On a method of proving the non-existence of modal formulae satisfying certain syntactic properties and defining a given class of frames

Petar Iliev
Institute of Mathematics and Informatics
and
Institute of Philosophy and Sociology
Bulgarian Academy of Sciences

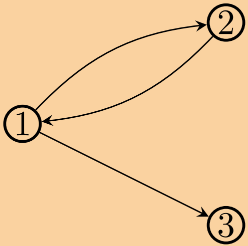
P. Balbiani, A. Herzig, D. Fernández-Duque, P. Iliev, *Frame-validity games and lower bounds on the complexity of modal axioms* (Logic Journal of IGPL, 2020)

P. Iliev, *On a method of proving the non-existence of modal formulae satisfying certain syntactic properties and defining a given class of frames*(submitted)

The unary modal language and its semantics

Formulae: $p_1, p_2, \ldots \mid \neg\varphi \mid \varphi \vee \psi \mid \varphi \wedge \psi \mid \varphi \to \psi \mid \Diamond\varphi \mid \Box\varphi$.

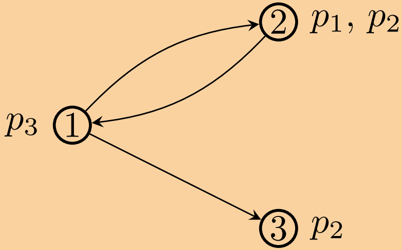
The unary modal language and its semantics



Kripke frame (i.e., directed graph).

The unary modal language and its semantics

Truth of a formula in a pointed Kripke model.
An example:

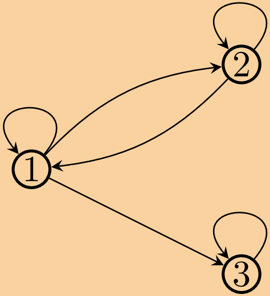


$(\mathcal{M}, 1) \vDash \Box p_2 \wedge \Diamond p_1 \wedge \Diamond\neg p_1$; $(\mathcal{M}, 2) \vDash p_1 \wedge p_2 \wedge \Diamond p_3 \wedge \Box p_3$;
$(\mathcal{M}, 1) \vDash p_1 \rightarrow p_2$; $(\mathcal{M}, 2) \vDash p_1 \rightarrow p_2$; $(\mathcal{M}, 3) \vDash p_1 \rightarrow p_2$.

The unary modal language and its semantics

Validity of a formula in a Kripke frame - truth in every possible pointed Kripke model based on the frame.

An example: $p_1 \to \Diamond p_1$ is valid on this reflexive frame.



A reflexive Kripke frame.

Frame definability

A modal formula defines a class of frames if and only if it is valid on all frames in the class and not valid on any frame not in the class. Example: $p_1 \to \Diamond p_1$ defines the class of reflexive frames.

Sahlqvist formulae

For a natural number $k \geq 0$, the expression $\Box^k \varphi$ ($\Diamond^k \varphi$) is an abbreviation of the formula $\underbrace{\Box \ldots \Box}_{k \text{ times}} \varphi$ ($\underbrace{\Diamond \ldots \Diamond}_{k \text{ times}} \varphi$).

Sahlqvist formulae

A Sahlqvist formula is a formula $\Box^{k_1}(\psi_1 \to \chi_1) \wedge \ldots \wedge \Box^{k_j}(\psi_j \to \chi_j)$, where for $1 \le i \le j$,

- both ψ_i and χ_i are $\to$-free;
- $\neg$ can appear only in front of propositional symbols in ψ_i whereas χ_i does not contain any $\neg$;
- ψ_i does not have a subformula $\Box(\ldots(\theta_1 \vee \theta_2)\ldots)$ where θ_1 or θ_2 contains a propositional symbol p not preceded by a $\neg$;
- ψ_i does not have a subformula $\Box(\ldots \Diamond\,\theta \ldots)$ where θ theta contains a propositional symbol p not preceded by a $\neg$.

