A Guest Lecture on Social Software

Jan van Eijck

Rohit Parikh, visitor to the project, has delivered a NIAS lecture on social software. On the next day, the project members discuss the contents and the reception of his talk. An ethicist (professor of ethics) has joined the project team, and a visiting political scientist is also present.

Logician: It is such a pity I had to miss Rohit's lecture. And Rohit himself has dashed off now, to a conference in Paris. On the day before the talk, one of the NIAS fellows asked me with a worried look on his face what the word "algorithm" meant that he had seen in the lecture announcement, and could we please make sure that our guest lecturer knew that part of the audience was unfamiliar—even uncomfortable—with the jargon of computer science and logic? So we forewarned Rohit, of course. Now you all understand why I am curious how it went. Can anyone tell me?

Philosopher: Yes, I did get the impression that part of the audience was a bit suspicious of logicians and computer scientists taking on problems in humanities. The NIAS audience consists of highly articulate opinion leaders in the field of humanities and social sciences, but some of them seemed wary about the methods of the exact sciences.

Logician: It seems to me that it is quite important to articulate our answers to the typical questions and worries of such an audience. The kind of objections that were raised on this particular occasion will no doubt be raised again and again when one tries to outline the task and goals of the social software enterprise.

Computer Scientist: Let's recall what went on, then. I will start, and maybe the others can all comment, so that we get at a reasonable reconstruction. The main theme of the talk, of course, was an outline of the conception of social

software as an interrelation of (i) logical structure, (ii) incentive structure, and (iii) knowledge transfer. By the choice of his examples, Rohit made clear that improving a social process involves analysis of what goes on (logical structure), understanding what makes the participants in the process "tick" (incentive structure), and understanding the flow of knowledge that takes place during the process (knowledge transfer).

Philosopher: Rohit took great care to explain his terminology and conceptual tool set, by the way. His explanation of what an algorithm is, for instance, used the example of Euclid's recipe for calculating the greatest common divisor of two positive whole numbers.

Computer Scientist: Yes, a nice illustration indeed. The calculation is based on the insight that if you have two positive whole numbers, A and B, with A larger than B, then replacing the larger number by $A - B$ does not affect the set of common divisors of the pair. As soon as this is clear, it is also clear that Euclid's procedure for finding the greatest common divisor has to be correct.

Logician: Yes, yes, but I suppose we can skip all that for now. How was the talk received? What were the questions?

Philosopher: The talk itself was very well attuned to the audience, it seemed to me. As for the questions, well, various people expressed doubts about the use of formal methods in trying to capture aspects of human interaction. Their main worry seemed to be that the essence of what goes on in the ways human beings behave towards one another and give meaning to their interactions might get lost in the mathematical analysis.

Computer Scientist: As I remember, Rohit had various things to say about this. One of the points he made was about the virtue of idealisation and abstraction. Analysis of the trajectories of moving bodies like flying cannon balls always starts by making some unwarranted but very useful assumptions: that there is no air resistance, or that there is no drag from the rotation of the earth. These assumptions are necessary to get started. Indeed, it takes great skill to find the right abstractions; this is what progress of the natural sciences is all about. In our understanding of the movement of cannon balls it turns out to be illuminating to disregard earth rotation, but for understanding the emergence of cyclones the drag from the rotation of the earth is an essential element.

Philosopher: Another thing that could be said—and if I remember well Rohit

touched on this too—is that there is no pretense that the abstractions fit the aspects of reality one tries to understand in every detail. Same for social software. A social software analysis might be useful despite the fact that it does not explain and is not meant to explain *all* that there is to explain about what goes on when human beings interact in institutions.

Computer Scientist: Can we elaborate on this still further? A related question was asked by Donald Light. He questioned the main paradigm of many mathematical approaches to economy, where the starting axiom is that human beings are always maximizing their interest. What is it that warrants this assumption? Selfish individuals surely are not the only possible paradigm? I cannot remember how Rohit handled this.

Political Scientist: For one thing, Rohit agreed that homo sapiens is not the same animal as homo economicus. As a matter of fact, the abstractions of economics were borrowed from a psychological fashion called behaviourism. Fortunately, psychologists have now abandoned this, and it is to be hoped that economics will follow suit. You might want to have a look at [1] if you are interested.

