Forgetting propositional formulas

David Fernández-Duque, Ángel Nepomuceno-Fernández, Enrique Sarrión-Morrillo,
Fernando Soler-Toscano, Fernando R. Velázquez-Quesada

Workshop on Logics for Resource-Bounded Agents
ESSLLI 2015

Epistemic attitudes and their dynamics

Epistemic attitudes are subject to the effect of different ***epistemic actions***.

For example, while ***beliefs*** can be affected by

- ***expansion*** (e.g., Rott 1989),
- ***contraction*** (e.g., Alchourrón et al. 1985),
- ***revision*** (e.g., Alchourrón et al. 1985, Rott 1989, Boutilier 1996, Leitgeb and Segerberg 2007, van Benthem 2007, Baltag and Smets 2008),
- ***merging*** (e.g., Konieczny and Pérez 2011) and
- diverse forms of ***inference*** (e.g., VQ 2014, VQ et al. 2013),

knowledge can be affected by

- ***deductive inference*** (VQ 2009, 2013),
- ***public*** (Plaza 1989, Gerbrandy and Groeneveld 1997) and ***other*** forms of ***announcements*** (Baltag et al. 1999).

Forgetting

An action that has not received much attention is that of ***forgetting*** and its effect on an agent's ***knowledge***.

A possible reason: it is in some sense similar to *belief contraction*.

But still, when *belief contraction* is represented semantically, it typically relies on an (plausibility) ordering among theories.

This work proposes a ***dynamic epistemic logic*** (van Ditmarsch et al. 2007, van Benthem 2011) representation for an action of ***forgetting***. (Source: Fernández-Duque et al. (2015).)

Some remarks

- Here, "*forgetting* π" is understood as "*now I do not know* π" (and not as "*now I am unaware of* π).
- This work focusses on ***forgetting whether*** ("***now I do not know whether*** π").
- This work uses ***relational models*** and represents the action with a ***model operation***.
- ***Related work***: forgetting *atoms* (van Ditmarsch et al. 2009), forgetting *set of atoms* (Lin and Reiter 1994, Zhang and Zhou 2009).

Semantic model and language

Definition (Relational model)

*A **relational model** M based on P is a tuple $\langle W, R, V\rangle$ where*

- $W \neq \emptyset$ *is a set of **possible worlds**;*
- $R \subseteq (W \times W)$ *is the agent's **indistinguishability relation**;*
- $V : \mathsf{P} \to \wp(W)$ *is an **atomic valuation**.*

*The pair (M, w) with $w \in W$ is a **possible worlds state** and w is the **evaluation point**.*

Definition (Language $\mathcal{L}_{\{\Box\}}$)

Formulas φ, ψ of the language $\mathcal{L}_{\{\Box\}}$ based on P are given by

$$\varphi, \psi ::= \top \mid p \mid \neg\varphi \mid \varphi \land \psi \mid \Box\varphi$$

with $p \in \mathsf{P}$. Other propositional constants ($\bot$), other propositional connectives ($\lor, \to, \leftrightarrow$) and the dual modal universal operator $\Diamond$ are defined as usual ($\Diamond\varphi := \neg\Box\neg\varphi$ for the latter).

Semantic interpretation

Definition (Semantic interpretation)

Given (M, w) *with* $M = \langle W, R, V \rangle$, *define* $\Vdash$ *as*

$$
\begin{array}{lll}
(M, w) \Vdash p & \mathit{iff}_{def} & w \in V(p) \\
(M, w) \Vdash \neg \varphi & \mathit{iff}_{def} & (M, w) \nVdash \varphi \\
(M, w) \Vdash \varphi \wedge \psi & \mathit{iff}_{def} & (M, w) \Vdash \varphi \text{ and } (M, w) \Vdash \psi \\
(M, w) \Vdash \Box \varphi & \mathit{iff}_{def} & \text{for all } u \in W, Rwu \text{ implies } (M, u) \Vdash \varphi
\end{array}
$$

Validity ($\Vdash \varphi$) *is defined as usual.*

Some concepts

It will be useful to represent propositional formulas π in conjunctive normal form.

