Logic and the dynamics of information

Johan van Benthem, Amsterdam & Stanford, March 2003

Abstract We discuss how issues of information and computation interact with logic today, and what might be a natural extended agenda of investigation.

1 The Dynamic Turn

Products in tandem with processes Modern computers find their origin in the design of logical deduction machines, and modern computer science in the logical foundations of these same machines. But over the past decades computer science has also begun to influence the research agenda of logic. Traditionally, logic is about propositions and inference. Its account of this is declarative, in terms of languages and semantic models that represent information. But inference is in the first place an information-generating *process*, and just one among many at that. Other mechanisms of information flow are just as crucial in intelligent activities, such as asking questions, giving answers, and communication in general. Modern computer science deals with this broader spectrum. And a major characteristic of its modus operandi, here and elsewhere, is the interplay of *static and dynamic* structure. Processes cannot work without representations, but equally, the design of a good representation depends on its use. Thus, representation of information cannot be separated from the processes which use and transform that information. These days, in the same spirit, modern logic is undergoing a *Dynamic Turn*, putting activities of inference, evaluation, belief revision or argumentation at centre stage, not just their products like proofs or propositions. In fact, even traditional terminology has this double aspect. The word 'statement' denotes both an activity and the thing resulting from it, and so do 'argument' or 'proof'.

Social systems Influences from computer science have been broad, as this field is evolving. The original focus was on the structure of machines, algorithms, programming languages and their semantics. The corresponding work in logic has close ties with core areas of mathematical logic, such as proof theory and model theory. But computer science has moved to more ambitious themes, spearheaded by AI. This is the computational study of reasoning, language use, planning, and intelligent behaviour generally – much of which resembles basic themes from philosophical logic. In Clausewitzian terms, AI might even be described as philosophical logic continued by other means! More recently, partly through the technology push of the Internet, even mainstream computer science has incorporated communication, transactions, networks, and information systems of diverse kinds (Wooldridge 2002). Much of the newer research is far removed from details of machines or programs, and deals with computational structures in about every sort of social activity. This has further influences in logic, affecting the Dynamic Turn. In particular, modern computational processes are *social*, involving many agents, with mixtures of informational moves and other types of action. And after all, key logical activities like argumentation, or asking and answering questions, are indeed social in this sense. Thus, the Dynamic Turn is about multi-agent processes. The lonesome thinker in an armchair is as marginal as he looks: most of our logical skills are displayed in interaction.

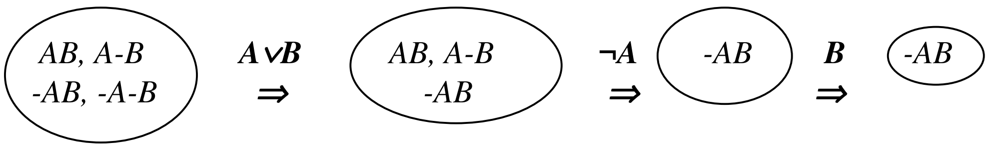
The dynamic disciplines Whether influences are just fashions or natural enrichments depends on the amount of resonance. Many of the above themes make sense precisely because they already existed in logic and related fields. Accordingly, the Dynamic Turn is fed by ideas from philosophy, linguistics, logic, probability theory, game theory, and other areas. Things that existed separately come together, but the pace and priorities of the convergence are stimulated by the focus of computation. Such an expansion in topics raises the issue of new natural boundaries. What is a coherent agenda for logic – or whatever better term one wants to pick – in this broader sense? This paper is a discussion of some contours, with a number of examples from current research.

2 The logic of communication

Inference and information update Logical processes can be brought to centre stage in several ways. For a start, existing practice has many dynamic features behind the scenes. To see this, consider propositional inference (van Benthem 1996). A person's information may be modelled as the set of all relevant possibilities that she entertains. Now consider the following ubiquitous inference:

from two data $\boldsymbol{A \vee B}$ *and* $\boldsymbol{\neg A}$, *draw the valid conclusion* $\boldsymbol{B}$

Without prior knowledge, two propositions *A, B* span *4* candidates for the real state of affairs. The premises then trigger updates restricting this set. In the limit, one option remains, and we know the actual situation. Here is an update video:



The two premises lead to an information state (in this special case, a complete one), where an update with the conclusion does not change anything. This is the hallmark of valid inference in a dynamic perspective: conclusions follow if they do not change the final information state produced by the successive premises.

Questions and answers Simple single-agent update mechanisms like the one pictured here work for captive audiences, or simple games like "Master Mind". But even the simplest episode of communication involves more than one person. Consider one question and an answer. Here is an example from van Benthem 2002. Suppose that I approach you with a question, and the following occurs

Q Is this the road to the Colosseum?
A Yes.

