The Future of AI

AI's 10 to Watch

IEEE Intelligent Systems has long helped promote young and aspiring AI scientists via its biennial "AI's 10 to Watch" special section. The 2018 group consists of 10 young stars who have demonstrated outstanding AI achievements. In 2017, *IEEE Intelligent Systems* called for nominations worldwide, with the requirement that nominees with doctorates must have received their PhDs in the past five years. The selection committee, made up of *IEEE Intelligent Systems* editorial and advisory board members, then voted on a short list of top candidates. They based their final selections on reputation, impact, expert endorsement, and diversity.

This year's 10 to Watch are as follows:

- **Bo An** has made fundamental contributions to the field of multiagent systems, particularly involving game-based algorithms for physical security, cybersecurity, and sustainability.
- **Erik Cambria** has championed a multidisciplinary approach to sentiment analysis that aims to bridge the gap between statistical natural language processing and other disciplines—such as linguistics, commonsense reasoning, and affective computing—necessary for understanding human language.
- **Yoav Goldberg** has contributed to the fields of lexical semantics (understanding the meaning of words), syntactic parsing (representing and automatically inferring the structure of sentences), and deep learning for language processing.
- **Akshat Kumar** has done seminal research on decision theoretic planning, constraint optimization, probabilistic reasoning, graphical models, and multiagent systems.
- **Wei Liu** has worked on large-scale learning, classification, search, and recommendation technology related to visual data.
- **Cynthia Matuszek** has combined robotics, natural language processing, and machine learning to build systems that nonspecialists can instruct, control, and interact with intuitively and naturally.
- **Sinno J. Pan** is a pioneer in the field of transfer learning. He has helped lay the early theoretical foundations and develop practical algorithms for transfer-learning techniques with diverse real-world applications.
- **B. Aditya Prakash** has won awards for his important work on understanding, reasoning about, and mining propagation of information over networks in diverse real-world systems. He has developed a suite of efficient, near-optimal immunization policies to control influenza spread in a variety of realistic settings. His work has also led to surprising applications in social media and cybersecurity.
- **Maria Vanina Martinez** has researched automated reasoning about rich knowledge bases that contain potentially uncertain information. She has proposed solutions for personalizing the integration of inconsistent and incomplete knowledge bases, and has developed inconsistency-management semantics for relational databases and logic-based knowledge bases.
- **Yang Yu** has helped build the theoretical foundation of evolutionary computation, addressing its basic questions by designing time-complexity analysis tools, proposing approximate analysis framework, and disclosing the effect of nature-inspired operators.

Published by the IEEE Computer Society
1541-1672/18/$33.00 ©2018 IEEE

IEEE Intelligent Systems thanks selection committee members Georg Gottlob, Sarit Kraus, and Ray Perrault, who devoted a lot of time studying the nomination materials and deliberating carefully about the best young members of our community. In the end, the magazine's editorial and advisory boards unanimously agreed on the top 10 candidates.

IEEE Intelligent Systems is proud to present these aspiring AI researchers in 2018's list of "AI's 10 to Watch."

—Qiang Yang, Selection Committee chair and Editorial Board member

AI'S 10 TO WATCH

BO AN

Nanyang Technological University

Bo An is an associate professor in Nanyang Technological University's School of Computer Science and Engineering. He received a PhD in computer science from the University of Massachusetts, Amherst. His research interests include AI, multiagent systems, game theory, and optimization. His work has received numerous awards including the International Foundation for Autonomous Agents and Multiagent Systems' (IFAAMAS') 2010 Victor Lesser Distinguished Dissertation Award and the Institute for Operations Research and the Management Sciences' 2012 Daniel H. Wagner Prize for Excellence in Operations Research Practice. He was invited to give an early-career spotlight talk at the 2017 International Joint Conference on Artificial Intelligence. An also led the team that won the 2017 Microsoft collaborative AI challenge. He is on the *Journal of Artificial Intelligence Research*'s editorial board, is the *Journal of Autonomous Agents and Multi-Agent Systems*' associate editor, and was elected to IFAAMAS' board of directors. Contact him at boan@ntu.edu.sg.

Bo An, Nanyang Technological University

Computational Game Theory Security and Sustainability

Security is a critical concern, but because of limited resources, we can't provide full security coverage at all times. Instead, we must deploy our limited resources intelligently, taking into account the differences in target priorities; adversaries' responses to our security posture; and the potential uncertainty about their capabilities, knowledge, and aims.