- A ***literal*** l is an atom (p) or its negation ($\neg p$).
- A ***clause*** C is a finite (possibly empty) set of literals interpreted disjunctively ($\widehat{C} := \bigvee C$).
- A propositional formula is in ***conjunctive normal form*** when it is given as a finite (possibly empty) set of clauses $\mathfrak{C}$ interpreted conjunctively ($\widehat{\mathfrak{C}} := \bigwedge_{C \in \mathfrak{C}} \bigvee C$).
- A clause C is ***tautological*** when there is p such that $\{p, \neg p\} \subseteq C$.
- A clause C is a ***consequence*** of π when $\Vdash \pi \to \widehat{C}$.
- A clause C is a ***minimal consequence*** of π when it is a consequence of π and there is no $C' \subset C$ such that $\Vdash \pi \to \widehat{C'}$.

Clausal form

Definition (Clausal form $\mathcal{C}(\pi)$)

Let π be propositional formula.

$$\mathcal{C}(\pi) := \{C \mid C \text{ is a clause which is a minimal non-tautological consequence of } \pi\}$$

Note how, for any π, the set $\mathcal{C}(\pi)$ is finite, its elements are finite, and $\Vdash \pi \leftrightarrow \widehat{\mathcal{C}}(\pi)$.

Some simple examples:

π	$\mathcal{C}(\pi)$	π	$\mathcal{C}(\pi)$
$p \wedge q$	$\{\{p\},\{q\}\}$	$\neg(p \wedge q)$	$\{\{\neg p, \neg q\}\}$
$p \vee q$	$\{\{p, q\}\}$	$\neg(p \vee q)$	$\{\{\neg p\},\{\neg q\}\}$
$p \rightarrow q$	$\{\{\neg p, q\}\}$	$\neg(p \rightarrow q)$	$\{\{p\},\{\neg q\}\}$
$p \leftrightarrow q$	$\{\{\neg p, q\},\{p, \neg q\}\}$	$\neg(p \leftrightarrow q)$	$\{\{p, q\},\{\neg p, \neg q\}\}$

The intuitive idea (1)

The initial observation.

- An agent ***knows*** φ when φ holds in ***all her epistemic alternatives***.
- Thus, in order to ***'forget'*** φ, she needs to consider as possible ***at least one*** world in which φ ***fails***.

First, ***how to falsify*** a propositional formula π ***in a world*** w?

- A given *contingent* propositional π can be falsified in different ways.
- If $\mathcal{C}(\pi) = \{C_1, \ldots, C_n\}$ is used, then there are $2^n - 1$ different forms of falsifying π.
- A simpler 'minimal' approach is ***to falsify only one clause*** in $\mathcal{C}(\pi)$.

The intuitive idea (2)

Second: which will be the valuation for other atoms? Third: how many new worlds should we add?

- For the third: we ***make a copy of*** the current ***epistemic possibilities***, ***falsifying the given clause*** in each one of them,
- For the second: we ***keep atoms not appearing in the clause as before***.

In the resulting model, the original π has been ***uniformly*** falsified.

Two final details.

- This work deals with ***forgetting whether***.
- To accommodate this, the operation works by ***falsifying any finite number of clauses***.

Operation and semantic interpretation

Definition

Let $M = \langle W, \leq, V \rangle$ be a relational model; let $\mathfrak{C} = \{C_i \mid i \in I\}$ be a finite set of non-tautological clauses ($0 \notin I$).

The relational model $M^{\mathfrak{C}} = \langle W', R', V' \rangle$ is given by

- $W' := W \times (\{0\} \cup I)$,
- *for all $w, u \in W$ and $i, j \in (\{0\} \cup I)$,*

$$R'(w,i)(u,j) \quad \text{iff}_{def} \quad Rwu$$

- *for every $p \in \mathbf{P}$, $w \in W$ and $i \in (\{0\} \cup I)$,*

$$(w,0) \in V'(p) \quad \text{iff}_{def} \quad w \in V(p)$$
$$(w,i) \in V'(p) \quad \text{iff}_{def} \quad \{p, \neg p\} \cap C_i = \varnothing \text{ and } w \in V(p), \text{ or } \neg p \in C_i;$$

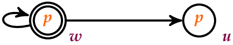
Definition (Semantic interpretation)

$$(M, w) \Vdash [\ddagger\pi]\,\varphi \quad \text{iff}_{def} \quad (M^{\{C_1, C_2\}}, (w,0)) \Vdash \varphi \quad \text{for all } C_1 \in \mathfrak{C}(\pi), C_2 \in \mathfrak{C}(\neg\pi)$$