A. Chagrov and M. Zakharyaschev, Modal Logic (Theorem 10.30).

The motivating question

D. Vakarelov, Modal definability in languages with a finite number of propositional variables and a new extension of the Sahlqvist's class. (Advances in modal logic 2003)

The motivating question

The class of frames that satisfy the condition

$$\mathbf{CR_4} = \forall x \forall y_1 \forall y_2 \forall y_3 \forall y_4 ((xRy_1 \wedge xRy_2 \wedge xRy_3 \wedge xRy_4) \to \exists z (\bigwedge_{1 \leq i \leq 4} y_i Rz))$$

is definable by both
$\Diamond \Box (p_1 \vee p_2) \wedge \Diamond \Box (p_1 \vee \neg p_2) \wedge \Diamond \Box (\neg p_1 \vee p_2) \to \Box \Diamond (p_1 \wedge p_2)$
and
$\Diamond \Box p_1 \wedge \Diamond \Box p_2 \wedge \Diamond \Box p_3 \to \Box \Diamond (p_1 \wedge p_2 \wedge p_3)$.
The first formula is not a Sahlqvist formula. The second is.

D. Vakarelov conjectured that there is no Sahlqvist formula with two different propositional variables that defines this class of frames.

The motivating question - let's generalise Vakarelov's conjecture

The class of frames that satisfy the condition

$$\mathbf{CR_{2^n}} = \forall x \forall y_1 \ldots \forall y_{2^n}((\bigwedge_{1\leq i\leq 2^n} xRy_i) \to \exists z(\bigwedge_{1\leq i\leq 2^n} y_iRz)).$$

is definable by the Sahlqvist formula
$\Diamond\Box p_1 \wedge \Diamond\Box p_2 \wedge \ldots \wedge \Diamond\Box p_{2^n-1} \to \Box\Diamond(p_1 \wedge p_2 \wedge \ldots \wedge p_{2^n-1})$
with $2^n - 1$ different propositional variables but there is no Sahlqvist formula with at most n different propositional variables that defines this class. Note that D. Vakarelov has proved that there is a formula with n different propositional variables that defines the class but this formula is not a Sahlqvist one.

Attacking the motivating question

We developed the needed tools in
Frame-validity games and lower bounds on the complexity of modal axioms (Logic Journal of IGPL, 2020)
Joint work with

Andreas Herzig
Toulouse University

Philippe Balbiani
Toulouse University

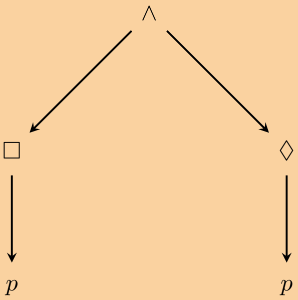
David Fernández-Duque
Ghent University

The main ideas

The general idea comes from Boolean function complexity. Karchmer, M., *On proving lower bounds on circuit size (1993)* (see the beginning of section 2 and Proposition 1)

For first-order and modal logics it was developed in Adler M., and Immerman N., *An n! lower bound on formula size (2003)*

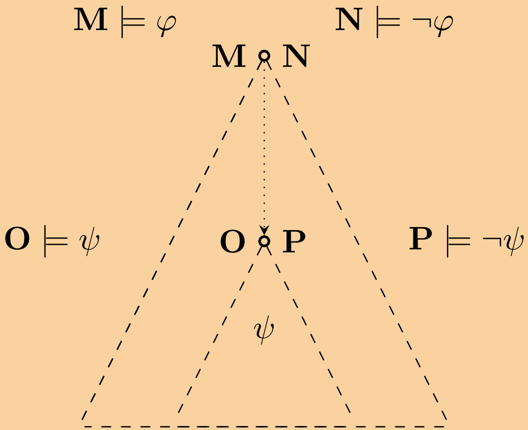
The main ideas



The syntax tree of $\Box p \wedge \Diamond p$.