What information has passed? I learn the physical fact that this is indeed the road to the Colosseum. But much more has gone on. Merely by asking the question, I convey to you that I do not know, but think it possible that you know. (Note 1.)

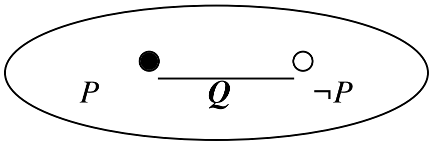
And by answering, you do not just convey the mere fact. You also bring it about that you know that I know, I know that you know that I know, etc. Indeed, we achieve so-called *common knowledge* of this being the road to the Colosseum, that is, mutual knowledge of arbitrary finite iteration depths. These epistemic overtones of communication can be crucial to further actions. Common knowledge is often taken to be a prerequisite to co-ordinated action in philosophy, linguistics, or game theory. But often we are conditioned by even finer epistemic differences. Even when I know your pin code, I may still not want to empty your bank account, unless I know that you do not know that I know your code.

These concerns have been around in linguistics and philosophy for quite a while, witness theories of speech acts. They have also reached computer science in the study of human-machine queries, and multi-agent information systems generally.

Epistemic logic dynamified From a logical point of view, a question–answer episode is about the smallest information-passing process. To describe it precisely, we need two things in tandem, as announced in Section 1:

statics	an account of multi-agent information models
dynamics	an analysis of natural updates transforming these, as triggered by assertions or other informational actions

Static models already exist in epistemic logic (Fagin et al. 1995), with possible worlds standing for all relevant total states of reality, related by uncertainty relations for all agents involved. E.g., a simple epistemic model for the 2-agent group *{**Q**, **A**}* in our question–answer episode has two states *'P', 'not-P'* – with *P* saying that this is the road to the Colosseum. The horizontal labelled line in the following picture indicates that ***Q*** cannot distinguish between the two:

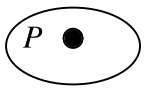


The black dot stands for the actual world. There are no uncertainty lines for ***A***, who knows the real situation. But to be completely precise, we would have to display looping lines for both ***Q***, ***A*** at both worlds. On such a model, the usual formulas of epistemic logic can be evaluated. In particular,

> knowledge $K_j\phi$ for an agent *j* in world *s* means
> that ϕ is true in all worlds accessible for *j* from *s*.

On our model, evaluating epistemic formulas in this way shows that, amongst other things, ***Q*** knows that ***A*** knows *whether P* is the case: $K_Q(K_AP \vee K_A\neg P)$.

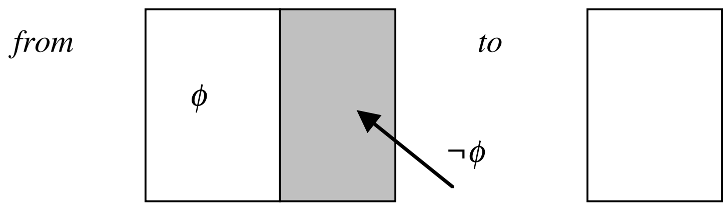
Next, ***A***'s answer triggers an *informational update* of this information model. This update eliminates the option *not-P*, turning the model into the one-point diagram



At this stage, *P* is common knowledge between ***Q***, ***A***. In epistemic semantics, this means that *P* holds at each world which can be accessed by following their uncertainty lines. Or, in standard notation, the following formula has become true:

$$C_{\{Q,\,A\}}P$$

The general dynamics here is this. A *public announcement* $\phi!$ of an assertion ϕ eliminates all those worlds from the current model which fail to satisfy ϕ:



Our example was extremely simple. With larger epistemic models, world elimination acquires much more striking effects. (Note 2.) It then clarifies, e.g., the famous Muddy Children puzzle and other group scenarios. These observations have been the starting point for a whole line of research on update mechanisms.

Dynamic-epistemic logic Even the above account still treats epistemic actions as second-class citizens. The language of epistemic logic does not display them – and as yet, we have no calculus for reasoning about them explicitly. A truly two-level static-dynamic system implementing the Dynamic Turn in this particular case arises when we import another idea from computer science. This is the coexistence of propositions and action expressions in so-called *dynamic logics*. In particular, such languages for describing behaviour of programs have expressions describing conditions which hold in states resulting from performing actions:

$[a]\phi$ $\quad$ ϕ holds after every successful execution of action a

Originally, one thought of a as an expression for a computational program or perhaps some physical action – but it can just as well be a communicative act. Now we can express statements about epistemic effects of communication, like

$[A!]K_j\phi$ $\quad$ after a true public announcement of A, j knows that ϕ

There are even complete and decidable calculi for this sort of statement. The dynamics then typically has to do with how static assertions relate before and after actions took place. As an illustration, here is a valid principle relating knowledge achieved after an announcement to what agents know beforehand:

$$[A!]\, K_j\phi \;\leftrightarrow\; (A \rightarrow K_j(A \rightarrow [A!]\, \phi))$$

