COGNITIVE COMPUTING IN ASTROBIOLOGY

Pauli E. Laine
University of Jyväskylä, Finland

UNIVERSITY OF JYVÄSKYLÄ

Introduction

Astrobiology is an interdisciplinary science, connecting results from e.g. biology, astrophysics, and geology. It is difficult for an individual researcher to have deep knowledge in all related disciplines, and to keep up with all information that is published all the time.

Cognitive computing is a system that learn at scale, reason with purpose and interact with humans naturally. Cognitive computing systems can be used to effectively access even unstructured large databases, such as in molecular biology, mineralogy, and exoplanet to retrieve relevant information with natural language inquiries.

Cognitive Computing

Cognitive computing is a term that IBM coined for its artificial intelligence (AI) system called Watson, designed to win *Jeopardy!* quiz in 2011. The system was designed to "understand" plain verbal questions, and to utilize computing power to go through large databases of structured and unstructured (e.g. articles, Wikipedia) information in looking for answer.

Cognitive computing combines the large information storage processing power of computers with the humans capability to understand semantics of the information. Technically it contains systems of natural language processing, machine learning, and semantics. Cognitive systems are thus a way to naturally extend the boundaries of human cognition.

Cognitive computing systems are not (yet) general answering systems to any kind of information. Instead they must be adapted to each domain with procedure called domain adaptation methodology. There are now at least medical applications of cognitive computing, but it can be adapted to any kind of domain that can benefit of fast searches from large information storages with natural interactions.

The System

Current cognitive computing systems have a question-answering (AQ) architecture as illustrated in (Fig. 1). Minimal deep QA pipeline consists of a) Question analysis b) Primary search c) Hypothesis generation d) Hypothesis and evidence scoring, and e) Final confidence merging and ranking.

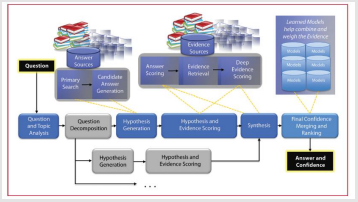

Figure 1 Question-answering architecture [1]

The first stage is a detailed analysis of the question (what it is asking for) with UIMA (Unstructured Information Management Architecture) (Fig. 2). Here CAS is common analysis structure (unstructured data), LAT is lexical answer type, and PAS predicate-argument structure.

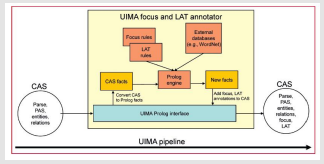

Figure 2 Unstructured Information Management Architecture [2]

Next stages are the analysis of the answer sources to find and hypothesize candidate answers. Candidate answers are then scored based on evidence sources, and finally answer confidence is merged and ranked.

Astrobiological Resources and Databases

Thera are many resources of unstructured information used in astrobiology research, including electronic articles, books, and internet resources. There are also many structured databases available, such as:

Biology
- RNA
- Protein sequence
- Microarray

Geology
- Minerals
- Geochemistry
- Geologic formation

Planetary Science
- Statistical information
- Atmospheric models
- Structural information

Astrophysics
- Astrochemistry
- Comets, meteorites

Exoplanets
- Mass, orbit etc.
- Spectra

Together these resources constitute the astrobiological answer sources.

Examples

First of all, this kind astrobiological cognitive computing system could be used on each constituting discipline for domain specific information retrieval. But the most beneficial application would be the adaptation to astrobiology domain general system, answering to questions like "What known microbes could survive in Mars conditions at 2 meters below the surface?" This would require the system to search data about the subsurface Mars conditions (temperature, radiation, water content etc., with some confidence level), and match them to limits of known organisms.

Another application could be the analysis of different exoplanet information sources. We have already discovered over 3000 exoplanets, and their information is in the limit of control. In the future more and more observations are coming from the new space and ground-based telescopes. How to handle this future information overflow? Currently cognitive computing systems are limited to textual and numeral information. As such, cognitive computing would add "only" natural interface to current database searches. However, in a short time cognitive computing systems will be able to do also image analysis. This will greatly expand the answer sources, and in this case could be used to e.g. compare exoplanet spectral images to known and modelled images for habitability assessment.

Conclusions

- Cognitive computing is a system that learn at scale, reason with purpose and interact with humans naturally.
- New domains can be adapted to any kind of domain that can benefit of fast searches from large information storages with natural interactions.
- Cognitive computing can be beneficial to a multidisciplinary sciences, like astrobiology.

References

1. Ferrucci, D.A. Introduction to "This is Watson", IBM J. RES. & DEV. VOL. 56 NO. 3/4 PAPER 1 MAY/JULY 2012

2. Lally, A. Prager, J.M., McCord, M.C., et al. Question Analysis: How Watson reads a clue, IBM J. RES. & DEV. VOL. 56 NO. 3/4 PAPER 2 MAY/JULY 2012

Cognitive computing in retail. Retailers who use Watson obtain a considerable competitive advantage and increased customer loyalty because they do not rely on business instincts and intuition alone, but can use insight obtained from effectively analysed unstructured customer data. Besides being capable of understanding and answering human questions, Watsonâ€™s budget-friendly solutions can act for companies in social networks, forums, and blogs. How Watson enhances life: A project by Azoft. Communication matters, and thatâ€™s what we strongly believe here at Azoft. Thatâ€™s also why weâ€™ve chosen thi We welcome astrobiology-relevant topics based on chemical observations from the large- to the nanoscale, nanofluidics or miniaturization of chemical analyses, organic analysis of specific targets with isotopic or enantiomeric specificity of complex mixtures with chemical resolution, observtion of biological systems (microbiomes of higher organisms) under space conditions and stresses.Â (This article belongs to the Special Issue Analytical Chemistry in Astrobiology). â–ºâ–¼ Show Figures. Graphical abstract. What Does Cognitive Computing Mean? Cognitive computing describes technologies that are based on the scientific principles behind artificial intelligence and signal processing, encompassing machine self-learning, human-computer interaction, natural language processing, data mining and more. Its aim is to solve complex problems characterized by uncertainty and ambiguity, which in other words means problems that are only solved by human cognitive thought. Advertisement. Techopedia Explains Cognitive Computing.Â Though there is currently no agreed-upon definition of cognitive computing in the industry or the academe, the term is often used to describe new technology that mimics the way that the human brain functions and how it approaches problem solving.