Part IV: Reminiscences

31

Questions and Answers

Nils J. Nilsson

Few people have contributed as much to artificial intelligence (AI) as has Judea Pearl. Among his several hundred publications, several stand out as among the historically most significant and influential in the theory and practice of AI. With my few pages in this celebratory volume, I join many of his colleagues and former students in showing our gratitude and respect for his inspiration and exemplary career. He is a towering figure in our field.

Certainly one key to Judea's many outstanding achievements (beyond dedication and hard work) is his keen ability to ask the right questions and follow them up with insightful intuitions and penetrating mathematical analyses. His overarching question, it seems to me, is "how is it that humans can do so much with simplistic, unreliable, and uncertain information?" The very name of his UCLA laboratory, the Cognitive Systems Laboratory, seems to proclaim his goal: understanding and automating the most cognitive of all systems, namely humans.

In this essay, I'll focus on the questions and inspirations that motivated his ground-breaking research in three major areas: heuristics, uncertain reasoning, and causality. He has collected and synthesized his work on each of these topics in three important books [Pearl 1984; Pearl 1988; Pearl 2000].

1 Heuristics

Pearl is explicit about what inspired his work on heuristics [Pearl 1984, p. xi]:

> The study of heuristics draws its inspiration from the ever-amazing observation of how much people can accomplish with that simplistic, unreliable information source known as *intuition*. We drive our cars with hardly any thought of how they function and only a vague mental picture of the road conditions ahead. We write complex computer programs while attending to only a fraction of the possibilities and interactions that may take place in the actual execution of these programs. Even more surprisingly, we maneuver our way successfully in intricate social situations having only a guesswork expectation of the behavior of other persons around and even less certainty of their expectations of us.

The question is "How do people do that?" The answer, according to Pearl, is that they use heuristics. He defines *heuristics* as "criteria, methods, or principles for deciding which among several alternative courses of action promises to be the most effective in order to achieve some goal." "For example," he writes, "a popular method for choosing [a] ripe cantaloupe involves pressing the spot on the candidate cantaloupe where it was attached to the plant, and then smelling the spot. If the spot smells like the inside of a cantaloupe, it is most probably ripe [Pearl 1984, p. 3]."

Although heuristics, in several forms, were used in AI before Pearl's book on the subject, no one had analyzed them as profitably and in as much detail as did Pearl. Besides focusing on several heuristic search procedures, including A*, his book beneficially tackles the question of how heuristics can be *discovered*. He proposes a method: consult "simplified models of the problem domain" particularly those "generated by *removing constraints* which forbid or penalize certain moves in the original problem [Pearl 1984, p. 115]."

2 Uncertain Reasoning

Pearl was puzzled by the contrast between, on the one hand, the ease with which humans reason and make inferences based on uncertain information and, on the other hand, the computational difficulties of duplicating those abilities using probability calculations. Again the question, "How do humans reason so effectively with uncertain information?" He was encouraged in his search for answers by the following observations [Pearl 1993]:

1. The consistent agreement between plausible reasoning and probability calculus could not be coincidental, but strongly suggests that human intuition invokes some crude form of probabilistic computation.
2. In light of the speed and effectiveness of human reasoning, the computational difficulties that plagued earlier probabilistic systems could not be very fundamental and should be overcome by making the right choice of simplifying assumptions.

Some ideas about how to proceed came to him in the late 1970s after reading a paper on reading comprehension by David Rumelhart [Rumelhart 1976]. In Pearl's words [Pearl 1988, p. 50]:

> In this paper, Rumelhart presented compelling evidence that text comprehension must be a distributed process that combines both top-down and bottom-up inferences. Strangely, this dual mode of inference, so characteristic of Bayesian analysis, did not match the capabilities of either the "certainty factors" calculus or the inference networks of PROSPECTOR – the two major contenders for uncertainty management in the 1970s. I thus began to explore the possibility of achieving distributed computation in a "pure" Bayesian framework, so as not to compromise its basic capacity to combine bi-directional inferences (i.e., predictive and abductive).

Previous work in probabilistic reasoning had used graphical structures to encode probabilistic information, and Pearl speculated that "it should be possible to use the links [in a graphical model] as message-passing channels, and [that] we could then update beliefs by parallel distributed computations, reminiscent of neural architectures [Pearl 1988, p. 51]." In the course of developing these ideas, Pearl says [Pearl 1988, p. 50]:

> it became clear that *conditional independence* is the most fundamental relation behind the organization of probabilistic knowledge and the most crucial factor facilitating distributed computations. I therefore decided to investigate systematically how directed and undirected graphs could be used as a language for encoding, decoding, and reasoning with such independencies.

