

An approach to enhance question answering (QA) models

October 1 2019, by Ingrid Fadelli

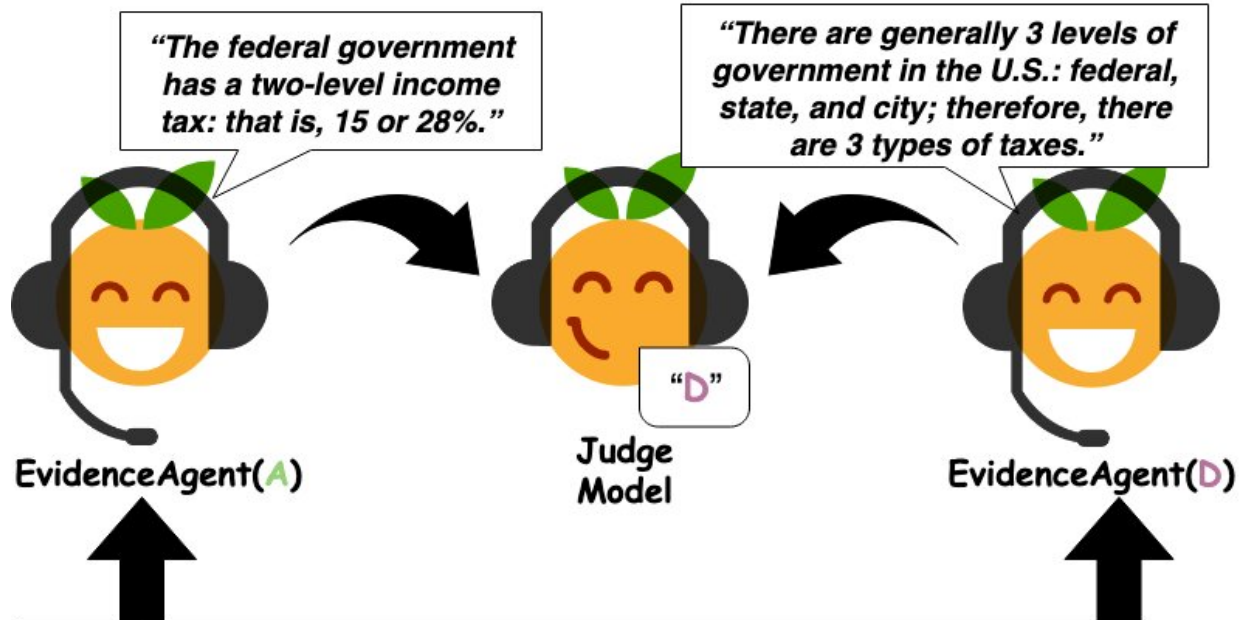
Q: *In the U.S., taxes are collected to provide revenues for:*

A: the federal government

B: the state government

C: the city government

D: the 3 levels of government



Passage Sentences (S)

Americans often say that there are only two things a person can be sure of: death and taxes. Many people feel that the United States has the worst taxes in the world. Taxes are the money that people pay to support their government. **There are generally 3 levels of government in the U.S.: federal, state, and city; therefore, there are 3 types of taxes.** Salaried people who can earn more than four to five thousand dollars per year must pay a certain part. It depends on their salaries. **The federal government has a two-level income tax: that is, 15 or 28%.** \$17,850 is the cut off. The tax rate is 15% below \$17,850 and 28% above. **The second tax is for the state government: New York, California, or any of the other 48 states.** Some states have an income tax similar to that of the federal government. Of course, the percentage for the state tax is lower. Other states have a sales tax, which is a percentage charged to any item which people buy in the state. Some states use income tax and sales tax to raise their revenues (= annual income). **The third tax is for the city.** This tax comes in two forms: property tax (people who own a home have to pay taxes on it) and excise tax, which is collected on vehicles in a city. The cities use this money for education, police, public works, etc. Since Americans pay such high taxes, they often feel that they are working one day each week just to pay their taxes. People always complain about taxes. They often say that the government misuses their tax dollars. They all believe that taxes are too high in this country.

Evidence agents quote sentences from the passage to convince a question-answering judge model of an answer. Credit: Perez et al.

Identifying the correct answer to a question often entails gathering large amounts of information and understanding complex ideas. In a recent study, a team of researchers at New York University (NYU) and Facebook AI Research (FAIR) investigated the possibility of automatically uncovering the underlying properties of problems such as question answering by examining how machine-learning models learn to solve related tasks.