Game theory is well suited to adversarial reasoning for security-resource-allocation and scheduling problems. With the problems cast as a Stackelberg game, novel algorithms efficiently provide solutions for randomized patrolling or inspection strategies. These algorithms have led to game-based decision aids for many real-world security domains including infrastructure security and wildlife protection. These initial successes point the way to important future research and applications in a wide range of security arenas.

Over the past few years, a major focus of our research has been studying the practical and challenging issues arising in realistic, complex security-resource allocation problems, including dynamic payoffs, protection externality, and strategic secrecy. Another important focus is addressing research challenges in new application domains such as cybersecurity, transportation-network security, adversarial machine learning, election protection, and nuclear smuggling. To solve these problems, we design exact algorithms—or approximate algorithms with quality guarantees—for solving large-scale games with continuous and infinite strategy space.

We also work on computational sustainability research, a vital field that applies computational techniques to address environmental, economic, and social problems arising from sustainable development. Humans play a central role in this by strategically deciding how to use public resources. We focus on providing optimal planning and policies in this area.

In the future, we must

- build more realistic models to make our proposed solutions effective in practice,
- develop new algorithms that efficiently compute optimal solutions of real-world security scenarios,
- further improve solutions' robustness,
- combine learning- and model-based approaches for resource allocation in data intensive domains, and
- integrate real-time information into the optimal allocation of security resources over a long period.

ERIK CAMBRIA

Nanyang Technological University

Erik Cambria is an assistant professor in Nanyang Technological University's School of Computer Science and Engineering. He's associate editor of several AI journals such as *IEEE Computational Intelligence Magazine*, *Artificial Intelligence Review*, *Knowledge-Based Systems*, and *Information Fusion*. Cambria also founded SenticNet, a Singapore-based university spin-off offering business-to-business sentiment-analysis services. Besides sentiment analysis, his research interests include natural language understanding, commonsense reasoning, and multimodal interaction. Cambria received a PhD in computing science and mathematics following the completion of a UK Engineering and Physical Sciences Research Council project born from a collaboration between the University of Stirling and the MIT Media Lab. Contact him at cambria@ntu.edu.sg.

Erik Cambria, Nanyang Technological University

Sentic Computing

With recent deep learning developments, AI research has gained new vigor and prominence. However, machine learning still faces three big challenges:

- It requires a lot of training data and is domain-dependent.
- Different types of training or parameter tweaking leads to inconsistent results.
- The use of black-box algorithms makes the reasoning process uninterpretable.

In my research group, we address such issues in the context of natural language processing (NLP) via our multidisciplinary approach, termed sentic computing. This approach aims to bridge the gap between statistical NLP and the many other disciplines necessary for understanding human language such as linguistics, commonsense reasoning, and affective computing.

Sentic computing enables the analysis of text not only at the document, page, or paragraph level but also at the sentence, clause, and concept level. This is possible thanks to an approach that is both top-down and bottom-up. It's top-down because it leverages symbolic models such as semantic networks and conceptual dependency representations to encode meaning. It's bottom-up because it uses subsymbolic methods such as deep neural networks and multiple kernel learning to infer syntactic patterns from data.

Coupling symbolic and subsymbolic AI is key for advancing from NLP to natural language understanding. Machine learning alone is useful just for making a good guess at understanding natural language based on experience because subsymbolic methods encode only correlation and use merely a probabilistic decision-making process. Natural language understanding, on the other hand, requires much more.

Sentic computing takes a holistic approach to sentiment analysis by handling the many problems involved in extracting meaning and polarity from text. Most other approaches address this as a simple categorization problem. Sentiment analysis, however, is actually a complex and composite task that requires leveraging many different NLP techniques at the same time.

Sentic computing addresses the compound nature of the problem via a three-layer structure that concomitantly handles tasks such as

- microtext analysis for handling abbreviations, acronyms, and emoticons;
- concept extraction for deconstructing text into words and multiword expressions;
- subjectivity detection for filtering out neutral content;
- sarcasm detection to identify and handle sarcastic expressions; and
- aspect extraction for enabling aspect-based sentiment analysis.

Our approach serves as a back end to many applications in areas such as e-business, e-commerce, e-governance, e-security, e-health, e-learning, e-tourism, e-mobility, e-entertainment, and more.

