Künstl Intell (2014) 28:215–221
DOI 10.1007/s13218-014-0309-0

INTERVIEW

Interview with Professor Sarit Kraus

Nils Bulling

Published online: 22 June 2014

Sarit Kraus is a Professor of Computer Science at Bar-Ilan University and an Adjunct Professor at the Institute for Advanced Computer Studies, University of Maryland. She obtained her B.Sc. in mathematics and computer science (1982), and her M.Sc. (1983) and Ph.D. (1989) in computer science from the Hebrew University in Jerusalem. She has focused her research on intelligent agents and multi-agent systems, in particular on the development of intelligent agents that can interact proficiently with people. She has also contributed to research on social networks, nonmonotonic reasoning, adversarial patrolling, machine learning and clustering and optimization.

Kraus was awarded the IJCAI Computers and Thought Award, the ACM SIGART Agents Research award, the EMET prize and was twice the winner of the IFAAMAS influential paper award. She is an AAAI and ECCAI fellow, the member of Academia Europaea and a recipient of the advanced ERC grant. She also received a special commendation from the city of Los Angeles, together with Prof. Tambe, Prof. Ordonez and their USC students, for the creation of the ARMOR security scheduling system. She has published over 300 papers in leading journals and major conferences. She is the author of the book Strategic Negotiation in Multiagent Environments (2001), a co-author of a book on Heterogeneous Active Agents (2000), both published by MIT Press, and a co-author of a forthcoming book on Principles of Automated Negotiation in Cambridge Press. Kraus is an associate editor of the Annals of Mathematics and Artificial Intelligence Journal and is on the editorial board of the Journal of Autonomous Agents and Multi-Agent Systems, the Journal of Applied Logic, and the Journal of Philosophical Logic (JPL).

N. Bulling (✉)
Department of Informatics, Clausthal University of Technology, Clausthal-Zellerfeld, Germany
e-mail: bulling@in.tu-clausthal.de

Interview The following interview was conducted on March 17, 2014.

KI: Prof. Kraus, you obtained your bachelor's, master's, and PhD degree in computer science in the 1980s. What led you to become interested in Computer Science?

SK: That's an interesting question. Originally, when I began my academic studies I wanted to become a speech therapist. But unfortunately by the time I made my decision, the registration for the speech therapists degree was already closed. On the other hand, I would be accepted without any problem to the computer science program, which was in its beginning stages and not much in demand. So, my father said that instead of waiting 1 year and then registering for the speech therapist program why not start in computer science. He said computer science is the future, and even if you become a speech therapist studying 1 year of computer science will be useful to you. So, I registered for the computer science program and the rest is history. Funny enough, recently I was involved in a project where we built an automated speech therapist. But that's a whole different story.

KI: Did you ever regret that choice?

SK: No, I really like my research and especially to develop new methods and algorithms for challenging problems. I believe that computer science is one of the most fascinating research areas. It is challenging and so wide; you can study almost anything. Whether you like theoretical algorithms, applications or whether you prefer working with people or graphical interfaces, everything is possible in the context of computer science. Computer science is not just one topic or one area it is really wide. The questions, quite often, are difficult to solve but if you solve a problem the solution will serve many applications. Computer science, I believe, has changed the world over the last 40 years.

KI: When did you realize that you would like to stay in academia?

SK: After I graduated I was so fascinated by my undergraduate studies that I said why not continue to a masters degree. From there it was a smooth transition to a PhD. The big decision was when I had to decide whether to do a postdoc abroad or to stay in Israel. In Israel, doing a postdoc abroad is a pre-requisite to getting a position in academia. This was a really tough decision for me. But after I spent 2-years at Maryland as a postdoc, academia was the obvious direction for me.

KI: In 2011 you received a prestigious ERC advanced grant for your work on "Computers Arguing with People". Can you tell us about the project and its goals?

