Academy of Mathematics and Systems Science
Chinese Academy of Sciences

2018 Symposium on Frontiers of Artificial Intelligence (Autumn)

August 21 (Tuesday), 2018, Beijing, China

Organizers: Songmao Zhang, Sanjiang Li, Ruqian Lu

Venue: Meeting Room 204, South Building of AMSS
55 Zhong-Guan-Cun East Road, Beijing 100190, China

Sponsors:

National Key Research and Development Program of China under grant 2016YFB1000902

National Natural Science Foundation of China under grant 61621003

Key Laboratory of Management, Decision and Information Systems (MADIS), Chinese Academy of Sciences

9:00 – 12:00 The morning session, chaired by Songmao Zhang

9:00 – 9:10 Opening

9:10 – 10:00 Mathematical Methods for Curating Big Knowledge in Biomedicine
Speaker: **GQ Zhang**
University of Kentucky, Lexington, Kentucky, USA

10:00 – 10:50 Introduction to Computational Argumentation
Speaker: **Anthony Hunter**
University College London, London, UK

10:50 – 11:10 **Coffee Break**

11:10 – 12:00 Hashing for Big Data
Speaker: **Xianglong Liu**
Behang University, Beijing, China

12:00 – 14:00 Lunch

14:00 – 17:00 The afternoon session, chaired by Ruqian Lu

14:00 – 14:50 What is Essential Unification?
Speaker: **Jörg Siekmann**
German Research Centre of Artificial Intelligence, Saarbrücken, Germany

14:50 – 15:40 Connecting Virtual-World and Real-World Reinforcement Learning
Speaker: **Yang Yu**
Nanjing University, Nanjing, China

15:40 – 16:00 **Coffee Break**

16:00 – 16:50 Constraint Satisfaction Problems and their Application in Spatial and Temporal Reasoning
Speaker: **Sanjiang Li**
University of Technology Sydney, Sydney, Australia

16:50 – 17:00 Closing

Introduction to the Talks

Talk: Mathematical Methods for Curating Big Knowledge in Biomedicine

Abstract: Ontologies are shared conceptualizations of a domain represented in a formal language. They represent not only the concepts used in scientific work, but just as importantly, the relationships between the concepts. Concepts in an ontology are organized into a concept hierarchy with more generic concepts modeled as parents of more specific concepts. Such hierarchical relationships allow us to think of an ontology as a mathematical structure of an ordered set (poset), where concepts are represented as nodes and relationships are represented as edges. In biomedicine, the Unified Medical Language System (UMLS), managed by the US National Library of Medicine, is perhaps the largest integrated repository of biomedical ontologies. The recent release of UMLS covers over 2.9 million concepts from more than 140 source ontologies. This represents a big knowledge resource indispensable for all intelligent applications that exploit big data. Visualization and continued quality assurance of such ontologies is an integral part of their lifecycle. The computational tasks involved in processing such large ordered structures are challenging and impactful.

This talk will focus on the following aspects of big knowledge curation: (1) Elastic parallel processing of posets, representing ontologies, using MapReduce, a cloud computing framework; (2) Scalable extraction and visualization of substructures from posets, such as lattice and non-lattice fragments; (3) A practically linear-time algorithm for finding all lowest common ancestors for all non-trivial pairs, achieving orders of magnitude in performance enhancement for extracting poset fragments; and (4) A non-lattice-based approach to detecting and mining potential errors in SNOMED CT, the most comprehensive clinical health careterminology worldwide, as well as other biomedical ontologies.

About the speaker: Professor GQ Zhang is the Director of the Institute for Biomedical Informatics (https://ibi.uky.edu); Chief of the Division of Biomedical Informatics; Associate Director, Center for Clinical & Translational Science (ccts.uky.edu) at the University of Kentucky. He is also the Director of data coordination center for NIH-NINDS-CWW Center for SUDEP Research (sudepresearch.org). Prior to joining the University of Kentucky, his role at Case Western Reserve University included Division Chief of Medical Informatics, Co-Director of Biomedical Research Information Management Core of the Case Western CTSA, and Associate Director for Case Comprehensive Cancer Center while performing duties as a tenured professor in the Case School of Engineering.

His research interests include data science and bigdata in biomedicine, large-scale, multi-center data integration, biomedical ontology development, information retrieval, query interface design, and agile, interface-driven, access-control grounded software development. These interests are reflected in his over 150 publications in Computer Science and Biomedical Informatics. Over more than a decade, Dr. Zhang has developed a range of clinical research informatics tools for data capturing, data management, cohort discovery, and clinical decision support, such as VISAGE (PMID: 21347154), MEDCIS (PMID: 23686934), OnWARD (PMID: 21924379), OPIC (PMID: 23304354), EpiDEA (PMID: 23304396), and Cloudwave (PMID: 23920671). Supported by multiple federal- and foundation-funded awards (NIH-NINDS, NIH-NHLBI, NIH-NCATS,

NIH-NCI, NSF, DOD) and acclimated in a multi-disciplinary team-science, collaborative setting, Prof. Zhang effectively brings cutting-edge computer science and informatics methodology to addressing biomedical data/big data challenges through the translation of theory, algorithms, methods and best practices to functional and usable tools impacting clinical and translational data science.

