Language, logic, and communication

Johan van Benthem
http://turing.wins.uva.nl/~johan/
Institute for Logic, Language and Computation. University of Amsterdam
& Center for the Study of Language and Information, Stanford University

Abstract

Natural language is the primary vehicle for human communication. But the information flow in questions, answers, or more general interactive discourse, is far from completely understood. Our paper presents a simple dynamic model for these phenomena, and explains some issues of 'informatics' raised by this at the interface of linguistics, philosophy, logic, and computer science.

1 A question–answer episode

In each linguistic episode, we convey information to each other. Consider the simplest building block of communication: a simple *question* and a truthful *answer*. Let us move to the good old times when the Dutch were non-nonsense Roman elite soldiers (mentioned by Tacitus), as opposed to timid UN peace-keepers. I approach you in a busy Roman street, A.D. 160, intent on contacting my general Maximus, now fighting for survival as a gladiator – and ask you for the road to the Coliseum:

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

What information has passed? I learn the physical fact that this is indeed the road to the Coliseum. 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. 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 I know, and indeed, we achieve so-called *common knowledge* of this being the road to the Coliseum. These epistemic overtones of communication are not side-effects, 'gone with the wind' once we part. They can be crucial to our further physical actions. Some bystanders' knowing that I know may lead them to run off and warn the Emperor Commodus – my knowing that they know I know may get me to prevent their doing just that.

Understanding what goes on in the above uses of language seems a typical concern for linguists. But as it happens, the first systematic theory of 'speech acts' is due to philosophers in the 1950s. Since the 1980s, similar questions have also come up in computer science, e.g., with human-machine queries. A question–answer episode is a small information-passing process. So, one would like to have (in computer science jargon) a 'specification' of its preconditions and complete effects. But perhaps our informal story was just that? Truthful answers to Yes/No questions establish common knowledge of the fact at stake? We will return to this issue later.

2 A dynamic turn: from knowledge to cognition

The standard semantic view of both natural and formal mathematical languages sees them as describing worlds, either the real world around us, or more abstract situations. Thus the meaning of an expression resides in its 'truth conditions', determining under which circumstances it is true. These truth conditions are guided by the syntactic form of the language. But in addition to syntax and semantics, languages also have a pragmatics of their use. Current studies of languages tend to put their informational and cognitive functions in focus, such as communication. Thus, activities and processes come to the fore, rather than static 'meanings' – and natural language comes to be more like the programming languages of computer science: an expression serves to trigger information changes in readers and hearers. This requires a precise account of *information states* and *updates* over these, which is a highly non-trivial research program. In particular, one needs a notion of meaningful information beyond the signal–coding provided by Information Theory. In this paper, we will show a few samples, moving freely between linguistics, philosophy, logic, and computer science – as is typical for modern developments in 'informatics'. Incidentally, this 'dynamic turn' from knowledge viewed as static information structure to *cognition* as a dynamic activity is not an isolated fashion. It reflects a much broader trend in current philosophy and computer science.

3 One-way update and individual knowledge

We start with one-way information flow, the simplest case. To a first approximation, a person's information about the world may be modeled as *the set of all relevant possibilities that she entertains*. An example is ordinary Boolean logic, now with a dynamic twist. Consider the following well-known inference:

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

This involves two propositions A, B, yielding 4 candidates for the real state of affairs. Information comes by utterances triggering updates, restricting this set. In the limit, only one option remains, and we know the actual situation. Here is a 'video' of the successive information states and the updates for the above inference:

AB, A-B -AB, -A-B	**A∨B** ⇒	AB, A-B -AB	**¬A** ⇒	-AB	**B** ⇒	-AB

Here the ovals stand for information states, with the relevant possibilities indicated. The two premises peforemd the changes indicated, leading to complete information (only one possibility '–AB' remained), and the valid conclusion did not add any. This update mechanism works for simple games like "Master Mind".

Players must guess the exact linear ordering of some coloured pegs. In each round, they are allowed one guess – which is evaluated with a black mark for correct colours at correct positions, and white marks for correct colours at incorrect positions. This eliminates all options incompatible with the information received. The aim is to narrow down one's options to one single situation. Consider a game with 4 colours and 3 pegs, with no colour repeated. Let the actual sequence be

<**red, white, blue**> with a total colour set also including **orange**.