SK: I'm very interested, in general, in computer systems that interact with people. First of all, developing systems that can interact with people is extremely difficult. As a joke I always say people are problematic; they always ruin my experiments. By nature I'm a logician. When I deal with multi-agent systems (MAS) where all agents are computers, in general I know what to expect in the experiments. With people it is much more difficult. Already from the beginning, even during my PhD studies, I worked with computers that interact with people. During my PhD studies, I developed DIPLOMAT[1] which is a computer system that plays the Diplomacy game. The main capability needed in Diplomacy is the ability to negotiate. Back then it seemed like science fiction. Over the years my work did not deal with people because it was very difficult to acquire funding and researchers were not interested in computer systems that interact with people. I was extremely pleased when I received the ERC grant for the project on systems that negotiate with people. Negotiation in this case is in a broad sense. One part of the project is about bargaining/negotiation. Another part of the project is about persuasion; for example, the system tries to persuade people to save energy. In both cases culture plays an important role. As a result we are dealing with culturally sensitive agents for negotiation, bargaining, and persuasion. The third part of the project includes an agent for interviewing, especially interview training. We built a virtual suspect to be used for training law enforcement officers in investigations. We incorporated a virtual human that is not only text-based but also includes a virtual character.

KI: The agent is supposed to support humans in their decision making?

SK: The agent can represent people, be used for training people or support them. In the interviewing domain, the computer system plays the role of a suspect used for training. Our studies show that negotiating with an agent is a good way to train people in negotiations [11, 14]. It is at least as good as having humans play the role of the other side. There are situations where we want the agent to replace a person in the negotiation. For example, in the persuasion setting we are currently conducting a project with General Motors where agents try to persuade people to save energy when using electrical cars [3]. In another setting we are dealing with crowd sourcing [2]. In this setting the agent negotiates on your behalf [7, 13]. I must tell you that my agents are better than me in negotiation and bargaining. I am really pleased that the ERC is supporting this project; it is a risky project so it is really great to have such support.

KI: You have received many outstanding awards for your work. Only recently you were again awarded the IFAAMAS Influential Paper Award 2014, jointly with Onn Shehory.[2] Congratulations! So,...

SK: Thank you!

KI: ...if you had to explain your most significant contributions to nonscientists what would you say?

SK: I think my main contribution is the development of computer systems that are autonomous, that are independent, and can interact with each other and with people. When I entered this field in the 80's the community was called Distributed Artificial Intelligence (DAI). I think I, myself as well as other people, added the issue of individualism, i.e., that agents can be self-interested, to the DAI community. Agents often represent organizations, and

[1] Paper [9].

[2] IFAAMAS Influential Paper 2014 [16].

organizations are self-interested while collaborating. For computer systems this was a lot harder to understand in the 80s and the 90s, but now with the internet it's so clear to everybody that computers do not exist in isolation. The idea that computer systems are self-interested was an important contribution, in particular, how self-interested agents should cooperate and should form teams. For example, the model of Shared Plans[3] that I developed with Barbara Grosz from Harvard was exactly about this: What are the necessary properties of teams of self-interested agents. The idea of self-interested agents is important for both teams and even more for competitive environments. For example, we use methods from game theory for competitive settings such as security in airports [15], a project with Milind Tambe. I also had a project with Gal Kaminka and Noa Agmon on patrolling robots [1]. We introduced the issue of adversaries to research on robotic patrolling.

KI: Do you use classical game theory in this work?

SK: When we build a computer system we would like the computer system to maximize the expected utility or to optimize some criteria. This is from the point of view of building the system. However, we need to take into consideration the agents we would like communicate with, both in competitive and collaborative settings. If we perceive it in this manner we need to understand that, for example, people don't always maximize expected utility. Well, some people from game theory say that they maximize some expected utility but we just don't know which utility function. But I won't go into these philosophical questions here. I'm simply saying that given my model I almost never see people who maximize expected utility. As such, my methodology is to: Try to model our adversary or counterpart and integrate his model into the optimization problem of our agent. I have a great example. You asked me whether I use game theory. Usually, when I want to build an agent that interacts with people the first thing I do is check whether I can use an equilibrium agent—mainly because I would like to find out what the equilibria are, for example subgame perfect equilibria. Usually it doesn't work. In a recent experiment with Kobi Gal, Galit Haim and Bo An, where we looked at a negotiation game, we ran the experiment in three countries: China, the US, and Israel. This experiment was really interesting. We found a subgame perfect equilibrium strategy. In one of the roles in the negotiation the equilibrium strategy did extremely well. On the other hand, in another role it did very poorly. When I say poorly I mean that it did worse than the people who played the game. When investigating this interesting finding, we realized that people playing one of the roles do not play rationally. Given this observation we were able to change the model in a way that incorporates this strange behavior and the agent in the updated setting still maximized expected utility. We ran this modified setting in the three countries and it did much better. So, this is an example that even if you do want to use game theory you must take into consideration the behavior of people. Note that, often also agents developed by people do not maximize expected utility.

