A Contributor and patent explorer reviews a recently-presented paper that suggests Google is grouping entities and using their relationships to listen for better answers to multipart questions.

At the Sixth International Conference on Learning Representations, Jannis Bulian and Neil Houlsby, researchers at Google AI, presented a paper that shed light on new methods they're testing to improve search results.

While publishing a paper certainly doesn't mean the methods are being used, or even will be, it likely increases the odds when the results are highly successful. And when those methods also combine with other actions Google is taking, one can be almost certain.

I believe this is happening, and the changes are significant for search engine optimisation specialists (SEOs) and content creators.
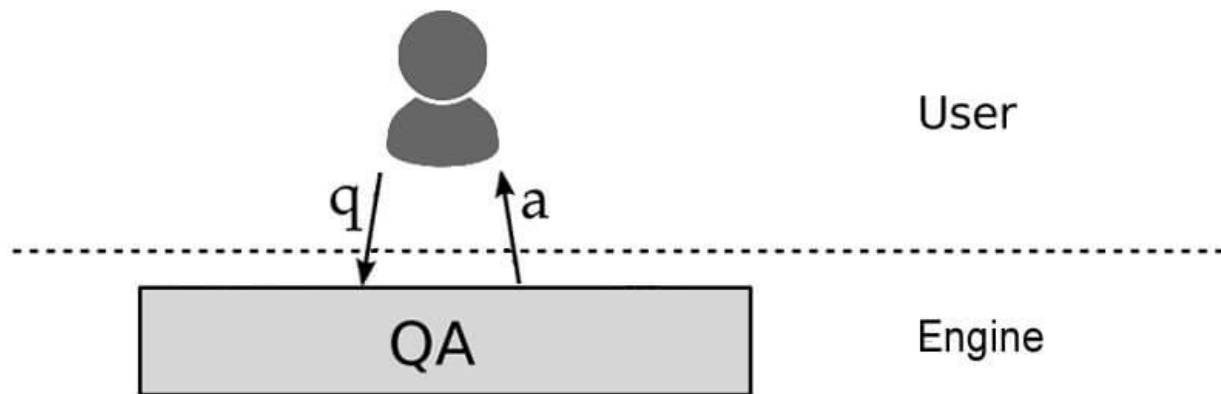
So, what's going on?

Let's start with the basics and look topically at what's being discussed.

A picture is said to be worth a thousand words, so let's start with the primary image from the paper.
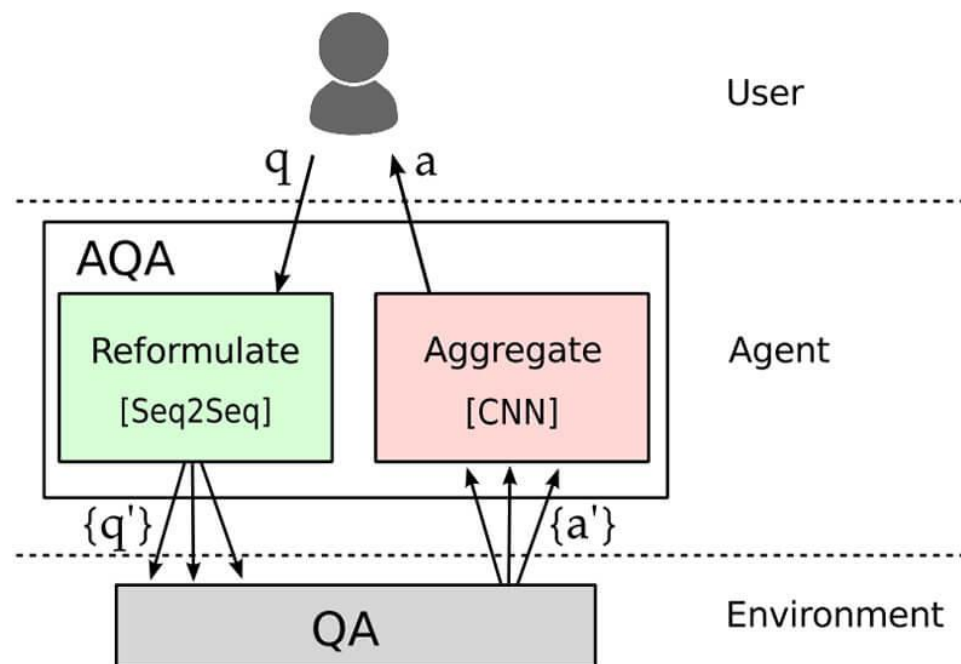
This image is definitely not worth a thousand words. In fact, without the words, you're probably pretty lost. You are probably visualizing a search system to look more like:

In the most basic form, a search system is:

- A user asks a question.
- The search algorithm interprets the question.
- The algorithm(s) are applied to the indexed data, and they provide an answer.

What we see in the first image, which illustrates the methods discussed in the paper, is very different.

In the middle stage, we see two parts: the Reformulate and the Aggregate. Basically, what's happening in this new process is:

- User asks a question to the "Reformulate" portion of the active question-answering (AQA) agent.
- The "Reformulate" stage takes this question and, using various methods discussed below, creates a series of new questions.
- Each of these questions is sent to the "Environment" (We can loosely think of this as the core algorithm as you would think of it today) for an answer.
- An answer for each generated query is provided back to the AQA at the "Aggregate" stage.
- A winning answer is selected and provided to the user.

Seems pretty straightforward, right? The only real difference here is the generation of multiple questions and a system figuring out which is the best, then providing that to the user.

Heck, one might argue that this is what goes on already with algorithms assessing a number of sites and working together to figure out the best match for a query. A slight twist, but nothing revolutionary, right?