Here is a game table, with numbers of open possibilities indicated.

guess	*answer*	*open options*
START		24
red, orange, white	• o	6
white, orange, blue	• o	2
blue, orange, red	o o	1

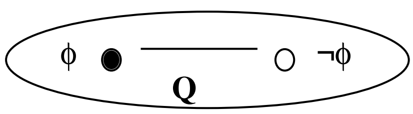
Real players also make useful *deductions*. After the first round in our Master Mind session, we conclude "blue must occur". After the second, red must occur, etc. Such inferences distinguish explicit *knowledge* from implicit *information*. Thus, in addition to update mechanisms, we are also interested in proof systems.

4 Two-way information flow in communication

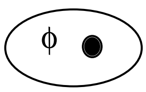
Our question–answer.example was a *social setting*, where two people update each other's information. Daily life is replete with coordination, planning, keeping secrets, or gossiping, involving such reckoning with other people's information at various 'depths' of reciprocity. Thus, information flow in communication is ubiquitous – and we are good at keeping track of it! Moreover, we are attuned to subtle distinctions. Groups generate levels of information beyond individual knowledge and ignorance. A celebrated example is universal knowledge versus *common knowledge* of a proposition ϕ . The former means all group members know ϕ – without implying that they know that the others know it, too. Common knowledge requires universal knowledge plus awareness that the others know that you know, you know that they know, etcetera – to any level of repetition. The distinction is not just academic. If it is universal knowledge that your partner is unfaithful, but people do not know that the others know, you may try to live the problem down. But with common knowledge, every meeting is an embarrassment, and you may want to restore your honour by a violent act. Common knowledge of rules is a cognitive 'equilibrium' which is essential to the emergence of stable linguistic and social behaviour, as has been argued by Lewis and other philosophers. The notion has also emerged in game theory and computer science, as a must for succesful interaction between agents.

5 Uncertainty models and update for group information

Despite our competent practice, it is not easy to specify precisely what information is passed in communication. One may try in words, but how to be sure the account is correct and complete: taking all relevant aspects into account? Pictures do better. An *information model* for the group {you, me} in our initial question–answer episode has states 'ϕ' , not-'ϕ', with ϕ saying "this is the road to the Coliseum". We draw these states as points in a diagram. In addition, we indicate agents' possibly different uncertainties between these various states. The horizontal labeled line in the following picture indicates that **Q** cannot distinguish between the two:



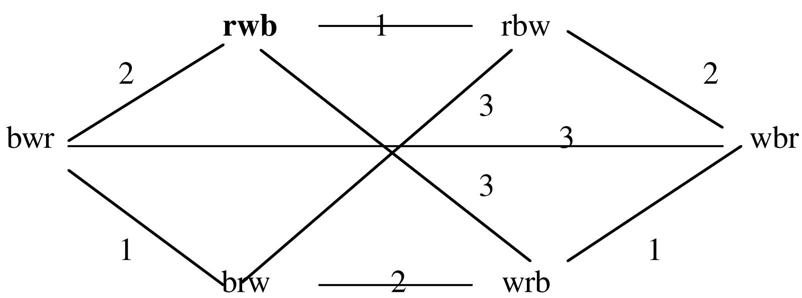
The black dot is our outside indication for the actual world where the agents live. There are no uncertainty lines for agent **A** . This reflects the fact that **A**, being a Roman local, know if this is the road to the Coliseum: no uncertainties there. Next, **A**'s answer triggers an *informational update* of this information state. It eliminates the option *not*-ϕ, turning the first information model into the one-point diagram



This picture has only one possible state of the world, where the proposition ϕ holds, and there is no uncertainty line for anyone. In particular, this indicates that ϕ is now common knowledge between you and me. To investigate this update mechanism more precisely, we will use a calculus of agents' information due to the philosopher Jaakko Hintikka in the1960s – and rediscovered in economics by Robert Aumann in the 1970s. But first, one more example.

