Introduction to the Special Issue:
AI at the Crossroads of NLP and Neurosciences

Guest Editor

Michael Zock

Aix-Marseille Université (LIS-CNRS)

michael.zock@lif.univ-mrs.fr

1. Problems and goals

The work presented here is to some extent the result of an AI workshop (IJCAI, Macao, 2019), that, alas, never took place. Realizing that the accelerating knowledge growth combined with its intimate linkage to an increasing number of fundamentally different domains make it harder and harder to stay on top of the wave, we've decided to organize a workshop. The idea was to get things under better control, to get a snapshot of the various disciplines, and to find out how to overcome some of the well-known bottlenecks (https://easychair.org/cfp/WS-928AIxRoads). In addition we were interested in reaching the following goals:

- explore the benefits of (using) big data for AI, NLP and neuroscience;
- create synergies between these communities, and
- explore whether big data and NLP can help us bridge the gap between such (seemingly) distant scientific fields.

While cross-fertilization is certainly one of our concerns, trying to understand the fundamental differences between natural and artificial

Journal of Cognitive Science 21-1:1-14, 2020

intelligence, and trying to see whether we can and should try to reduce the gap is another one.

This being so, we would like to gain a clearer picture concerning the type of knowledge acquired by machines. For example, what do artificial neural networks really learn from language regarding its nature and structure? We would also like to discuss whether the usage of knowledge about the human brain can enable engineers to produce better software. In order to get some answers to our questions, we planned to use the following strategy: (a) take a broader perspective in order to attract the critical mass of papers needed; (b) take a more focused view, by organizing a round-table next to the invited talk. This round-table could be devoted to one or more specific topics, for example

- Should AI take into account the human factor (psychology, neuroscience)?
- What does it give us by doing so? (benefits, price, trade-offs)
- Is it unrealistic, in other words, a waste of time to put the human in the loop, since we get already so good results without 'him?'
- Do current deep learning models really simulate something plausible about the human mind?

2. Evolution of AI

After half a century of experimentation in research labs, Artificial Intelligence (AI) has moved into the arena of the real world (Lungarella, et al. 2007, Mitchell, 2019). AI systems are currently used in many domains (e.g., medicine, finances and communication), outperforming humans in a broad range of tasks (acoustic-, visual- and natural language processing). Several reasons may explain this success: (a) the growth of power of computational resources, (b) the capacity growth of storage devices, (c) the development of smart learning algorithms and, last but not least, (d) the existence of huge amounts of data, annotated or not, available in various modes (audio, text, video). Progress is also due to the fact that researchers have managed to leverage and integrate discoveries made in disciplines that seemingly had little in common (Linguistics, Psychology, Mathematics, Machine Learning, Information Science, Neuroscience). Yet, different as they may be, these disciplines turn out to be complementary, yielding a virtuous circle, which is an asset, allowing us not only to build 'smart' artifacts, but also to improve our understanding of the human mind. Finally, nature played an important role, as it inspired researchers by providing a model that, in order to be turned into sophisticated working solutions had to be understood, formalized and recast in engineering terms. Artificial neural networks are a good example of this process as they represent an effective, yet very loose imitation of the neural system. State- of-the-art neural architectures, such as Convolutional Neural Networks and Transformers, are directly inspired by biological (e.g., visual cortex) and cognitive (e.g., attention processes) models. In sum, it seems that AI and neuroscience (i.e., our knowledge of the human brain) can benefit from each other, and that NLP, big data, and Machine Learning can provide additional elements likely to help us to bridge the still existing gap. This being said, we are still in search of what Dan Norman has astutely called a ‘conceptual model’ (Norman, 2013) meaning by that the knowledge we need to have concerning a technology in order to use it effectively.

3. Motivation and Topics of interest

The goal of this workshop is to stimulate cross-fertilization between the different communities of the AI universe (e.g., Mathematicians, Linguists, Cognitive Scientists, Neuroscientists, Information Scientists) in order to identify the knowledge needed to bridge the gap between Natural and Artificial Intelligence.

More precisely, we would like to discuss whether and how the usage of knowledge concerning the human brain may help engineers to produce better software, including software for school education, that is, software making learning in the classroom more fun while helping students to learn something new. We may refer to this as brain-compatible learning software, or brain compatible school.

In addition to this, we would like to explore whether the use of artificial neural networks and our understanding of their workings may shed some new light on cognitive processes. It should be noted that despite the performance boost of NLP due to the advances of deep learning, there are still a number of problems with this approach, the lack of interpretability being one of them. While progress in performance can be measured, and possibly even explained, it is hard to translate it in terms of human competence (the knowledge people have to come to grips to carry out a given task). For example, what can translators learn from a successful system like GoogleTranslate?

