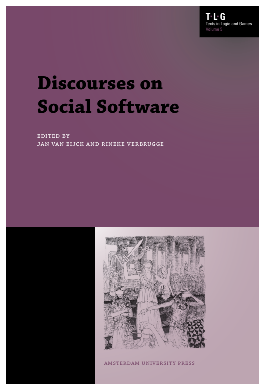
Reasoning About Communication

Jan van Eijck

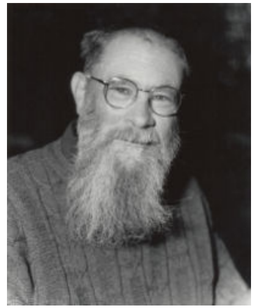
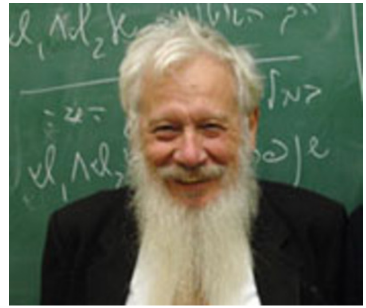
CWI, 26 November 2010



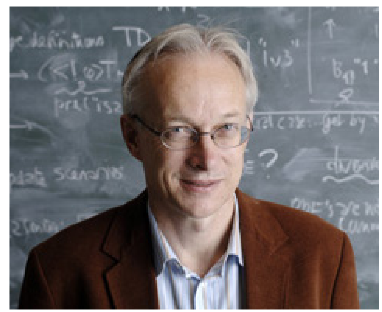
Abstract

The communicative effect of a collective message from the Dutch former minister of finance Wouter Bos to inform all his contacts about his new email address is completely different from that of a set of individual messages to the same list. The talk will explain how differences of this kind can be modelled in epistemic logic (the logic of knowledge). A central notion here is common knowledge. We will explain the general framework for describing update effects of messages as mappings on epistemic models ("knowledge models"), and we will give a sketch of some recent work in this area.

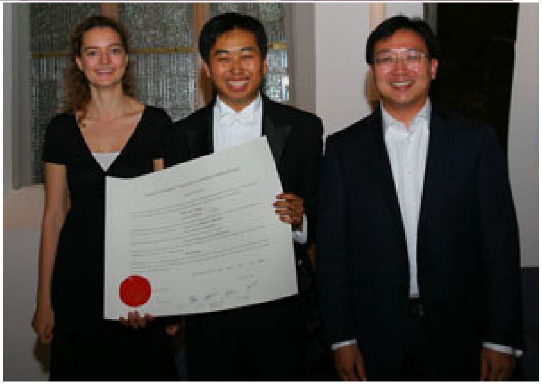
Very Brief History





At CWI



drawing by
Marco Swaen

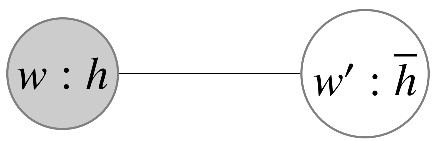
The Muddy Children Puzzle

a (1) clean, b (2), c (3) and d (4) muddy.

a	b	c	d
○	●	●	●
?	?	?	?
?	?	?	?
?	!	!	!
!	!	!	!

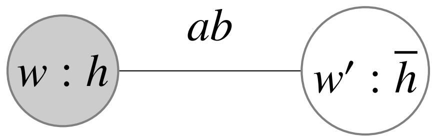
Individual Ignorance

You have to finish a paper, and you are faced with a choice: do it today, or put it off until tomorrow. You decide to play a little game with yourself. You will flip a coin, and you promise yourself: “If it lands heads I will have to do it now, if it lands tails I can postpone it until tomorrow. But wait, I don’t have to know right away, do I? I will toss the coin in a cup.” And this is what you do. Now the cup is upside down on the table, covering the coin. The coin is showing heads up, but you cannot see that.



Multi Agent Ignorance

Suppose Alice and Bob are present, and Alice tosses a coin under a cup. We will use a for Alice and b for Bob. The result of a hidden coin toss with the coin heads up:

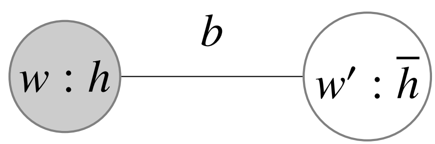


After Alice has taken a look

Assume that Alice is taking a look under the cup, while Bob is watching.