Cognitive Scientist: Unfortunately, this was only mentioned in passing, for the death blow to psychological behaviourism was dealt by cognitive science. An important development in our field is that subjective feelings of happiness and despair can be correlated to objective happenings in the brain. 'Feeling good' turns out to have a physical basis, and what is more, the way people report on how they feel corresponds quite well with the findings of fMRI scans [3; 4].

Political Scientist: A recent plea to take these findings into account in public policy making was made by economist Richard Layard [5; 6]. Layard argues that the key question economists should ask themselves is this. How can we explain that since 1950, despite a huge increase in income, average happiness among people in the West has not increased? For more information I can recommend the World Database of Happiness on the web [7].

Computer Scientist: Ah, the site maintained by our Dutch happiness professor Ruut Veenhoven, right? Yes, a visit to his website always cheers me up. At the very least such information makes clear that money is not everything.

Philosopher: None of this came up in yesterday's discussion, but Rohit mentioned the fact that social procedures often have parameters that can be

adjusted to reflect participant attitude. For instance, Steven Brams and his co-workers have developed algorithms for mediating the property settlement in a divorce. Suppose you and your ex-partner want to use this software. Then the starting point is for each of you to divide 100 points over the common property items, reflecting your individual valuation of the items. Next, use the algorithm to decide what is the agreement that will maximize happiness of each of you, and is most fair. Now suppose you want to do your ex-partner a good turn. Then you can decide that she is allowed to divide 150 points, and you are content with 50 points. This shows that the social software algorithm is really just a tool; it is completely up to you to decide just how greedy you want to be.

Logician: A topic that strikes me as relevant in this context are the game-theoretical paradoxes, such as the Allais paradox, the Elsberg paradox, or the St. Petersburg betting paradox. Were any of those mentioned?

Computer Scientist: I don't think so.

Philosopher: Can anyone explain, please?

Logician: Rohit and I discussed the St. Petersburg paradox at some other occasion. The St. Petersburg game is played by flipping a fair coin until it comes up tails, and the total number of flips, n, determines the prize, which equals 2^n euros. If the coin comes up tails the first time, the prize is $2^1 = 2$ euros, and the game ends. If the coin comes up heads, it is flipped again until it shows tails. So if the coin comes up heads the first and the second time, and tails the third time, the prize is $2^3 = 8$ euros, and so on. The relevant events are sequences of head flips followed by a tail flip, and the probability of the sequence of $n - 1$ head flips followed by a tail flip is 1 over 2^n. The prize for this event is 2^n euros, so the expected payoff (prize times likelihood) is 1 euro. Now the space of possible events is infinite, and each of these has an expected payoff of 1 euro. So the value of the game is infinite. A rational gambler would enter the game if the price of entry was less than the value. Still, most people would be reluctant to offer even 25 euros for playing the game.

Computer Scientist: What does this show? That most people are irrational? Or that there is something wrong with the underlying concept of rationality?

Logician: Daniel Bernoulli, who invented the paradox—he was a mathematics professor in St. Petersburg for some time—believed the latter. He observed

that the calculation of expected value does not take into account that money has a decreasing marginal utility: money means less to the rich than it does to the poor. However greedy an individual is, an extra assumption of diminishing marginal utility will explain why human beings tend to reject the bet.

Computer Scientist: But it is well known that this does not resolve the paradox. For if you give me a function for calculating the decrease in utility, then I can use that function for constructing a new version of the game, and the paradox reappears. Discussion and links can easily be found on the internet, by the way. `http://econlog.econlib.org/archives/2003/11/st_petersburg_p.html`

Philosopher: I can think of a different and very rational reason for refusing to play the game. Ask yourself who is supposed to act as bank, if the game is played? The problem is that it takes infinite wealth to underwrite it. If Yukos or Gazprom invite offers to play the St Petersburg game, I will abstain. Chances are they are not rich enough to pay up just when I am about to collect real money.

