23

Modal Definability in Languages with a Finite Number of Propositional variables and a New Extension of the Sahlqvist's Class

DIMITER VAKARELOV

ABSTRACT. Modal definability in languages with a finite number of propositional variables is introduced and the related notion of minimal modal equivalent is studied. Some applications to the Sahlqvist theory are given: it is proved that the class of Sahlqvist formulas needs infinitely many propositional variables and that there are infinitely many Sahlqvist formulas with minimal modal equivalents which are not Sahlqvist formulas. This suggests a new extension of the class of Sahlqvist formulas, called complex Sahlqvist formulas.

Introduction

In order to keep things as simpler as possible we will work in the standard modal language, denoted here by L_∞, containing an infinite number of propositional variables, the logical constants and operations $\bot, \top, \wedge, \vee$ and the modalities $\Box$ and $\Diamond$ all taken as primitives. We assume also that the notions of modal definability (local and global) for this language are as in [1]. In this paper we will be interested only with the local version of modal definability and from now on "modal definability" will mean "local modal definability".

Let L_n, $n = 0, 1, \ldots$ be the restriction of L_∞ with n different propositional variables. Our first aim in this paper is to study the notion of modal definability in L_n and its connection with the famous Sahlqvist's theorem. For this theorem the reader may consult [7] and also [1]. A very short topological proof of the Sahlqvist's theorem is given in [8]. Nice presentation of the same proof is given also in the book [3]. Extensions for polyadic languages are given in [6] and in a very readable form in the book [2]. The last generalization which extends those given in [6, 2] is in [5].

Probably the only result in modal definability with restricted number of propositional variables is that the first-order condition

Advances in Modal Logic, Volume 4, 499–518.

$F(x): \quad (\forall yz)(xRy \& xRz \to (yRz \vee zRy \vee y = z)$

is modally definable in $L2$ by the formula with two propositional variables

$$\Box(\Box p \wedge p \to q) \vee \Box(\Box q \to p)$$

but it is not modally definable in L_1 [1].

Let F be a first-order formula with one free variable in the first-order language with the only predicate letters R and $=$. If F is modally definable by some modal formula B then F is said to be the *first-order equivalent* of B and B is said to be the *modal equivalent* of F. If F is modally definable then obviously there exists a natural number n such that F is modally definable in L_n but not in L_{n-1}. This means that there exists a modal formula $A(p_1, \ldots, p_n)$ with n different propositional variables $p_1, \ldots, p_n$ such that $A(p_1, \ldots, p_n)$ is a modal equivalent of F and that there is no modal formula with less than n variables which is a modal equivalent of F. The modal formula $A(p_1, \ldots, p_n)$, which modally defines F in L_n, will be called a *minimal modal equivalent* of F. If B is any modal formula which modally defines F then $A(p_1, \ldots, p_n)$ will also be called a minimal modal equivalent of B.

We will be interested in the following general problem: given a first-order formula F corresponding to some Sahlqvist formula B to find a minimal modal equivalent A of B.

The following two theorems are in this direction and are among the main results in the paper.

THEOREM 1 *For each natural number n there exists a Sahlqvist formula A_n, containing $n+1$ different propositional variables, such that A_n is a minimal modal equivalent of the first-order condition corresponding to A_n.*

THEOREM 2 *There exist infinitely many Sahlqvist formulas with minimal modal equivalents which are not Sahlqvist formulas.*

As a corollary of Theorem 1 we obtain that the Sahlqvist's class of formulas needs infinitely many propositional variables. Theorem 2 has a more interesting consequence: the Sahlqvist's class is not closed under taking minimal equivalents of their members. This fact suggests to consider a new extension of the class of Sahlqvist formulas, called now *complex Sahlqvist formulas*, which is the second aim of this paper. Complex Sahlqvist formulas violate some of the forbidden combinations in (classical) Sahlqvist formulas: *boxes over disjunctions* in (simple) Sahlqvist antecedents. Note that violating this condition leads in general outside of first-order definability ([1, 2]).

We found an effective translation τ transforming each complex Sahlqvist formula A into a deductively equivalent in K (classical) Sahlqvist formula $\tau(A)$ (Translation Lemma). In general $\tau(A)$ has much more propositional variables than A. As a corollary of the Translation Lemma we obtain an analog of the Sahlqvist's theorem: complex Sahlqvist formulas are (locally) first-order definable and canonical (the Main Theorem). Let us note that the Sahlqvist-van Benthem substitution algorithm for computing the first-order conditions of Sahlqvist formulas is not applicable to complex Sahqvist formulas. At the end of the paper we give an extension of this algorithm for complex Sahlqvist formulas.

1 Some preliminary technical lemmas

In this section we will list some technical lemmas needed later on.