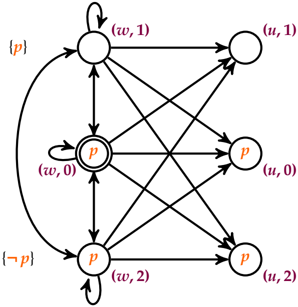
Example 1

Recall: $\mathcal{C}(p) = \{\,\{p\}\,\}$ (so $C_1 = \{p\}$)

$\mathcal{C}(\neg p) = \{\,\{\neg p\}\,\}$ (so $C_2 = \{\neg p\}$)



$\ddagger p$



$(M, w) \Vdash \Box p$

$(M, w) \Vdash [\ddagger p]\,(\neg \Box p \wedge \neg \Box \neg p)$

$(M^{\{\{p\},\{\neg p\}\}}, (w, 0)) \Vdash \neg \Box p \wedge \neg \Box \neg p$

Basic result

Lemma

Let $M = \langle W, \leq, V \rangle$ be a relational model; let $\mathcal{C} = \{C_i \mid i \in I\}$ be a finite (possibly empty) set of clauses ($0 \notin I$).

For any $w \in W$ and any $i \in I$,

$$(M^{\mathcal{C}}, (w, i)) \nVdash \widehat{C_i}$$

Proposition

For any contingent propositional formula π,

$$\Vdash \langle \ddagger\pi \rangle (\Box \neg \pi \lor \Box \pi) \leftrightarrow \Box \bot$$

(i.e., $\mathsf{S} \Vdash [\ddagger\pi] (\neg \Box \pi \land \neg \Box \neg \pi)$)

Tautologies and contradictions

If π is a (propositional) tautology $\top$,

- $\mathcal{C}(\top) = \varnothing$ so,
- by vacuity, $(M^{\{C_1, C_2\}}, (w, \mathbf{0})) \Vdash \varphi$ for all $C_1 \in \mathcal{C}(\top), C_2 \in \mathcal{C}(\neg \top)$.
- Thus, $\Vdash [\ddagger\top]\, \varphi$ (but $\Vdash \neg \langle\ddagger\top\rangle\, \varphi$).

If π is a (propositional) contradiction $\bot$,

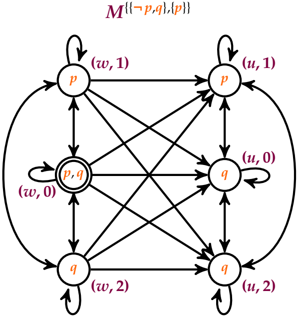
- $\mathcal{C}(\neg \bot) = \varnothing$ so,
- by vacuity, $(M^{\{C_1, C_2\}}, (w, \mathbf{0})) \Vdash \varphi$ for all $C_1 \in \mathcal{C}(\bot), C_2 \in \mathcal{C}(\neg \bot)$.
- Thus, $\Vdash [\ddagger\bot]\, \varphi$ (but $\Vdash \neg \langle\ddagger\bot\rangle\, \varphi$).

Example 2

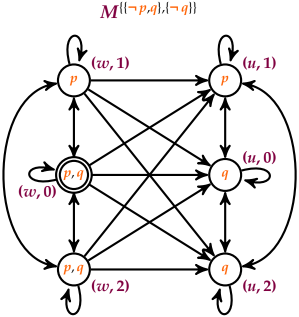
$w \Vdash \Box (p \to q)$

$\mathfrak{C}(p \to q) = \{\{\neg p, q\}\}$ (so $C_1 = \{\neg p, q\}$)

$\mathfrak{C}(\neg(p \to q)) = \{\{p\}, \{\neg q\}\}$ (so $C_2 = \{p\}$ or $C_2 = \{\neg q\}$)



$(w, 0) \Vdash \Diamond (\neg p \wedge q \wedge \Diamond (p \wedge q))$



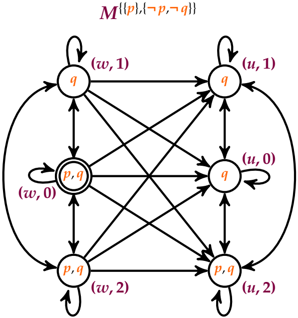
$(w, 0) \nVdash \Diamond (\neg p \wedge q \wedge \Diamond (p \wedge q))$