Pearl's key insight was that beliefs about propositions and other quantities could often be regarded as "direct causes" of other beliefs and that these causal linkages could be used to construct the graphical structures he was interested in. Most importantly, this method of constructing them would automatically encode the key conditional independence assumptions among probabilities which he regarded as so important for simplifying probabilistic reasoning.

Out of these insights, and after much hard work by Pearl and others, we get one of the most important sets of inventions in all of AI – Bayesian networks and their progeny.

3 Causality

Pearl's work on causality was inspired by his notion that beliefs could be regarded as causes of other beliefs. He came to regard "causal relationships [as] the fundamental building blocks both of physical reality and of human understanding of that reality" and that "probabilistic relationships [were] but the surface phenomena of the causal machinery that underlies and propels our understanding of the world." [Pearl 2000, p. xiii]

In a Web page describing the genesis of his ideas about causality, Pearl writes [Pearl 2000]:

> I got my first hint of the dark world of causality during my junior year of high school.
>
> My science teacher, Dr. Feuchtwanger, introduced us to the study of logic by discussing the 19th century finding that more people died from smallpox inoculations than from smallpox itself. Some people used this information to argue that inoculation was harmful when, in fact, the data proved the opposite, that inoculation was saving lives by eradicating smallpox.

> "And here is where logic comes in," concluded Dr. Feuchtwanger, "To protect us from cause-effect fallacies of this sort." We were all enchanted by the marvels of logic, even though Dr. Feuchtwanger never actually showed us how logic protects us from such fallacies.
>
> It doesn't, I realized years later as an artificial intelligence researcher. Neither logic, nor any branch of mathematics had developed adequate tools for managing problems, such as the smallpox inoculations, involving cause-effect relationships.

So, the question is "How are we to understand causality?" Even though, as Pearl noted, most of his colleagues "considered causal vocabulary to be dangerous, avoidable, ill-defined, and nonscientific," he felt that his intuitions about causality should be "expressed, not suppressed." He writes that once he "got past a few mental blocks, I found causality to be smiling with clarity, bursting with new ideas and new possibilities." The key, again, was the use of graphical causal models.

Pearl's work on causality, the subject of his third book, has had major impacts even beyond the normal boundaries of AI. It has influenced work in philosophy, psychology, statistics, econometrics, epidemiology, and social science. Judging by citations and quotations from the literature, it is hard to identify another body of AI research that has been as influential on these related disciplines as has Pearl's work on causality.

One must be mathematically proficient to understand and to benefit from Pearl's work. Some have criticized him for "substituting mathematics for clarity." But, as Pearl points out [Pearl 1993, p. 51], "...it was precisely this conversion of networks and diagrams to mathematically defined objects that led to their current acceptance in practical reasoning systems." Indeed AI practitioners now acknowledge that successful applications depend increasingly on skillful use of AI's mathematically deep technology. Pearl, along with others in "modern AI," have made it so.

I'll close with a non-mathematical, but none-the-less important, topic. As we all know, Judea and Ruth Pearl's son, Danny, a *Wall Street Journal* reporter, was kidnapped and murdered by terrorists in Pakistan. In their grief, Judea and Ruth asked the question "How could people do this to to someone like Danny who 'exuded compassion and joy wherever he went'?" To help diffuse the hatred that led to this and other tragedies, Danny's family and friends formed the Daniel Pearl Foundation. Among the principles that the foundation hopes to promote are ones Judea himself has long exemplified: "uncompromised objectivity and integrity; insightful and unconventional perspective; tolerance and respect for people of all cultures; unshaken belief in the effectiveness of education and communication; and the love of music, humor, and friendship [Daniel Pearl Foundation]."

Shalom!

References

Pearl, J. (1984). *Heuristics: Intelligent Search Strategies for Computer Problem Solving*, Reading, MA: Addison-Wesley Publishing Company.

Pearl, J. (1988). *Probabilistic Reasoning Systems: Networks of Plausible Inference*, San Francisco: Morgan Kaufmann Publishers.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*, New York: Cambridge University Press (second edition, 2009).

Pearl, J. (1993). Belief networks revisited. *Artificial Intelligence 59*, 49–56.

