INTERACTION AS A FORM OF COGNITION

'Pre-dinner lecture', Johan van Benthem, 27 October 2004

KNAW Workshop Cognitive Foundations of Interpretation

1 Are there social cognitive phenomena beyond our individual heads?

2 'Dynamics' in semantics and logic: actions, mutual information \(\Rightarrow\) interaction

3 Starting from a simple conversation

\[Q\] "Is this the road to the KNAW?" \[A\] "Yes."

What does the questioner convey – and what about the answerer?

4 Processes involved:

* Interpretation (speaker's and hearer's meaning, equilibrium in interpretative strategies, one-shot versus repeated games)

* Communication: pre- and postconditions of speech acts, which information is passed exactly when we say something?

* Strategic games: behind every question, there is a "Why" question: intentions, strategies for asking and answering, imperfection/hiding.

* Protocols over time, reliability of agents, what have we learnt about them?

* Long-term phenomena, practices and societies, dynamical systems.

My own current interests mostly in the middle of this spectrum:

….. logics for update, belief revision, games, temporal evolution …..

5 Disciplines involved: linguistics, philosophy, logic, computer science, economics, mathematics, cognitive psychology – and so on.

6 Concrete examples: information update, multi-agents, and games

6.1 * Epistemic logic and questions/answers:

\(Q\) asks a factual question "\(P\)?", and \(A\) answers truly: "Yes". Presupposition for a truthful answer: \(A\) knows that \(P\): \(K_A P\). A normal cooperative question has presuppositions (i) \(\neg K_Q P \land \neg K_Q \neg P\)

\(\langle Q\rangle K_Q P \lor K_Q \neg P\) (\(Q\) thinks it possible that \(A\) knows the answer).

After the episode, \(P\) has become common knowledge among \(Q, A: C_{\{Q, A\}} P\).

* Updates = model change:

from \(P \bullet \bigcirc \neg P\) by assertion \(P!\) to

Briefly, public announcement changes models by world elimination.
Dynamic-epistemic logic then adds formulas with action modalities:

\[ M, s \models [P!] \phi \text{ iff } \begin{cases} M, s \models P, \text{ then } M \upharpoonright P, s \models \phi \end{cases} \]

Formulas \([P!]K\phi\) or \([P!]C_G\phi\): what agents know after announcements.

Logic: complete, decidable. 'Two Cultures' (3?) in such formulas: ling, phil, CS.

* New issues: specifying effects of speech-acts, 'learning problem'.

\([P!]C_G P\)?

Public statements do not always lead to common knowledge! "You don't know it, but P". Or, the last ignorance statement of the Muddy Children.

Moore-Fitch Paradox of the Knower: 'every true statement is learnable'?

More complex forms of communication: hiding, security, lying, ...

6.2 * Information games

Conversation space of admissible assertions, players say something non-trivial at every turn, who is the first to know the true situation? "Cluedo".

What we say depends on what we have heard before, plus our current plan.

Game theory: Global Nash equilibria exist. Most game values unknown…

* Dynamic-epistemic logic provides fine-structure to the game theory: update.

But also belief revision needed! Consider a game of the famous Centipede type:

\[
\begin{array}{c|c|c}
A & E & A \\
0,1 & 3,0 & 2,4 \\
\end{array}
\]

Backward Induction: \(A\) should play down at the start – so both players are worse off than toward the far right. But will this happen? 'Rational agents' always chooses the action that is to their own greatest advantage. But this is not the only possible kind, \(E\) may change her beliefs when \(A\) plays across: \(A\) is stupid, generous, adventurous...

Links with belief revision theory: Friday's Symposium "Changing Minds".

6.3 Diversity of agents: observational powers, strategies, societies of agents.

Types have to be learnt. "Memento": do finite automata always lose?

6.4 The long term: finite and infinite processes co-exist: special-purpose terminating programs versus the 'operating system' of a language. Need: integrate 'dynamic logic' and mathematical theory of dynamical systems.

7 But can we do experiments, 'real cognitive science' here?

* Solving puzzles: all around us, just check students' GSMs.
* Observing real games, experimental GT, e.g. 'complexity vs difficulty'.
Threshold behaviour, like from Chess to Kriegsspiel, new card games.

8 General questions, if all this makes sense:

From inference to interaction as a logical paradigm?

Redraw boundaries of disciplines having a common agenda?

Where does all this theory meet cognitive experiment?