6 Updates in card games

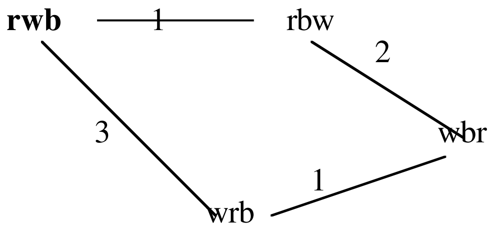
Consider *games*, an area with vivid intuitions from first-hand experience. Three cards 'red', 'white', 'blue' are distributed over 3 players: 1, 2, 3, who get one each. Each can see their own card, but not that of the others. The real distribution over the players 1, 2, 3 is red, white, blue. Here is the resulting information state:



This pictures the 6 relevant states of the world (the distributions of the cards), with the appropriate uncertainty lines between them. E.g., the single 1-line between *rwb* and *rbw* indicates that player 1 cannot distinguish these two situations, whereas players 2 and 3 can (since they have different cards in them). In particular, the diagram says the following, intuitively. Though they are in **rwb** (as an outside observer might notice), no player knows this. As they ponder their group situation, they must take into account all 6 worlds. Now *2 asks 1 "Do you have the blue card?",* and *1 answers truthfully "No"*. Here is the effect in words:

> Everyone learns 1 does not have the blue card. But other effects are quite different. Player 2 can figure out the facts: he has white, so 3 must have blue, and 1 red. Player 3 already knew 1 did not have blue, so he learns less – and player 1 even less by telling the truth about himself. But player 3 does understand that 2 has figured out where he is: an advance in 3's 'higher-order' knowledge. By contrast, 1 only learns that one of 2 or 3 must be the lucky knower – but cannot tell which one, because of the obvious symmetry.

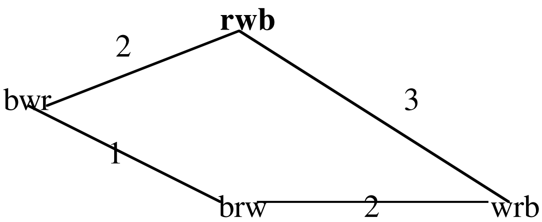
We can describe this episode in ever subtler assertions, but mistakes threaten. Pictures do better: 1's answer *eliminates* all worlds whose first position is 'b':



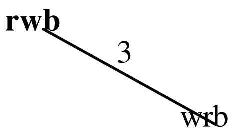
The resulting diagram is read just like the preceding one. In particular, at **rwb** , we see the truth of all the above assertions at once. We can also make predictions. E.g., only one player can make the real distribution of cards *common knowledge* by mentioning his own card, viz. player 2. The resulting information state is

rwb

In this scenario, questions themselves had no informative value about the questioner's information – as with rhetorical questions by teachers to their students. But the model can also deal with other scenarios, with linguistically more common types of question. Suppose now that 2's question conveys *ignorance of the answer*. Then the question *itself* produces an update to:



after which just 1's answer would update to the much more informative diagram

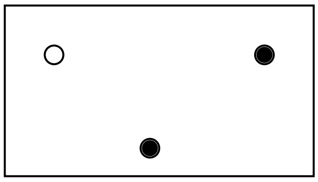


The model can also be used to *plan assertions*. The reader may try the following puzzle, reputedly from an entrance exam to Moscow University. A, B get 3 cards each, and C gets a seventh. Now, A and B must communicate in the presence of C so that they themselves find out the distribution of the cards, while C does not.

7 A classical puzzle: 'Muddy Children'

Here is an influential puzzle with a long pedigree in the literature:

> "After playing, two of three children have got mud on their foreheads. They all see the others, but not themselves, so they do not know their own status. Father comes along and says: "At least one of you is dirty". He then asks: "does anyone know if he is dirty?" The children answer truthfully. After the first round, he asks again. What will happen?"



The quick answer is: "*Nothing*: for everybody knew already what the father said." But the correct answer is that *the muddy children will both know in the second round,* since each of them can argue as follows.

> "If I were clean, the one dirty child that I see would have seen clean children around her, and would have known she was dirty in the first round. But she did not. So I am dirty!"

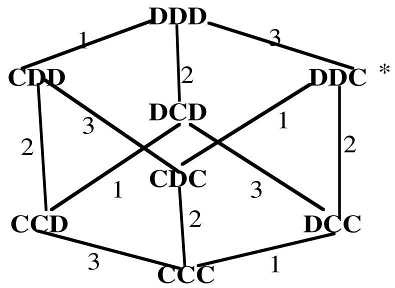
This is symmetric for both muddy children. The third child knows it is clean in the following round. This example highlights some of our earlier themes. The point of the father's announcement is that it turns a fact from universal knowledge into common knowledge – and in the further update process, learning about other people's ignorance rather than their knowledge may provide positive information.

