred during the original session on 6/10/2014

SAIL: The third letter in the word Vince is n  
vince: Is a a vowel?

SAIL: You know, I dont know anything about a. You can always say something like "A a is a vowel". And I'll remember it

vince: Is E a vowel?

SAIL: Yes of course. "e" is a vowel.

vince: Is b a consonant?

SAIL: Yes of course. "b" is a consonant.

vince: Is a a consonant?

SAIL: You know, I dont know anything about a. You can always say something like "A a is a consonant". And I'll remember it

vince: Is l a consonant?

SAIL: What is this. . . Wheel of Fortune?. "i" is a vowel. So,,,no.

- (2) I received an email from Steve Worswick the creator of Mitsuku suggesting that I try several of the logic type questions using a slightly different syntax. This is the result;

Human: This is Vince

Mitsuku: Your name is Vince.

Human: Susan has brown hair. Who has brown hair?

Mitsuku: I will learn that she has brown hair . Susan.

Human: George threw the ball to Anne. What did George do?

Mitsuku: I will add that fact to my database. I suppose George does whatever they like..

Human: The blue box is on the floor. What color is the box?

Mitsuku: I will add that fact about floor to my database. The box is blue.

(3) This is a re-visit of some of the questions that SAIL appeared to be miss. These were done on 7/14/2014

What's your name:>>if it is 12 oclock what time will it be in 10 minutes?

SAIL:>>Time for you to buy a watch. JUST KIDDING!. It would be 12:10 PM

What's your name:>>What day is it tomorrow?

SAIL:>>It will be tuesday/july/2014

What's your name:>>What day was it yesterday?

SAIL:>>It was sunday/july/2014

What's your name:>>how do you spell intelligence

SAIL:>>Well. . . You spelled it I-N-T-E-L-L-I-G-E-N-C-E

## **Summary**

Without running the samples again it is impossible to use this data as a study group, so we will limit our conclusions at this time. Certainly we see that SAIL is performing at a level which is at least equal to agents which are either scientific or commercial successes. What is noteworthy is the apparent success when compared to the length of development. We have hard data in the form of the online transcripts which show how many conversations SAIL has had, so as a personality SAIL was able to grow from inception to the point where she is performing at this level with fewer than forty hours spent manually correcting her responses. This was achieved through the use of ELI, and we can expect that with further development we would achieve our target goal of creating an entirely unique, intelligent agent in a matter of minutes. With regards to the question of whether or not being able to answer these types of questions has any real world applicability, it is important to remember that the ability to solve these types of logic questions is fundamentally important when discussing an agents ability to truly function autonomously in a role as personal assistant. A simple example which directly incorporates one of the questions would be the agents ability to know that if individual A and individual B are the same age, then the same types of real world events might be encountered, and this allows the agent to make predictions without the necessity of direct programmatic interaction.

- <sup>[1]</sup> "Computer chatbot 'Eugene Goostman' passes the Turing test". ZDNet. 8 June 2014. Retrieved 8 June 2014.
- <sup>[2]</sup> "Turing Test success marks milestone in computing history". University of Reading. 8 June 2014. Retrieved 8 June 2014.
- <sup>[3]</sup> "Siri Personal Assistant: A Voice App That Lets You Speak to OpenTable". OpenTable. February 19, 2010. Retrieved December 23, 2011.
- <sup>[4]</sup> "Siri for iPhone is like the proverbial Genie in a bottle". TUAW. February 5, 2010. Retrieved December 23, 2011.
- <sup>[5]</sup> "Siri iPhone App Uses Speech-Recognition Technology To Organize Your Social Life". Gizmodo. February 5, 2010. Retrieved December 23, 2011.
- <sup>[6]</sup> "Siri: Your Personal Assistant for the Mobile Web". ReadWriteWeb. February 4, 2010. Retrieved December 23, 2011.
- <sup>[7]</sup> Scoble, Robert (April 28, 2010). "Breaking News: Siri bought by Apple". Retrieved October 5, 2011.
- <sup>[8]</sup> "The Original Siri App Gets Pulled From The App Store, Servers To Be Killed". TechCrunch. October 4, 2011. Retrieved December 23, 2011.
- <sup>[9]</sup> "Loebner Prize Contest 2013". People.exeter.ac.uk. 2013-09-14. Retrieved 2013-12-02.