Analyzing speech acts These systems provide a grip on more general issues. Suppose we do as computer scientists do in programming, and ask for complete general *specifications* of speech acts. Say, what do we learn from a public announcement? Is the point of an action $\phi!$ for a proposition ϕ in epistemic logic that it always produces *common knowledge* of ϕ? After all, we report such events as: "I learnt that ϕ". In the perspicuous notation of dynamic-epistemic logic, this would read as follows, referring to public announcement in a group of agents G:

$$[\phi!]C_G\phi$$

But this 'learning principle' is false. E.g., if ***A*** had said ϕ = "You don't know it, but this is the road to the Colosseum", this would have been true, the same update would have occurred, but the assertion ϕ itself would become false! Philosophers will recognize Moore's Paradox here, now as an issue in dynamic epistemic logic. It is an interesting open question which forms of epistemic assertion do produce common knowledge when announced. Thus, update logic takes up issues from speech act theories, but with techniques unknown in the early days. (Note 3.)

General communication Not all communication is public. There are many forms of more private information transfer, hidden wholly or in part from other agents. E.g., van Ditmarsch 2000 gives a complete analysis of all communicative moves in the game "Cluedo". The corresponding updates produce more complex changes in epistemic models than just world elimination. The most sophisticated system to date is that of Baltag, Moss & Solecki 1998/2003, which deals with mixtures of public and private information, and even hiding and cheating, where agents may become systematically misinformed. Our daily lives contain many subtle communicative settings of this sort. Update logic promises a more systematic logical taxonomy and understanding of these phenomena.

The dynamic stance This section is meant as an existence proof, not a course in update logic. It shows that the Dynamic Turn is not just a metaphor: it can be made to work in a concrete technical sense. Also, it is good to step back, and realize the long intellectual history encapsuled in a simple update formula like $[A!]K_j\phi$. It brings together linguistic speech acts, philosophical epistemology, and program logics from computer science. Such links help propagate insights in one area to others. Finally, update analysis is a mind-set which, once acquired, changes one's perception all around. In particular, cognitive processes are behind just about every topic in epistemology, and the Dynamic Turn makes these central concerns. Thus, the counterpoint to the traditional study of knowledge or belief and their links to reality is a vigorous account of cognitive activities of acquisition and revision of beliefs. (Note 4.) Let us now turn to these.

3 Further cognitive actions: belief revision, learning

From update to revision Information update is just one important activity that we engage in. Many other characteristic processes exist with logical features. Another key source for the Dynamic Turn is the theory of belief revision in the 1980s (cf. Gärdenfors 1987), which highlighted the interplay of three processes:

(a) information *update*, adding certain propositions
(b) information *contraction* leaving out propositions
(c) *belief revision* changing prior beliefs to accommodate new ones.

All three are ubiquitous in life, as we confront our expectations with observations, and have to rearrange them. They also occur on a grander scale in science, when we change theories that contradict the facts, or even *themselves*.

Belief revision theory There is nothing mysterious about these processes, and it makes sense to search for their logic. Taking the dual computational stance again, what are the relevant data structures, and what are natural transformation steps? Belief revision theory proposes syntactic and semantic representations of theories plus an account of the revision process via basic postulates, and optional ones reflecting more conservative or more radical policies for changing one's beliefs. Moreover, there is not just transformation of propositional information. One can also change agents' plausibility orderings between worlds, or their preferences, or indeed any parameter in logical semantics that admits of meaningful variation over time. Revision dynamics may even change one's language or conceptual framework. Rott 2001 is a modern treatment of the state of the art.

From a philosophical point of view, belief revision theory filled a gap. The classical foundations of mathematics picture reasoning as serene accumulation of truths, and sometimes even proven guarantees for consistency. Then Kuhn's account of paradigm shifts in science washed the dirty linen of this aristocratic family, and made theory changes look like palace revolutions. Belief revision theory shows there is a task for logic in describing change as well as continuity.

Learning theory Evidently, people have various strategies for revising theories, or just our ordinary opinions. In a sense, belief revision theory is not out-and-out dynamics yet, as those processes themselves are not manipulated as first-class citizens in the calculus. An example of the latter move is the explicit theory of *learning mechanisms* in Kelly 1996, merging ideas from the philosophy of science, mathematical topology, and computer science. Hendricks 2002 makes an extensive plea for the broad epistemological relevance of this move. Update, revision, and learning form a coherent family of issues, going upward from short-term to long-term behaviour.