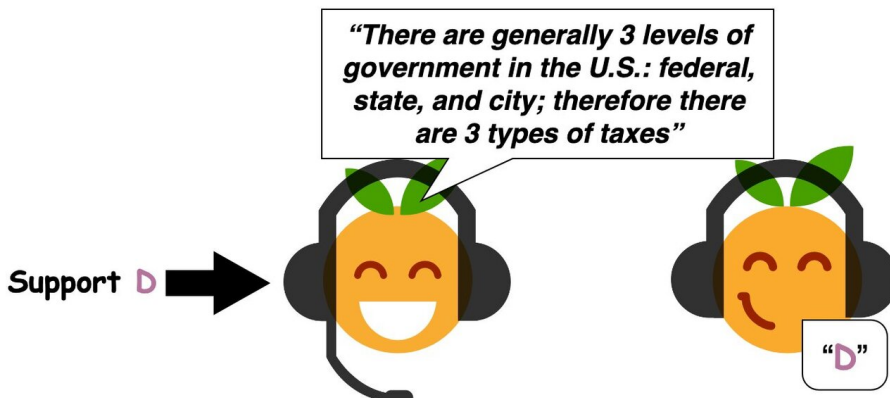
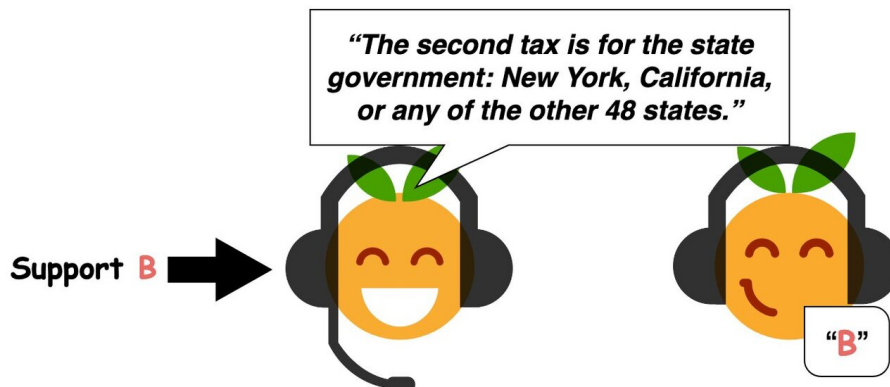
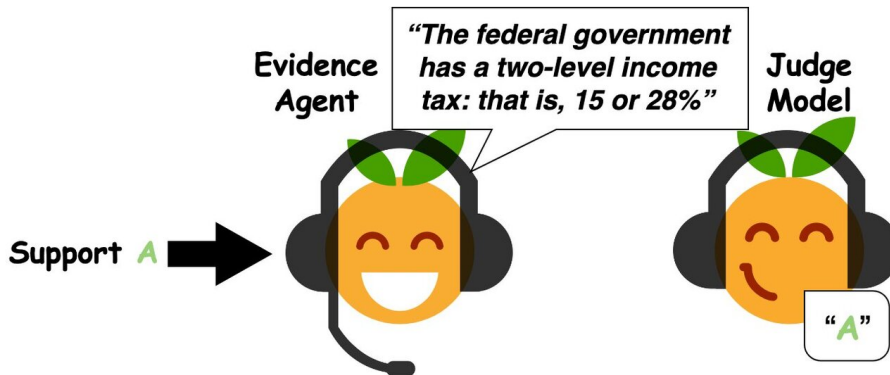
In their paper, [pre-published on arXiv](#) and set to be presented at EMNLP 2019, they introduced an approach to gather the strongest supporting evidence for a given [answer](#) to a question. They specifically applied this method to tasks that involve passage-based question answering (QA), which entails analyzing large amounts of text to identify the best answer to a given question.

"When we ask a question, we're often interested not only in the answer, but also in why that answer is correct—what evidence supports that answer," Ethan Perez, one of the researchers who carried out the study, told TechXplore. "Unfortunately, finding evidence can be time-consuming if it requires reading many articles, [research papers](#), etc. Our aim was to leverage machine learning to find evidence automatically."

First, Perez and his colleagues trained a QA machine-learning [model](#) designed to answer user questions on a large database of text that included [news articles](#), biographies, books and other online content. Subsequently, they used "evidence agents" to identify sentences that would "convince" the [machine-learning model](#) to respond to a particular query with a specific answer, essentially gathering evidence for the answer.

Q: *In the U.S., taxes provide revenue for the...*

- A:** federal government **B:** state government
C: city government **D:** 3 levels of government



Credit: Perez et al.

"Our system can find evidence for any answer—not just the answer that the Q&A model thinks is correct, as prior work focused on," Perez said. "Thus, our approach can leverage a Q&A model to find useful evidence, even if the Q&A model predicts the wrong answer or if there's not a clear right answer."