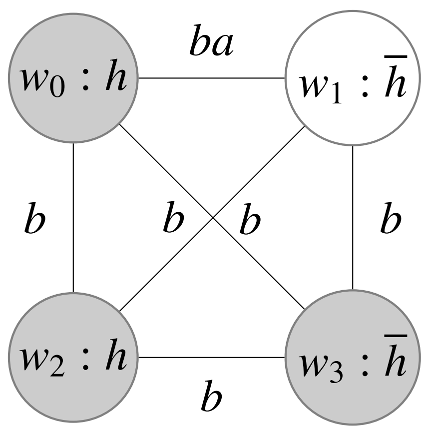
Now Alice knows whether the coin shows heads or tails.

Bob knows that Alice knows the outcome of the toss, but he does not know the outcome himself.



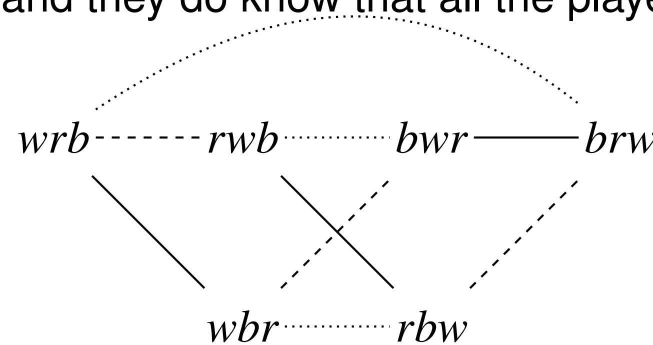
Bob leaves the room and returns

The coin gets tossed under the cup and lands heads up. Alice and Bob are present. Now Bob leaves the room for an instant. After he comes back, the cup is still over the coin. Bob realizes that Alice might have taken a look. She might even have reversed the coin! Alice also realizes that Bob considers this possible.

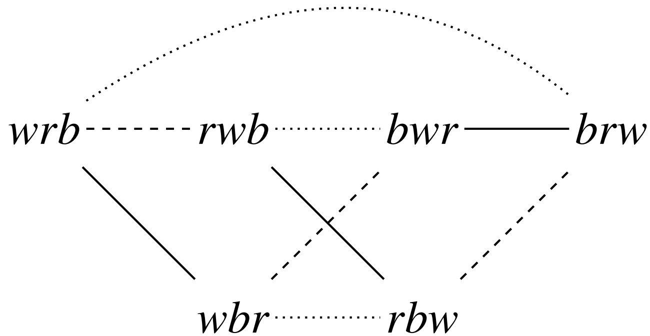


Epistemic Situations: Card Deals

Three players, Alice, Bob and Carol, each draw a card from a stack of three cards. The cards are red, white and blue and their backsides are indistinguishable. Let rwb stand for the deal of cards where Alice holds the red card, Bob the white card, and Carol the blue card. There are six different card deals. Players can only see their own card, but not the cards of other players. They do see that other players also only hold a single card, and that this cannot be their own card, and they do know that all the players know that.



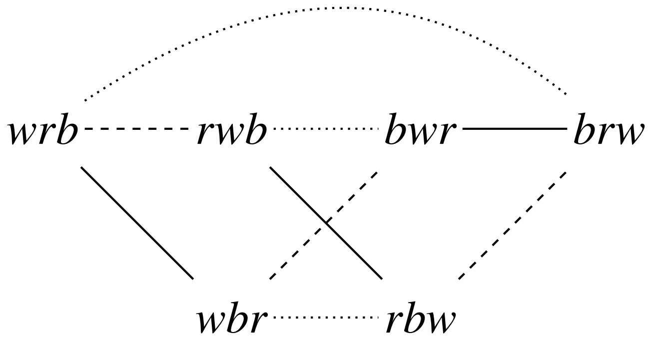
Communication



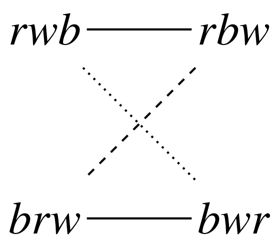
Alice says: 'I hold the red card'. What is the effect of this?

$$rwb \text{ ——— } rbw$$

Communication 2



Alice says: ’I do not hold the white card’. What is the effect of this?



Epistemic Model Checking: Muddy Children

- initMuddy: model where children cannot see their own state
- m1: model after the public announcement that at least one child is muddy.
- m2: model after public announcement that none of them knows their state.
- m3: model after public announcement that none of them knows their state.
- m4: model after public announcement that b, c, d know their state.