This is not to say that all manifestations of the Dynamic Turn are one coherent family. For instance, update logics and belief revision theory still call for a merge. E.g., there are no generally accepted multi-agent dynamic logics yet dealing with belief revision and learning in the perspicuous style of our update calculus for communication. And conversely, epistemic update logic without revision is just half of the story of our lives, which constantly mix update of information with revision of expectations. Indeed, as one more motivation for the Dynamic Turn, it seems that the most characteristic human cognitive ability is not the static virtue of *being right*, but the dynamic one of being able to *correct* ourselves. This is an old Popperian and Quinean point of course, but now with a logical twist.

4 From single actions to games

Scale levels The preceding section was about diversity of cognitive processes. But dynamics also raises an important issue of *scale*. Public announcements are just building blocks for larger activities, such as arguments or conversations. But to understand what is going on there, we do not just ask *what* people are telling us, but also *why*. It is hard to make sense of even a single question without understanding what setting we are in. Is the questioner a high-minded Gricean trying to be helpful? Recall the Colosseum example. Was the question perhaps rhetorical, with some ulterior motive, and do I have best options for responding serving my own aims in that scenario? Competent language users are good at sensing where they stand, and planning their communicative moves accordingly.

What are natural scale levels in linguistic and logical activities? The usual emphasis in logic has been on the micro-level of single propositions and their meaning, with an occasional interest – though an important one! – in meso-level structures like proofs. Finally, logic has a minimalist account of epistemic macro-structures, treating theories as sets of formulas. (Notes 5, 6.) The Dynamic Turn has a natural interest in higher levels of aggregation, as these provide much of the point of separate assertions or inference steps in the first place.

Games One difficulty in extending logic from the sentence level to a discourse level has been the scarcity of mathematical paradigms satisfying the standards that one has become used to at the sentence level. In recent years, a congenial mid-size level has been found in *game theory*. Games are typically a model for a group of agents trying to achieve certain goals through interaction. They involve two new notions compared with what we had before: agents' *preferences* among possible outcome states, and their longer-term *strategies* providing successive responses to the others' actions over time. In particular, strategies take us from the micro-level to a description of longer-term behaviour.

Game theory adds an ambitious agenda to what we saw so far. Update or revision steps are just single steps of a cognitive machine which might run any kind of program. It is only a game which provides a purpose and sense to such moves. Why am I asking? What am I trying to prove, for what? Our task now becomes to devise optimal strategies, i.e., ways of asking, answering, or proving that will serve the purpose. Indeed, game theory is still more ambitious, as it tries to predict strategic equilibria that reflect stable long-term behaviour for agents interacting in a group. This may be applied to concrete information games, but also to generic games standing for types of social activity – including language use or logical reasoning. Much of the mathematics of the field is about equilibria and their properties, for players having more or less information at their disposal.

Games and logic Despite these differences in scope and aims, game theory and logic have natural connections. Van Benthem 1999–2002 presents a panorama of games inside logic for semantic evaluation, argumentation and other key activities. (Hintikka 1973 is an early source.) The other side of the contact are current logical investigations of deliberation and decision making by players in general games, as an underpinning to the mathematics strategic equilibrium. For instance, Stalnaker 1996, 1999 show how the study of rational behaviour in game theory ties in to mutual benefit with basic concerns of philosophical logic.

But moving to games make sense across the whole community involved in the Dynamic Turn. In linguistics, interpretation of utterances is like a game where preferences of speakers and hearers determine what is said and how it is taken (Parikh 2002). Following Lewis' work on conventions and the game theory of signalling games, van Rooy 2002 takes this to a game-theoretic analysis of Gricean maxims. And from a computational viewpoint, games are distributed processes of interaction and communication. Various authors have begun merging ideas from logics in computer science with ideas from game theory, witness Parikh 1985, Abramsky 1998. The resulting mix of interests is already creating a new community in conferences like TARK, LOFT, or GAMES.

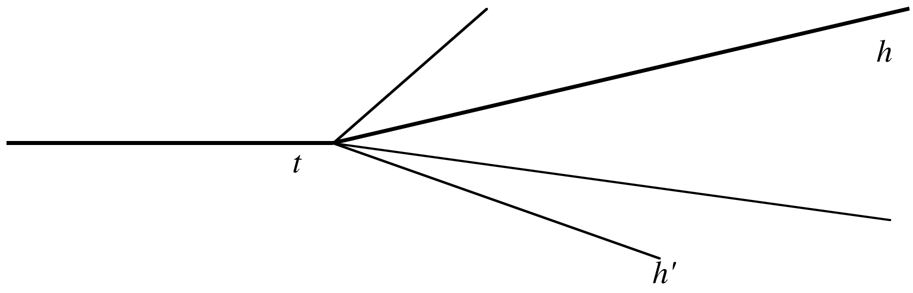
Given the coherence in aims and the emergence of sound technical connections, some mix of dynamic-epistemic logics, belief revision theories, and game theory might be the best engine for achieving the aims of the Dynamic Turn in logic. But in making this work, we still have one more stage to go!