YOAV GOLDBERG

Bar Ilan University

Yoav Goldberg is a senior lecturer in Bar Ilan University's Computer Science Department. He received his PhD at Ben-Gurion University and did postdoctoral work as a research scientist at Google Research in New York City. He currently works on problems related to natural language processing (NLP) and machine learning. In particular, he's interested in syntactic structures, structured prediction models, learning for greedy decoding algorithms, multilingual language understanding, and cross-domain learning and semi-supervised representations. Lately, he has worked on neural-network-based methods for NLP.

Yoav Goldberg, Bar Ilan University

Natural Language Processing

NLP, which enables computers to generate and understand text in human languages, is an important challenge because human language is our primary means for communication and information transfer.

Language understanding today relies on advanced machine-learning techniques. In particular, deep-learning methods have been surprisingly effective in identifying the intricate patterns and regularities that govern sentence structure and meaning. We now have effective models for mapping from sentences to their underlying linguistic structure, as well as from sentences in one language to their translations.

Unfortunately, while we know how to apply deep networks so that they work well for many natural language tasks, we know little about interpreting the networks' behavior. We don't know which factors govern a trained model's behavior (why it predicted what it did) and which kinds of patterns it can learn. We also don't know the network's blind spots or how and when it will fail.

Interpreting and understanding deep neural networks' behavior and dynamics for language processing, and more rigidly characterizing their capabilities and limitations is a fascinating and crucial challenge, which my research lab is exploring.

We also work on the building blocks of analyzing and understanding text, from learning to understand how the words in a sentence connect to each other (syntax) to the behavior of the word "and" and the meaning of prepositions such as "on" or "for." For example, when we say, "We sat there for hours," "We sat there for lunch," and "He paid for me," the word "for" indicates duration, purpose, or beneficiary, respectively. How could a computer understand which is which?

Another example is our work on recovering missing elements of speech. When I say, "I'll give you 50 for it," you guess that 50 refers to a currency. When I say, "She just turned 50," you know its 50 years of age. When I say, "No thanks, I've had three already," you look back at the conversation to figure out what exactly I had three of. How can a computer do these things?

Perhaps the biggest question we face is how we move from understanding NLP's pieces to a more global and coherent view. NLP has progressed tremendously over the past decades. But to use the available tools and techniques, you must be an NLP expert. How can we change this so that we expose our findings in a unified way that non-experts can use? I think that's the big question we're facing today.

AKSHAT KUMAR

Singapore Management University

Akshat Kumar is an assistant professor in Singapore Management University's School of Information Systems. His research interests include planning and decision making under uncertainty with a focus on multiagent systems and urban-system optimization. Kumar received a PhD in computer science from the University of Massachusetts, Amherst. His thesis received best dissertation award at the 2014 International Conference on Automated Planning and Scheduling (ICAPS). His work has also received the outstanding application paper award at ICAPS 2014 and the best paper award in the 2017 AAAI Conference on Artificial Intelligence's computational sustainability track. Contact him at akshatkumar@smu.edu.sg.

Akshat Kumar, Singapore Management University

Planning and Inference for Multiagent Systems

As our society and urban environments rapidly become connected, they present an opportunity to deploy autonomous agents, from personal digital assistants to self-driving cars. Such agent-based systems promise to radically improve productivity and safety, while reducing human effort and risk. My research aims to develop computationally efficient techniques to help realize this promise by enabling large autonomous-agent teams to make coordinated decisions toward achieving a common goal, while coping with uncertainty and limited information.

This could apply, for example, to the optimization of autonomous taxi fleets, for which a key problem is positioning cabs strategically based on their changing local environment, such as demand and the availability of other taxis in the same area. My research addresses several challenges such as scalability to thousands of agents, uncertainty, partial observability, and resource-constrained optimization in urban settings.

We are developing scalable algorithmic techniques for control and planning in rich, formal models of multiagent coordination such as distributed constraint optimization and decentralized, partially observable Markov decision processes. Exact planning in such models quickly becomes intractable even for small multiagent systems. Previous approaches suffered from poor scalability or reliance on assumptions that limited their deployment. My PhD dissertation introduced numerous general scalable techniques and frameworks for a large class of practical multiagent planning problems.

In addition to standard techniques based on search and dynamic programming, my work develops close connections between multiagent planning and machine learning using graphical models and probabilistic inference. Thanks to these connections, I showed how numerous machine-learning approaches could also help in multiagent planning. Such connections have provided several new insights and scalable techniques for multiagent decision making.