Common Knowledge (See Lewis [9])

ϕ is common knowledge if everyone knows that ϕ and, moreover, everyone knows that ϕ is common knowledge.

$$C\varphi \leftrightarrow (E\varphi \wedge EC\varphi).$$

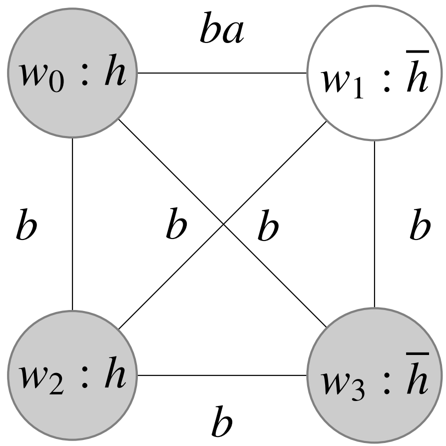
Compare:

zeros = 0 : zeros

If a denotes Alice's accessibility relation and b Bob's, and we use $\cup$ for union of relations, and * for reflexive transitive closure, then $(a \cup b)^*$ expresses the common knowledge of Alice and Bob.

Cashiers, ATMs, and the Creation of Common Knowledge

Sample Question about Common Knowledge

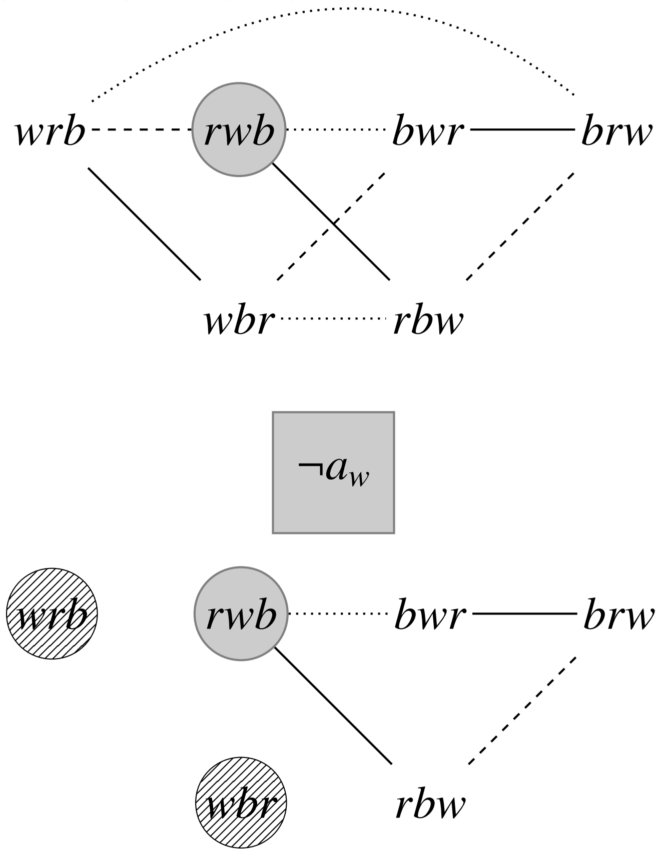


Is it common knowledge between a and b in w_0 that Alice knows that Bob does not know the outcome of the toss?

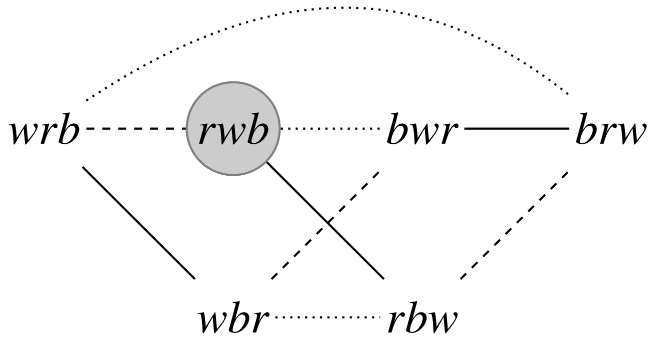
Formula for this: $C_{a,b}K_a\neg(K_bh \vee K_b\neg h)$.

Expressed in epistemic PDL: $[(a \cup b)^*; a]\neg([b]h \vee [b]\neg h)$.