Last but not least, we hope that the outcome of this workshop may help us reduce the gap between our understanding of how humans use their knowledge (David Marr's level 2) and how this knowledge is implemented in machines and in the human brain (Marr's level 3).

Even if we cannot expect from an event like this a set of full-blown solutions, the brainstorming among leading experts of the various fields may nevertheless yield new insights or promising leads.

Here is a categorized list of questions for which we would like to find answers, list which has been improved with the help of Vito Pirrelli:

3.1 Mutual benefits, knowledge transfer

- In what ways can the techniques developed in AI inspire cognitive scientists to get new ideas/models/theories, or, to help them refine existing ones?

- Can the way machines learn shed any light on the way humans do, or, conversely, can the techniques developed to study children's learning also be used to understand the way machines learn?
- Can the knowledge of the brain mechanisms involved in intelligent behavior (interpretation or production of sound, images and language) help us to develop better software or architectures?
- What are the benefits for AI to mimic humans or the human mind while processing language?
- How can AI help us solve problems in other disciplines, for example, NLP, CALL?

3.2 Mimic/extend human cognitive capacities

- Can we build machines that learn as ordinary people do? Life is short, and biological design (evolution) is still quite different from human design (constraints, time scales, ...). For example, children induce rules on the basis of very few examples, containing even noisy data (few-shot learning; learning of abstractions). They are also able to reuse the learned knowledge for new tasks (transfer learning). Can we replicate these two 'cognitive skills' by a machine?
- Can we impose order and logic on an unordered set of ideas, by detecting the nature of the links between them automatically, to help authors in producing coherent texts?
- Is there a way for AI to make use of embodied representations?
- Can we make Natural and Artificial Intelligence cooperate in problem-solving, or, should the two be applied separately? If ever there is an interaction between the two, what should it look like? What are the interfaces and workflows?

3.3 "How" questions, i.e., engineering problems

- How to build AI increasing human intelligence, or, how to use human

intelligence to enhance AI?
- How to improve the interaction between humans and machines?
- Where in the development cycle and how shall AI engineers consider specific human aspects, such as the human brain/mind?

3.4 "Why" questions: the search for explanations

- Is it possible to build a glass box and open the neural network black box?
- What can NLP practitioners learn from network science (complex graphs)?
- How relevant is deep learning for modeling human thought?
- Do we still need models and theories in the age of deep learning? Are there ways to interpret their results?

3.5 Ragbag

- Can machines liberate us from the boring and mechanical aspects of problem-solving (logical proofs), to allow us to focus more on the creative aspects of the task?
- How can AI help us to solve problems in other disciplines, for example, NLP, CALL?
- Specificities of humans and machines: how relevant is deep learning for modeling human thought?

4. Contributions for this Special Issue

The following papers were chosen for inclusion

4.1 Katja Artemova et al. Data-Driven Models and Computational Tools for Neurolinguistics: a Language Technology Perspective

The paper provides a good overview of recent developments in the field of neurolinguistics, striking a good balance between influential studies from the past, and their connection to more recent work. This kind of bridge is especially useful in a domain that evolves so rapidly. Another contribution of this paper is its discussion of how to use such models in medical applications. A third contribution of this work lies in its revelation of the link between language technologies and neurolinguistics. More precisely, the authors offer a review of brain imaging-based neurolinguistic studies with a focus on the natural language representations, such as word embeddings and pre-trained language models.

The authors also show that mutual enrichment of neurolinguistics and language technologies leads to development of brain-aware natural language representations. The importance of this kind of research is emphasized with respect to medical applications.

Finally, the authors point out several directions of future research and development. First of all, the advances of multimodal models, blending different types of language and neuroimaging data, measured in different time scales. The second direction of future developments concerns multilingual and cross-lingual aspects. Last, but not least, the authors point out awareness rising. We need to acknowledge the shortcomings of data-driven approaches, namely the absence of solid theoretical grounding. For example, experiments suggest only a correlation between language model prediction and measured brain activity, yet there is very little theoretical understanding of how a human brain processes semantics which limits the applicability for medical purposes significantly.