Additionally, my work incorporates techniques from lifted-inference and graphical models to enable scalable planning by exploiting the fact that urban systems are often composed of a large number of nearly identical agents, such as taxis. A critical question for urban planners is how to design prescriptive policies for agents that result in the emergence of a desired collective behavior, such as reducing traffic congestion or increasing safety.

We are developing models and algorithms for such collective cooperative decision making and are focusing on designing multiagent reinforcement-learning-based approaches that optimize policies by using samples from domain simulators. Multiagent planning and control in our increasingly connected urban environment provides many exciting research directions and tremendous opportunities for practical impact.

WEI LIU

Tencent AI Lab

Wei Liu is the director of the Computer Vision Center at China's Tencent AI Lab and pursued his PhD at Columbia University. His thesis addresses large-scale machine-learning algorithms for classification and search. Liu's research interests are machine learning, data mining, information retrieval, and computer vision. He is particularly interested in developing binary coding and hashing techniques to handle big data. Liu is currently working on deep spatio-temporal models to tackle a variety of multimedia AI problems. Contact him at wl2223@columbia.edu.

Wei Liu, Tencent AI Lab

Multimedia AI

With the vast popularity of intelligent devices such as smartphones, large amounts of multimedia data have been produced and uploaded to the Internet. The volume of multimedia data has significantly outpaced the capability of conventional methods to effectively process, analyze, and understand it. My research focuses on developing powerful AI algorithms to create and process multimedia content, analyze and understand semantic meanings conveyed by diverse multimedia data, and precisely match multimedia content with a specific user's preferences.

My Tencent AI Lab research group is working to use deep-learning models to capture multimedia data's complicated characteristics and behaviors. We're focusing on three major multimedia-AI research topics: content generation, understanding, and distribution.

Generative adversarial networks have been widely used for image, text, and video generation. However, multimedia data consists of different modalities with different characteristics, which makes generation particularly challenging. Fortunately, some types of multimedia data, such as images and text, usually appear simultaneously and express or share the same meanings. Thus, correlated modalities can benefit one another and yield pleasant and vivid multimedia content during the generation process.

Multimedia data can be noisy and extremely large in volume, and is often generated and uploaded by many users. It can also be of multiple modalities, attributes, and labels. Thus, meaningfully understanding such data can be challenging. However, user behaviors (such as clicking or sharing) when accessing multimedia data on the Internet provide certain noisy labels of the data, which we can leverage to perform robust and scalable machine-learning methods.

Given the semantic correlations between multiple modalities, we can employ one modality to interpret the others. In doing so, multiple labels and attributes can be transferred from one modality to another. This will improve our interpretations of multimedia data. My research team is planning to build a deep-learning platform to understand massive amounts of multimedia data.

For multimedia content distribution, we must consider not only the data's characteristics but also user profiles. Via multimedia-content understanding, we learn about multimedia data's labels, attributes, and even semantic meanings. We incorporate user behaviors to portray the users' characteristics and preferences. My research team is utilizing this capability to develop a novel deep-learning architecture to sensibly match users' preferences with multimedia content in which they're interested.

The aforementioned paradigms have fostered several key AI techniques that use deep learning, including image and video editing, image and video content-based recommendation, video summarization, and video classification.

CYNTHIA MATUSZEK

University of Maryland, Baltimore County

Cynthia Matuszek is an assistant professor of computer science and electrical engineering at the University of Maryland, Baltimore County. Her research occurs at in the intersection of robotics, natural language processing, and machine learning, and their application to human-robot interaction. She works on building robotic systems that nonspecialists can instruct, control, and interact with intuitively and naturally. Matuszek received her PhD in computer science and engineering from the University of Washington. Contact her at cmat@umbc.edu.

Cynthia Matuszek, University of Maryland, Baltimore County

Robot Learning from Language and Context

As robots become more powerful, capable, and autonomous, they are moving from controlled industrial settings to human-centric spaces such as medical environments, workplaces, and homes. As physical agents, they will soon be able help with entirely new categories of tasks that require intelligence. Before that can happen, though, robots must be able to interact gracefully with people and the noisy, unpredictable world they occupy.