Rumelhart, D. (1976). Toward an interactive model of reading. *Tech. Rept. #CHIP-56*. University of California at San Diego, La Jolla, CA.

Pearl, J. (2000). http://bayes.cs.ucla.edu/BOOK-2K/why.html.

Daniel Pearl Foundation. http://www.danielpearl.org/.

32

Fond Memories From an Old Student

Edward T. Purcell

I was very lucky to have been Professor Judea Pearl's first graduate student advisee in the UCLA Computer Science Department. Now I am further honored to be invited to contribute – in distinguished company – some fond memories of those early days studying under Professor Pearl.

In January 1972, after completing the core coursework for the M.S. degree, I took my first class in Artificial Intelligence from Professor Pearl. Thirty-eight calendar years seems like cyber centuries ago, such has been the incredible pace of growth of computer technologies and Computer Science and AI as academic disciplines.

The ARPAnet maps posted on the Boelter Hall corridor walls only showed a few dozen nodes, and AI was still considered an "ad hoc" major field of study, requiring additional administrative paperwork of prospective students. (Some jested, unfairly, this was because AI was one step ahead of AH — ad hoc.)

The UCLA Computer Science Department had become a separate Department in the School of Engineering only two and a half years earlier, in the Fall of 1969, at the same time it became the birthplace of the Internet with the deployment of the first ARPAnet Interface Message Processor node in room 3420 of Boelter Hall.

The computers available were "big and blue," IBM S/360 and S/370 mainframes of the Campus Computing Network, located on the fourth floor of the Mathematical Sciences Building, access tightly controlled. Some campus laboratories were fortunate to have their own DEC PDP minicomputers.

Programming was coded in languages like Assembly Language, Fortran, APL, PL/1, and Pascal, delimited by Job Control Language commands. Programs were communicated via decks of punched cards fed to card readers at the Campus Computing Network facility. A few hours later, the user could examine the program's output on print-out paper. LISP was not available at the Campus Computing Network. Time-sharing terminals and computers were just beginning to introduce a radical change in human-computer interaction: on screen programming, both input and output.

Professor Pearl's first "Introduction to AI" course was based on Nils Nilsson's *Problem-Solving Methods in AI,* a classic 1971 textbook focusing on the then two core (definitely non-ad-hoc) problem-solving methodologies in AI: search and logic. (As with the spectacular growth of computer technology, it is wondrous to regard how much Judea's research has extended and fortified these foundations of AI.) Supplemental study material included Edward Feigenbaum's 1963 compilation of articles on early AI systems, *Computers and Thought,* and a 1965 book by Nils Nilsson, *Learning Machines.*

In class I was immediately impressed and enchanted by Judea's knowledge, intelligence, brilliance, warmth and humor. His teaching style engaging, interactive, informative and fun. My interest in AI, dating back to pre-Computer Science undergraduate days, was much stimulated.

After enjoying this first AI class, I asked Professor Pearl if he would serve as my M.S. Advisor, and was very happy when he agreed.

Other textbooks Professor Pearl used in subsequent AI classes and seminars included Howard Raiffa's 1968 *Decision Analysis: Introductory Lectures on Choices under Uncertainty,* Duncan Luce and Howard Raiffa's 1957 *Games and Decisions*, and George Polya's *How to Solve it*, and the challenging 1971 three-volume *Foundations of Measurement,* by David Krantz, Duncan Luce, Patrick Suppes and Amos Tversky. The subtitles and chapter headings in this three-volume opus hint at Professor Pearl's future research on Bayesian networks: *Volume I: Additive and Polynomial Representations; Volume II: Geometrical, Threshold, and Probabilistic Representations; and Volume III: Representation, Axiomatization, and Invariance.*

It was always fun to visit Professor Pearl in his office. Along with the academic consultation, Judea had time to talk about assorted extra-curricular topics, and became like a family friend. One time, I found Judea strumming a guitar in his office, singing a South American folk song, *"Carnavalito,"* which I happend to know because of my U.S. diplomat's son upbringing in South America. I was happy to help with the pronunciation of the song's lyrics. It was nice to discover that we shared a love of music, Judea more in tune with classical music, myself more a jazz fan. Now and then I would see Judea and his wife Ruth at Royce Hall concerts, for example, a recital by the classical guitarist Narciso Yepes.