LEMMA 3 (Combinatorial Lemma) *Let $\|W\| > 2^n$ and let $A_1, \ldots, A_n$ be subsets of W. Then there exist two different elements $x, y \in W$ such that for any $j = 1, \ldots, n$ either $x, y \in A_j$ or $x, y \notin A_j$.*

Proof. Let $A^0 =_{def} W \setminus A$ and $A^1 =_{def} A$. By a pure Boolean reasoning we have that

$\bigcup_{i_1,\ldots,i_n} A_1^{i_1} \cap \ldots \cap A_n^{i_n} = W$

Note that the number of the summands in this union is equal to 2^n and that the different summands are disjoint. Then if for all indices $i_1 \ldots i_n$ we have $\|A_1^{i_1} \cap \ldots \cap A_n^{i_n}\| \leq 1$, we will have $\|W\| \leq 2^n$ contrary to the assumption that $\|W\| > 2^n$. Hence there exist indices $i_1 \ldots i_n$ such that $\|A_1^{i_1} \cap \ldots \cap A_n^{i_n}\| > 1$. This means that there exist two different elements $x, y \in A_1^{i_1} \cap \ldots \cap A_n^{i_n}$. Let $j \in \{1, \ldots, n\}$. If $i_j = 1$ then both x, y are in A_j and if $i_j = 0$ then both x, y are not in A_j, which proves the Combinatorial Lemma. ∎

In order to formulate the next lemma we will remind the reader of some definitions and notations from the classical propositional logic PL. If A, B are formulas of PL then $A \vdash B$ means that the implication $A \Rightarrow B$ is a theorem of PL. We define $A \equiv B$ iff $A \vdash B$ and $B \vdash A$. By a substitution in the set of propositional variables of PL we mean any homomorphism S in the set of formulas of PL.

Let p be a propositional variable. Then $p^0 =_{def} \neg p$ and $p^1 =_{def} p$. Let $p_1, \ldots, p_n$ be a list of different variables. Formulas of the form $D = p_1^{i_1} \vee \ldots \vee p_n^{i_n}$ (in this order of the variables) will be called elementary disjunctions. In a similar way one can define elementary conjunctions. It is known that there are exactly 2^n non-equivalent elementary disjunctions (conjunctions) of $p_1, \ldots, p_n$. Sometimes we will consider a fixed order of all elementary disjunctions (conjunctions) $D_1, \ldots, D_{2^n}$. If we consider the strings $(i_1, \ldots, i_n)$ as binary codes of the integers then there are two natural orderings of the sequence D_i: the *increasing order* – starting from $(0, \ldots, 0) = 0$ and ending with $(1, \ldots, 1) = 2^n - 1$, and the *decreasing order* which is just the opposite of the increasing order.

LEMMA 4 (Substitution lemma 1) *Let $p_1 \ldots p_n$ be a list of different propositional variables and $D_1 \ldots D_{2^n}$ be a fixed list of all elementary disjunctions of them. Let $A_1 \ldots A_{2^n}$ be an arbitrary list of propositional formulas. Then:*

(A) The following two conditions are equivalent:

(i) There exists a substitution S acting on the variables $p_1 \ldots p_n$ such that the following equations hold:

$$(\#1) \quad \left\| \begin{array}{lcl} A_1 & \equiv & S(D_1) \\ A_2 & \equiv & S(D_2) \\ \ldots & & \ldots \\ A_{2^n} & \equiv & S(D_{2^n}). \end{array} \right.$$

(ii) The following two conditions hold for any $i, j \leq 2^n$:

(a) *If $i \neq j$ then $A_i \vee A_j \equiv \top$,*

(b) $\bigwedge_{i=1}^{2^n} A_i = \bot$.

(B) If (ii) is fulfilled then the substitution S is uniquely determined by the following equations:

$$(\#2) \quad S(p_k) = \bigwedge\{A_l : p_k \in D_l, l \leq 2^n\},\ k = 1, \ldots, n.$$

We consider this lemma as a folklore result and left it without proof. The statements for $n = 2, 3$ are quite easy and can be obtained by standard Boolean manipulations. We invite the reader to extend these proofs for arbitrary n.

The following lemma can be considered as a non-trivial statement in the Propositional Logic. We do not know if it is a new one.

LEMMA 5 (Substition lemma 2) *Let $p_1, \ldots, p_n$, be a sequence of different propositional variables, let $D_1, \ldots, D_{2^n}$ be a sequence of all elementary disjunctions of these variables and let $q_1, \ldots, q_{2^n-1}$ be a sequence of propositional formulas. Then:*

(i) There exists a substitution S (depending on the given order of the disjunctions D_i) satisfying the following conditions:

$$(*) \quad \left\| \begin{array}{lcl} q_1 & \equiv & S(D_1) \\ q_1 \wedge q_2 & \equiv & S(D_1 \wedge D_2) \\ \ldots & & \ldots \\ q_1 \wedge q_2 \wedge \ldots \wedge q_{2^n-1} & \equiv & S(D_1 \wedge D_2 \wedge \ldots \wedge D_{2^n-1}). \end{array} \right.$$

(ii) The conditions $(*)$ *uniquely determine* S *(up to logical equivalence) and* S *can be effectively computed from them.*