The main ideas

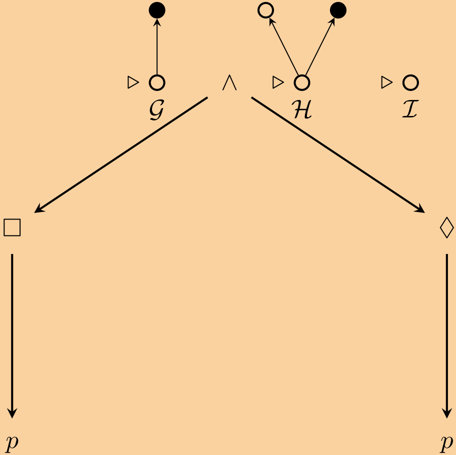
Let **M** and **N** be sets of pointed Kripke models and let φ be a formula such that $\mathbf{M} \vDash \varphi$ whereas $\mathbf{N} \vDash \neg\varphi$. It is known that there is a specific recursive labelling of T_φ.



An example

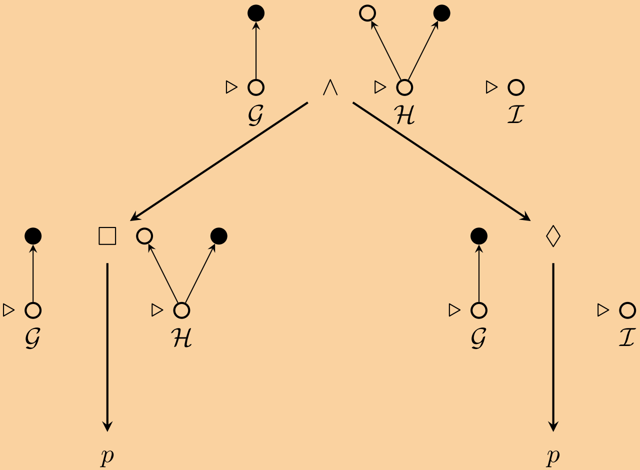
Consider $\Box p \wedge \Diamond p$ and the pair of pointed models where the proposition p is true only in the black nodes.

An example



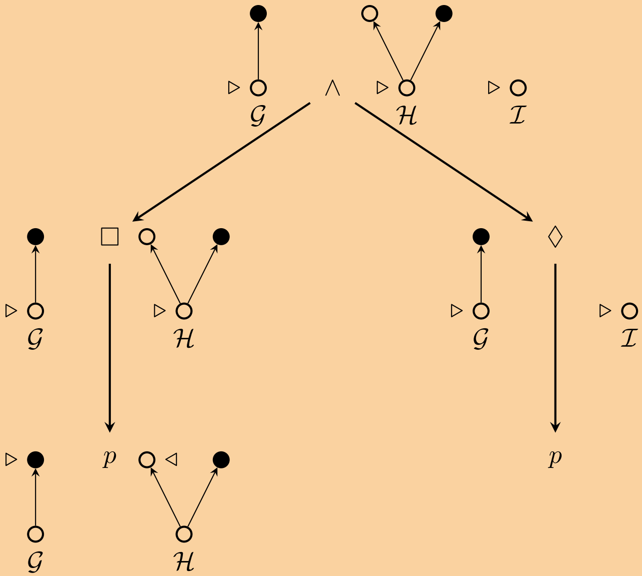
$\triangleright$
$\wedge$
$\triangleright$
$\triangleright$
$\mathcal{G}$
$\mathcal{H}$
$\mathcal{I}$
$\Box$
$\Diamond$
p
p