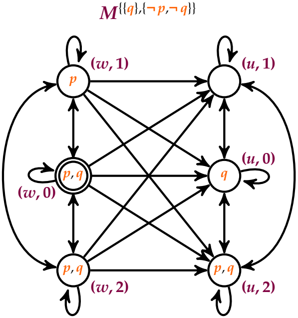
Example 3

$w \Vdash \neg\Box(p \wedge q) \wedge \neg\Box\neg(p \wedge q)$

$\mathcal{C}(p \wedge q) = \{\{p\}, \{q\}\}$ (so $C_1 = \{p\}$ or $C_1 = \{q\}$)

$\mathcal{C}(\neg(p \wedge q)) = \{\{\neg p, \neg q\}\}$ (so $C_2 = \{\neg p, \neg q\}$)





$(w, 0) \Vdash \neg\Box(p \wedge q) \wedge \neg\Box\neg(p \wedge q)$

$(w, 0) \Vdash \Diamond(\neg p \wedge \Diamond p)$

$(w, 0) \Vdash \neg\Box(p \wedge q) \wedge \neg\Box\neg(p \wedge q)$

$(w, 0) \Vdash \Diamond(\neg p \wedge \neg q)$

Semantic interpretation and basic result

A simpler "*forgetting that*" action.

Definition (Semantic interpretation)

$(M, w) \Vdash [\dagger\pi]\,\varphi \qquad \text{iff}_{def} \qquad (M^{\{C\}}, (w, 0)) \Vdash \varphi$ *for all* $C \in \mathcal{C}(\pi)$

Proposition

For any contingent propositional formula π,

$$\Vdash \langle\dagger\pi\rangle \Box\pi \leftrightarrow \Box\bot$$

(i.e., $\mathsf{S} \Vdash [\dagger\pi]\neg\Box\pi$*)*

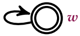
Forgetting whether and forgetting that

Fact

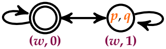
The formula $[\ddagger\pi]\,\varphi \leftrightarrow [\dagger\pi]\,[\dagger\neg\pi]\,\varphi$ *is not valid.*

Proof *Take* $\pi := \neg(p \wedge q)$, *so* $\mathcal{C}(\neg(p \wedge q)) = \{\{\neg p, \neg q\}\}$ *and* $\mathcal{C}(p \wedge q) = \{\{p\}, \{q\}\}$.

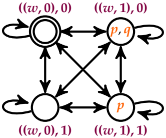
M



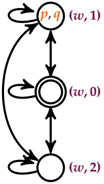
$M^{\{\{\neg p, \neg q\}\}}$



$(M^{\{\{\neg p, \neg q\}\}})^{\{\{q\}\}}$



$M^{\{\{\neg p, \neg q\}, \{p\}\}} = M^{\{\{\neg p, \neg q\}, \{q\}\}}$



$((M^{\{\{\neg p, \neg q\}\}})^{\{\{q\}\}}, ((w,0),0)) \Vdash \Diamond(p \wedge \neg q)$ | $(M^{\{\{\neg p, \neg q\}, \{q\}\}}, (w,0)) \nVdash \Diamond(p \wedge \neg q)$

Tautologies and contradictions

As before, if π is a (propositional) tautology $\top$,

- $\mathcal{C}(\top) = \emptyset$ so,
- by vacuity, $(M^{\{C\}}, (w, \mathbf{0})) \Vdash \varphi$ for all $C \in \mathcal{C}(\top)$.
- Thus, $\Vdash [\dagger\top]\,\varphi$ (but $\Vdash \neg\langle\dagger\top\rangle\,\varphi$).

But now, if π is a (propositional) contradiction $\bot$,

- $\mathcal{C}(\bot) = \{\emptyset\}$ so
- Thus, $\Vdash [\dagger\bot]\,\varphi \leftrightarrow \langle\dagger\bot\rangle\,\varphi$.
- Nevertheless, $(M^{\{\emptyset\}}, (w, \mathbf{0})) \,\underline{\leftrightarrow}\, (M, w)$, so $\Vdash \varphi \leftrightarrow [\dagger\bot]\,\varphi$.