Computer Scientist: Let's get back to the discussion after Rohit's lecture. A completely different issue was brought up by Gül Ozyegin. Is the attempt to describe human behaviour in abstract (and maybe quantified) terms in any way related to the attempts of damage insurance lawyers who tried to *calculate* the monetary value of their clients who were killed in the 9/11 disaster, for use by the heirs of the life insurance policy owners? The dishwashers in the WTC restaurants were worth much less than the high profile chief executives that were killed. Gül described these lawyers as a kind of vultures, I remember, and expressed moral qualms about any attempts to describe the worth of a human life in terms of money.

Philosopher: I have no notes of how the actual discussion went on this point, and I must admit I got lost.

Political Scientist: The fact that people make money because others have died is a fact of life. That cannot be the moral issue. The same holds for the fact that some life insurance policies are worth more than others.

Ethicist: If these damage insurance lawyers work on a 'no win no fee' basis then there is a moral issue, I suppose. For then they may induce their clients to engage in endless litigation, and this—it has been argued—generates a claims culture that is clearly not in the interest of the community. In many

countries of European Union—including the Netherlands—'no win no fee' is against the law.

Political Scientist: The case is not clear cut. Other countries allow what is known as conditional fee agreements (cfas). Under such agreements, if you win your case, you must pay your solicitor's fees and any expenses for items such as experts' reports, so-called disbursements. If you lose, you need pay no fees to your solicitor. However, you may have to pay your opponent's legal costs and both sides' disbursements. So also in these cases there is a mechanism to discourage pursuing weak cases. Conditional fees are subject to regulations which set out what a solicitor must tell the client. A solicitor who does not abide by the regulations runs the risk of not getting paid at all, win or lose.

Computer Scientist: I suppose that finding out what is the effect of the fee structure for attorneys on patterns of litigation in a country is also social software analysis?

Philosopher: But it seems to me that we need have no qualms about expressing what someone's life is worth in quantitative terms. Quality Adjusted Life Year (QALY) has been proposed in the medical profession as a measure for combined quantity and quality of life. QALY calculations are useful for measuring efficacy of medical treatment. QALY calculation takes one year of perfect health-life expectancy to be worth 1, but regards one year of less than perfect life expectancy as less than 1. Suppose the prognosis for a patient is to die within one year, with quality of life fallen from 1 to 0.4. Then an intervention which keeps the patient alive for an additional four years rather than die within the year, and where quality of life falls from 1 to 0.6 on the continuum will generate 2.0 QALYs: 4×0.6 for the extra life years at quality 0.6, minus 1 year at quality 0.4 which would have been the result of no intervention. The definition is in any medical dictionary, or you can look it up in Wikipedia. What is the moral worry?

Computer Scientist: In my notes, there is also an entry on how to take the non-rational into account in rational analysis of human behaviour. In this context, Michael Suk-Young Chwe's *Rational Ritual* was mentioned [2].

Philosopher: Yes, that is a beautiful book. I happen to have it with me. It illustrates that rituals that appear at first sight to be completely non-rational turn out to have a strong rational element. The rationale of many public rituals comes to light if one views them as procedures for creating common

knowledge.

Political Scientist: Ah, common knowledge is what is generated when I send out emails with long `cc:` lists. I only use those for invitations, when I want to generate common knowledge of who is also invited, so that everybody who gets invited knows what kind of party to expect.

Philosopher: Chwe argues that the wish to create common knowledge is behind many social rituals. It explains Apple's decision to introduce their (then) new Macintosh computer during the 1984 Super Bowl TV show. Look here (*quotes from Chapter 1 of the book [2]*):

> By airing the commercial during the Super Bowl, Apple did not simply inform each viewer about the Macintosh; Apple also told each viewer that many other viewers were informed about the Macintosh.

Creating common knowledge in this case was important, for prospective buyers knew that getting a Macintosh was a good investment only in case the Macintosh would turn out a success. The book has many more examples, of course.

Computer Scientist: The issue of common knowledge and how it is created is a topic in its own right. Let's get back to it at another occasion (see page **??**).