(iii) The substitution S *satisfies also the following conditions:*

$$(**)\quad \left\|\begin{array}{lll} q_1 & \vdash & S(D_1) \\ q_2 & \vdash & S(D_2) \\ \dots & & \dots \\ q_{2^n-1} & \vdash & S(D_{2^n-1}). \end{array}\right.$$

Proof. (i). Define the following sequence of formulas: $A_1 = q_1$, $A_2 = \neg q_1 \vee q_2$, $A_3 = \neg q_1 \vee \neg q_2 \vee q_3$,..., $A_{2^n-1} = \neg q_1 \vee \ldots \vee \neg q_{2^n-2} \vee q_{2^n-1}$ and $A_{2^n} = \neg q_1 \vee \ldots \vee \neg q_{2^n-2} \vee \neg q_{2^n-1}$. It is easy to see that the conditions (a) and (b) for $A_1, \ldots, A_{2^n}$ from the Substitution lemma 1 are fulfilled and hence there exists a substitution S satisfying the equations (#1).

One can see by pure propositional reasoning that $A_1, \ldots, A_{2^n}$ satisfy also the following equations:

(#3) $\quad A_1 \wedge \ldots \wedge A_k \equiv B_1 \wedge \ldots \wedge B_k$, $k = 1, \ldots, 2^n$.

Now from (#1), (#3) and the fact that S is a homomorphism one can easily derive the equations $(*)$.

(ii). Suppose now that S exists and satisfies $(*)$. Using the special structure of the elementary disjunctions one can see that they satisfy the conditions (a) and (b):

(a) $\quad$ If $i \neq j$ then $D_i \vee D_j \equiv \top$, $i, j \leq 2^n$,

(b) $\quad \bigwedge_{i=1}^{2^n} D_i = \bot$.

Applying the condition (a) one can see that the following equations are true

(#4) $\quad \neg(D_1 \wedge \ldots \wedge D_k) \vee D_{k+1} \equiv D_{k+1}$, $k < 2^n$.

Now we transform the equations $(*)$ as follows – negate both sides of the $k-$th row and add it disjunctively to the $(k+1)$-th row:

$\neg(q_1 \wedge \ldots \wedge q_k) \vee (q_1 \wedge \ldots \wedge q_k \wedge q_{k+1}) \equiv S(\neg(D_1 \wedge \ldots \wedge D_k) \vee (D_1 \wedge \ldots \wedge D_k \wedge D_{k+1}))$.

The left side of this equation is equivalent to A_{k+1} and by (#4) the right side is equivalent to $S(D_{k+1})$. In this way, using also (b), we equivalently transform the system $(*)$ into the system (#1). Then by the part B of Substitution Lemma 1 the substitution S is uniquely determined and can be computed effectively.

(iii). From the definition of A_i, $i < 2^n$ one can see that the following is true:

(#5) $\quad q_i \vdash A_i$.

Then applying (#5) to the system (#1) we obtain the conditions $(**)$. ■

Let for brevity AB and $\overline{A}$ denote the conjunction $A \wedge B$ and the negation $\neg A$ respectively and consider the decreasing order of the elementary disjunctions of the variables p_1, p_2 (p_1, p_2, p_3). In the next two examples we show the explicit form of the substitution S for $n = 2$ and $n = 3$.

Example 1: $(n = 2)$.

$S(p_1) = q_1 q_2$,
$S(p_2) = (q_1 \overline{q_2}) \vee (q_1 q_3)$.

Example 2: $(n = 3)$.

$S(p_1) = q_1 q_2 q_3 q_4$,
$S(p_2) = q_1 q_2 \overline{q_3 q_4} \vee q_1 q_2 q_5 q_6$,
$S(p_3) = q_1 \overline{q_2} \vee q_1 q_3 \overline{q_4} \vee q_1 q_3 q_5 \overline{q_6} \vee q_1 q_3 q_5 q_7$.

In the next lemma we will consider systems of equations over an arbitrary Boolean algebra B in the following forms:

$$(+) \quad \left\| \begin{array}{lcl} q_1 & = & D_1 \\ q_1 \wedge q_2 & = & D_1 \wedge D_2 \\ \dots & & \dots \\ q_1 \wedge q_2 \wedge \dots \wedge q_k & = & D_1 \wedge D_2 \wedge \dots \wedge D_k \end{array} \right.$$

and

$$(++) \quad \left\| \begin{array}{lcl} q_1 & \leq & D_1 \\ q_2 & \leq & D_2 \\ \dots & & \dots \\ q_k & \leq & D_k. \end{array} \right.$$

In $(+)$ and $(++)$ $q_1, \dots, q_k$ denote given elements from B. A solution of the system $(+)$ (system $(++)$ or both together) is any valuation of the variables $p_1, \dots, p_n$ in B which satisfies the system.