8 Uncertainty models as group information states

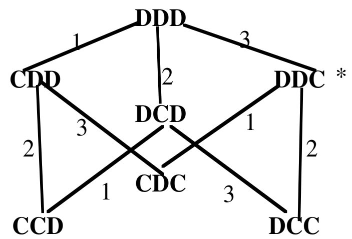
More technically, information models for groups of agents are of the form

$$\mathbf{M} = (S, \{\sim_a \mid a \in A\}, V)$$

Here S is a set of *states of the world*, which make atomic propositions true or false according to the *valuation* V , and each agent $a \in A$ has an *uncertainty relation* $\sim_a$, relating world states s to all states that a cannot distinguish from it. Each $\sim_a$ is an *equivalence relations*: reflexive, symmetric, and transitive. As an example, consider the earlier card examples, or the Muddy Children. A state of the world assigns **D** (dirty) or **C** (clean) to each child: 8 in total. In any of these, a child has one uncertainty. It is fully informed about the others' faces, but it cannot distinguish the state from one where its own **D/C** value is different:



The updates in the above story may be computed by elimination as before, starting with Father's initial announcement eliminating the world **CCC** :



Here intuitively, things have changed at the lower points. E.g., in state **CCD** , we want to say that 3 now knows her status, whereas she did not in the initial state. Next, we introduce a formal notation making this more precise and perspicuous.

9 A language of facts, knowledge and ignorance

Here is a precise notation that talks about knowledge explicitly:

$K_a \phi$	agent a *knows that* ϕ

With such a symbolism, we can also analyse further patterns of statement:

$\neg K_a \neg\phi$	agent a considers it *possible that* ϕ (also: $\langle a\rangle\phi$)
$K_a \phi \vee K_a \neg\phi$	agent a knows *whether* ϕ

Gradually, this formal notation wins out over natural language in terms of clarity. We can also iterate operators in a perspicuous manner, as in

$K_a \neg K_b \phi$	a knows that b does not know that ϕ

For an illustration, consider once again the basic question-answer episode:

" **Q** : ϕ? **A**: Yes. "

By asking the question, **Q** conveys he does not know if ϕ :

$$\neg K_{\mathbf{Q}}\phi \ \& \ \neg K_{\mathbf{Q}}\neg\phi$$

and also that he expects **A** to know: that is, at least

$$\langle \mathbf{Q} \rangle (K_{\mathbf{A}}\phi \vee K_{\mathbf{A}}\neg\phi)$$

By answering affirmatively, **A** conveys she knows that ϕ, but also makes **Q** know this fact ($K_{\mathbf{Q}} K_{\mathbf{A}} \phi$) – leading to common knowledge, written as follows:

$$C_{\{\mathbf{Q},\ \mathbf{A}\}}\ \phi$$

10 Knowledge as truth in all open alternatives

The key idea in reading the earlier diagrams was this. An agent a *knows* those propositions that are *true in all worlds she cannot distinguish from the current one.* Thus, the knowledge operator of our language means the following:

$\mathbf{M}$ makes $K_a\,\phi$ true at s iff ϕ is true in $\mathbf{M}$ *at all* t *with* $s \sim_a t$

This describes information models in terms of what agents know at each world-state in them. The related notation $\neg K_a\neg\phi$, for which natural language has no single verb, works out to the related condition

ϕ is true in $\mathbf{M}$ *for some* t *with* $s \sim_a t$

This says “a finds it conceivable that ϕ". Finally, the common knowledge operator

$$C_{\mathbf{G}}\ \phi$$

at some state s says assertion ϕ is true in every state of the model reachable from s through some finite path of uncertainly links for any members of the group G . Repeating these clause as needed, one computes for any epistemic formula in our notation if it holds in any state of a given model. This determines the whole 'zone' where ϕ holds, identifying the worlds to be eliminated in an update step.

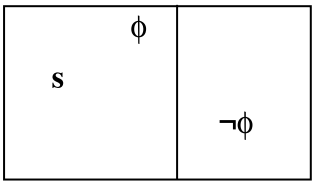
11 Public update defined

A public announcement of a proposition ϕ changes the current group information state **M**, **s** (with an 'actual world' **s**) via the following simple instruction:

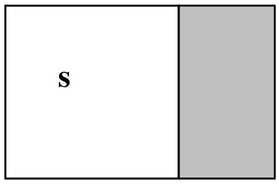
eliminate all worlds which currently do not satisfy ϕ

Thus we go

from



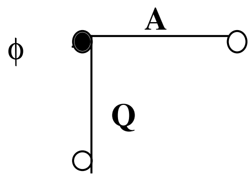
to



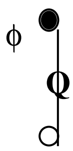
The idea is that the whole group understands that the eliminated worlds have been ruled out as candidates for the actual situation, whence this update is uniform, with the same repercussions for everyone. This procedure is extremely simple, and it will not work for more delicate actions of communication. Still, it is the basic case upon which to build – and it has its surprising subtleties.