An example

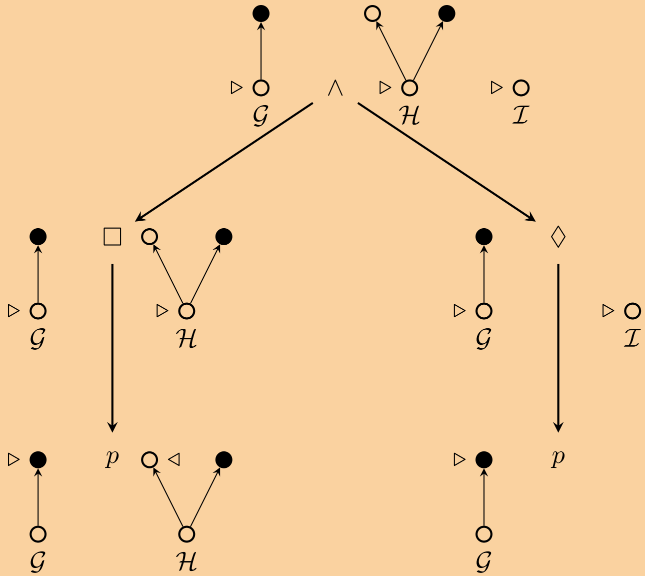


$\mathcal{G}$
$\wedge$
$\mathcal{H}$
$\mathcal{I}$
$\mathcal{G}$
$\Box$
$\mathcal{H}$
$\mathcal{G}$
$\Diamond$
$\mathcal{I}$
p
p

An example



An example



$\mathcal{G}$
$\wedge$
$\mathcal{H}$
$\mathcal{I}$
$\Box$
$\mathcal{G}$
$\mathcal{H}$
$\Diamond$
$\mathcal{G}$
$\mathcal{I}$
p
$\mathcal{G}$
$\mathcal{H}$
p
$\mathcal{G}$

Back to our motivating question

How do we prove that there is no Sahlqvist formula with two different propositional symbols that defines the class of frames that satisfy the condition

$$\mathbf{CR_4} = \forall x \forall y_1 \forall y_2 \forall y_3 \forall y_4((xRy_1 \wedge xRy_2 \wedge xRy_3 \wedge xRy_4) \to \exists z(\bigwedge_{1 \leq i \leq 4} y_i Rz))?$$

P. Iliev, *On a method of proving the non-existence of modal formulae satisfying certain syntactic properties and defining a given class of frames* (submitted).

Back to our motivating question

- We can try to construct two sets of frames $\mathbb{M}$ and $\mathbb{N}$ such that
  - all frames in $\mathbb{M}$ satisfy $\mathbf{CR_4}$;
  - no frame in $\mathbb{N}$ satisfies $\mathbf{CR_4}$.

Back to our motivating question

- We can try to construct two sets of frames $\mathbb{M}$ and $\mathbb{N}$ such that
  - all frames in $\mathbb{M}$ satisfy $\mathbf{CR_4}$;
  - no frame in $\mathbb{N}$ satisfies $\mathbf{CR_4}$.

Obviously, every formula φ that defines $\mathbf{CR_4}$ is valid on $\mathbb{M}$ and not valid on any frame in $\mathbb{N}$.

Back to our motivating question

- We can try to construct two sets of frames $\mathbb{M}$ and $\mathbb{N}$ such that
  - all frames in $\mathbb{M}$ satisfy $\mathbf{CR_4}$;
  - no frame in $\mathbb{N}$ satisfies $\mathbf{CR_4}$.

Obviously, every formula φ that defines $\mathbf{CR_4}$ is valid on $\mathbb{M}$ and not valid on any frame in $\mathbb{N}$.

- For every frame in $\mathbb{N}$, there must be a φ-falsifying pointed model based on that frame. Let $\mathbf{N}$ be the set of falsifying pointed models. Clearly, $\mathbf{N} \vDash \neg\varphi$

Back to our motivating question

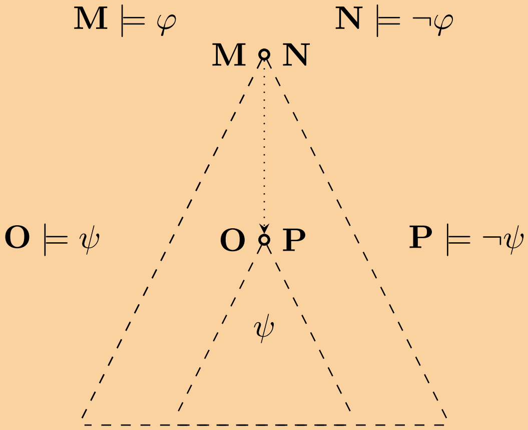
- We can try to construct two sets of frames $\mathbb{M}$ and $\mathbb{N}$ such that
  - all frames in $\mathbb{M}$ satisfy $\mathbf{CR_4}$;
  - no frame in $\mathbb{N}$ satisfies $\mathbf{CR_4}$.