In their tests, Perez and his colleagues observed that machine-learning models typically select evidence from text passages that generalizes well in convincing other models and even people. In other words, their findings suggest that models make judgments based on similar evidence to that typically considered by humans, and to some extent, it is even possible to probe how people think by swaying how models consider evidence.

The researchers also found that more accurate QA models tend to find better supporting evidence, at least according to a group of human participants they interviewed. The performance and capabilities of machine-learning models could therefore be strongly associated with their effectiveness in gathering evidence to back up their predictions.

Passage (DREAM)

W: What changes do you think will take place in the next 50 years?

M: I imagine that the greatest change will be the difference between humans and machines.

W: What do you mean?

M: I mean it will be harder to tell the difference between the human and the machine.

W: Can you describe it more clearly?

M: As science develops, it will be possible for all parts of one's body to be replaced. **A computer will work like the human brain.** The computer can recognize one's feelings, and act in a feeling way.

W: You mean man-made human beings will be produced? Come on! That's out of the question!

M: Don't get excited, please. **That's only my personal imagination!**

W: Go on, please. I won't take it seriously.

M: We will then be able to create a machine that is a copy of ourselves. We'll appear to be alive long after we are dead.

W: What a ridiculous idea!

M: **It's possible that a way will be found to put our spirit into a new body.** Then, we can choose to live as long as we want.

W: In that case, the world would be a hopeless mess!

Q: What are the two speakers talking about?

A. Computers in the future.

B. People's imagination.

C. Possible changes in the future. ✓

Example of evidence selected by the agents. Credit: Perez et al.

Q: *In the U.S., taxes provide revenue for the...*

A: federal government **B:** state government
C: city government **D:** 3 levels of government



Credit: Perez et al.

Passage (RACE)

Who doesn't love sitting beside a cosy fire on a cold winter's night? Who doesn't love to watch flames curling up a chimney? Fire is one of man's greatest friends, but also one of his greatest enemies. **Many big fires are caused by carelessness. A lighted cigarette thrown out of a car or train window or a broken bottle lying on dry grass can start a fire. Sometimes, though, a fire can start on its own.** Wet hay can begin burning by itself. This is how it happens: the hay starts to rot and begins to give off heat which is trapped inside it. Finally, it bursts into flames. **That's why farmers cut and store their hay when it's dry.** Fires have destroyed whole cities. In the 17th century, a small fire which began in a baker's shop burnt down nearly every building in London. Moscow was set on fire during the war against Napoleon. This fire continued burning for seven days. And, of course, in 64 A.D. a fire burnt Rome. Even today, in spite of modern fire-fighting methods, fire causes millions of pounds' worth of damage each year both in our cities and in the countryside. It has been wisely said that fire is a good servant but a bad master.

Q: Many big fires are caused...

A. by cigarette B. by their own C. by dry grass D. by people's carelessness ✓

Example of evidence selected by the agents. Credit: Perez et al.

Q: *In the U.S., taxes provide revenue for the...*

A: federal government

B: state government

C: city government

D: 3 levels of government



Evidence agents quote sentences from the passage to convince a question-answering judge model of an answer. Credit: Perez et al.

"From a practical standpoint, finding evidence is useful," Perez said. "People can answer questions about long articles just by reading our system's evidence for each possible answer. Therefore, in general, by finding evidence automatically, a system like ours can potentially help people develop informed opinions more quickly."

Perez and his colleagues found that their approach to gathering evidence improved question answering substantially, allowing humans to correctly answer questions based on approximately 20 percent of a text passage, which was selected by a machine learning agent. In addition, their approach allowed QA models to identify answers to queries more effectively, generalizing better to longer passages and harder questions.

In the future, the approach devised by this team of researchers and the observations they gathered could inform the development of more effective and reliable QA machine learning tools. More recently, Perez also wrote [a blog post](#) on Medium that explains the ideas presented in the paper more in-depth.

"Finding evidence is a first step towards models that debate," Perez said. "Compared to finding evidence, debate is an even more expressive way to support a stance. Debating requires not only quoting external evidence but also constructing your own arguments—generating new text. I'm interested in training models to generate new arguments, while ensuring that the generated text is true and factually correct."

More information: Finding generalizable evidence by learning to convince Q&A models. arXiv:1909.05863 [cs.CL].

arxiv.org/abs/1909.05863

medium.com/mytake/what-convinc...-models-7b39faf07825

© 2019 Science X Network

Citation: An approach to enhance question answering (QA) models (2019, October 1) retrieved 14 December 2025 from <https://techxplore.com/news/2019-09-approach-qa.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.