12 Fine-structure of communication

Simple diagrams model basic communicative situations. Our basic question-answer example had two states. But already a 3-state model can be a much more complex setting, where questions themselves may provide hard information:



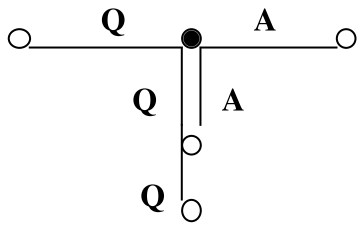
In the black actual state, **Q** thinks **A** might know the answer, as this is so in the bottom-most state. But in the right-most state **Q** *knows* that **A** does not know, and so would not ask there. The question therefore tells **A** he is in the black state:



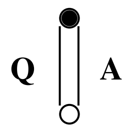
But then, **A** knows, and can update **Q** to the right state as before! Thus, agents can pool their uncertainties, and by simple logical analysis get hard knowledge: a miracle which usually costs large sums in psychotherapy. The more general issue is this. Agents in an information state **M** , **s** can keep telling each other things they know, thereby cutting down the model:

> *What is the maximal correct information they can give each other via successive updates – and what does the resulting 'communicative core model' look like?*

E.g., what is the best to be achieved in the following model?



Geometrical intuition suggests the following 'communicative core':



This is correct. It can be proved that the ‘communicative core' of a model consists of all states tied to the real one by every uncertainty link. These states encode what has been called the ‘implicit knowledge’ of a group of agents. A more detailed analysis shows *two agents need only 2 rounds of communication to convert implicit knowledge into common knowledge.* The reader may try to find two succinct assertions that cut down the preceding model to its communicative core. This is just one of many interesting logical properties of our information models.

13 Valid laws of reasoning with knowledge: the statics

As we said earlier, agents do not just update: they also reason about their situation. Appropriately, our information models validate a ‘logic’ that can describe and automate reasoning with knowledge and ignorance. Its assertions describe valid principles about what agents know and do not know at ambient states. These work for both for outside observers describing group information and for the agents inside the group themselves Here are some major laws of epistemic reasoning – true everywhere in our models – displayed here not for their mathematical aspects, but rather for their interesting intuitive meaning:

$K_a(\phi \to \psi) \to (K_a\phi \to K_a\psi)$	*Knowledge Distribution*
$K_a\phi \to \phi$	*Veridicality*
$K_a\phi \to K_aK_a\phi$	*Positive Introspection*
$\neg K_a\phi \to K_a\neg K_a\phi$	*Negative Introspection*

In particular, rational agents know when they know, but also know about their ignorance. That these axioms are always true is easy to see. They also suffice for *deriving all valid principles* in our language – but this is a difficult 'completeness theorem' in abstract logic. Moreover, validity in the logic is *decidable*, and hence mechanizable. Here are the key additional axioms for *common knowledge*, involving the 'universal knowledge' operator $E_G\ \phi$, defined as the conjunction of all assertions $K_a\ \phi$ where the index 'a' runs over all agents in the group:

$C_G\ \phi \leftrightarrow \phi\ \&\ E_G\ C_G\ \phi$ *Equilibrium Axiom*

$\phi\ \&\ C_G\ (\phi \rightarrow E_G\ \phi) \rightarrow C_G\ \phi$ *Induction Axiom*

These two more abstract principles reflect the 'cognitive equilibrium' achieved by a group of agents enjoying common knowledge of some proposition.

14 Dynamic logic of information update

The preceding calculus is of a widely used logical sort. But it lacks an essential element. Its statements describe what is true at given information states, statically. But they do not describe the update actions that transform one information model into another. Statics and dynamics occur intertwined, witness this crucial issue. Recall the putative specification in Section 1.