Attempt 1: condition for *where* to evaluate φ

Definition (Semantic interpretation)

$$(M, w) \Vdash [\dagger'\pi]\,\varphi \quad \textit{iff}_{def} \quad \begin{cases} (M^{\{C\}}, (w, 0)) \Vdash \varphi \text{ for all } C \in \mathcal{C}(\pi) & \text{if } (M, w) \Vdash \Box\,\pi \\ (M, w) \Vdash \varphi & \text{otherwise} \end{cases}$$

Note how, from $\langle\dagger'\pi\rangle\,\varphi := \neg\,[\dagger'\pi]\,\neg\,\varphi$,

$$(M, w) \Vdash \langle\dagger'\pi\rangle\,\varphi \quad \text{iff} \quad \begin{cases} (M, w) \Vdash \Box\,\pi \text{ and } (M^{\{C\}}, (w, 0)) \Vdash \varphi \text{ for some } C \in \mathcal{C}(\pi), \text{ or} \\ (M, w) \Vdash \neg\,\Box\,\pi \wedge \varphi \end{cases}$$

Proposition

For any contingent propositional formula π,

$$\Vdash [\dagger'\pi]\,\varphi \;\leftrightarrow\; ((\Box\,\pi \rightarrow [\dagger\pi]\,\varphi) \wedge (\neg\Box\,\pi \rightarrow \varphi))$$

Attempt 1: relation with *public announcement*

Assuming the standard definition for $M_{!\chi}$ and $[!\chi]\,\varphi$,

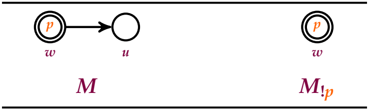
Fact

The formula $\varphi \rightarrow [\dagger'\pi]\,[!\pi]\,\varphi$ *is* ***not*** *valid.*

Proof *Take* $\pi := p$ *and* $\varphi := \Diamond\neg p$. *By previous proposition,*

$$(\Diamond\neg p \rightarrow [\dagger' p]\,[!p]\,\Diamond\neg p) \;\leftrightarrow\; (\Diamond\neg p \rightarrow ((\Box p \rightarrow [\dagger p]\,[!p]\,\Diamond\neg p) \wedge (\neg\Box p \rightarrow [!p]\,\Diamond\neg p)))$$

But consider



$(M, w) \Vdash \Diamond\neg p$ *but also* $(M, w) \Vdash \neg\Box p \wedge \langle !p\rangle\neg\Diamond\neg p$, *i.e.* $(M, w) \nVdash \neg\Box p \rightarrow [!p]\,\Diamond\neg p$. ∎

Attempt 2: condition for *whether* to evaluate φ

Definition (Semantic interpretation)

$(M, w) \Vdash [\dagger'\pi]\,\varphi$ $\quad$ *iff*$_{def}$ $\quad$ $(M, w) \Vdash \Box\,\pi$ *implies* $(M^{\{C\}}, (w, 0)) \Vdash \varphi$ *for all* $C \in \mathcal{C}(\pi)$

Note how, from $\langle\dagger'\pi\rangle\,\varphi := \neg\,[\dagger'\pi]\,\neg\,\varphi$,

$(M, w) \Vdash \langle\dagger'\pi\rangle\,\varphi$ $\quad$ iff $\quad$ $(M, w) \Vdash \Box\,\pi$ and $(M^{\{C\}}, (w, 0)) \Vdash \varphi$ for some $C \in \mathcal{C}(\pi)$

Proposition

For any contingent propositional formula π,

$$\Vdash [\dagger'\pi]\,\varphi \;\leftrightarrow\; (\Box\,\pi \rightarrow [\dagger\pi]\,\varphi)$$

Attempt 2: relation with *public announcement*

Assuming the standard definition for $M_{!\chi}$ and $[!\chi]\,\varphi$,

Proposition

$$\mathsf{T} \Vdash \varphi \rightarrow [\dagger'\pi]\,[!\pi]\,\varphi$$

Semantic interpretation and a property

Definition (Semantic interpretation)

$$(M, w) \Vdash [\dagger^{\bullet}\pi]\,\varphi \quad \textit{iff}_{def} \quad (M^{\mathcal{C}(\pi)}, (w, 0)) \Vdash \varphi$$

Note how, from $\langle\dagger^{\bullet}\pi\rangle\,\varphi := \neg\,[\dagger^{\bullet}\pi]\,\neg\,\varphi$,

$$(M, w) \Vdash \langle\dagger^{\bullet}\pi\rangle\,\varphi \quad \text{iff} \quad (M^{\mathcal{C}(\pi)}, (w, 0)) \Vdash \varphi$$

Fact

The formula $[\dagger(p \wedge q)]\,(\neg\Box\,p \wedge \neg\Box\,q)$ *is not valid.*

Proposition

$$\Vdash [\dagger^{\bullet}(p \wedge q)]\,(\neg\Box\,p \wedge \neg\Box\,q)$$

Up to now . . .