[3] IFAAMAS Influential Paper 2007 [6].

KI: From a theoretical point of view, do you think it makes sense to use classical game theory in theoretical foundations of MAS?

SK: It's a wonderful tool. I think it gives us some perspective on the problem. There are situations where game theory—in particular, the research area of mechanism design—has a lot of interesting applications. Even though, applying mechanism design to people is problematic, the use of mechanism design on computer systems is quite beneficial.

KI: Now, I would like to talk with you more specifically about multi-agent decision making. How would you explain decision making to, say, computer science students?

SK: A decision making problem involves several options—which I'll call actions—to choose from. Each action has several possible outcomes associated with them. In an ideal world the outcome will be deterministic. But usually this is not the case and each action may result in several outcomes where there is some probability distribution associated with the actions over the outcomes. We can compute the expected benefit for each of the actions. A rational player would choose the action with the highest expected benefit. For example, we would like to buy a lottery ticket. The question is how to decide whether to buy or not to buy a ticket. If we buy a ticket there are two possible outcomes, either we win or we lose. The probability that we'll win, say, 2 million euro is extremely low and the probability that we'll get nothing is extremely high. There are different concepts that need to be examined. If we are simply interested in the expected monetary value it's clear that we shouldn't buy a lottery ticket because the expected monetary value is negative. However, if we are prepared to take risks or if we enjoy the game we may buy a lottery ticket. Decision making is looking into modeling this kind of problem.

KI: ...and what about multi-agent decision making?

SK: Ok, in decision making the question is how do we take other players into consideration? If our decision to buy a lottery ticket doesn't influence the other peoples' decisions then the other peoples' strategies are just another factor in our decision making. However, if our decision influences the actions that will be chosen by others then we go from classical decision making to game theory. Suppose we need to schedule a meeting time and let's assume that we cannot communicate with each other. I need to decide whether to meet you at 2 pm or at 4 pm. My main goal is that we'll meet. I prefer 2 pm over 4 pm. You may prefer 4 pm over 2 pm. But if we don't meet our reward is zero. If we meet there is some positive reward. If you know that I'll go to the meeting point at 2 pm and you are sure about it then you'll go at 2 pm, too. So, my decision influences your decision. When this is the case then we go from decision theory to game theory. Not only are we trying to maximize expected utility but we are also attempting to reach stable outcomes.

KI: 25 years ago you developed DIPLOMAT (as you already mentioned), a Diplomacy-playing agent. What types of decision making techniques did you use?

SK: In the Diplomacy game first some strategizing capability is needed which is related to search, expected utility, and so forth. Then, we need to model our negotiators and collaborators. In Diplomacy, on the one hand, we must form coalitions with other players to reach the best deal in the negotiation. On the other hand, since we eventually want to win the game, at some point in time we need to breach the agreement.

One of the main innovations of the DIPLOMATagent, is the fact that it consists of five different modules: a Prime Minister, a Ministry of Defense, a ForeignOffice, a Headquarters, and Intelligence. They all worktogether to achieve a common goal winning the game.Different personality traits such as aggressiveness, willingness to take chances and loyalty are implemented inthe different modules. These personality traits affect thebehavior of the agent and can be changed during eachrun, which allows DIPLOMAT to change its behaviorfrom one game to another. In addition, the agent has alimited learning capability which allows it to try to estimate the personality traits of its rivals (e.g., their riskattitude). Based on this, DIPLOMAT assesses whetheror not the other players will keep their promises. In addition, DIPLOMAT incorporates randomization in itsdecision making component. This randomization, influenced by DIPLOMATs personality traits, determineswhether some agreements will be breached or fulfilled.