Judea's musical orientation (and humor) appeared in the title of a presentation a few years later at a Decision Analysis workshop, with the title acronym *"AIDA"* as Artificial Intelligence and Decision Analysis. The titles of other Pearl papers also revealed wry humor: *"How to Do with Probabilities What People Say You Can't,"* and *"Reverend Bayes on Inference Engines: a Distributed Hierarchical Approach."*

My M.S. thesis title was *"'A Game-Playing Procedure for a Game of Induction,"* and included results from a (PL/1) program for the induction game Patterns, a pattern sampling and guessing game introduced by Martin Gardner in his November 1969 *Scientific American* "Mathematical Games" column. (After sending Martin Gardner a copy of my M.S. thesis, I received a letter of appreciation from the game wizard himself.)

At a small public demonstration of the Patterns game-playing program in early 1973, a distinguished elderly scholar was very interested and asked many questions. After the presentation Professor Pearl asked if I knew who the inquisitive gentleman was. "No," I said. "That was Jacob Marschak," said Judea. Whenever I attend a Marschak Colloquium presentation at the UCLA Anderson School of Management, including several talks by Judea, I remember Professor Marschak's interest in my modest game-playing program.

Then, as now, seminars at Boelter Hall 3400 were an integral part of the UCLA Computer Science education. I remember several distinguished presentations there, for example, a seminar on coding theory given by Professor Andrew Viterbi, then still at UCLA, whom Professor Pearl engaged in an animated discussion, and another standing-room-only seminar on algorithms given by Donald Knuth, who listened attentively to Judea at a smaller, post-seminar gathering.

Soon enough, in June 1973, I was very proud and happy to receive my M.S. degree in Computer Science.

When I began my graduate studies in Computer Science at UCLA, I had only hoped to study for a Masters' degree. Though I was having a lot of fun studying AI and being mentored by Professor Pearl, I was not sure of my ability to pursue the doctorate degree. Encouraged and approved by Judea, I applied for and was accepted as a Ph.D. candidate, with Professor Pearl as my Advisor.

The early Ph.D. qualifying exams were challenging, because of the depth and breadth of topics covered, some topics beyond those covered in my classes. Thanks to Judea's guidance and support, I was able to overcome these challenges.

Professor Pearl's support extended beyond academic issues. On one lean occasion, I remember Judea lending me some funds to cover my registration fees. Fortunately, UCLA tuition fees were very modest in those days (unlike today's costs), and I was soon able to repay Judea's kind loan.

My classes were now mostly individual study seminars led by Professor Pearl. Despite a variety of readings and studies, I was stumped for a good dissertation topic. Judea suggested a very interesting topic: learning of heuristics for search algorithms.

I was immediately piqued by this topic, and soon formulated a perceptron-like learning-while-searching procedure for A*-like heuristic search algorithms. The unsupervised learning consisted of adjusting the weight vector $\mathbf{w}$ of a heuristic vector function $\mathbf{h}$, trying to satisfy, on a local scale, necessary (but not sufficient) metric and order consistency properties of the perfect knowledge heuristic function h^*. The learning samples derived from search observations of problem graph edge costs and node orderings, obtained as the search algorithm progressed.

The topic of learning heuristics for search algorithms was well received by the Ph.D. dissertation qualifying committee. I remember Professor Pearl telling me committee member Dr. Ken Colby of the UCLA School of Medicine expressed a favorable review of this topic and of my introductory overview of the topic.

I was able to complement and support my UCLA Computer Science studies with interesting part-time work, near campus and related to my studies. During 1974 and 1975 I worked part-time at Technology Service Corporation for William Meisel and Leo Breiman, and was invited to be co-author of a 1977 paper (*"Variable-Kernel Estimates of Multi-Variate Densities,"* **Technometrics**, vol. 19, no. 2, pp.

135-144, 1977), whose experimental results were based on my programming. (Many years later I learned this paper earned me an Erdős 4 number.)

In late 1976 and early 1977 I worked part-time for System Development Corporation, and was tasked by Drs. Jeff Barnett and Mort Bernstein with writing summaries of papers, reports and other documents on the emerging technology of knowledge-based systems, which contributed to a June 1977 System Development Corporation report (ADA044883), *"Knowledge-Based Systems: A Tutorial."*

Many of the early expert systems implemented the MYCIN - Prospector certainty factor calculus. Probabilities were dismissed because of the exponential number of joint probabilities presumed to be required. I remember Professor Pearl discussing the topic of uncertainty calculus with colleagues at a Workshop on Decision Analysis held at a hotel in Bel Air in the summer of 1977.