This last point has its equivalence to machine learning; It is well known that machines often perform very well, better than humans, but generally we do not know why. Here is what Domingos (2015) writes concerning this problem. « These seemingly magical technologies work because, at its core, machine learning is about prediction. »…” Paradoxically, even as they open new windows on nature and human behavior, learning algorithms themselves have remained shrouded in mystery »…”… in each case the learning algorithm driving the story is a black box. Even books on big data skirt around what really happens when the computer swallows all those terabytes and magically comes up with new insights. At best, we're left with the impression that learning algorithms just find correlations between pairs of events, such as googling "flu medicine" and having the flu. But finding correlations is to machine learning no more than bricks are to houses, and people don't live in bricks." (Domingo, 2015: page XV)

4.2 Juyang Weng : Unified Hierarchy for AI and Natural Intelligence through Auto-Programming for General Purposes (APFGP)

Starting from the observation that there is still a huge gap between artificial intelligence (AI) and natural intelligence (NI) the author tries to explain what it takes to bridge this gap. More precisely, he proposes to describe the way how an autonomous agent, natural or artificial, could be enabled to develop a unified intelligence in his brain. The term "unified" refers here not only to AI, NI or their combination, but also to the various sensory- and motor modalities, as well as to other actions like perception, representation/ reasoning, learning.

Even though supported by rigorous mathematical proofs and initial experimental verification, this is still largely an idea paper. The author embarks the reader to find an answer to the following general question: "How can a machine, natural or artificial, Autonomously develop a Program for General Purposes (APfGP) based exclusively on the input received from its environment, i.e., the physical world?" He tries to answer this question by providing a theoretical, practical and mathematical solution. This paper addresses many difficult problems, and while not everyone may be inclined to subscribe to all of the author's conclusions, the paper does contain a lot of food for thought.

Actually, the author's quest can somehow be seen to be parallel to one of the problems posed by the author of the 'Master Algorithm'. According to (Domingos, 2015: 52-53) scientists belong to one of the following tribes: symbolizers, connectionists, evolutionaries, bayesians, analogizers. “Each tribe’s solution to its central problem is a brilliant, hard-won advance. But the true Master Algorithm must solve all five problems, not just one.” Each tribe has a different part of the puzzle which they must gather, just like the blind men looking at the elephant have to accommodate their knowledge in order to come up with a coherent whole or satisfying solution.

4.3 Takumi Ito, et al. Assisting Authors to Convert Raw Products into Polished Prose.

If speaking or writing is already a difficult task, to do so in a foreign language is a real challenge. The authors of this paper present a new task, called Sentence-level Revision (SentRev). Its goal is to produce high quality sentences in the context of academic writing based on the draft versions (typically) produced by inexperienced non-native speakers. The task takes place in the context of learning to write and producing well-formed sentences (or even text) in a foreign language. The writing assistant has been built for the production of scientific papers. A new dataset, called SMITH, is presented from supporting the task along with the methodology for building it, taking into account the fact that the access to true draft sentences is difficult in practice. Finally, three high-level baseline approaches are presented for performing the task and evaluated on the SMITH dataset according to various measures. In addition, two methods for producing synthetic data for training two of the baseline approaches are presented.

This article addresses an important task in writing. In the recent years, a lot of work has addressed related tasks devoted to non-native English speakers, in particular, detection and correction of grammatical errors. It should be noted though that among the existing systems the following points are hardly addressed: errors related to meaning distortion, incorrect lexical choice, poor style, lack of coherence and information gaps. Some of these points will be part of the authors' future work.

4.4 Tseng & Hsieh Computational Representation of Chinese Characters: Comparison between Singular Value Decomposition and Variational AutoEncoding

Chinese have developed a sophisticated system to encode information in their writing system. Actually, the characters encode an enormous amount of world knowledge in compact visual forms, i.e., patterns. In order to extract this knowledge, researchers resort to computational models. For example, using a variational autoencoder, a popular technique in deep learning (in particular for face recognition), allowed the authors of this paper to extract variational features from Chinese characters. Next they compared them to the eigencharacters extracted from singular value decomposition. Comparing both techniques showed that both of them are able to reveal important aspects of Chinese characters (e.g., radical or components information), but eigencharacters, facilitate training process in a form-to-meaning task. Since both representations can easily be incorporated into modern computational models, future studies should show their added value.

4.5 Benjamin Heinzerling : NLP's Clever Hans Moment has Arrived.

While being relatively short this paper contains nevertheless quite a bit of food for thought. For example, the author reveals the shortcomings of the current deep neural approaches in NLP, concluding that performance as such, or superior performance does not necessarily say very much about the system's competency. Indeed, the results may be based on superficial cues. The author reviews work by Niven & Kao (2019) of this so-called Clever Hans effect on an argument reasoning task to discuss then possible solutions for its prevention. He suggests that in order to avoid this kind of failure we should create not only better and more flexible models, . i. e., models that are able to accommodate to new data, be they more complex, but also create models that are more robust. Hence, in order to reach this goal creators should adopt a ‘Build_It and Break_It’ mentality’, which values robustness just as much as sheer performance, i.e., a high score on a certain dataset.