Does public announcements of a statement ϕ *always*
make ϕ *common knowledge in the group?*

This seems plausible. Is not the purpose of public communication to make all group members learn that ϕ? Indeed, the principle often holds, as in our Coliseum story. But it may fail in general! A good example is the earlier Muddy Children puzzle. The muddy childrens' last public statement of ignorance led precisely to their knowing their status. Thus, the update reversed the very statement it was reporting. The solution to this 'paradox' is this. A public announcement reports how things were *before* it occurred. But that very action may change those things. A solution requires an idea from computer science. We have processes changing information states, plus assertions made at various 'snapshots' of these processes. Keeping track of this, without confusing truth values changing en route, requires a logical calculus of a well-known sort combining static assertions about states with dynamic expressions for communicative actions:

[A!] ϕ saying that "after a truthful public update with A , ϕ holds"

For some types of statement A, we will indeed have production of common knowledge $[A!]C_G A$ – but not for all. And a logical calculus of this sort will derive what is what, as well as many other assertions about effects of untruthful updates.

There are complete and decidable logical calculi for this richer language, too. Here are some typical valid axioms, together with their intuitive interpretation:

$$[A!]p \quad \leftrightarrow \quad p$$

basic facts p are not affected by updates

$$[A!]\neg\phi \quad \leftrightarrow \quad \neg[A!]\phi$$

updates are deterministic: only one new state results

$$[A!]\phi\vee\psi \quad \leftrightarrow \quad [A!]\phi \vee [A!]\psi$$

another effect of determinism

$$[A!]K_a\phi \quad \leftrightarrow \quad K_a(A \to [A!]\phi)$$

update effects can be precomputed from existing knowledge

$$[A!]\, C_G\phi \quad \leftrightarrow \quad C_G(A, [A!]\phi)$$

a similar principle for common knowledge

General 'dynamic logics' include further appealing principles, such as

$([A!](\phi\to\psi) \;\&\; [A!]\phi) \;\to [A!]\psi$ effects may be combined

$([A!]\,;\,[B!]]\,!]\phi \;\leftrightarrow\; [A!][B!]\phi$ successive actions decompose

Systems like this are currently developed for many types of communication, using computer science ideas on dynamic logics for program and process behaviour.

15 More complex types of communication

Communication is not always public, and not all participants will get the same information from an episode. Moving from public to more private information flow leads to a whole range of phenomena, not all of them understood so far. Examples abound in *games*, which stipulate different rights and duties for players. You may ask a public question, and I may answer to you – but others may not be allowed to see the content of my answer. Or, I may not be paying attention, and some relevant information may or may not have been passed. At the extreme end are current cryptographic protocols, where agents hide their communication – leading to the thriving practical area of *security*. At some stage, there is also a moral boundary: logics must also describe *lying* and *cheating*: updates are not always truthful.

16 Screened actions and semi-private updates

Here is small episode that might occur in a game:

> *We have both just drawn a closed envelope. It is common knowledge between us that one envelope holds an invitation to a lecture on logic, the other to a wild night out in Amsterdam. Clearly, we are both ignorant of the fate in store for us. Now I open my envelope, and read the contents, without showing them to you. Yours remains closed. Which information has passed exactly because of my action? I certainly know now which fate is in store for me. But you have also learnt something, viz. that I know – though not what I know. Likewise, I did not just learn what is in my envelope. I also learnt something about you, viz. that you know that I know. The latter fact has even become common knowledge between us. And all these 'cognitive overtones' can be relevant to our further interactions. What is a general principle behind this?*

This calls for updates that *eliminate links* instead of worlds. The initial information state is one of collective ignorance, while the update removes my uncertainty link:

In the resulting state, all observations from our little card story come out right. More generally, the new update procedure is this extension of that in Section 11.

> When a public question of yours is being answered, without everyone observing the precise answer, the public update effect is the cutting of all your uncertainty links between zones corresponding to different answers.

'Knowledge games' like this are a good methodology for studying different types of communication. More general communicative update procedures may even have to make information models *bigger* by adding structure – as happens when you are sure if a certain statement was announced. General calculi for these updates, in games and other settings, have been proposed recently by a line of Ph.D. students and postdocs in informatics, including the logical challenges of lying and cheating.

17 The broader setting

The broader challenge here is a general classification of types of communication, their information models and reasoning systems, giving us a systematic theory of language use. Moreover, we want to know the different computational complexities showing the costs of conveying types of meaningful information, from public to more private. Such a theory would also lead to a better understanding of the borderlines between publicly- and privately-induced ignorance, and between truthful and misleading communication. Moreover, it could be used in the design of optimal information-passing mechanisms, perhaps in the form of games. Some people use the term 'social software': the formal study and rational design of fair procedures for communication and action with tools from computer science, making the computer scientists 'philosophers in action'. These are mainly fundamental questions. But there are also practical angles to this type of research.

