An Evolutionary Machine Learning Algorithm to Enhance Language Processing

Martin W. Kinch, Wim J.C. Melis, Simeon Keates

M.W.Kinch@greenwich.ac.uk, Wim.J.C.Melis@greenwich.ac.uk, S.Keates@greenwich.ac.uk

University of Greenwich - Department of Engineering and Science

Introduction

In the current age of Artificial Intelligence (AI) there have been numerous inroads into the topic [2][6][7], many of which have made unprecedented leaps in completing tasks that were once thought impossible; notably professional level game playing; such as: Watson[5], Deep Blue[3] or AlphaGO[1]. However, despite these great leaps, nearly all these AI systems[1][3] focus on the use of processor intensive algorithms to achieve their goals; while not necessarily an issue it requires all data to be mapped towards an arithmetic world. Since the world is not necessarily mathematical in all aspects and can only be represented in this manner as far as our current level of mathematics has been developed, which leads to the problem that any mathematical model will be limited by the development of mathematics in relation to representing the real world. The better one can make this representation, the more likely the system will operate with a good accuracy. However, as complexity increases to deal with the real world in a mathematical way, data and processing requirements increase likewise and so eventually one needs massively parallel machines consuming large amounts of power to provide for only a limited amount of intelligent functionality. Additionally, the data will be represented in a format that is optimal for that particular task, but does not necessarily allow for generalisation towards other processes, and so the scope of the application is connected with the mathematical ruleset used for the AI to work from [1][3]; something that then implicitly limits the AI's ability to move from one environment to another. Which poses the question, how can something be intelligent if it cannot transfer its knowledge from one application to another, as such transference capabilities are clearly seen in humans.

Methods

To address the questions posed previously, effort will be directed towards the development of new data management models, primarily, focusing on how to best store and capture data and its complex inter-relationships and make this useful towards an intelligent system. This will require the stored data to be relational, so that meaning can be inferred more than determined. The important part is rather than perform extensive calculations to determine the importance and relevance of the information, much of this can be determined by the relationships between data, meaning that the context is of particular importance towards the meaning of the actual data. Some of this data complexity can obviously be grasped within an artificial neural network, but these suffer from the limitation that these networks then often become black boxes making it challenging to ensure the reliability of the system. On the other hand, using mathematical approaches for machine learning requires massively parallel machines, which are great for arithmetic functionality, but struggle to efficiently deliver the intelligent functionality necessary for these type of networks.

Within the context of learning data and, storing and structuring it, it is important to understand the level of detail that is required. In a machine learning context this is often referred to as generalisation and can lead to over/under-fitting, while essentially the purpose is to find an invariant representation as mentioned in [4], which has investigated this topic from a neuroscientific perspective. The more important challenge in this context is to identify the minimum required amount of data to be able to

recognise something, and consequently the amount of data that really needs to be stored for recognition, even though they may both depend on the amount of context information available.
Being able to use contextual information should provide the system with a far greater understanding of the data provided to it, while also becoming more able to determine the intention behind the message that is being “translated” from human language into computational terms. The most important forms of context are probably the “hidden” elements of communication that people take for granted, such as: non-verbal cues, locational cues, situational cues and possibly even tonal or pitch cues, to name a few. The aim is to study the importance of individual elements and then select the best matches for human understanding and computational ease, to ensure that this extra information will ensure a higher accuracy rate.
While it is obviously easier to try and structure everything in a deterministic/organised world, it seems unlikely that the data and its organisation in the brain may be anywhere close to organised, and is more likely based on approximations rather than exact outcomes. This then also reflects on the data being invariant rather than very specific, although further details may be required to go from an initial “estimate” to a more detailed “result”. Even though this context suddenly becomes very unsure and unknown, due to it not being exact in nature, it is likely to be in this uncertainty that may lie the answers to some of the challenges faced. It will therefore be essential to take a more holistic approach and include the field of AI as well as other fields such as psychology, philosophy and biology to obtain the required answers.

Expected Results

This research aims to develop an algorithm that is biologically inspired, and understands and uses approximate data to achieve transferability in its applications. Considering that all human learning is an evolutionary process, it is expected that this algorithm will have to incorporate evolutionary aspects to “fine tune” any learned data. It is in this context that language processing seems most suitable as an application as it allows for the exploration of general elements as a building block, to master one aspect of natural language while invariance and evolution can be found in secondary aspects such as interpretations and so on.

Discussion

In order to overcome some of the challenges of current research, this work will focus on developing from a data centred model, rather than a processor based one, with the aim to achieve significant savings in data storage and processing requirements for AI systems. Savings of this nature will also help improve the efficiency of such systems, while using a less deterministic approach is expected to make AI’s transferability more effective; allowing them to be competent in several applications rather than just one.

References

[1] Silver, D., Huang, A., Maddison, C. J., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 484-489.

[2] Vinyals, O., Blundell, C., Lillicrap, T., Kavukcuoglu, K., & Wierstra, D. (2016). Matching Networks for One Shot Learning. Available at: https://arxiv.org/pdf/1606.04080v1.pdf, Accessed on: 21/10/2016

[3] Campbell, M., Hoane, A. J., & Feng-hsiung, H. (2002). Deep Blue. *Artificial Intelligence*, 57-83.

[4] Hawkins, J., & Blakeslee, S. (2004). *On Intelligence*. New York: St. Martin's Griffin.

[5] Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. , et al. (2010). Building Watson: An Overview of the DeepQA Project. *AI Magazine*, 59-79.

[6] Chan, W., Jaitly, N., Le, Q., & Vinyals, O. (2016). Listen, attend and spell: a neural network for large vocabulary conversational speech recognition. *ICASSP* IEEE, 4960-4964.

[7] Bahdanau, D., Chorowski, J., Serdyuk, D., Brakel, P., & Bengio, Y. (2016). End-to-end attention-based large vocabulary speech recognition. *ICASSP* IEEE, 4945-4949.