.

Talk: Introduction to Computational Argumentation

Abstract: The field of computational argument is emerging as an important part of artificial intelligence research. The reason for this is based on the recognition that if we are to develop robust and interactive intelligent systems, then it is imperative that they can handle incomplete and inconsistent information in a way that somehow emulates the human ability to tackle such information. And one of the key ways that humans do this is to use argumentation, either internally, by evaluating arguments and counterarguments, or externally, by for instance entering into a discussion or debate where arguments are exchanged. Much research on computational argument focuses on one or more of the following layers: the structural layer (How are arguments constructed?); the relational layer (What are the relationships between arguments?); the dialogical layer (How can argumentation be undertaken in dialogues?); the assessment layer (How can a constellation of interacting arguments be evaluated and conclusions drawn?); and the rhetorical layer (How can argumentation be tailored for an audience so that it is persuasive?). This has led to the development of a number of formalisms for aspects of argumentation and some promising application areas.

About the speaker: Anthony Hunter has a BSc (1984) from the University of Bristol, and an MSc (1987) and a PhD (1992) from Imperial College, London which was supervised by Dov Gabbay. He was a research fellow at the IT Research Institute at the University of Brighton from 1987 to 1989, and a research associate in the Department of Computing at Imperial College, London, from 1989 to 1996. Since 1996, Hunter has been in the Department of Computer Science at University College London. Currently, he is Professor of Artificial Intelligence, and head of the Intelligent Systems Research Group, in the UCL Department of Computer Science. Prof. Hunter's research is in the area of knowledge representation and reasoning which is a branch of artificial intelligence. More specifically, he is interested in the inter-related topics of computational models of argumentation; knowledge merging and aggregation; and measuring and analyzing inconsistency.

Talk: Hashing for Big Data

Abstract: Nearest neighbor search plays an important role in many areas like large-scale visual search, machine learning, data mining, etc. Nowadays, as the amount of data and information explodes, to solve the problem over gigantic database, the hashing based approximate nearest neighbors search technique has been widely studied and successfully applied in practice, owing to its compressed storage and efficient computation. This talk will briefly survey the literature of hashing research for big data,

and introduce our recent work including structure sensitive quantization, complementary hash indexing and machine learning speedup, respectively addressing the efficient representation, indexing and analysis of big data.

About the speaker: Xianglong Liu received the BS and PhD degrees in computer science from Beihang University, in 2008 and 2014. From 2011 to 2012, he visited the Digital Video and Multimedia (DVMM) Lab, Columbia University as a joint PhD student. He is currently an Associate Professor with the School of Computer Science and Engineering, Beihang University. He has published over 50 research papers at top venues like IEEE CVPR/ICCV/TIP, AAAI, IJCAI, etc. His research interests include deep learning, computer vision, and multimedia information retrieval. He was a StarTrack Program Visiting Young Faculty in Microsoft Research Asia and was selected into the 2015 CCF Young Talents Development Program. He also received a number of awards including IEEE ICME 2011 best paper candidate, IEEE CVPR 2014/2016 Young Researcher Support and the 2015 CCF Outstanding Doctoral Dissertation Award.

Talk: What is Essential Unification?

Abstract: Ordering is a well established concept in mathematics and also plays an important role in many areas of computer science, where quasi orderings, most notably well founded quasi orderings and well quasi orderings, are of particular interest. This talk deals with quasi orderings on first order terms and introduces a new notion of unification based on a special quasi order, known as homeomorphic tree embedding.

Historically, the development of unification theory began with the central notion of a most general unifier based on the subsumption order. A unifier σ is most general, if it subsumes any other unifier τ, that is, if there is a substitution λ with $\tau =_{\mathrm{E}} \sigma\lambda$, where E is an equational theory and $=_{\mathrm{E}}$ denotes equality under E. Since there is in general more than one most general unifier for unification problems under equational theories E, called E-Unification, we have the notion of a complete and minimal set of unifiers under E for a unification problem Γ, denoted as $\mu\mathcal{U}\Sigma_{\mathrm{E}}(\Gamma)$. This set is still the basic notion in unification theory today.

But, unfortunately, the subsumption quasi order is not a well founded quasi order, which is the reason why for certain equational theories there are solvable E-unification problems, but the set $\mu\mathcal{U}\Sigma_{\mathrm{E}}(\Gamma)$ does not exist, i.e. they are of type nullary in the unification hierarchy.