I thoroughly enjoyed those lean student days, commuting to campus on bicycle, studying Computer Science and AI under Professor Pearl. I remember many fun activities: a barbecue dinner hosted by Judea and Ruth Pearl for Donald Michie in May 1976, participating in experiments with Norman Dalkey's Delphi group decision-making system, attending Royce Hall concerts, playing perhaps too much soccer and rugby. (But I had good company in these sports activities: fellow UCLA Computer Science graduate student David Patterson was also a UCLA rugby teammate.)

The final hurdles on the doctoral track were more logistical and administrative rather than technical, and included scheduling (in pre-email days) five busy dissertation committee members to a common time and place, applying (in pre-PC days) for additional computer run time from Campus Computing Network, obtaining the approval of the UCLA School of Engineering bibliography checker, finding (in pre-TEXdays) a good typist, making copies of the dissertation, etc.

In June 1978, thanks to much encouragement, guidance and nurturing from Professor Pearl, I completed my Ph.D. dissertation, *"Machine Learning of Heuristics for Ordered-Search Algorithms."*

The fun memories associated with Professor Pearl continued after my graduation. During an AI conference in Miami in December 1984, a dinner with Judea at a restaurant in little Havana. Other AI conference dinners hosted by Professor Pearl for his graduate students. One day in 1985, when I visited Judea in his office enroute to a Computer Science Seminar, I remember him asking me which designation I liked better: "Bayes net" or "Bayesian network." I voted for the latter as more poetic. In November 1996 I was invited by Judea to attend his University of California Faculty Research Lecture at Schoenberg Auditorium. A capacity crowd listened attentively as Judea discussed *"The Art and Science of Cause and Effect."* Afterward, Judea and his family celebrated at a tea reception at the Chancellor's Residence. A special seminar for the publication of *"Causality"* in 2000. And the fond memories continue.

Many colleagues ask me, "Did you study under Judea Pearl?" "Yes!" I answer proudly. I am very proud to have been Professor Pearl's first student, even though

I was probably not worthy.

I cherish the memories of those student days in the UCLA Computer Science Department, studying under and learning from Professor Pearl.

With deep appreciation, I would like to thank you very much, Judea, for all your kindness, help, guidance and education through the years.

God bless you!

33

Reverend Bayes and inference engines

DAVID SPIEGELHALTER

I first met Judea in 1986 at a conference in Paris on the "management of uncertainty in knowledge-based systems": this topic, which now sounds rather dated, was of consuming interest at the time and I was anxious about coming face-to-face with someone who might be considered a competitor in the field – what would he be like? I need not have worried.

This was an unusual research area for a statistician, but since the early 1980s I had been part of a group working on decision-support systems in medicine which used explicit probabilities for diagnosis and prognosis. There was a strong and (usually) good-natured rivalry between techniques based on formal probabilistic methods for so-called 'expert' or 'knowledge-based systems' and those arising in the computer science community that were more rooted in artificial intelligence and attempted to represent expert knowledge through a network of rules, with a separate 'inference engine' to control the reasoning process. The MYCIN team from Stanford were international leaders in this field with their system of 'certainty factors' to handle uncertainty in the rules. There was no rigid distinction between the approaches: everything was in the pot, from purely qualitative reasoning, through rather ad-hoc quantification, fuzzy logic, and attempts to use probabilities in a network of rules. It was a very exciting and competitive time, in which many disciplines were trying to establish authority.

A Royal Statistical Society discussion paper [Spiegelhalter and Knill-Jones, 1984] brought the controversy into the statistical world, but this was outdated even by the time of its publication. Struggling with the problem of handling probability models on directed graphs of arbitrary size and complexity (and, crucially, being totally unaware of Judea's work in this area), I realised that Wermuth and Lauritzen [1983] provided a connection between probabilistic models on directed graphs and unconnected graphs: the latter had been established for some time as an elegant way of representing conditional independence relations in contingency tables [Darroch, Lauritzen and Speed, 1980]. In April 1984 I met Steffen Lauritzen at the 150th Anniversary celebrations of the Royal Statistical Society and told him that his work on graphical models was of fundamental importance to artificial intelligence. He did not seem to believe me, but he was kind enough to take me seriously enough to invite me to visit.