This undertaking requires insight from multiple areas of AI. Useful robots will need to be flexible in dynamic environments with evolving tasks, meaning they must learn and must also be able to communicate effectively with people. Building advanced intelligent agents that interact robustly with nonspecialists in various domains requires insights from robotics, machine learning, and natural language processing.

My research focuses on developing statistical learning approaches that let robots gain knowledge about the world from multimodal interactions with users, while simultaneously learning to understand the language surrounding novel objects and tasks. Rather than considering these problems separately, we can efficiently handle them concurrently by employing joint learning models that treat language, perception, and task understanding as strongly associated training inputs. This lets each of these channels provide mutually reinforcing inductive bias, constraining an otherwise unmanageable search space and allowing robots to learn from a reasonable number of ongoing interactions.

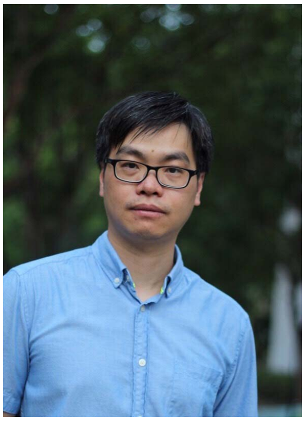
Combining natural language processing and robotic understanding of environments improves the efficiency and efficacy of both approaches. Intuitively, learning language is easier in the physical context of the world it describes. And robots are more useful and helpful if people can talk naturally to them and teach them about the world. We've used this insight to demonstrate that robots can learn unanticipated language that describes completely novel objects. They can also learn to follow instructions for performing tasks and interpret unscripted human gestures, all from interactions with nonspecialist users.

Bringing together these disparate research areas enables the creation of learning methods that let robots use language to learn, adapt, and follow instructions. Understanding humans' needs and communications is a long-standing AI problem, which fits within the larger context of understanding how to interact gracefully in primarily human environments. Incorporating these capabilities will let us develop flexible, inexpensive robots that can integrate into real-world settings such as the workplace and home.

SINNO JIALIN PAN

Nanyang Technological University

Sinno Jialin Pan is an assistant professor in Nanyang Technological University's School of Computer Science and Engineering. He received a PhD in computer science from Hong Kong University of Science and Technology. His research interests include machine learning and its applications to opinion mining, social-media analysis, information extraction in text, Wi-Fi-based indoor mobile localization, recommender systems, and software engineering. Contact him at sinnopan@ntu.edu.sg.



Sinno Jialin Pan, Nanyang Technological University

Transfer Learning

Recently, supervised-learning algorithms such as deep learning have made a great impact on our society, but it has become clear that they also have important limitations. First, the learning of supervised models relies heavily on the size and quality of the annotated training data. However, in many real-world applications, there is a serious lack of annotation, making it impossible to obtain high-quality models. Second, models trained by many of today's supervised-learning algorithms are domain specific, causing them to perform poorly when the domains change.

My long-term research goal is to develop transfer-learning techniques to build learning systems that can adapt to new and related domains with little or no human supervision. In the past, I've focused on discovering new features for different domains so that we can reduce domain discrepancies while preserving the original data's important properties. With the learned features, we can reuse standard supervised learning methods for cross-domain prediction. I have developed several effective transfer-learning methods using kernel embedding and deep-learning techniques, and have applied them to interdisciplinary application areas including wireless sensor networks, sentiment analysis, and software engineering.

Though transfer learning has attracted attention, it's still in its early stages. First, transfer-learning research primarily focuses on classification or regression problems. However, many real-world applications can't be modeled simply as classification or regression problems. Advanced machine-learning paradigms—such as deep reinforcement learning—might be more suitable to complicated learning problems. The goal of transferring knowledge across different deep reinforcement learning problems is not only to improve results but also to reduce the amount of work necessary. This is because finding an optimal policy in an end-to-end manner is extremely time consuming, which is a bottleneck for deep reinforcement learning.

Most current transfer-learning methods assume we can centralize data from different tasks for extracting commonality among tasks. In practice, though, data from different tasks might be geodistributed on local machines. Due to expensive communication costs and concerns about data privacy and security, centralizing different tasks' data to a master machine to perform transfer learning is impossible.

Therefore, during the next five to 10 years, I plan to explore two new transfer-learning research directions: transfer learning for deep-reinforcement learning and distributed optimization for transfer learning. I believe that if our research is successful, it will make transfer learning more practical and important for real-world applications.

