Patterns of Intelligent Interaction: Games, Action, and Social Software

Johan van Benthem, Amsterdam & Stanford, http://staff.science.uva.nl/~johan

Sitting in the office of a distinguished philosopher of language recently, I watched him lean back (somewhat precariously) in his chair, look at the ceiling, and sigh: "Johan, we both write all this stuff about information, context, and communication – but is not the only time you *really* feel that you are making progress, when you resolutely close your eyes, and shut out the world and the others?" I appreciated his point, and indeed, in most spheres of life on this planet, "l'Enfer" is most definitely "Les Autres".

But is splendid isolation right in *theory*? Indeed, many core notions of classical philosophy, and logic, and linguistics live in Platonic Heaven, where they have been sanitized to mathematical 'propositions', 'meanings', or 'consequence relations' which involve no human agency at all. But this suppresses the crucial dynamic and social character of such notions as statements, information, and argumentation, which involve agents engaging in various patterns of interaction. Take my own field of logic. Should our ideal be Euclid's *Elements* with mathematical proof as strings of symbols and diagrams, or should it rather be the equally classical Socratic *Dialogues*, with their argumentative dynamics between different agents pursuing different goals?
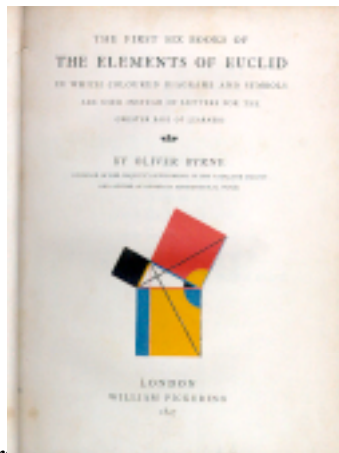
In recent years, many disciplines in my corner of the academic world are turning toward the latter stance – explaining linguistic meaning as equilibrium in successful communication, and logical proofs as winning strategies in successful argumentation. Some philosophers have even suggested that the traditional search for 'true knowledge' as a 24+ carat property of abstract propositions is misguided – it is human *interaction* that is the grindstone of false belief, and human knowledge is that which survives it.

One discipline where this move is particularly striking is contemporary *informatics*: 'computer science' if you wish (but please don't). Modern computers are no longer autistic devices striving (in delusional spells of grandeur) to emulate humans with the techniques of AI. They live in distributed multi-agent societies, they communicate, collaborate and compete for scarce resources, all very much like us. Accordingly, computer scientists have developed a rich arsenal of new techniques for modeling information flow in such interactive settings, occasionally joining forces with philosophers, linguists, economists, and logicians (the jet set of inter-disciplinarity). Topics in this mixture cross borders between disciplines everywhere you look. A simple NIAS conversation involves speech acts (a philosopher's notion), information update (an informatics classic), and even longer-term program structures like flattery: "Keep saying the right nice things to our Rector until he releases the funding" which, well-understood, involve major action structures like conditional choice, sequential execution, and recursion until some goal is reached. We are talking logic here, and probability, computation, and general learning theory.

Once on this road, the *social sciences* are on the horizon. Modern logic and informatics have already founds links with economic game theory, and with social choice theory, analyzing the detailed mechanics of procedures for voting, judgment aggregation, fair distribution, or about every meaningful group activity. So, why don't 'we' just acknowledge that others have been there before, give up our safe tenured university jobs to the social scientists, and retire in righteous poverty? This is not how the contact looked to our social scientists at NIAS, who seemed to feel a threat the other way around – with technical upstarts without much erudition telling them their business. But the intellectual realities are more complex. And they *are* that many topics in the broader area of what might be called 'intelligent interaction' no longer fit the standard academic geography. Insights travel between different disciplines: humanities, natural and social sciences, and they have done so for quite a while now. Ideas in my own research on dynamic logics for information flow in games come from philosophy, linguistics, game theory, and sociology – and it is very hard to tag them for their original provenance.