LEMMA 6 (Boolean lemma) *Let $B = (B, 0, 1, \wedge, \vee, \neg)$ be a Boolean algebra and $D_1, \dots, D_{2^n}$ be a list of all elementary disjunctions of the different variables $p_1, \dots, p_n$ and let $q_1, \dots, q_k$, $1 \leq k \leq 2^n - 1$, be given elements of B. Then the system consisting of $(+)$ and $(++)$ together has a solution in B and if $k = 2^n - 1$ then this solution is unique.*

Proof. Using the connection between Boolean algebras and PL one can easily derive the Boolean Lemma from the Substitution Lemma 2. ■

2 Proof of Theorem 1

Let (W, R) be a frame and consider the following first-order condition:

$Alt_{k-1}(x) \quad (\forall y_1 \dots y_k)(xRy_1 \& \dots \& xRy_k \rightarrow \|y_1, \dots, y_k\| < k)$, $k > 1$.

It is known that this condition is modally definable by the following formula ([9]):

$$\mathbf{Alt}_{k-1} \quad \Box q_1 \vee \Box(q_1 \Rightarrow q_2) \vee \ldots \vee \Box(q_1 \wedge \ldots \wedge q_{k-1} \Rightarrow q_k).$$

We will us this formula in the following equivalent form:

$$\Box q_1 \vee \Box(\neg q_1 \vee q_2) \vee \ldots \vee \Box(\neg q_1 \vee \ldots \vee \neg q_{k-1} \vee q_k).$$

Note. If in the above formula we substitute q_k with $\perp$ we obtain an equivalent in K formula to $\mathbf{Alt}_{k-1}$ which we denote by $\mathbf{Alt}'_{k-1}$.
Note also that $\mathbf{Alt}_{k-1}$ and $\mathbf{Alt}'_{k-1}$ can easily be rewritten into a Sahlqvist form. However neither of these equivalents of $\mathbf{Alt}_{k-1}$ is a minimal modal equivalent of $Alt_{k-1}(x)$. Now we will construct another Sahlqvist formula which is a minimal equivalent of $Alt_{k-1}(x)$.

Let $p_1, \ldots, p_{n+1}$ be a sequence of $n+1$ different propositional variables and let $Q_1, \ldots, Q_{2^{n+1}}$ be a fixed sequence of all 2^{n+1} elementary conjunctions of the variables p_i (for instance in the increasing order). We introduce the following Sahlqvist formula of $n+1$ propositional variables:

$$\mathbf{ALT}^{min}_{k-1} \quad \Diamond Q_1 \wedge \ldots \wedge \Diamond Q_k \Rightarrow \perp.$$

We will use this formula also in the following equivalent form:

$$\Box D_1 \vee \ldots \vee \Box D_k,$$

where the formula D_i $(i = 1, \ldots, k)$ is the elementary disjunction equivalent to $\neg Q_i$.

Our first aim now is to show that $\mathbf{ALT}^{min}_{k-1}$ is a modal equivalent of $Alt_{k-1}(x)$. For each concrete k and n one may use for this purpose the Sahlqvist's Theorem but for the general case it is more convenient just to prove that $\mathbf{ALT}^{min}_{k-1}$ and $\mathbf{Alt}_{k-1}$, are deductively equivalent in K. So we have the following lemma.

LEMMA 7 *Let $k \leq 2^{n+1}$. Then:*

(i) The formulas $\mathbf{ALT}^{min}_{k-1}$ and $\mathbf{Alt}_{k-1}$ are deductively equivalent in K.

(ii) $\mathbf{ALT}^{min}_{k-1}$ modally defines $Alt_{k-1}(x)$.

Proof. (i). Let $k \leq 2^{n+1}$ and let $D_1, \ldots, D_{2^{n+1}}$ be the list of the elementary disjunctions equivalent to the negations of the elementary conjunctions $Q_1, \ldots, Q_{2^{n+1}}$.

For the implication from $\mathbf{Alt}_{k-1}$ to $\mathbf{ALT}^{min}_{k-1}$ substitute q_i in $\mathbf{Alt}_{k-1}$ with D_i, $i = 1, \ldots, k$. Using the fact that $D_i \vee D_j \equiv \top$ for $i \neq j$ one can easily prove that the following equivalences are true in the Propositional Logic:

$$\neg D_1 \vee D_2 \equiv D_2,\ \neg D_1 \vee \neg D_2 \vee D_3 \equiv D_3,\ \ldots \neg D_1 \vee \ldots \vee \neg D_{k-1} \vee D_k \equiv D_k.$$

Then the result of the substitution is just equivalent to $\mathbf{ALT}^{min}_{k-1}$.

For the implication from $\mathbf{ALT}^{min}_{k-1}$ to $\mathbf{Alt}_{k-1}$ define the following sequence of formulas:

$A_1 = q_1$, $A_2 = \neg q_1 \vee q_2$,...,$A_{k-1} = \neg q_1 \vee \ldots \vee \neg q_{k-2} \vee q_{k-1}$,

$A_k = \neg q_1 \vee \ldots \vee \neg q_{k-2} \vee \neg q_{k-1}$, $A_{k+1} = \ldots = A_{2^{n+1}} = \top$.