B. ADITYA PRAKASH

Virginia Tech

B. Aditya Prakash is an assistant professor in Virginia Tech's Computer Science Department and a member of the school's Discovery Analytics Center. He received a PhD in computer science from Carnegie Mellon University. He has published one book, written more than 60 refereed papers, and obtained two US patents. His work received a best paper award at ACM's 2012 International Conference on Information and Knowledge Management, and best-of-conference selections at IEEE's 2011, 2012, and 2017 International Conference on Data Mining and IEEE/ACM's 2013 International Conference on Advances in Social Networks Analysis and Mining. Prakash's research interests include data mining, applied machine learning, and databases, with an emphasis on big data problems in large real-world networks and sequences. Contact him at badityap@cs.vt.edu.

B. Aditya Prakash, Virginia Tech

Network and Propagation Analysis

Networks such as gene-regulatory, communication, hospital, and online social networks are ubiquitous. They effectively model many real-life phenomena—including in cybersecurity, epidemiology, social-systems, and biology—because they expose both local dependencies and large-scale structure. Also, contagion-like propagation processes on networks (intuitively, where the state or action of an agent depends on the states or actions of its neighbors) can yield a variety of macroscopic behavior. They lead to challenging and exciting research problems such as modeling diseases—like Ebola and influenza—spreading through population networks, finding failure hotspots across power networks, and quickly detecting online viral memes.

These big-data problems typically are very challenging, as they entail high-impact real-world applications, as well as technical issues such as scalability and the need for ingesting multiple data sources. Understanding such propagation processes will eventually enable us to utilize them for our benefit. For example, understanding how epidemics spread helps us design more robust immunization policies.

To address these challenges, my interdisciplinary research draws concepts and techniques from theory and algorithms (combinatorial and stochastic optimization), systems (asynchronous computation), machine learning and statistics (minimum description length, deep learning, and graphical models), and nonlinear dynamics.

I've worked on characterizing epidemic spread in a unified fashion for an array of models. Using these novel characterizations, my group is designing realistic and implementable immunization policies that take into account the often noisy and uncertain quality of available surveillance data such as anonymized medical claims. For some of these models, we've developed the first provably near-optimal policies.

Eventually, we want to develop data-driven frameworks for controlling disease spread. These frameworks could use fine-grained surveillance data and models of people-to-people contact to develop robust immunization algorithms, seamlessly factoring in constraints and fusing noisy heterogeneous data sources. In contrast, most prior work has relied on a pre-emptive modeling approach, which is good for baseline strategies but isn't responsive to the increasing availability of surveillance data.

My work has also demonstrated that propagation over networks generates surprising connections among disparate areas. For example, working with cybersecurity researchers, we developed ensemble models of spreading malware attacks that also produced insights into unsafe human cyberbehavior. In fact, our study was among the first to use terabyte scale virus telemetry datasets to map global infection patterns.

MARIA VANINA MARTINEZ

Universidad Nacional del Sur

Maria Vanina Martinez is an adjunct researcher in Argentina's National Scientific and Technical Research Council (CONICET) and a member of the Universidad Nacional del Sur's Artificial Intelligence Research and Development Laboratory. Martinez earned her PhD in computer science from the University of Maryland, College Park, and did postdoctoral research at the University of Oxford's Department of Computer Science. Her research interests include reasoning under uncertainty, inconsistency management in relational databases and knowledge bases, defeasible reasoning, and argumentation. Her current research focuses on the development of a family of inconsistency- and uncertainty-tolerant data models for knowledge bases extracted from the Web, and on scalable query-answering algorithms. Contact her at mvm@cs.uns.edu.ar.

Maria Vanina Martinez, Universidad Nacional del Sur

Toward Hybrid Data- and Knowledge-Driven Decision Support

Most AI work has focused on either the development of automated ways of finding and collecting data or the construction of knowledge-representation models. I believe the next big step is delving deeper into how automated systems could use available knowledge to enhance humans' decision-making capabilities in real-world applications.

For the development of intelligent decision-making support systems (IDMSSs), AI could automatically combine acquired data with domain expertise from the users that the tool is supposed to help. This could happen by combining data-driven AI (machine learning), which would extract as much information as possible from raw data, with knowledge-based models, which would handle more abstract and complex cognitive tasks.

I've worked on formalizing solutions to several central IDMSS-related problems. To enable such systems to understand and work with people, I focused on preference modeling, the management of inconsistent or conflicting information, understanding the dynamics of knowledge and beliefs, and modeling social interactions' effects on reasoning tasks.