5 Longer-term processes

Finite versus infinite processes Games seem terminating activities, similar to proofs or talks. But computer science suggests a different perspective. After all, programs come in two broad varieties. Some are instructions for terminating computational tasks, and infinite continuation amounts to failure. This is the way most people think about Turing machines, or plans, or talks. But other programs are designed to go on forever, and it is finite termination that would be a problem. The operating system of a computer is a good example of the latter kind. The same dichotomy occurs in the cognitive processes involved in the Dynamic Turn. Some activities are meant to terminate, others provide the operating system for short-term tasks to succeed. Examples of the latter are logical calculi in the functioning of proof, or Gricean maxims in running conversation. Game theory has the same dichotomy, witness the importance of infinite games like repeated Prisoner's Dilemma in understanding social co-operation (Axelrod 1984).

Protocols Long-term issues also come up in the analysis of communication or other logical activities. Computer scientists model regularities in potentially infinite behaviour in *protocols* (Fagin et al. 1995). A typical protocol may restrict moves available to players at any stage, like restricting your choice of things to say. A protocol can also encode more global regularities, like a server's making sure that every request gets answered eventually. Thus, knowing that some protocol is being followed excludes certain courses of events when thinking about the total development of a process. In the same way, one may have other long-term information about other agents. Perhaps my interlocutor is a person who lies and speaks the truth alternatively – or even, a person who speaks the truth about 50% of the time on average. This may be highly relevant to understanding the true effect of an assertion. If you are an alternative liar and truth-teller, I need to maintain a parity count to determine how to take your next assertion. It has even been argued that individual assertions do not determine meaning at all without all this protocol information. This would go far beyond the usual, more conventional linguistic notion of context in determining the content of an utterance. Be this as it may, there certainly seems to be a case for placing cognitive activities on a larger, potentially infinite temporal stage. How can this be achieved?

Branching temporal logic As it happens, a common model seems to be emerging in the literature on processes and protocols. *Branching temporal models* reflect the well-known intuitive tree picture of forking world-lines (Note 7):



Such models describe the temporal evolution of a system over time, but they can also include agents' knowledge and beliefs. E.g., expectations over time can be modelled as subsets of the possible future histories, while there may also be indistinguishability relations all across the tree to model limited knowledge of where agents are. Roughly this temporal universe underlies the computational run model of Fagin et. 1995, the infinite game model of Abramsky 1996, the protocol model for messages in Parikh & Ramanujam 2002, the universe for learning mechanisms in Kelly 1996, and the philosophical analysis of action, choice, and deliberation in Belnap et al. 2001. This seems the appropriate stage for putting together individual update steps, games, and other lower- to mid-scale logical activities, while allowing for infinite processes running in the background. Of course, the real work will be in the detailed description of logical activities over time including the possible role of long-term protocols.

Evolution and dynamical systems In modern game theory, the infinite perspective has come up most forcefully in connection with *evolution*. Many properties of social behaviour can be analyzed as equilibrium features of infinite dynamical systems, often with a state-transition function of some biological sort (cf. Osborne & Rubinstein 1994). Strategic equilibria then derive their stability from repeated simple encounters between different types of agents, rather than some deductive justification to be gone through every time they meet. This is a very different style of thinking about long-term behaviour, where stable structures emerge as statistical properties of populations. On this view, social, linguistic, and perhaps even logical behaviour may consist of emergent properties of simple steps repeated in bulk. This view is also becoming felt in philosophy as an alternative to classical justification-based explanations of behaviour.

A natural border at last? Compared to our motivations for the Dynamic Turn, the evolutionary paradigm represents a very different way of thinking. But it cannot be denied that some cognitive phenomena seem emergent statistical features of large-scale human behaviour, rather then logic-driven ones. Examples are the spread of gossip, where the initial information tends to evaporate, or the dynamics of mass opinion, which can be modelled quite well in terms of physical state equations (Mouwen 1998). Even closer to logic, evidence for the importance of an emergent statistical level comes from the interesting phenomena discovered in automated deduction, with complexity thresholds for repeated tasks behaving like phase transitions (Kirkpatrick & Selman 1994). Some recent architectures for language understanding even mix logical rule-based components for creating partial representations with statistics-based memories (Bod 1998). (Note 8.)

Life at this border is exciting. E.g., Skyrms 2003 uses evolutionary models, amongst many other things, to explain the emergence of social structures regulating interaction and communication. These tie in with earlier concerns: think of communication in structured groups with prescribed channels. (Note 9.) Also, there are mathematical challenges of integrating dynamic logic and the theory of dynamical systems. Even so, the statistical and biological phenomena outlined here seem a natural frontier for the Dynamic Turn as initially conceived.