18 Hidden manoevres and security

Consider modern information technology. Communication on the Internet, in itself a public channel, involves secrecy. Which mechanisms guarantee the required privacy? In *cryptographic protocols*, agents communicating over a public broadcast system wish to encode messages that achieve common knowledge just among them without others learning anything relevant. The well-known method of 'public key cryptography' can be analysed by uncertainty models. (In ordinary conversation, we have other ways of screening information, or course.) But we still lack the means to fully define what it means for even widely used Internet protocols to be *safe* against hostile prying, let alone prove it. Just to exhibit the challenge, here is a widely used protocol whose safety w.r.t hostile attack is still in some doubt.

Needham-Schröder Public-Key Protocol Agents A and B wish to exchange private keys for a session of communication which others should not be able to penetrate.

(1) A sends B message A (name) + $\#_A$ (session key), encrypted with B's public key: $ENCR_B(A, \#_A)$.
(2) B acknowledges by sending back $ENCR_A(\#_A, \#_B)$, where $\#_B$ is a new number only known to him so far.
(3) A acknowledges that by sending $ENCR_B(\#_B)$.

19 Updating other aspects of information

Evidently, actual use of natural language does not just modify straight information, but also other things that are crucial to how we communicate. There are also update systems for this, involving more delicate notions of 'state'. Examples include:

‘issues’ in questioning, or general discourse, ‘perspective’ in temporal reasoning, ‘preferences’ in default reasoning, probability distributions in decision making.

20 Unreliable communication

Here is another realistic feature, and challenge to the above model. If we cannot be sure that a public broadcast reaches everybody, it may be impossible to achieve common knowledge at all! The following popular story dramatizes this:

> *Two generals want to co-ordinate an attack in the morning against an enemy whom they can only defeat in conjunction. Their camps are on different hilltops. The area in between is infested by enemy soldiers. General 1 sends a messenger proposing a joint attack at dawn. As he cannot be sure that the message arrived, he asks for confirmation. General 2 receives the message, and sends the confirmation. But he cannot be sure either that the confirmation arrived, so he asks for a confirmation, etc. Even if all messengers arrive safely, no coordination and common knowledge will be achieved!*

A problem of this sort occurs in Arrian's *Life of Alexander the Great*. When campaigning in present-day Pakistan, the Macedonian army was facing an Indian force at the 'Rock of Aornus' in the Punjaab. Alexander was down-hill with the bulk of his army, his lieutenant Ptolemy – the later king of Egypt – with a smaller force on a rise above the Indian camp. They needed to co-ordinate their attack, without attracting the attention of their enemies. Only one messenger was sent to achieve this – but that was an Indian deserter (Arrian tells us) who had every reason not to fall into the hands of his countrymen... In real life, we trust that 'enough' mutual knowledge has been achieved if the communication channel has *sufficiently high reliability*. So, our logical model needs additional probabilistic features explaining how communication often gets away with less than perfect information.

21 Conclusion

This paper has addressed some basic types of communication, in a light manner. We have shown how the relevant information flow can be modeled, and how agents' reasoning can be described formally, using simple diagrams that a layman can use. In all this, we do not claim great originality. Most of our tools have long existed in different fields – but they come together in new ways. The resulting notions are very teachable, and might fit into an informatics core curriculum even at high school level. We see this topic as a typical instance of 'informatics', the more general science of information structure and information flow, based on common concerns between the various disciplines involved in our story, including linguistics, logic, philosophy, computer science, AI, economics, and mathematics. There is also an experimental angle. How do humans really represent information and perform updates in communication – given their processing limitations?

22 References

The following publications contain many references to the work of older and newer relevant authors in the area. (a) J. van Benthem & A. ter Meulen, eds., 1997, *Handbook of Logic and Language*, Elsevier, Amsterdam. (b) J. van Benthem, 2001, 'Update Delights', downloadable from http://turing.wins.uva.nl/~johan/ – or its tutorial version in Proceedings TARK VIII, Siena 2001, Morgan Kaufmann, San Francisco. Good general links to organisations in the world of informatics are http://www.folli.org/, http://www.tark.org/, http://www.ifcolog.com.