This definition is possible because of the inequality $k \leq 2^{n+1}$. It can easily be seen that the formulas A_i satisfy the conditions (a) and (b) of the Substitution Lemma 1. Then there is a substitution S such that

$$\begin{array}{lcl} A_1 & \equiv & S(D_1) \\ A_2 & \equiv & S(D_2) \\ \ldots & & \ldots \\ A_k & \equiv & S(D_k). \end{array}$$

By this substitution we obtain

$S(\mathbf{ALT}^{min}_{k-1}) = \Box A_1 \vee \ldots \vee \Box A_k$

which is just the formula $\mathbf{Alt}'_{k-1}$.

(ii). Condition (ii) is a corollary of (i). ■

PROPOSITION 8 *Let* $2^n < k \leq 2^{n+1}$. *Then:*

(i) There is no modal formula of n *variables which modally defines the condition* $Alt_{k-1}(x)$.

(ii) The Sahlqvist formula $\mathbf{ALT}^{min}_{k-1}$ *of* $n+1$ *variables is a minimal modal equivalent of* $Alt_{k-1}(x)$.

Proof. (i). Let $2^n < k \leq 2^{n+1}$ and let $p_1, \ldots, p_n$ be a list of different propositional variables. Suppose that such a formula $\alpha = \alpha(p_1, \ldots, p_n)$ exists and proceed to obtain a contradiction. This can be done realizing the following

PROGRAM

(1) First we construct a frame $F = (W, R)$ which is not a frame for α and consequently we define a model $M = (W, R, v)$ falsifying α.

The next steps are:

(2) to construct a frame $F' = (W', R')$ which is a frame for α and

(3) to construct a p-morphism f from F onto F'.

(4) Then we define a valuation v' in F' such that f is a p-morphism from the model M onto the model $M' = (W', R', v')$.

(5) Now the contradiction can be obtained as follows. By (1), (4) and p-morphism lemma M' falsifies α which contradicts (2).

Let us start realizing the above program.

Step (1). Define the frame $F = (W, R)$ as follows. Let $0, a_1, \ldots, a_k$ be $k+1$ different elements. Then put

$W = \{0, a_1, \ldots, a_k\}$ and $R = \{(0, a_1), \ldots, (0, a_k)\}$.

Obviously F does not satisfy the condition $Alt_{k-1}(x)$ because 0 has just k different $R-$successors and hence F is not a frame for α. Then there exists a valuation v in F falsifying α.

Step (2). Define $A_i = v(p_i) \cap \{a_1, \ldots, a_k\}$, $i = 1, \ldots, n$. Since $k > 2^n$ then by the Combinatorial Lemma there exist $x \neq y \in \{a_1, \ldots, a_k\}$ such that for any $i = 1, \ldots, n$ we have either $x, y \in A_i$ or $x, y \notin A_i$. Then for some $l \neq m \in \{1, \ldots, k\}$ we have $x = a_l$ and $y = a_m$. Consequently we obtain the following equivalence:

(#) $a_l \in v(p_i)$ iff $a_m \in v(p_i)$, $i = 1, \ldots, n$.

Now define F' in the same way as F assuming only that $a_l = a_m$. Now $\|a_1, \ldots, a_k\| = k - 1$ and consequently F' satisfies $Alt_{k-1}(x)$. Hence F' is a frame for α.

Step (3). Define f from W onto W' as follows: $f(0) = 0$ and for $i \neq l$ put $f(a_i) = a_i$ and finally put $f(a_l) = a_m$. Obviously f is a p-morphism from W onto W' "gluing" a_l and a_m.

Step (4). The valuation v' in W' is defined by $v'(p_i) = v(p_i) \cap W'$. We will prove that f is a p-morphism from the model M onto the model $M' = (W', R', v')$. The only thing we have to show is the equivalence:

$(\forall x \in W)(x \in v(p_i)$ iff $f(x) \in v'(p_i))$, $i = 1, \ldots, n$.

For $x \neq a_l$ this is obvious by the definition of f and v'.
For $x = a_l$ we proceed as follows:

$a_l \in v(p_i)$ iff (by (#)) $a_m \in v(p_i)$ iff $a_m \in v(p_i) \cap W'$ iff (by the definitions of f and v') $f(a_l) \in v'(p_i)$, $i = 1, \ldots, n$.

The proof of (i) is finished.

(ii). Let $2^n < k \leq 2^{n+1}$. By (i) there is no modal formula of n variables which modally defines $Alt_{k-1}(x)$. By 7 the formula $\mathbf{ALT}_{k-1}^{min}$ of $n+1$ propositional variables modally defines $Alt_{k-1}(x)$. Hence $\mathbf{ALT}_{k-1}^{min}$ is a minimal modal equivalent of $Alt_{k-1}(x)$. ■

Proof of Theorem 1. Let n be a given natural number and take any k such that $2^n < k \leq 2^{n+1}$. Then by Proposition 8 the Sahlqvist formula $A_n = \mathbf{ALT}_{k-1}^{min}$ contains $n+1$ propositional variables and is a minimal modal equivalent of the corresponding first-order condition $Alt_{k-1}(x)$. ■

COROLLARY 9 *The class of Sahlqvist formulas needs an infinite number of propositional variables.*

REMARK 10 (i) The formula $\mathbf{ALT}_{k-1}^{min}$ depends on the chosen order of the string $< Q_1, \ldots, Q_{2^{n+1}} >$ of the elementary conjunctions. Different orders present different formulas but on the base of lemma 7 all are deductively equivalent in K.