7 From description to design

The final relevant aspect of computation that we wish to mention in this paper strikes out in a different direction. Computer science does not just describe reality as it is, it also creates a new reality conforming to its theories by *designing* systems and virtual realities. This more activist perspective also applies to the Dynamic Turn. For instance, consider the update logic of Section 2. This formalism may be used to *analyze* given assertions, or given communicative practices. But it can also help synthetize new statements for certain purposes. An example is the 'Moscow Puzzle' (van Ditmarsch 2002):

> "*A* gets *1* card, *B* and *C* get *3* cards each. What should *B, C* tell each other in *A*'s hearing so that they find out the distribution, while *A* does not?".

Going beyond such puzzles, one might even think about creating whole new practices. This is the thrust of 'mechanism design' in game theory, or more generally, the 'social software' program of Parikh 2002. The latter proposes to merge mechanism design and programming techniques to design optimal social procedures for communication, voting, and other practical purposes. Many of these will have to do with action and communication. Here is a simple example, proposed by a student in a recent mathematics course in Amsterdam:

> In the preparation for 'Sint Nicolaas Night' on December 5th in The Netherlands, each family member is assigned one person for whom they have to write a poem, and make some kind of surprise gift. The specifications are that no one should be assigned to himself, and that no one should know anything about 'who has whom'. Some families sometimes draw lots, and repeat this if people draw themselves. But other practices abound. Is there a way of assigning people to people which works without probabilities, and which involves only public actions that are observable by everyone? (Note 10.)

The same move from understanding to creating can be made in the Dynamic Turn. The dominant mode in logic or philosophy has been descriptive. Accordingly, discussions often have to do with adequacy of some proposed logical system for the initial descriptive purpose. If counter-examples emerge, the system is modified or discarded. But in science, systems often find their true use by discarding the original application, finding new ones. Likewise, nothing prevents us from turning logics of action and communication that fall short of existing reality into design systems for new languages or argumentation practices! Indeed, this ambition has been around in AI for a long time. But it need not just apply to machines: it can also enrich our own fund of human practices.

8 Conclusion

The Dynamic Turn adds a second focus to the agenda of logic: the systematic study of the cognitive actions producing the static representational objects that were studied mainly so far. This turn involves a mixture of logic, philosophy, linguistics, computer science, and even economics, and it may lead to quite different border lines between these fields in the future.

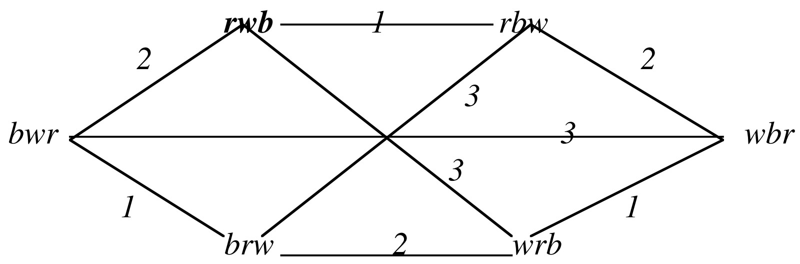
But it is good to reflect on the agenda that we get in this way, as ambitions may be running wild. Logic as it is describes mainly the building blocks of inference and conversation. But as we proceeded through the sections of this paper, ambitions became much higher. In Section 2, we wanted a taxonomy for natural styles of communication. In Section 3, in addition to describing a steady state of competence, we wanted to explain how logical systems can be learnt. In Section 4, we wanted to explain the functioning of linguistic rules and conventions, thereby intertwining semantic issues with pragmatic ones. And subsequent sections added yet further goals, such as reasoning and information flow in groups and organisations, and integration of short-term tasks with long-term processes. Much of this has to do with not just describing logical phenomena, but also explaining them. (Note 11.) And finally, in addition to description or explanation, we wanted to use all these insights to change the world.

Does this paper present a realistic set of goals for the Dynamic Turn, with natural frontiers, or does it just display over-extended imperialism? We think that the broader set of goals outlined here forms a coherent enterprise, even though it lies scattered across different fields so far. But even conservatives and sceptics of the Great Leap Ahead proposed for logic in this paper might agree that it is worthwhile rethinking the agenda of a field every now and then.

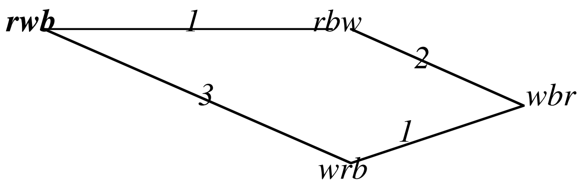
9 Notes

1 Questions come in genres. Of course, neither presupposition of the question would hold if I were a teacher, and you a student. More on that below.