Philosopher: At some point the discussion also touched on cultural relativism. One of Rohit's examples was about queueing for buses. He had noticed that shelters at bus stops near Wassenaar have advertisements on all sides except for the side where you see the bus approaching if you are inside the shelter, and where the driver sees you, of course. Also, in his talk he had mentioned signs in London with 'queue this side please,' as examples of social software. In connection with this, Sadik Al-Azm remarked that boarding a bus in Cairo or Damascus is rather different than boarding a bus in London or Wassenaar. It involves different skills: London bus boarding habits would simply fail to get you a place. If social software designers were hoping to come up with proposals to improve bus boarding procedures for Cairo or Damascus, he wished them good luck.

Ethicist: Here Rohit's reply was that social software analysis should always take how people actually behave and what they actually believe as given, and propose small adjustments to improve a given situation.

Computer Scientist: Another thing that raised questions was the issue of strategic behaviour. What is it? Is it good, is it bad, or are moral qualifications of it beside the point? Should social software be designed in such a way that possibilities for strategic behaviour get minimized? If so, why?

Philosopher: Yes, this issue came up in an amusing way during the discussion. At the start of the discussion many hands were raised, and someone proposed to chairman Wim Blockmans that people who wanted to ask follow-up questions should raise a single finger, while people who wanted to address a different aspect of the lecture should raise their whole hand.

Computer Scientist: Yes, and someone then remarked that when Sadik Al-Azm raised a single finger and asked a question about something completely different, this was a nice example of strategic behaviour. Sadik was only mildly amused, it seemed. Others thought his reaction was quite funny.

Philosopher: Maybe we should try to make explicit—or at least say something illuminating about—how social software relates to moral debate. The moral aspects of some of Rohit's examples seemed to baffle the audience.

Computer Scientist: One of Rohit's examples during the lecture revolved around the notions of agency, ability, and responsibility. Rohit asked us to imagine a trolley moving downhill along a track, with malfunctioning brakes. You are standing beside the track. The only way you can prevent it to kill five people standing on the track is by switching a lever, to divert the trolley to a different track. The trouble is that there is also a person standing on that other track, who will certainly get killed as a result of the diversion. The point of the example is the distinction between moral responsibility for one's action and moral responsibility for one's inaction. Not touching the lever makes one guilty through inaction, switching the lever makes one guilty through active involvement. Guilt through inaction presupposes the ability to act, of course.

Political Scientist: I can see why this made the audience uneasy. If fine metaphysical distinctions like the difference between sins of commission and sins of omission are relevant for a 'science' of social software, one might reasonably ask whether one is doing exact science at all. Science can only flourish where one has learned ways to put metaphysical worries on hold.

Philosopher: Anyway, when time was up, there were still many questions left unanswered. Is analysis of how procedures are incorporated in social institutions also part of the task of social software? What does social software

have to say about how we interact with banks, schools and churches? Does social software provide analysis, or does it make recommendations? And if both: how do the two relate? Does social software take a moral stance? If so, what is the foundation? If not, how can it still make recommendations? I seem to remember that some part of the audience thought that it was strange that value judgements like "order is better than chaos", or "it is better if less people get hurt" seemed to play a role.

Ethicist: Questions, questions. Does anyone care for a coffee?

References

[1] Colin F. Camerer, George Loewenstein, and Matthew Rabin, editors. *Advances in Behavioral Economics*. Princeton University Press, 2003.

[2] Michael Suk-Young Chwe. *Rational Ritual*. Princeton University Press, Princeton and Oxford, 2001.

[3] R.J. Davidson. Affective style, psychopathology and resilience: Brain mechanisms and plasticity. *American Psychologist*, 55:1196–1214, 2000.

[4] R.J. Davidson. Well-being and affective style: Neural substrates and biobehavioural correlates. *Philosophical Transactions of the Royal Society*, 359:1395–411, 2004.

[5] Richard Layard. Happiness — has social science a clue? Lionel Robbins Memorial Lectures delivered at LSE on 3rd, 4th and 5th March 2003: available from `http://cep.lse.ac.uk/layard/`, March 2003.

[6] Richard Layard. *Happiness — Lessons from a New Science*. The Penguin Press, 2005.

[7] R. Veenhoven. World database of happiness, distributional findings in nations. Technical report, Erasmus University, Rotterdam, Last revision: 2007. Available at `www.worlddatabaseofhappiness.eur.nl`.