Can computers understand words?

18th June 2020

We know computers can be trained to use human language, but will they ever be able to handle ambiguity?

The verb *run* has 606 different meanings. It's the largest single entry in the Oxford English Dictionary, placing it ahead of *set*, at 546 meanings.

Although words with multiple meanings give English a linguistic richness, they can also create ambiguity: drawing a gun could mean pulling out a weapon, or simply illustrating one.

We humans can generally avoid this confusion because our brain takes into account the context surrounding words and sentences. But, for computers, lexical ambiguity **poses** a major challenge.

"Computers are **hopeless** at disambiguation because they don't have our **world knowledge**" explains Dr Stephen Clark, who leads two large-scale research projects that hope to **overcome** this difficulty. Applications of the research include improved internet searching, machine translation, and automated essay **marking** and summarisation.

"Many online translation tools are based on statistical models that 'learn' the relationship between words in different languages. But if we want the computer to really understand text, a new way of processing language is needed," says Clark. "Humans are able to generate an unlimited number of sentences using a limited vocabulary," he continues. "We would like computers to have a similar capacity to humans."

Until now, two main **approaches** have been taken by computer scientists to model the meaning of language. The first is based on the principle that the meaning of a phrase can be determined from the meanings of its parts and how those parts are combined. The second **approach** focuses on the principle that the meaning of a word can be **worked out** by considering the various contexts in which words appear in text, and uses word "clouds" to show which words are frequently associated with one another.

By **drawing on** the mathematics of quantum mechanics, working with researchers at several UK universities, Clark plans to exploit the strengths of these two methods through a single mathematical model.

Clark has spent the past decade developing a sophisticated parser – a programme that takes a sentence in English and identifies the grammatical relationships between the words. The next step is to combine this tool with the word clouds **so as to** provide a new meaning representation that has never been available to a computer before. All of this, he hopes, will help solve the ambiguity problem.

Adapted from www.cam.ac.uk by ECP coach Alison Keable

Before you read the article, find this vocabulary in the text:

to pose: to constitute, to represent (a problem, a challenge, a danger)

hopeless: very bad, useless

world knowledge: non-linguistic information, such as culture and experience, that helps us understand words and sentences

to overcome: to solve a problem or get past an obstacle

marking: correction of academic exams or texts

approach: method, way of doing something

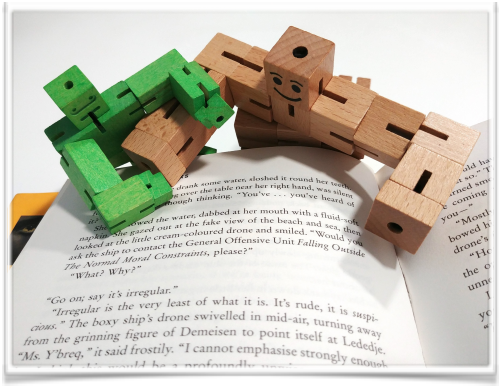
to work sth out: to find the result of a calculation

to draw on sth: to use sth

so as to: in order to

English Coaching Projects S.Coop.

www.englishcoachingprojects.net
info@englishcoachingprojects.net



Let's chat about that!

Write your opinions in an email and send them to your ECP coach!

Do you use machine translation or similar tools?

What is Dr Stephen Clark's goal?

Is this kind of research important in your opinion?

Do you think computers will be good at languages in the future?

How can technology help humans to learn languages?

Weekly English Practice

Social Events In English

Real-Life English in Vitoria-Gasteiz

English Coaching Projects S.Coop.

Practical English Stuff

Real Language to Express Yourself Better

Thank you to everyone who attended
Coffee Saturdays 2019-20!
It was great fun, and great conversation.
See you after the summer!

Have you ever volunteered?
What did you do?

Found in Translation

When the first COVID-19 cases were confirmed in Africa, Cambridge researcher Dr Ebele Mogo and colleagues were worried that a dangerous **language gap** in public health information would soon develop across the continent. She asked on social media if anyone could help them fill the gap.

Thirty young Africans **came forward**.

Examples of resources available in languages that include Xhosa, Igbo, and Tigrinya.

Language can be a major barrier to transmitting good public health information at any time. But when a pandemic like coronavirus is **sweeping** the globe, it becomes even more crucial to promote preventative messages like handwashing and social distancing in the language that people are most familiar with.

With this in mind, a project involving 30 crowdsourced volunteers has completed the translation of COVID-19 public health **guidelines** into 18 of the most spoken languages across the African continent. The resources are freely **available** for download and have been disseminated through community networks and international organisations across Africa.

Now draw a line to match the words with their definitions:

to be available	to offer to help
to come forward	recommended steps
guidelines	a difference in knowledge between two groups of people
a language gap	to be at people's disposition
to sweep	to quickly move through an area

Read more here: https://www.cam.ac.uk/stories/Translations-for-Africa