Retrospectively, when I think about the methods we used, DIPLOMAT was a rule based agent that based its decision making on a social utility function. A social utility function takes into consideration the agent's own utility and the utility of the other side. Because in negotiations where agreements are not enforceable, as in Diplomacy, if you sign an agreement that is extremely good for you the probability that the other side will keep it is quite low. You need to balance between the utility for yourself and the utility for the other side. Thus, already in DIPLOMAT we considered the issue of modeling the other side, the technology of social utility and some rules that came from experts, in this case from me. This methodology should be used when you do not have data on human behavior. We used a similar methodology more recently in PURB that negotiated well with people from the US, Israel and Lebanon [4].

KI: Was the agent able to win against human players?

SK: Setting up a game with people was really challenging. 25 years ago there was no Internet and the only people who had emails were people in academia. So, DIPLOMAT negotiated with people only in two games because it was extremely difficult to organize such games. If I would have done it today I could find hundreds of people who would play Diplomacy over the internet. But back then it was extremely difficult; so, we ran only two games and it played quite well. I think the main point I wanted to get across to the community with DIPLOMAT was that if you build and develop an agent that should proficiently interact with people you must run experiments with people. But, as I said at the beginning, running experiments with people is very distressing but is necessary to evaluate your agent.

KI: In one of your articles, which was published about 4 years ago in the Communications of the ACM,[4] you compared automatic negotiators. One of the observations you made was that you "...cannot find one specific feature that connects them [the negotiators] or can account for their good negotiation skills". So, what makes a good automatic negotiator and why is it so difficult to identify good techniques in general?

SK: Mainly because there is a lot of uncertainty concerning people's behavior. I think over the last 4 years I have actually formed a methodology which, I believe, should be used when building a computer system that interacts with people. If we can collect data about peoples' behavior then we are in good shape. If we are able to do this we can build a model to predict how people will respond to offers. Trying to predict which offers people will make is very difficult. But at least predicting how people will respond to an offer, and in agreements that are not enforceable, whether they'll keep their agreements or not, I believe, can be done. Then we can integrate the model in the decision making of the agent. The problem is that if we do not have enough data to build a prediction model of people then we are in bad shape. But, in general, a good automatic negotiator will offer agreements that are good for both sides.

[4] Paper: [12].

KI: It seems to be very hard to develop a general purpose negotiator. Is it always domain specific?

SK: I think currently such negotiators are mainly domain specific. Well, they are not really domain specific in the sense that often we can keep the same agent and just replace the human prediction model. So, it is not that we need to build a specific agent for each domain. Sometimes we can use an agent from one domain to another but the model of peoples' behavior is necessary for each new domain. In particular, we can use an agent that negotiates over a multi-issue agreement (take for example an employer-candidate negotiation) for new multi-issue negotiations (for instance, neighbor disputes). However, building human prediction for each setting can improve the agent. This is what we can do today. That's why we are working on allowing more general methods.

KI: In the comparison presented in your ACM article, some agents used quantitative decision making approaches and others qualitative ones. Are there general settings in which one approach is more suitable than the other?

SK: If I would build an agent that acts in situations where everybody is rational and maximizes expected utility, I would use quantitative decision making. However, if agents interact with other agents or people that do not maximize expected utility then qualitative approaches come to mind. For example, the idea of using social utility functions or trying to model the other side by using decision trees. So, the use of qualitative decision making procedures becomes necessary when facing other agents, either automated or humans, that do not maximize expected utility.

KI: What about logic-based approaches?

SK: Well, as I mentioned I am a logician by nature—actually, my most cited paper is about non-monotonic logic [10]. I think that logical methods give us a formal methodology to accurately define criteria and requirements for agents. For example, in my work on shared plans with Barbara Grosz we use logic for specification. But logical methods are still problematic from a computational point of view and also with respect to maximizing expected utility. So, I would say that some integration of logic-based methods with decision making—qualitative or quantitative—is necessary. Interestingly, recently I have been participating in a project partially funded by Intel on argumentation for supporting people in argumentation in deliberation. Here I'm back to logic. In this project we use logic as a basic principle. While theoretical argumentation is a really nice theory people do not act according to theory. So, we use features from the logical model to predict the next argument a person will use in the deliberation. In this case we use machine learning to predict what people will do. As the project progresses I am quite sure that eventually the agent will somewhat maximize expected utility given the prediction of people and the logical foundation. I believe that only the integration of all models will help us advance in the development of agents that can interact well with people.