2 Here is an example from van Ditmarsch 2000, van Benthem 2002. Three players *1, 2, 3* get a card from 'red', 'white', 'blue'. Each can see their own card, but not that of the others. The real distribution over *1, 2, 3* is *red, white, blue*. Here is the resulting information state pictured as an epistemic model:



The diagram says the following. Though they are in ***rwb***, no player knows this. As they ponder their group situation, they must take into account all 6 worlds. Now *1* says: *"I do not have the blue card"*. What do players know about the cards after this? Solving this in words is a bit complicated, but here is the correct update, removing the two worlds starting with ***b*** :



This shows at once that *2* knows the distribution, *3* knows that *1* knows, and *1* knows only that *2* or *3* knows. But, e.g., it is not common knowledge that *2* knows! For, *1* thinks it possible that *2* has the blue card, in which case the first assertion would not have helped her. The diagram shows the effects of further assertions. E.g., if *3* now were to say *"I still don't know",* only the left-most worlds would remain, and *2* would find out the correct distribution.

3 Another typical aspect of communication are complex actions constructed out of basic actions such as assertions, questions, and the like. Just as in computer

programs constructed out of primitive instructions, one encounters *composition* (e.g., "first say this, then say that"), *conditional choice* (e.g., "if you know it, then say so, else ask"), and *guarded iteration* (e.g., "as long as they do not know, keep telling them…"). Update logic relates effects of complex actions to those of their components, just as happens in the dynamic logic of computer programs.

4 Here is a toy example. Consider the much-discussed Knower's Paradox in verificationism (Wansing 2002). The following elementary derivation undermines the equation of truth of assertions ϕ with 'possible knowledge' $<>K\phi$:

$$P \mathbin{\&} \neg KP \rightarrow <>K(P \mathbin{\&} \neg KP), \qquad P \mathbin{\&} \neg KP \rightarrow <>(KP \mathbin{\&} K\neg KP)$$

$$P \mathbin{\&} \neg KP \rightarrow <>(KP \mathbin{\&} \neg KP), \qquad P \mathbin{\&} \neg KP \rightarrow \bot \text{, i.e., } P \text{ implies } KP!$$

Note how the troublesome substitution instance is Moore's formula $P \mathbin{\&} \neg KP$ again. Now, presumably, verificationism means something like: "if proposition ϕ is true, we might come to learn it ". But learning involves an action. In update terms, one of the simplest actions of this sort would be hearing ϕ from some authority. Thus, the principle would say, in quasi-formal jargon, that

$$P \rightarrow \exists\phi\colon \phi \mathbin{\&} [\phi\,!]P$$

The operator $\exists\phi\colon \phi \mathbin{\&} [\phi\,!]P$ on the right-hand side is not in the above update language as it stands. But even so, the earlier analysis of dynamic phenomena would lead us to predict that this principle must be false, as some assertions P are affected by the announcement of their truth – Moore-like ones in particular. Determining the logic of this new operator would generalize the earlier learning problem. Thus, we turn a problem to be cured into an object of constructive study.

5 More sophisticated views of theory structure and inter-theory relations occur in the philosophy of science – and to some extent also in computer science.

6 Higher scale levels have further emergent phenomena. E.g., at a macro-level, it makes sense to look at the logical structure and collective actions in *organisations* (cf. Kamps 2000). Pioneers here are computer scientists analyzing complex information systems, or game theorists involved in 'mechanism design'.

7 Mathematical options are set differently from author to author in this area: cf. Reynolds 2002, Zanardo 2002.

8 Adriaans 2002 looks at machine learning in various realistic settings, proposing a taxonomy of types of information systems calling for either logical or statistical approaches, depending on scale size and available expert knowledge. The scope of the approach includes learning and self-adapting organisations.

9 Van Benthem 2002 takes a first look at what happens to basic update logic for groups with prescribed communication channels.

10 The student's family solved this as follows, in full view of the entire group. Everyone's name was written on a card and an envelope. The cards were put in the envelopes, with the names facing the same way. First the envelopes were shuffled. (Shuffling is an interesting logical action removing information.) Then they were put in a circle on the table. Now the cards were drawn out, face down, and shifted one position around the circle of envelopes. They were then put in the envelope next to them. Afterwards, the envelopes were shuffled again.

11 There are even further possible ambitions, not covered in this paper. For instance, what is the connection between the Dynamic Turn and the psychological and neuro-biological insights coming to light in experimental cognitive science?

10 References

S. Abramsky, 1996, Semantics of Interaction: an Introduction to Game Semantics, in P. Dybjer & A. Pitts, eds., Proceedings 1996 CLiCS Summer School, Isaac Newton Institute, Cambridge University Press, Cambridge, 1–31.