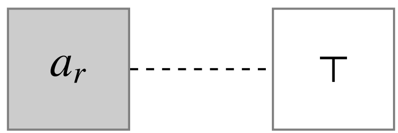
Effect of Public Announcement



Private Message



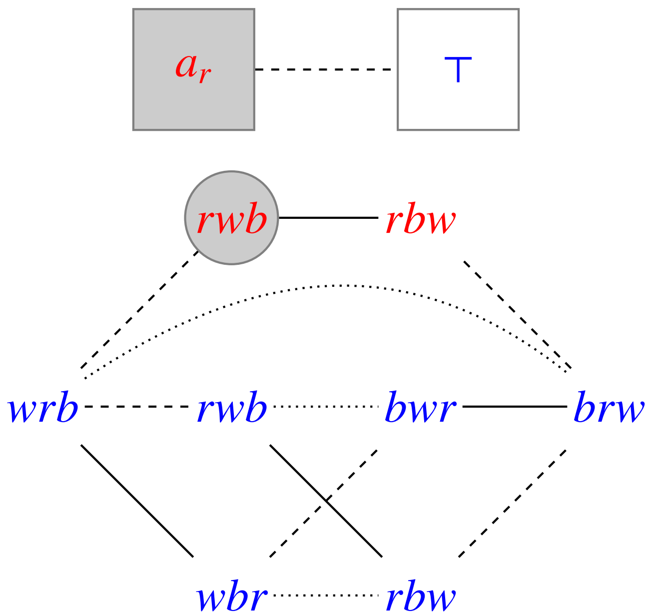
Alice says “I hold the red card” privately to Bob.



Carol cannot distinguish this from the action where nothing happens.

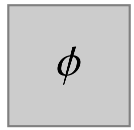
Effect of This

Compute the result with a model product construction (Baltag cs., [3]):

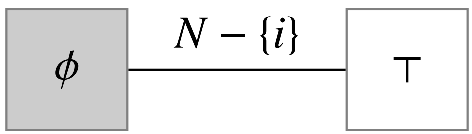


Sending Email Messages

"Wouter Bos email": message where all can see the recipient list. This is like a public announcement.



Private message ϕ to agent i: all other agents cannot distinguish this from the action where nothing happens:

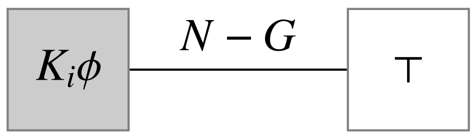


Recent Work at CWI

Apt, Witzel, Zvesper [2], Sietsma and Apt [1], Sietsma and Van Eijck [6]. Sietsma and Van Eijck [7].

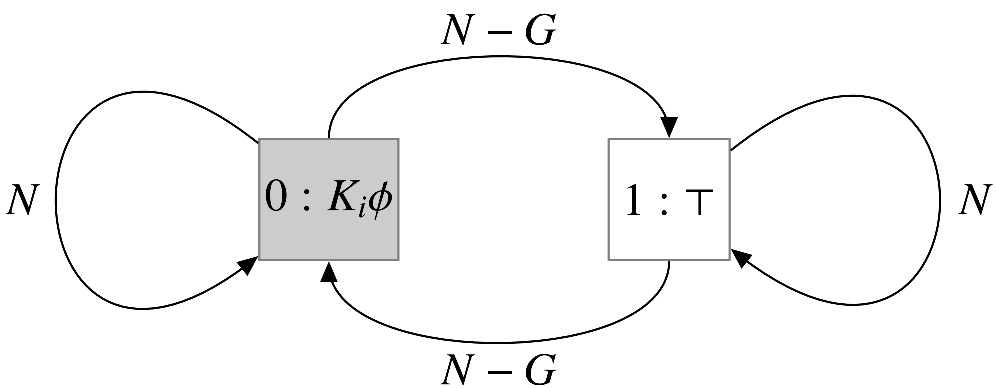
Assumption: messages are always sent by agents who know the message is true.

General action model of the action of sending a message from i to group of agents G, with contents ϕ. Assume $i \in G$, i.e., agent i sends cc to herself:



Axiomatisation

Conside the action model for sending a message as a finite automaton:



Define functions $T^m_{00}, T^m_{01}, T^m_{11}, T^m_{10}$ on the set of regular epistemic expressions, where m indicates the message (i, ϕ, G), and where

$$T^m_{00}(\pi)$$

gives the result of “moving through the automaton in synch with the expression π”, from state 0 to state 0.

$$\begin{array}{rcl}
T^m_{00}(a) & := & ?K_i\phi; a \\
T^m_{01}(a) & := & \begin{cases} ?\bot & \text{if } a \in G \\ ?K_i\phi; a & \text{otherwise} \end{cases} \\
T^m_{11}(a) & := & a \\
T^m_{10}(a) & := & \begin{cases} ?\bot & \text{if } a \in G \\ a & \text{otherwise} \end{cases}
\end{array}$$

Then the following reduction axioms define the effects of message passing:

$$
\begin{array}{rcl}
[m]\phi & \leftrightarrow & [m,0]\phi \\
[m,0][\pi]\phi & \leftrightarrow & [T^m_{00}(\pi)][m,0]\phi \wedge [T^m_{01}(\pi)][m,1]\phi \\
[m,1][\pi]\phi & \leftrightarrow & [T^m_{11}(\pi)][m,1]\phi \wedge [T^m_{10}(\pi)][m,0]\phi
\end{array}
$$

The general technique is from [4].