(ii) The Segerberg's formula $\mathbf{Alt}_{k-1}$ is based on the fact that the difference between k points in a frame can be expressed by means of k (or $k-1$) different propositional variables. In $\mathbf{ALT}^{min}_{k-1}$ the same thing is done by means of a smaller number of variables, namely by the smallest number $n+1$ such that $k \leq 2^{n+1}$.

3 Proof of Theorem 2

In order to prove Theorem 2 we introduce the following sequence of first-order conditions which generalize the well-known Churc-Rosser confluence condition:

$$CR_{k-1}(x) \quad (\forall y_1 \ldots y_k)(xRy_1 \& \ldots \& xRy_k \rightarrow (\exists z)(y_1 Rz \& \ldots \& y_k R_k z)).$$

LEMMA 11 *The following Sahlqvist formula is a modal equivalent of the Churc-Rosser confluence condition $CR_{k-1}(x)$:*

$$\mathbf{CR}_{k-1} \quad \Diamond \Box q_1 \wedge \ldots \wedge \Diamond \Box q_{k-1} \Rightarrow \Box \Diamond (q_1 \wedge \ldots \wedge q_{k-1}).$$

Proof. Use directly the Sahlqvist's theorem. ■

To define a minimal equivalent of the Churc-Rosser confluence condition $CR_{k-1}(x)$ let $2^n < k \leq 2^{n+1}$, let $p_1, \ldots, p_{n+1}$ be a sequence of different propositional variables and let $D_1, \ldots, D_{2^{n+1}}$ be a sequence of all non-equivalent elementary disjunctions of these variables. Then we define the formula:

$$\mathbf{CR}^{min}_{k-1} \quad \Diamond \Box D_1 \wedge \ldots \wedge \Diamond \Box D_{k-1} \Rightarrow \Box \Diamond (D_1 \wedge \ldots \wedge D_{k-1}).$$

Example 3:

$$\Diamond \Box (p_1 \vee p_2) \wedge \Diamond \Box (p_1 \vee \neg p_2) \wedge \Diamond \Box (\neg p_1 \vee p_2) \Rightarrow \Box \Diamond (p_1 \wedge p_2).$$

(Here $p_1 \wedge p_2$ is substituted for its equivalent $(p_1 \vee p_2) \wedge (p_1 \vee \neg p_2) \wedge (\neg p_1 \vee p_2)$.

LEMMA 12 *Let $2^n < k \leq 2^{n+1}$. Then:*

(i) The formulas $\mathbf{CR}_{k-1}$ and $\mathbf{CR}^{min}_{k-1}$ are deductively equivalent in K.

(ii) The formula $\mathbf{CR}^{min}_{k-1}$ is canonical and a modal equivalent of the first-order formula $CR_{k-1}(x)$.

Proof. (i). The direction from $\mathbf{CR}_{k-1}$ to $\mathbf{CR}^{min}_{k-1}$ is easy – simply substitute q_i with D_i, $i = 1, \ldots, k-1$.

For the converse direction let S be the substitution guaranteed by the Substitution Lemma 2. Then we obtain:

$$(\#) \quad \Diamond \Box S(D_1) \wedge \ldots \wedge \Diamond \Box S(D_{k-1}) \Rightarrow \Box \Diamond S(D_1 \wedge \ldots \wedge D_{k-1}).$$

By the Substitution Lemma 2 we have:

$$(**)\quad \begin{array}{lcl} q_1 & \vdash & S(D_1) \\ q_2 & \vdash & S(D_2) \\ \ldots & & \ldots \\ q_{k-1} & \vdash & S(D_{k-1}) \end{array}$$

and also

$(*)\qquad q_1 \wedge q_2 \wedge \ldots \wedge q_{k-1} \equiv S(D_1 \wedge D_2 \wedge \ldots \wedge D_{k-1}).$

Using the monotonicity of $\Diamond$ and $\Box$ and the replacement of equivalents we obtain from (#), $(**)$ and $(*)$ the formula

$\Diamond\Box q_1 \wedge \ldots \wedge \Diamond\Box q_{k-1} \Rightarrow \Box\Diamond(q_1 \wedge \ldots \wedge q_{k-1}).$

This is just the formula $\mathbf{CR}_{k-1}$ and the proof is finished.

(ii). The proof of (ii) follows from (i). ■

PROPOSITION 13 *Let $2^n < k \leq 2^{n+1}$. Then:*

(i) There is no modal formula of n variables which modally defines the condition $CR_{k-1}(x)$.