Wrong. There's a lot more to this paper and the method than just this image. So let's push forward. It's time to add some…

Machine learning

Where the REAL power of this method comes in is in the application of machine learning. Here are the questions we need to ask about our initial breakdown:

How does the system select from the various questions asked? Which question has produced the best answer?

This is where it gets very interesting and the results, fascinating.

In their testing, Bulian and Houlsby began with a set of "Jeopardy!-like questions (which, if you watch the show, you know are really answers).

They did this to mimic scenarios where the human mind is required to extrapolate a right or wrong response.

If you're not familiar with the game show "Jeopardy!," here's a quick clip to help you understand the "question/answer" concept:

From the paper:

In the face of complex information needs, humans overcome uncertainty by reformulating questions, issuing multiple searches, and aggregating responses. Inspired by humans' ability to ask the right questions, we present an agent that learns to carry out this process for the user.

Here is one of the "Jeopardy!" questions/answers posed to the algorithm. We can see how the question can be turned into a query string:

Travel doesn't seem to be an issue for this sorcerer and onetime surgeon; astral projection and teleportation are no problem.

Not an easy question to answer, given it requires collecting various pieces of data and also interpreting the format and context of often cryptic questions themselves. In fact, without people posting "Jeopardy!"- like questions, I don't think Google's current algorithms would be able to return the right results, which is exactly the problem they were seeking to address.

Bulian and Houlsby programmed their algorithm with "Jeopardy!"-like questions and calculated a successful answer as one that gave a right or wrong answer. The algorithm was never made aware of *why* an answer was right or wrong, so it wasn't given any other information to process.

Because of the lack of feedback, the algo couldn't learn success metrics by anything more than when it got a correct answer. This is like learning in a black box which is akin to the real world.

Where did they get the questions?

Where did the questions used in the test come from? They were fed to a "user" in the Reformulate stage. Once the questions were added, the process:

- Removed stop words from the query.
- Put the query to lowercase.
- Added wh-phrases (who, what, where, when, why).
- Added paraphrasing possibilities.

For paraphrasing, the system uses the United Nations Parallel Corpus, which is basically a dataset of over 11 million phrases fully aligned with six languages. They produced various English-to-English translators that would adjust the query but maintain the context.

Results

So here's where this all landed us:

| | | Baseline | | MI-SubQuery | | Base-NMT | | AQA | | | | |
		ASR	BiDAF	TopHyp	CNN	TopHyp	CNN	TopHyp	Voting	MaxConf	CNN	Human
Dev	EM	-	31.7	24.1	37.5	26.0	37.5	32.0	33.6	35.5	**40.5**	-
	F1	24.2	37.9	29.9	44.5	32.2	44.8	38.2	40.5	42.0	**47.4**	-
Test	EM	-	28.6	23.2	35.8	24.8	35.7	30.6	33.3	33.8	**38.7**	43.9
	F1	22.8	34.6	29.0	42.8	31.0	42.9	36.8	39.3	40.2	**45.6**	-

Source	a boy holding a glove while on a baseball field.
Reference	a young boy bending over holding a catchers mitt.
Base-NMT	how do you hold a glove on a baseball field?
Base-NMT-NoParalex	a cell holding a globe at a later field.
Base-NMT+Quora	is it possible to hold a glove while on a baseball field?
AQA-QR	what is name an boy make up a glove while on baseball field?
Source	a street light with a red arrow pointing left.
Reference	a tagged traffic light displays a red arrow.
Base-NMT	a street light with red arrow pointing?
Base-NMT-NoParalex	a strength with a skilling point.
Base-NMT+Quora	what is the street light of a red arrow pointing?
AQA-QR	what is name a street light with red arrow pointing left?

What they have developed is a system which can accurately understand complex and convoluted questions and, with training, produce the correct answer with a surprising degree of accuracy.

So what, Dave? What does this get me?

You might be asking why this matters. After all, there are constant evolutions in search and constant improvements. Why would this be any different?

The biggest difference is what it means for search results. Google also recently published a paper for the ICLR Conference that suggested Google can produce its own content based on data provided by other content producers.

We all know that just because a paper is written, it doesn't mean a search engine is actually implementing the concept, but let's pause a minute for the following scenario:

- Google has the capabilities of providing its own content, and that content is well-written.
- Google has a high confidence in its capabilities of determining the right answer. In fact, by tweaking its capabilities, it may surpass humans.
- There are multiple examples of Google working to keep users on its site and clicking on its search results with layout and content changes.

With this all stacked up, we need to ask:

- Will this impact search results? (It probably will.)
- Will it hinder a webmaster's content production efforts?

- Will it restrict the exposure of our content to a greater public?

Again, just because a paper is published, it does not mean the contents will be implemented; but Google **is** gaining the capability of understanding complex nuances in a language in a way that surpasses humans. Google is also interested in keeping users on Google properties because, at the end of the day, they are a publishing company, first and foremost.

What can you do?

You do the same thing you've always done. Market your website.

Whether you are optimizing to be in the top 10 of the organic results or optimising for voice searchor virtual reality, the same number of blue widgets is being sold. You just need to adapt, since search engine result pages (SERPs) change quickly.

The methods we're seeing used here raise an important subject everyone interested in search engine optimisation (SEO) should be paying close attention to, and that's the use of entities.

If you look at the query sets above that were generated by the systems Bulian and Houlsby created, you'll notice that in general, the closer they are to accurately understanding the relationship between entities, the better the answer.

The specific wording is irrelevant, in fact. Fully deployed, the system would not be required to use words you or I understand. Thankfully, they enable us to see that success is attained through grouping entities and their relationships in a way that makes giving an answer based on those relationships more reliable.

If you're just getting your feet wet in understanding entities, there's a piece here that introduces the concept and covers of the ins and outs. I guarantee that you'll quickly see how they relate, and you need to focus on this area as we head into the next generation of search.