This viewpoint is perfectly in line with the mindset of philosophers like Popper (2002), Kuhn (2012) or Richard Feynman (https://www.youtube.com/watch?v=0KmimDq4cSU) who writes: “In general, we look for a new law by the following process: First we guess it; then we compute the consequences of the guess to see what would be implied if this law that we guessed is right; then we compare the result of the computation to nature, with experiments or experience, compare it directly with observation, to see if it works. If it disagrees with the experiment, it is wrong.” This is the key to science, it’s as simple as that. Likewise, according to Popper (2002) we can never prove the veracity of a law. All we can do is to show that it is wrong, and where it is wrong. Yet, rather than being simply a shortcoming there are virtues to the possibility to show that something is wrong, and where in particular lies the mistake. Perhaps this is a point to be taken more seriously by the NLP community.

5. Conclusion and thanks

Having reached the end of this introduction, we can conclude that there are many questions we would not even have thought about a decade ago, or so. This is a vibrant field with many ramifications, huge potential, but also confronted with a large number of very complex problems (Dupoux, 2018; Schmidhuber, 2007). Hoping that this issue will reach the readers, and inspire them in their future work.

I would like to take this opportunity to thank the following persons for their feedback during the various stages of drafting the Call for Papers of this workshop (https://sites. google.com/view/ijcai-ws928-aixroads/home):

Emmanuele Chersoni (The Hong Kong Polytechnic University, China), *Yoed Kennett* (University of Pennsylvania, Philadelphia, PA, USA), *Vito Pirrelli* (ILC, Pisa, Italy), *Enrico Santus* (CSAIL, MIT, Boston, USA), and *Mingyu Wan (*The Hong Kong Polytechnic University, China).

Likewise, I would like to thank all the reviewers (here below in alphabetical order) for their efforts to provide useful feedback, allowing the authors to improve considerably their original document:

Patrice Bellot (LPL, Aix-Marseille Université, France), *Jean Pierre Briot* (LIP6, CNRS, Paris, France), *Michael Carl* (Kent State University, Cleveland, Ohio, USA), *Gerard de Melo* (Rutgers University, USA), *Simon de Deyne* (School of Psychological Sciences, University of Melbourne, Australia), *Olivier Ferret* (CEA LIST, France), *Michael Flor* (Educational Testing Service, Princeton, NJ, USA), *Chu-Ren Huang* (The Hong Kong Polytechnic University, China), *Aurélie Herbelot* (University of Trento, Italy), *Markus J. Hofmann* (Bergische Universität Wuppertal, Germany), *Ekaterina Kochmar* (University of Cambridge), *Patrick Paroubek* (LIMSI, Orsay, France), *Pease Adam* (InfoSys, Palo Alto,CA, USA), *Vito Pirrelli* (ILC, Pisa, Italy), *Alessandro Raganato* (University of Helsinki, Helsinki, Finland), *Massimo Stella* (Complex Science Consulting, Lecce, Italy), *Carlo Strapparava* (Fondazione Bruno Kessler, Trento, Italy), *Dan Tufis* (RACAI, Bucharest, Romania).

References

Domingos, P. (2015). The master algorithm: How the quest for the ultimate learning machine will remake our world. Basic Books.

Dupoux, E. (2018). Cognitive science in the era of artificial intelligence: A roadmap for reverse-engineering the infant language-learner. *Cognition, 173*, 43-59.

Kuhn, T. S. 2012.The structure of scientific revolutions. University of Chicago Press.

Lungarella, M., Iida, F., Bongard, J. & R. Pfeifer. eds. 2007. *50 years of artificial intelligence: essays dedicated to the 50th anniversary of artificial intelligence.* Springer.

Mitchell, M. (2019). *Artificial Intelligence: A Guide for Thinking Humans.* Penguin UK

Niven, T., and Kao, H. Y. 2019. Probing Neural Network Comprehension of Natural Language Arguments. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 4658-4664.

Popper, K. 2002.*The logic of scientific discovery.* New York, NY: Routledge Classics.

Schmidhuber,J. 2007.2006: celebrating 75 years of AI-history and outlook: the next 25 years.InLungarella,et al.2007.*50 years of artificial intelligence.* Springer,Berlin, Heidelberg. 29-41.