S. Abramsky, 1998, 'From Computation to Interaction, towards a science of information', BCS/IEE Turing Lecture.

P. Adriaans, 2002, *The Robosail Project*, keynote lecture, ECAI-2002, Lyon.

R. Axelrod, 1984, *The Evolution of Co-operation*, Basic Books, New York.

A. Baltag, L. Moss & S. Solecki, 1998, 'The Logic of Public Announcements, Common Knowledge and Private Suspicions, *Proceedings TARK 1998*, 43–56, Morgan Kaufmann Publishers, Los Altos. Updated version, 2003, department of cognitive science, Indiana University, Bloomington, and department of computing, Oxford University.

N. Belnap, M. Perloff & M. Xu, 2001, *Facing the Future*, Oxford University Press, Oxford.

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford.

J. van Benthem, 2002, 'One is a Lonely Number', invited lecture, Colloquium Logicum, Muenster. Report PP-2003-07, ILLC, University of Amsterdam.

R. Bod, 1998, *Beyond Grammar*, CSLI Publications, Stanford.

H. van Ditmarsch, 2000, *Knowledge Games*, dissertation DS-2000-06, ILLC Amsterdam and University of Groningen.

H. van Ditmarsch, 2002, 'Keeping Secrets with Public Communication', department of computer science, University of Otago.

R. Fagin, J. Halpern, Y. Moses & M. Vardi, 1995, *Reasoning about Knowledge*, The MIT Press, Cambridge (Mass.).

P. Gärdenfors, 1987, *Knowledge in Flux: on the dynamics of epistemic states,* The MIT Press, Cambridge (Mass.).

V. Hendricks, 2002, 'Active Agents', PHILOG Newsletter, Roskilde. To appear in J. van Benthem & R. van Rooy, eds., special issue on Information Theories of the *Journal of Logic, Language and Information*.

J. Hintikka, 1973, *Logic, Language Games, and Information*, Clarendon Press, Oxford.

J. Kamps, 2000, *A Logical Approach to Computational Theory Building, with applications to sociology*, dissertation, ILLC Amsterdam.

K. Kelly, 1996, *The Logic of Reliable Inquiry,* Oxford University Press, Oxford.

S. Kirkpatrick & B. Selman, 1994, ' Critical Behavior in the Satisfiability of Random Boolean Expressions', *Science* 264, 1297–1301.

C. Mouwen, 1998, *The Dynamics of Opinion Change*, Tilburg University Press, Tilburg.

M. Osborne & A. Rubinstein, 1994, *A Course in Game Theory*, The MIT Press, Cambridge (Mass.).

P. Parikh, 2001, *The Use of Language*, CSLI Publications, Stanford.

R. Parikh, 1985, ‘The Logic of Games and its Applications’, *Annals of Discrete Mathematics* 24, 111–140.

R. Parikh, 2002, 'Social Software', *Synthese* 132, 187–211.

R. Parikh & R. Ramanujam, 2002, 'A Knowledge Based Semantics of Messages', CUNY New York & Chennai, India, to appear in J. van Benthem & R. van Rooy, eds., special issue on Information Theories of the *Journal of Logic, Language and Information*.

M. Pauly, 2001, *Logic for Social Software*, dissertation DS-2001-10, Institute for Logic, Language and Computation, University of Amsterdam.

M. Reynolds, 2002, 'Axioms for Branching Time', *Journal of Logic and Computation* 12, 679–697.

R. van Rooy, 2002, ''Quality and Quantity of Information Exchange', to appear in J. van Benthem & R. van Rooy, eds., special issue on Information Theories of the *Journal of Logic, Language and Information*.

H. Rott, 2001, *Change, Choice and Inference*, Clarendon Press, Oxford.

Y. Shoham, 2002, *Multi-Agent Systems*, lecture notes, Department of computer science, Stanford University.

B. Skyrms, 2002, *The Stag Hunt: evolution of social structure*, Department of philosophy, UC Irvine.

R. Stalnaker, 1996, 'Knowledge, belief, and Counterfactual Reasoning in Games', *Economics and Philosophy* 12, 133–163.

R. Stalnaker, 1999, ‘Extensive and Strategic Form: Games and Models for Games’, Research in Economics 53:2, 93-291.

H. Wansing, 2002, 'Diamonds are a Philosopher's Best Friends', *Journal of Philosophical Logic* 31*, 591–612.*

M. Wooldridge, 2002, *An Introduction to Multi-Agent Systems*, John Wiley, Colchester.

A. Zanardo, 2002, 'First-Order and Second-Order Aspects of Branching-Time Semantics', Department of mathematics, University of Padova, to appear in *Proceedings HPLMC–02*, San Sebastian.