So I took the long journey to Aalborg in the north of Jutland in Denmark. By a remarkable stroke of good fortune Steffen shared a department with a research team who were working on uncertainty in expert systems, and together we persuaded him that he should get involved. And he is still working in this area after 25 years.

As a bonus, the Danish group finally introduced me to Pearl [1982] and Kim and Pearl [1983]. These came as a shock: looking beneath the poor typography revealed fundamental and beautiful ideas on local computation that made me doubt we could contribute more. But Judea was working solely with directed graphs, and we felt the connection with undirected graphs was worth pursuing in the search for a general algorithm for probability propagation in arbitrary graphs.

I wrote to Judea who replied in a typically enthusiastic and encouraging way, and so at a 1985 workshop at Bell Labs I was able to try and put together his work with our current focus on triangulated graphs, clique separations, potential representations and so on [Spiegelhalter, 1986]. Then in July 1986 we finally met in Paris at the conference mentioned at the start of this article, where Judea was introducing the audience to d-separation. I have mentioned that I was nervous, but Judea was as embracing as ever. We ended up in a pavement café in the Latin quarter, with Judea drawing graphs on the paper napkin and loudly claiming that anyone could see that observations on a particular node rendered two others independent – grabbing a passer-by, Judea demanded to know whether this unfortunate Frenchman could recognise this obvious property, but the poor innocent man just muttered something and walked briskly away, pleased to have escaped these lunatics.

We continued to meet at conferences as he developed his propagation techniques based on directed graphs [Pearl, 1986] and we published our algorithm based on embedding the directed graph in a triangulated undirected graph that could be represented as a tree of cliques [Lauritzen and Spiegelhalter, 1988]. We even jointly presented a tutorial on probabilistic reasoning at the 1989 IJCAI meeting in Detroit, which I particularly remember as my bus got stuck in traffic and I was late arriving, but Judea had just carried on, extemporising from a massive pile of overhead slides from which he would apparently draw specimens at random.

Then I started on MCMC on graphical models, and he began on causality, which was too difficult for me. But I look back on that time in the mid 1980s as perhaps the most exciting and creative period of my working life, continually engaged in a certain amount of friendly rivalry with Judea, who always responded with characteristic generosity of spirit.

References

Darroch, J. N., Lauritzen, S. L. and Speed, T. P. (1980) Markov Helds and log-linear models for contingency tables. *Ann. Statist*., 8, 522-539.

Kim, J. H. and Pearl, J. (1983) A computational model for causal and diagnostic reasoning in inference systems. In *Proc. 8th International Joint Conference on Artificial Intelligence*, Karlsruhe, pp. 190-193.

Lauritzen, S. L. and Spiegelhalter, D. J. (1988) Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 50, 157-224.

Pearl, J. (1982) Reverend Bayes on inference engines: a distributed hierarchical approach. *Proc. AAAI National Conference on AI, Pittsburgh*, pp. 133-136.

Pearl. J. (1986) Fusion, propagation and structuring in belief networks. *Artificial Intelligence*, 29, 241-288.

Spiegelhalter, D. J. (1986) A statistical view of uncertainty in expert systems. In *Artificial Intelligence and Statistics (ed. W. Gale)*, pp. 17-56. Reading: Addison-Wesley.

34

An old-fashioned scientist shaping a modern discipline

Hector Geffner

I took a course with Judea in September of 1984, while I was finishing my MS in Systems Science at UCLA. Right after, I switched to Computer Science, became his PhD student, and started working as a research assistant in his group. I finished my PhD five years later, a time during which I learned from him how science is done and how a scientist works, two things that were extremely valuable to me for at least two reasons. The first is that I was a pure science 'consumer', enthusiastic and well-informed but more inclined to read than to produce science. The second is that, unknown to me, AI was being redefined, with no one playing a larger role in the change than Judea.

While Judea published regularly in AI conferences from the late 70s on and the *Heuristics* book was about to be published, he still felt very much like an outsider in AI, even at UCLA, where the AI Lab, headed by former students of Roger Schank, used to get the spotligth, lavish funding, and most of the aspiring AI students. Judea, on the other hand, directed the Cognitive Systems Lab, which to my surprise was no more than a sign on the door of a secretary, whose main task, although not an easy one, was to input Judea's handwritings into the computer.