(ii) The non-Sahlqvist formula $\mathbf{CR}^{min}_{k-1}$ (with $n+1$ variables) is a canonical minimal modal equivalent of $CR_{k-1}(x)$.

Proof. (i). The proof is similar to the proof of Proposition 8. Let $2^n < k \leq 2^{n+1}$ and suppose that such a formula $\alpha = \alpha(p_1, \ldots, p_n)$ exists and proceed to obtain a contradiction. The contradiction can be obtained realizing the same PROGRAM of 5 steps as in Proposition 8 but now the constructions are more complicated.

Step (1). The frame $F = (W, R)$ from which the program starts for this case is the following:

$W = \{0, 1, \ldots, k, \omega, a_1, \ldots, a_k\}$, where $a_1, \ldots, a_k$ are different elements and different from $0, 1, \ldots, k, \omega$.

$R = \{(0, i)/i = 1, \ldots, k\} \cup \{(a_i, \omega)/i = 1, \ldots, k\} \cup \{(i, a_j)/i \neq j, i, j = 1, \ldots, k\} \cup \{(\omega, \omega)\}$.

We invite the reader to draw this frame in order to follow more easily the next reasonings.

It can be seen that F does not satisfy the condition $CR_{k-1}(x)$, because we have $0R1, \ldots, 0Rk$ but there is no $z \in W$ such that $1Rz, \ldots, kRz$. Hence F is not a frame for α. Then there is a valuation v in F falsifying α.

Step (2). Define $A_i = v(p_i) \cap \{a_1, \ldots, a_k\}$, $i = 1, \ldots, n$. Since $k > 2^n$ then by the Combinatorial Lemma there exist $x \neq y \in \{a_1, \ldots, a_k\}$ such that for any $i = 1, \ldots, n$ we have either $x, y \in A_i$ or $x, y \notin A_i$. Then for some $l \neq m \in \{1, \ldots, k\}$ we have $x = a_l$ and $y = a_m$. Consequently we obtain the following equivalence:

(#) $\quad a_l \in v(p_i)$ iff $a_m \in v(p_i)$, $i = 1, \ldots, n$.

Define F' in the same way as F assuming only that $a_l = a_m$. It is easy to see now that F' satisfies the condition $CR_{k-1}(x)$. Hence F' is a frame for α.

Step (3). Define f from W into W' as follows: $f(i) = i$ for $i = 0, 1, \ldots, k, \omega$ and for $i \neq l$ put $f(a_i) = a_i$ and finally put $f(a_l) = a_m$. Obviously f is a p-morphism from W onto W' "gluing" a_l and a_m.

Step (4). The valuation v' in W' is defined by $v'(p_i) = v(p_i) \cap W'$. The proof that f is a p-morphism from the model M onto the model $M' = (W', R', v')$ is similar to the corresponding proof in Proposition 8.

The PROGRAM has been realized and the expected contradiction can be obtained as in Proposition 8: since α is falsified in the model (F, v) then by the p-morphism lemma it is falsified also in the model (F', v'), which is impossible, because F' is a frame for α. The proof of (i) is finished.

(ii). Let $2^n < k \leq 2^{n+1}$. By Lemma 12 the formula $\mathbf{CR}^{min}_{k-1}$ containing $n+1$ propositional variables is a modal equivalent of $CR_{k-1}(x)$. It follows from (i) that it is a canonical minimal modal equivalent of $CR_{k-1}(x)$. ■

Proof of Theorem 2. Theorem 2 follows just from the Proposition 13 and the remark that $\mathbf{CR}^{min}_{k-1}$ are not Sahlqvist formulas. ■

REMARK 14 (i) Note that $\mathbf{CR}^{min}_{k-1}$ are not Sahlqvist formulas. They violate one of the forbidden conditions for the Sahlqvist antecedents: disjunctions in the scope of boxes.

(ii) The formula $\mathbf{CR}^{min}_{k-1}$ depends on the given order of the elementary disjunctions. But by Lemma 12 any order will give a formula deductively equivalent in K to the Sahlqvist formula $\mathbf{CR}_{k-1}$. Note that this equivalence is not based on easy and obvious transformations.

(iii) The formula $\mathbf{CR}_{k-1}$ is based on the fact that common R-successor of k points in a frame can be expressed by means of $k-1$ different propositional variables. The formula $\mathbf{CR}^{min}_{k-1}$ is based on the observation that the same thing can be done by a smaller number of variables, namely by the smallest number $n+1$ such that $k \leq 2^{n+1}$. Proposition 13 shows that this number is optimal.

4 A new extension of the Sahlqvist's class

The non-Sahlqvist formulas $\mathbf{CR}^{min}_{k-1}$ which are canonical and local correspondents of some first-order conditions – features typical for the formulas of the Sahlqvist's class, suggest this class to be extended. These formulas violate the forbidden condition disjunctions to be preceded by a box in the Sahlqvist antecedents. But in this case the disjunctions are of a special kind – elementary disjunctions. If we compare the equivalent formulas $\mathbf{CR}_{k-1}$ and $\mathbf{CR}^{min}_{k-1}$, we can see the following things:

(i) The number of the variables in the first formula is $k-1$, the number in the second – $n+1$, and the numbers k, n satisfy the following inequality $2^n < k \leq 2^{n+1}$. It follows from this inequality that for $n > 1$ the second formula always has a smaller number of variables.