KI: Where do you see promising industrial application areas for automatic negotiators in general?

SK: I have been thinking about agents interacting with people since the 80s. But it was extremely frustrating because researchers, people from industry, weren't interested. For many years I used only agent settings, with multi-agents but without people. But in the past 5 years many have become interested in the interaction of agents with humans. This is very exciting. I have a project with General Motors, with a hospital about an automated speech therapist, Intel is interested and security institutions are extremely interested. Even the US army has realized that it is important for soldiers to know how to shoot but knowing how to interact with a villager in Afghanistan is probably even more important. So, training people who must interact with people from different cultures is very important [5]. But currently, I think that the most important domain for application is health. We are all getting older and living much longer. The progress in healthcare is demented. Even if we are ill and have some disabilities we still live. So, the main issue is how to improve our quality of life. The problem is, for example, if people need physiotherapy or cognitive training, or some care, there are not enough people to provide such support and care. It seems that people now realize that computer systems must play an important role in such settings. As you can imagine one of the ideas is to use serious games for physiotherapy, speech therapy, cognitive therapy and other situations. The problem is that although people in health have tried to use computer systems for these important tasks over the last 10 years it has not caught on. I think that the reason is because it's not enough to tell a person to play a game as part of his physiotherapy. Usually, there should be a person who is in the loop who designs a personalized training plan and monitors the behavior of the patient, and especially encourages and incentivizes the patient. Given our progress in the last few years, I think that agents that can negotiate and interact proficiently with people can now be used for this purpose. I believe that some of the tasks that are currently done or not done, because there are not enough people, by the caregiver or specialist can now be replaced by an intelligent agent. Today, intelligent agents can monitor and encourage humans. I believe that this is an extremely important application and I hope to be involved in these developments [8].

KI: I would briefly like to touch upon decision making in social networks, another topic you are interested in. Could you give us an idea about social networks and how social factors affect multi-agent decision making?

SK: I have two projects related to social networks. In one we try to figure out how to find missing information. We look at the nodes of the social networks as the agents and we are trying to identify missing nodes [17].

In another work that will be reported in this coming AAMAS [18] we consider a small social network and people in the network need to form teams to work together. When I form a team I prefer to be with my friends—that is what social networks are all about—and I would like not to be with people that I don't like or do not know. One question is what are the algorithms to form these teams given the relationship between the nodes in the social networks. You can think about the nodes as people and the edges, which have costs, indicate the relation between people. Then one question is whether we can find algorithms to form a coalition that maximize social welfare, and the other question is, can we form coalition that are stable. We had another issue. Let's assume I am looking at the company and there is a social network inside the company. I would like to form teams to perform some tasks. I can influence the teams to be created by introducing people to each other in order to increase the social welfare. Of course, introducing people to each other is expensive. So, how do I decide which people to connect? I think this is again an instance of the problem of optimization and stability. For the agent it is always some kind of optimizing problem.

KI: In the next 10 years where do you see the main challenges in the field of multi-agent decision making?

SK: I think that we are proceeding with many more applications and multi-agent decision making has reached a rather mature point. We are moving on to more practical domains and, from my point of view, we face two challenges. First, I believe that looking at more realistic deployments of multi-agent systems will involve a great deal of post theoretical and practical research questions. This is very exciting. Second, we need to develop an experimental research culture. There should be very strict evaluation methodologies of how to evaluate experiments and the deployment of systems to ensure the preservation of high standards of research. Evaluation is very difficult in the real world. There are many opportunities out there. I can think about almost any domain where multi-agent decision making could play an important role, starting with cars, health, banks, and economics. I can hardly think of a domain where there are no intelligent computer systems that interact with themselves and with people. There are many possible applications and many open questions. In particular, you asked me if we can move from domain specific agents to general agents. Our dream is that eventually there will be an agent that will be able to do many tasks, learn from others and be a general purpose agent. Well, I don't think that we will accomplish this in the next 10 years. Currently, if we have some specific problem and we'll spend enough money and effort I am certain we'll be able to design and implement an agent that acts extremely well, usually better than people in this specific decision making scenario. What we don't know is how to build general purpose decision making software or agents like this. I think in the next 10 years, while deploying the methodologies, methods and algorithms that we have developed, we can proceed with the development of more general purpose agents.