Obviously, every formula φ that defines $\mathbf{CR_4}$ is valid on $\mathbb{M}$ and not valid on any frame in $\mathbb{N}$.

- For every frame in $\mathbb{N}$, there must be a φ-falsifying pointed model based on that frame. Let $\mathbf{N}$ be the set of falsifying pointed models. Clearly, $\mathbf{N} \models \neg\varphi$
- Let us pick a set of pointed models $\mathbf{M}$ based on frames in $\mathbb{M}$. Obviously, $\mathbf{M} \models \varphi$.

Back to our motivating question

Because φ defines $\mathbf{CR_4}$, we must be able to apply the labelling to the syntax tree T_φ of φ so that the property bellow is preserved.

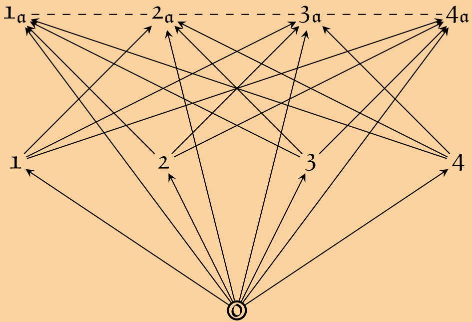


Back to our motivating question

The trick is to construct such sets of frames $\mathbb{M}$ and $\mathbb{N}$ so that

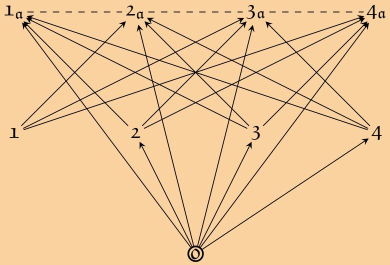
no matter what the set $\mathbf{N}$ *of* φ*-falsifying pointed models is, we can pick a set of pointed models* $\mathbf{M}$ *for which we can prove that we cannot apply the labelling* $\mathcal{L}$ *to the syntax tree of a Sahlqvist formula that contains at most two different variables and that defines the condition* $\mathbf{CR_4}$.

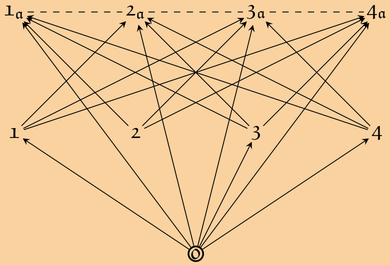
Back to our motivating question



This frame that does not have $\mathbf{CR_4}$. Hence, for every modal formula with at most two different propositional symbols that defines $\mathbf{CR_4}$, there must be a falsifying pointed model based on it. We can prove about such pointed models that they must be based on the point 0 and that the points $1_a, 2_a, 3_a, 4_a$ must satisfy different subsets of the set $\{p_1, p_2\}$. Let $(\mathcal{N}, 0)$ be one such pointed model.

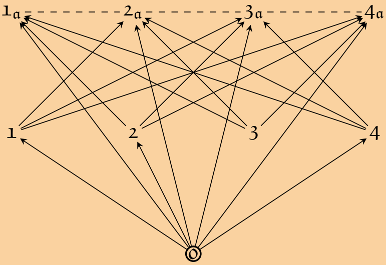
Back to our motivating question

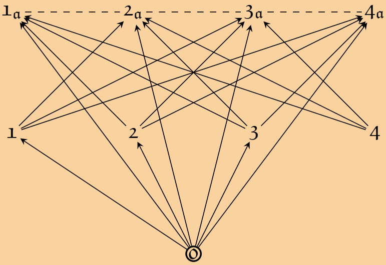




These frames satisfy $\mathbf{CR_4}$. Let $(\mathcal{M}_1, 0)$, $(\mathcal{M}_2, 0)$ be pointed models based on these frames that mimic the valuation function of $\mathcal{N}$.