In order to overcome this problem and also to substantially reduce the number of most general unifiers, we extended the well known encompassment order on terms to an encompassment order on substitutions (modulo E). Unification under the encompassment order is called essential unification and if $\mu\mathcal{U}\Sigma_{\mathrm{E}}(\Gamma)$ exists, then the complete set of essential unifiers is a subset of $\mu\mathcal{U}\Sigma_{\mathrm{E}}(\Gamma)$. An interesting effect is, that many E-unification problems with an infinite set of most general unifiers (under the subsumption order) reduce to a problem with only finitely many essential unifiers. Moreover there are cases of an equational theory E, for which the complete set of most general unifiers does not exist, the minimal and complete set of essential unifiers however does exist.

Unfortunately again, the encompassment order is not a well founded quasi ordering,

that is, there are still theories with a solvable unification problem, for which a minimal and complete set of essential unifiers does not exist.

This talk deals with a third approach, namely the extension of the well known homeomorphic embedding of terms to a homeomorphic embedding of substitutions (modulo E). We examine the set of E-unifiers under the quasi order of homeomorphic embedment modulo an equational theory E and propose an appropriate definitional framework based on the standard notions of unification theory extended by notions for the seminal tree embedding theorem or "Kruskal Theorem" as it is called.

The main result of this approach is, that for any solvable E-unification problem the minimal and complete set of E-unifiers always exists under the new ordering introduced in this talk, it is usually even smaller than the set of essential unifiers and with some additional special conditions it is always finite.

About the speaker: Prof. Jörg H. Siekmann is one of the pioneers and founders of artificial intelligence research in Germany. He studied mathematics and physics at the university of Göttingen. He was professor at university of Keiseslautern from 1983 to 1990 and at univerity of Saarland in Saarbrücken since 1991. Prof. Siekmann is an internationaly leading expert in formal methods, automatic theorem proving, knowledge representation, muti-agent systems and e-learning. He has published numerous excellent papers in top international journals and conferences. Prof. Siekmann has made great contributions to the developement of artifitial intelligence in Germany as a major discipline, both in the universities and in the German Computer Society where he is the founder and first chairman of the technical committee for AI. He founded in 1989 with colleagues the German Research Centre for Artificial Intelligence (DFKI) and became one of its directors. He is the founding editor of the *Springer Lecture Notes in AI* (LNAI) and other scientific journals and book series in logic and AI. In 1983 he organized in Karlsruhe the first large conference on AI, *IJCAI,* in Germany. He has been serving on several international advisory boards of research institutions as well as government committees.

Talk: Connecting Virtual-World and Real-World Reinforcement Learning

Abstract: Reinforcement learning achieved significant successes include being part of the AlphaGo system and playing Atari games. However, reinforcement learning is also criticized for applicability only in virtual worlds due to the requirement of huge amount of interaction data. In this talk, we will report our recent progress towards real-world reinforcement learning, including virtualizing real-world tasks and drawing virtual-world policies out to the real world.

About the speaker: Yang Yu is an associate professor of computer science in Nanjing University, China. He joined the LAMDA Group as a faculty since he got his Ph.D. degree in 2011. His research area is in machine learning and reinforcement learning. He was recommended as AI's 10 to Watch by IEEE Intelligent Systems in 2018, invited to have an Early Career Spotlight talk in IJCAI'18 on reinforcement learning, and received the Early Career Award of PAKDD in 2018.

Talk: Constraint Satisfaction Problems and their Application in Spatial and Temporal Reasoning

Abstract: Constraint satisfaction problems (CSPs) are mathematical problems defined as a set of variables whose value must jointly satisfy a collection of constraints. CSPs have been intensively studied in artificial intelligence and operations research and have been widely used in areas such as automated planning, resource allocation, and spatial and temporal reasoning. In this talk, I will give a concise introduction of constraint solving techniques and then focus on the application of CSPs in spatial and temporal reasoning.

About the speaker: Sanjiang Li received his B.Sc. and Ph.D. degrees in mathematics from, respectively, Shaanxi Normal University, in 1996, and Sichuan University, in 2001. He is now a full professor in Centre for Quantum Software & Information (QSI), Faculty of Engineering & Information Technology, University of Technology Sydney (UTS). Before joining UTS, he worked in the Department of Computer Science and Technology, Tsinghua University from September 2001 to December 2008. He was an Alexander von Humboldt research fellow at Freiburg University from January 2005 to June 2006; held a Microsoft Research Asia Young Professorship from July 2006 to June 2009; and held an ARC Future Fellowship from January 2010 to December 2013. His research interests are mainly in spatial reasoning and artificial intelligence. Some of his most important work has been published in international journals AIJ and JAIR and international conferences AAMAS, IJCAI, AAAI, KR, and ECAI.