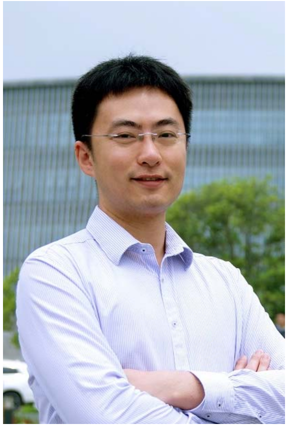
During my PhD research, I developed a framework for policies for managing inconsistencies in relational databases. Later, I deepened my research on inconsistency management by focusing on more-expressive languages that let us model uncertainty in different ways. They let us define reasonable and computationally efficient semantics for inconsistency-tolerant query answering. I've also worked on reasoning with preferences, especially in combination with probabilistic inference models, which relate to my work on inconsistency management. Furthermore, I've researched several aspects of belief-revision theory and its connection to inconsistency management. In the process, I've developed frameworks for merging and revising knowledge that focus on finding meaningful results by considering conflicts from a more global perspective than that used in classical approaches, leveraging information about the context in which they must be resolved. Lately, I've been working on the dynamics of knowledge in social media, particularly on modeling the diffusion of beliefs in complex social networks.

So far, I've tackled this research from a knowledge-representation point of view, defining formal models for complex reasoning with uncertain knowledge. However, there's a gap between the formalization of knowledge-based systems and the actual construction of IDMSSs. My next goal involves working toward shrinking this gap by engineering hybrid data- and knowledge-driven approaches, effectively finding a sweet spot among raw heterogeneous data, semantics, and abstract models for reasoning.

YANG YU

Nanjing University

Yang Yu is an associate professor in Nanjing University's Department of Computer Science and Technology. He received a PhD in computer science from Nanjing University. His dissertation received the National Outstanding Doctoral Dissertation Award of China and the China Computer Federation Outstanding Doctoral Dissertation Award. Yu received the best-paper award at 2008's Pacific-Asia Conference on Knowledge Discovery and Data Mining, 2011's Genetic and Evolutionary Computation Conference (theory track), and 2016's International Conference on Intelligent Data Engineering and Automated Learning. Contact him at yuy@nju.edu.cn.



Yang Yu, Nanjing University

Derivative-Free Optimization in Machine Learning

Machine learning requires identification of the model best fitting the circumstances. To search for that model, a learning system's designers must consider the capabilities of the optimization algorithm, which lies at the heart of such learning tasks. The algorithm's limitations restrict the choice of model types, data representations, and learning objectives. Gradient-based optimization is widely employed, but learning tasks are not always simple enough for this approach. Most learning problems are highly sophisticated and require stronger optimization.

There are useful derivative-free optimization tools, such as evolutionary algorithms. However, many believe you need some luck to use these tools successfully because there are still unanswered questions about how long they will take to yield good results, how good the results will be, and how to choose the proper parameters.

My research has focused on establishing derivative-free optimization's theoretical foundation. Our studies have covered time-complexity analysis, approximation ratio analysis, operators' effects, and more recently, optimization in a high-dimensional space and noisy environment. We now can explore designing better learning systems without worrying about nonconvexity and nondifferentiability, which cause problems for gradient-based approaches.

For combinatorial optimization, we developed the Pareto optimization approach—which we call ParetoOpt—for constrained Boolean functions. ParetoOpt has yielded the best polynomial-time approximation ratio for the subset selection problem so far and thus has enabled us to create a state-of-the-art sparse regression-learning algorithm. ParetoOpt has also achieved the best theoretical and empirical performance in more settings, including subset selection with general cost constraints and minimization of the ratio of monotone set functions.

For continuous domain optimization, we developed the sampling-and-classification (SAC) framework. We've shown that it can approximate the global optimum of any function in polynomial time for the approximation gap, dimensionality, and local Lipschitz constants. Using SAC, we can solve the highly nonconvex ramp-loss learning problem better than state-of-the-art gradient methods.

Currently, derivative-free optimization has limitations in performing machine-learning tasks. We are working on ways to overcome this for applications in large-scale, noisy, and sophisticated tasks, particularly for real-world reinforcement learning.

We believe that derivative-free optimization will play an even more important role when machine learning integrates with other AI components such as planning and logic representation, for achieving the next level of intelligence.