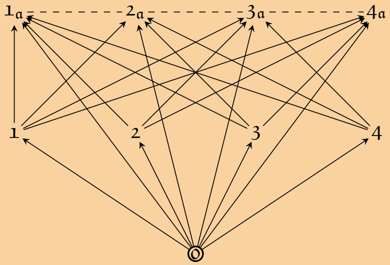
Back to our motivating question





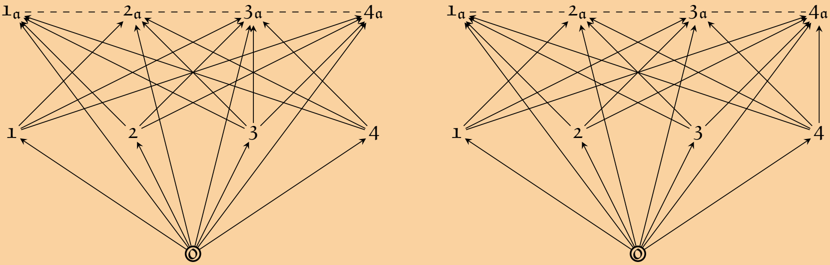
Again, these frames satisfy $\mathbf{CR_4}$. Let $(\mathcal{M}_3, 0)$, $(\mathcal{M}_4, 0)$ be pointed models that mimic the valuation function of $\mathcal{N}$.

Back to our motivating question



Let $(\mathcal{M}_5, 0)$, $(\mathcal{M}_6, 0)$ be pointed models that mimic the valuation function of $\mathcal{N}$.

Back to our motivating question



Let $(\mathcal{M}_7, 0)$, $(\mathcal{M}_8, 0)$ be pointed models that mimic the valuation function of $\mathcal{N}$.

The proof

Let φ be a modal formula with not more than 2 different propositional variables that is true in $\{(\mathcal{M}_1, 0), \ldots, (\mathcal{M}_8, 0)\}$ and false in $(\mathcal{N}, 0)$. Then we can label the syntax tree T_φ as described. However, we can show that, in order to label it as described, then T_φ is not the syntax tree of a Sahlqvist formula.

The general results

For any $n \geq 2$,

- there is no Sahlqvist formula containing n different propositional symbols that defines $\mathbf{CR}_{2^n}$;
- $\mathbf{CR}_{2^n}$ cannot be defined by a modal formula with n different propositional variables that contains only $\Box$'s or only $\Diamond$'s, i.e, every formula with not more than n different propositional variables that defines $\mathbf{CR}_{2^n}$ has alternation depth at least 2 or in other words every such formula has a $\Diamond$ in the scope of a $\Box$ or a $\Box$ in the scope of a $\Diamond$;
- if φ contains not more than n propositional variables and is an $\rightarrow$-free formula that defines $\mathbf{CR}_{2^n}$, then
  - there at least $2^n - 1$ occurrences of $\vee$ in φ;
  - the combination of $(n-1)$ $\wedge$'s in the scope of a $\Diamond$ which is in the scope of a $\Box$ occurs at least 2^n times in φ.

More such results

P. Balbiani, A. Herzig, D. Fernández-Duque, P. Iliev, *Frame-validity games and lower bounds on the complexity of modal axioms* (Logic Journal of IGPL, 2020)

More such results

- for $m, n \geq 0$, the formula $\Diamond^m p \to \Diamond^n p$ is (essentially) an optimal modal formula defining the first-order property

  $$xR^m y \to xR^n y$$

  (note that this result applies to the transitivity, reflexivity, and density axioms);

More such results

- for $m, n \geq 0$, the formula $\Diamond^m p \to \Diamond^n p$ is (essentially) an optimal modal formula defining the first-order property

  $$xR^m y \to xR^n y$$

  (note that this result applies to the transitivity, reflexivity, and density axioms);
- the Löb axiom $\Box(\Box p \to p) \to \Box p$ is an optimal modal formula defining transitivity plus the second-order property of converse well-foundedness.