- A model operation representing ***forgetting whether*** for ***propositional*** formulas.
- ***Minimal conjunctive normal form*** is used.
- Some ***variations*** explored.

. . . and yet to do/finish

- A model operation representing the ***forgetting*** of ***modal*** formulas.
- ***Derivation system*** still missing for some variations.
- ***Multiagent versions***, as, e.g., public and private *individual* forgetting, or *collective* forgetting.
- Proper ***comparison*** of proposal with related approaches (e.g., belief contraction).

C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50(2):510–530, 1985. doi: 10.2307/2274239.

A. Baltag and S. Smets. A qualitative theory of dynamic interactive belief revision. In *Logic and the Foundations of Game and Decision Theory (LOFT7)*, pages 13–60. Amsterdam University Press, 2008. ISBN 978-90 8964 026 0.

A. Baltag, L. S. Moss, and S. Solecki. The logic of public announcements, common knowledge and private suspicions. Technical Report SEN-R9922, CWI, Amsterdam, 1999.

J. van Benthem. Dynamic logic for belief revision. *Journal of Applied Non-Classical Logics*, 17(2):129–155, 2007. doi: 10.3166/jancl.17.129-155.

J. van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, 2011. ISBN 978-0-521-76579-4.

C. Boutilier. Iterated revision and minimal change of conditional beliefs. *Journal of Philosophical Logic*, 25(3):263–305, 1996. doi: 10.1007/BF00248151.

H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, 2007. ISBN 978-1-4020-5838-7.

H. van Ditmarsch, A. Herzig, J. Lang, and P. Marquis. Introspective forgetting. *Synthese (KRA)*, 169(2):405–423, 2009. doi: 10.1007/s11229-009-9554-4.

D. Fernández-Duque, Á. Nepomuceno-Fernández, E. Sarrión-Morrillo, F. Soler-Toscano, and F. R. Velázquez-Quesada. Forgetting complex propositions. Under submission, 2015. url: http://arxiv.org/abs/1507.01111.

J. Gerbrandy and W. Groeneveld. Reasoning about information change. *Journal of Logic, Language, and Information*, 6(2):147–196, 1997. doi: 10.1023/A:1008222603071.

S. Konieczny and R. P. Pérez. Logic based merging. *Journal of Philosophical Logic*, 40(2):239–270, 2011. doi: 10.1007/s10992-011-9175-5.

H. Leitgeb and K. Segerberg. Dynamic doxastic logic: why, how, and where to? *Synthese*, 155(2):167–190, 2007. doi: 10.1007/s11229-006-9143-8.

F. Lin and R. Reiter. Forget it! In *Proceedings of the AAAI Fall Symposium on Relevance*, pages 154–159, 1994.

J. A. Plaza. Logics of public communications. In *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, Tennessee, USA, 1989. Oak Ridge National Laboratory, ORNL/DSRD-24.

H. Rott. Conditionals and theory change: revisions, expansions, and additions. *Synthese*, 81(1):91–113, 1989. doi: 10.1007/BF00869346.

F. R. Velázquez-Quesada. Inference and update. *Synthese (KRA)*, 169(2): 283–300, 2009. doi: 10.1007/s11229-009-9556-2.

F. R. Velázquez-Quesada. Explicit and implicit knowledge in neighbourhood models. In *LORI*, volume 8196 of *Lecture Notes in Computer Science*, pages 239–252. Springer, 2013. ISBN 978-3-642-40947-9. doi: 10.1007/978-3-642-40948-6_19.

F. R. Velázquez-Quesada. Dynamic epistemic logic for implicit and explicit beliefs. *Journal of Logic, Language and Information*, 23(2): 107–140, 2014. doi: 10.1007/s10849-014-9193-0.

F. R. Velázquez-Quesada, F. Soler-Toscano, and Á. Nepomuceno-Fernández. An epistemic and dynamic approach to abductive reasoning: abductive problem and abductive solution. *Journal of Applied Logic*, 11(4):505–522, 2013. doi: 10.1016/j.jal.2013.07.002.

Y. Zhang and Y. Zhou. Knowledge forgetting: Properties and applications. *Artificial Intelligence*, 173(16–17):1525–1537, 2009. doi: 10.1016/j.artint.2009.07.005.