Enter the NIAS pilot project on 'Games, Action, and Social Software', a pioneering initiative of NIAS and the Lorentz Center for creating trans-disciplinary 'Meetings of the Minds': (http://www.nias.knaw.nl/en/research_group_2006_07/nucleus/)

New for NIAS' scholarly orientation, but in keeping with its own ambitions, it was a complex dynamic group effort. Naturally, its core fellows did not fit into fixed academic pigeon-holes: Jan van Eijck (philosopher, linguist, computer scientist), Barbara Dunin-Keplicz (sociologist, computer scientist), Martin van Hees (philosopher, decision theorist), Krister Segerberg (philosopher, logician), and Rineke Verbrugge (mathematician, computer scientist, and cognitive scientist). Together with a lively galaxy of short-term visitors, and comet-like day visitors to special events (such as the undersigned with his Amsterdam students making the long 'Leidensweg' to NIAS), they organized half a year of intensive work on core themes in the above general area. These included belief revision, higher-order knowledge about others, reasoning about preference, intentions and plans in decision making, and the dynamics of duties and obligations. The project generated three international workshops. One held at the Lorentz Centre concentrated mainly on formal machinery: designing logics for all of the above activities. A second workshop at NIAS concentrated on the facts rather than theory, viz. experimental backgrounds in cognitive psychology – where by the way, interaction is also becoming more and more of a crucial feature in understanding human intelligence. The event was graced by the presence of the first recipient of the KNAW Heineken Award for Cognitive Science, John Anderson (Pittsburgh), himself a role model for the mixture of mixture of computational and experimental strands which drives so much progress today. This public aspect of the project reached its apotheosis in the KNAW Conference on Games and Interaction (February 2007), which brought together this community with some of today's most prominent game theorists.

In the aftermath, the project will produce a technical monograph on the many strands it has pursued in the area of intelligent interaction. It will also produce a lively book of '*Dialogues on Social Software*', many of them already written at NIAS, for a general audience explaining what the new mixtures of topics and disciplines are about.

Ah, but here, at last, a word of caution is needed. Understanding formal structures in interaction actually has two different aspects. One is the formal *analysis* of given social phenomena, modeling them, and trying to understand what makes them tick. But there is also the more ambitious activist stance of *designing* new, better forms of interaction, using ideas from logic, computer science, game theory, and what have you. Both, but especially also the latter, go under the heading of 'Social Software', a term coined by the eminent logician/philosopher/computer scientist Rohit Parikh from New York, who was an inspiring short-term NIAS fellow last autumn. So how *does* the public respond?

'Social Software' is a challenging 'upper' in some areas, and an immediate 'downer' in others. Having lunch at NIAS one beautiful September day at the start of the project, I mentioned to my single table companion that our project was about 'social software'. Her spontaneous response was "How terrible!" She felt that this was all about the wrong way of looking at real human behaviour: fixed automata, rigid rules, and robots right in Wassenaar, the secluded forest sanctuary of the humanities, wit and elegance.

Well, I beg to disagree. Logical and computational tools do not supplant our natural sense of interaction, they rather offer new ways of looking at its structure – and new perspectives on what is after all a highly mysterious, complex, and difficult feat which we humans perform. Moreover, in doing so, all of our lives are filled with designed procedures, be they laws, supermarket queues, or rules of debate. This is a reality created for many different purposes: efficiency, safety, and often just sheer enjoyment. I think we need all the help we can get from disciplines across the board to get to grips with what is going on here. And who is to say we cannot improve our daily procedures (some of them pretty creaky) for debate, or decision, or divorce, by applying some mathematical clarity and formal tools from disciplines who have developed these to great maturity? I would go even further and say that we need to bring in even further mathematical and natural sciences for the total picture of 'truly human' behaviour – with neuroscience and evolutionary biology high on the list.

My lunch was not the success I might have expected, but one can always hope that life at NIAS will offer second chances, at other tables (and maybe with other ladies). What I do know for sure is that the project 'Games, Action, and Social Software' has a future. The European Science Foundation in Strassburg just started a Eurocores strategic research

initiative *LogiCCC* across all major disciplines, and with participation from some twenty countries, from the British Isles to Cyprus:

I am sure that the GA&SS community, now that it has been formed, will be active on that wider scene as well, and that its last interactive word has not been spoken.