Judea's door was in front of the Lab with no sign revealing his name so that unwanted intrusions would be discouraged. Years later he added a sign, "Don't knock. Experiments in Progress" that remained there for more than 20 years. Judea liked to work at home early in the day, showing up by his office at 3pm, for meeting students and the secretary, for answering mail, and of course, for thinking, which is what he liked and needed the most. He kept administration to a minimum, and since the 80s at least, has not taught undergraduates (I still don't know how he got away with this). He also used to wear a pair of earplugs, and you could often discover that you said something interesting when you saw Judea taking them off.

What struck me first about Judea was not his research – I couldn't say much about it then – nor his classes, which I certainly liked but were not typical of the 'best teachers' (I still remember Judea correcting a slide in class with his finger, after dipping it into the coffee!), but his attitude toward students, toward science, and toward life in general. He was humble, fun, unassuming, respectful, intelligent, enthusiastic, full of life, very easy to get along with, and driven by a pure and uncorrupted *passion for understanding.* Judea doesn't just seek understanding, he needs it; it's something personal. I'm sure that this is the way scientists and philosophers like Hume, Newton, and Leibniz felt centuries ago, although I doubt that they were as much fun to be with.

In the late 80s, Judea had a small group of students, and we all used to meet weekly for the seminars. Judea got alone well with everyone, and had a lot of patience, in particular with me, who was a mix of rebel and dilettante, and couldn't get my research focused as Judea expected (and much less on the topics he was interested in, even if he was paying my research assistantship!). I remember telling him during the first couple of years that I didn't feel I was ready for research and preferred to learn more AI first. His answer was characteristic: "you do research now, you learn later — after your PhD". I told him also that I wanted to do something closer to the mainstream, something closer to Schankian AI for example, then in fashion. Judea wouldn't get offended at all. He would answer with calm "We will get there eventually", and he certainly meant it. Judea was probably a bit frustrated with me, but he never showed it; quite the opposite, he was sympathetic to my explorations, gave me full confidence and support, and eventually let me do my thesis in the area of non-monotonic reasoning using ideas from probability theory, something that actually attracted his interest at the time.

Since Judea was not an expert in this area (although, unsurprisingly, he quickly became one), I didn't get much technical guidance from him in my specific dissertation research. By that time, however, I had learned from him something much more important: I learned how science is done, and the passion and attitude that go into it. Well, may be I didn't learn this at all, and rather he managed to infect me with the 'virus'; in seminars, in conversations, by watching him work and ask questions, by osmosis. If so, by now, I'm a proud and grateful carrier. In any case, I have been extremely privileged and fortunate to have had the chance to benefit from Judea's generosity, passion, and wisdom, and from his example in both science and life. I know I wouldn't be the same person if I hadn't met him.

35

Sticking With the Crowd of Four

Rina Dechter

I joined Judea's lab at UCLA at about the same time that Hector did, and his words echo my experience and impressions so very well. In particular, I know I wouldn't be the same person, scientist, and educator if I hadn't met Judea.

Interestingly, when I started this journey I was working in industry (with a company named Perceptronics). We had just come to the U.S. then, my husband Avi started his Ph.D. studies, and I was the breadwinner in our family. When I discussed my plans to go back to school for a PhD, I was given a warning by three former students of Judea who worked in that company (Chrolotte, Saleh, and Leal). They all said that working with Judea was fun, but not practical. "If you want a really good and lucrative career," they said, "you should work with Len Kleinrock." This was precisely what I did. I was a student of Kleinrock for three years (and even wrote a paper with him), and took AI only as a minor. During my 3rd year, I decided to ignore practical considerations and follow my interests. I switched to working with Judea.

At that time, Judea was giving talks about games and heuristic search to whoever was willing to listen. I remember one talk that he gave at UCLA where the audience consisted of me, Avi, and two professors from the math department. Judea spoke enthusiastically just like he was speaking in front of the Hollywood Bowl. Even the two math professors were mesmerized.

Kleinrock was a star already, and his students were getting lucrative positions in Internet companies. I congratulate myself for sticking with the crowd of four, fascinated by how machines can generate their own heuristics. Who could tell that those modest seminars would eventually give birth to the theories of heuristics, Bayesian networks, and causal reasoning?

Judea once told me that when he faces a really hard decision, a crossroad, he asks himself "What would Rabbi Akiva do?". Today, when I face a hard decision, I ask "What would Judea do?".

Thanks Judea for being such a wonderful (though quite a challenging) role model!