KI: So, there still remains a lot to do...

SK: Yes, there is no question that I'll be busy for the next 20 years.

KI: Before we finish, would you like to add anything?

SK: I would like to say that as my agents are team players and they interact, negotiate and collaborate with other agents and with people, I deem that my research is also always done by a team. I like to collaborate with people. In this interview I mentioned a few of my collaborators but I have many more and I strongly believe that we need to work together if we would like to make progress in MAS. I find it very rewarding to collaborate with people. I collaborate with researchers from the US, Europe, Lebanon and China, and all around the world...; Of course I have collaborators from Computer Science, but also from Psychology, Political Science and other research areas—I suppose over the years I have had more than 285 coauthors. Of course I could not mention them all in this interview and I apologize for that. So, I would like to thank my collaborators and hope that they'll keep collaborating with me in the future to continue to jointly develop multiagent systems.

KI: Many thanks for your time and for giving this very interesting interview!

References

1. Agmon N, Kraus S, Kaminka GA (2008) Multi-robot perimeter patrol in adversarial settings. In: ICRA, pp 2339–2345
2. Azaria A, Aumann Y, Kraus S (2013) Automated agents for reward determination for human work in crowdsourcing applications. Auton Agents Multi-Agent Syst, pp 1–22
3. Azaria A, Kraus S, Goldman C, Tsimhoni, O (2014) Advice provision for energy saving in automobile climate control systems. In: Proceedings of the Twenty-Sixth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-14)

4. Gal Y, Grosz BJ, Kraus S, Pfeffer A, Shieber SM (2010) Agent decision-making in open mixed networks. Artif Intell 174(18): 1460–1480
5. Gal Y, Kraus S, Gelfand M, Khashan H, Salmon E (2011) An adaptive agent for negotiating with people in different cultures. ACM TIST 3(1):8
6. Grosz BJ, Kraus S (1996) Collaborative plans for complex group action. Artif Intell 86(2):269–357
7. Haim G, Gal YK, Gelfand M, Kraus S (2012) A cultural sensitive agent for human-computer negotiation. In: AAMAS, pp 451–458
8. Kraus S (2014) Intelligent agents for rehabilitation and care of disabled and chronic patients (extended abstract). In: Towards better and more affordable healthcare: incentives, game theory, and artificial intelligence workshop
9. Kraus S, Ephrati E, Lehmann D (1994) Negotiation in a non-cooperative environment. J Exp Theor Artif Intell 3(4):255–281
10. Kraus S, Lehmann DJ, Magidor M (1990) Nonmonotonic reasoning, preferential models and cumulative logics. Artif Intell 44(1–2):167–207
11. Lin R, Gal YK, Kraus S, Mazliah Y (2014) Training with automated agents improves people's behavior in negotiation and coordination tasks. Decision Support Syst 60:1–9
12. Lin R, Kraus S (2010) Can automated agents proficiently negotiate with humans? Commun ACM 53(1), 78–88 (2010). doi:10.1145/1629175.1629199. URL http://doi.acm.org/10.1145/1629175.1629199
13. Lin R, Kraus S, Wilkenfeld J, Barry J (2008) Negotiating with bounded rational agents in environments with incomplete information using an automated agent. Artif Intell 172(6–7):823–851
14. Lin R, Oshrat Y, Kraus S (2009) Investigating the benefits of automated negotiations in enhancing people's negotiation skills. In: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1, International Foundation for Autonomous Agents and Multiagent Systems, pp 345–352
15. Paruchuri P, Pearce JP, Marecki J, Tambe M, Ordóñez F, Kraus S (2008) Playing games for security: an efficient exact algorithm for solving bayesian stackelberg games. AAMAS 2:895–902
16. Shehory O, Kraus S (1998) Methods for task allocation via agent coalition formation. Artif Intell 101(1):165–200
17. Sina S, Rosenfeld A, Kraus S (2013) Solving the missing node problem using structure and attribute information. In: ASONAM, pp 744–751
18. Sless L, Hazon N, Kraus S, Wooldridge M (2014) Forming coalitions and facilitating relationships for completing tasks in social networks. In: Proceeding of the AAMAS