More such results

- for $m, n \geq 0$, the formula $\Diamond^m p \to \Diamond^n p$ is (essentially) an optimal modal formula defining the first-order property

$$xR^m y \to xR^n y$$

  (note that this result applies to the transitivity, reflexivity, and density axioms);
- the Löb axiom $\Box(\Box p \to p) \to \Box p$ is an optimal modal formula defining transitivity plus the second-order property of converse well-foundedness.
- $(p \vee \Diamond \Diamond p) \to \Diamond p$ is optimal among those defining reflexivity plus transitivity (no need to take T and 4);

More such results

- for $m, n \geq 0$, the formula $\Diamond^m p \to \Diamond^n p$ is (essentially) an optimal modal formula defining the first-order property

  $$xR^m y \to xR^n y$$

  (note that this result applies to the transitivity, reflexivity, and density axioms);
- the Löb axiom $\Box(\Box p \to p) \to \Box p$ is an optimal modal formula defining transitivity plus the second-order property of converse well-foundedness.
- $(p \vee \Diamond \Diamond p) \to \Diamond p$ is optimal among those defining reflexivity plus transitivity (no need to take T and 4);
- $p \to \Box \Diamond p$ is optimal among the modal formulae that define symmetry.

The general results

(Upper Bound) for every n, there is a formula in the language with the universal box and diamond with length $O(n \times \log_2 n)$ and $\lceil \log_2 n \rceil$ variables that modaly defines n-non-colourability;

(Lower Bound) $\lceil \log_2 n \rceil$ variables are necessary, but we were able to show only a linear lower bound n on the formula-size.

Open problems

- I was not able to show that we need $n-1$ different propositional variables in order to have a Sahlqvist formula that defines $\mathbf{CR_n}$.
- In the meantime, prof. Vakarelov has conjectured that adding the backward looking modality can lead to an exponential decrease of the number of variables needed to define certain frames.
- Are the Lemmon-Scott's axioms $\Diamond^m \Box^i p \rightarrow \Box^j \Diamond^n p$ optimal among those defining the first-order condition

$$xR^m y \wedge xR^j z \rightarrow \exists t(yR^i t \wedge zR^n t)?$$

Open problems - the infamous one

Prove or disprove that there is an exponential succinctness gap between LTL with non-strict until and LTL formulated in a language with a strict until-operator.

Open problems - the infamous one

The future according to our naive everyday intuition

$$F\varphi = \langle < \rangle \varphi$$

Open problems - the infamous one

The future according to our naive everyday intuition

$$F\varphi = \langle < \rangle \varphi$$

The future according to people working on formal verification of software.

or

$$F\varphi = \langle \leq \rangle \varphi$$

Open problems - the infamous one

The future according to our naive everyday intuition

$$F\varphi = \langle < \rangle \varphi$$

The future according to people working on formal verification of software.

or

$$F\varphi = \langle \leq \rangle \varphi$$

$$\langle \leq \rangle \varphi \equiv \varphi \vee \langle < \rangle \varphi.$$

The infamous one

I. Hodkinson and M. Reynolds, *Separation - past, present, and future* (2005).

Conjecture: there is an infinite sequence of
formulae with $\langle \leq \rangle$: $\varphi_1, \quad \varphi_2, \quad \dots \quad, \varphi_n, \quad \dots,$
$|\varphi_n| = k.n$ such that the lengths of formulae in any sequence of equivalent
formulae with $\langle < \rangle$: $\geq 2^1, \quad \geq 2^2, \quad \dots \quad, \geq 2^n, \quad \dots$

Open problems: the infamous one

I conjecture that one such sequence is:

$$
\begin{aligned}
\varphi_1 &= \langle\leq\rangle p_1, \\
\varphi_2 &= \langle\leq\rangle(p_2 \wedge \langle\leq\rangle p_1), \\
&\vdots \\
\varphi_n &= \langle\leq\rangle(p_n \wedge \varphi_{n-1}), \\
&\vdots
\end{aligned}
$$

Thank you!