Example Axiom: Common Knowledge after Message Passing

Suppose there are three agents a, b, c, and m is a message with contents ϕ from a to b, with cc to a.

$$m : \qquad \boxed{0 : K_a\phi} \overset{c}{\rule[0.5ex]{6em}{0.4pt}} \boxed{1 : \top}$$

Calculation of what happens to the common knowledge of b and c:

$$\begin{aligned}
& [m][(b \cup c)^*]\psi \\
\leftrightarrow\ & [m, 0][(b \cup c)^*]\psi \\
\leftrightarrow\ & T^m_{00}((b \cup c)^*)[m, 0]\psi \wedge T^m_{01}((b \cup c)^*)[m, 1]\psi \\
\leftrightarrow\ & [(?K_a\phi; b \cup c)^*; (?K_a\phi; c; (b \cup c)^*; c; (?K_a\phi; b \cup c)^*)^*][m, 0]\psi \\
& \ \wedge[(?K_a\phi; b \cup c)^*; ?K_a\phi; c; (b \cup c)^*; \\
& \quad (c; (?K_a\phi; b \cup c)^*; ?K_a\phi; c; (b \cup c)^*)^*][m, 1]\psi.
\end{aligned}$$

A Riddle and A Protocol

A group of 100 prisoners,

all together in the prison dining area, are told that they will be all put in isolation cells and then will be interrogated one by one in a room containing a light with an on/off switch. The prisoners may communicate with one another by toggling the light-switch (and in no other way). The light is initially switched off. There is no fixed order of interrogation. Every day one prisoner will get interrogated. At any stage every prisoner will be interrogated again sometime.

When interrogated, a prisoner can either do nothing, or toggle the light-switch, or announce that all prisoners have been interrogated. If that announcement is true, the prisoners will (all) be set free, but if it is false, they will all be executed. Can the prisoners agree on a protocol that will set them free?

A Protocol for Solving the Riddle

The set of prisoners is $\{0, \ldots, n-1\}$, with $n \geq 2$.

The prisoners appoint one among them as the counter. We will assume prisoner 0 is appointed as counter.

All prisoners except the counter act as follows: the first time they enter the room when the light is off, they switch it on; on all next occasions, they do nothing.

The counter acts as follows: The first $n-2$ times that the light is on when he enters the interrogation room, he turns it off. Then the next time he enters the room when the light is on, he announces that everybody has been interrogated.

This protocol is proved correct in [5].

Further Work

- Study the connection between Sietsma and Apt [1] and Sietsma and Van Eijck [6]. Two different ways of modelling bcc's: are they equivalent or not?
- Next step: add networks plus assumptions about knowledge of networks.
- Extend existing epistemic model checking tools to handle the effects of message passing over networks.
- Connect up with work in network analysis in economics and social science [8].

References

[1] K.R. Apt and F. Sietsma. On email exchanges. CWI manuscript, November 2010.

[2] K.R. Apt, A. Witzel, and J.A. Zvesper. Common knowledge in interaction structures. In Proceedings of the 12th Conference on Theoretical Aspects of Rationality and Knowledge (TARK XII), pages 4–13, 2009.

[3] A. Baltag, L.S. Moss, and S. Solecki. The logic of public announcements, common knowledge, and private suspicions. In I. Bilboa, editor, Proceedings of TARK'98, pages 43–56, 1998.

[4] J. van Benthem, J. van Eijck, and B. Kooi. Logics of communication and change. Information and Computation, 204(11):1620–1662, 2006.

[5] Hans van Ditmarsch, Jan van Eijck, and William Wu. Verifying one hundred prisoners and a lightbulb. Journal of Applied Non-classical Logics, 2010 (to appear).

[6] Jan van Eijck and Floor Sietsma. Message generated Kripke semantics. CWI manuscript, October 2010.

[7] Jan van Eijck and Floor Sietsma. A proof system for message passing. CWI manuscript, November 2010.

[8] M. Jackson. Social and Economic Networks. Princeton University Press, 2008.

[9] D.K. Lewis. Convention: A Philosophical Study. Harvard University Press, 1969.