(ii) The second thing we can see is that the elementary disjunctions D_i in the second formula just "code" the variables q_i from the first formula. This suggests to look at the elementary disjunctions D_i as a new kind of "complex" variables. Another important block in the second formula is the conjunction $D_1 \wedge \ldots \wedge D_{k-1}$ in the consequent. Looking at the Substitution Lemma 2 we can see similar blocks D_1, $D_1 \wedge D_2$, $D_1 \wedge D_2 \wedge D_3, \ldots$ and they will be also used in the generalization.

Let us start building the new definition.

Let $p =< p_1, \ldots, p_n >$ be a string of different propositional variables and let the elementary disjunctions built from p have the same order of the variables. Let $D(p) =< D_1, \ldots, D_{2^n-1} >$ be a fixed sequence of different elementary disjunctions (the last one D_{2^n} is missing). Consider also the sequence $D^\wedge(p) =< D_1, D_1 \wedge D_2, \ldots, D_1 \wedge \ldots \wedge D_{2^n-1} >$. Then the pair $P =< D(p), D^\wedge(p) >$ is called *propositional complex of dimension n*, the disjunctions D_i from $D(p)$ are called *simple blocks* or *complex variables* and the elements from $D^\wedge(p)$ are called *compound blocks*. We say that the complex $< D(p), D^\wedge(p) >$ is *disjoint* from the complex $< D(q), D^\wedge(q) >$ if the strings p and q do not share common variables.

By a *constant formula* we mean any formula not containing propositional variables. By a *positive formula* we mean any formula built from variables and constant formulas by means of the operations $\wedge, \vee, \Diamond, \Box$. Negation of a positive formula is called a *negative formula*.

Let Σ be a set of pairwise disjoined propositional complexes. The expression $\Box^n A$ where A is a simple block from Σ is called *complex boxed atom* ($n = 0$ is included: $\Box^0 A = A$). Let $A = A(p_1, \ldots, p_k)$ be a positive (negative) formula and $p_1, \ldots, p_k$ be the list of all variables of A. If $B_1, \ldots, B_k$ is a list of compound blocks from Σ then $A(B_1, \ldots, B_k)$ is called *complex positive (negative) atom* in Σ. By a **complex Sahlqvist antecedent** in Σ we mean any formula composed from complex boxed atoms and negative complex atoms from Σ by means of $\wedge$ and $\Diamond$. Formulas of the form $\alpha \Rightarrow \beta$ where α is a complex Sahlqvist antecedent and β is a complex positive atom are called **complex Sahlqvist implications** in Σ.

Now the class of **complex Sahlqvist formulas** in Σ can be obtained from the complex Sahlqvist implications in Σ by freely applying boxes and conjunctions, and disjunctions only to formulas sharing no common variables. We allow Boolean simplifications of compound blocks as in the example 3.

Suppose that we have a string $p =< p_1 >$ of dimension 1. Then D_1 is either p_1 or $\neg p_1$. So in the first case we have $D(p) = D^\wedge(p) =< p_1 >$ which shows that if all complex variables are of dimension 1 then the notion of complex Sahlqvist formula coincides with the notion of Sahlqvist formula. Thus the class of complex Sahlqvist formulas contains the original Sahlqvist's class. The next lemma however shows that the new class do not extend the corresponding class of first-order equivalents.

LEMMA 15 (Translation Lemma) *There exists an effective translation τ such that every complex Sahlqvist formula A is deductively equivalent in K to the Sahlqvist formula $\tau(A)$.*

Proof. The translation τ is very simple – just consider complex variables in A as notations of new propositional variables ("new" here means "not included in A"). It is clear that if A is a complex Sahlqvist formula then $\tau(A)$ is an (ordinary) Sahlqvist formula. Note that if we prove the lemma for the complex Sahlqvist implications then the general case is straightforward.

Let $A = \alpha \Rightarrow \beta$ be a complex Sahlqvist implication. The proof for this case is quite similar to the proof of lemma 12. Since the propositional complexes in A are disjoined we may realized the translation τ step-by-step substituting the complex variables one by one (in arbitrary order). This means that we consider a "partial" translation τ_p for each propositional complex $P =< D(p), D^\wedge(p) >$ just substituting in A only occurrences of complex variables from P.

So we will prove that A is deductively equivalent in K to $\tau_p(A)$.

For the implication from $\tau_p(A)$ to A note that $\tau_p(A)$ can be obtained from A by ordinary substitution, so this case is obvious.

For the implication from A to $\tau_p(A)$ we proceed as follows.

Let $p =< p_1, \ldots, p_n >$ and let $D(p) =< D_1, \ldots, D_{2^n-1} >$ be the string of the simple blocks from p. For the translation we need a string of new variables $< q_1, \ldots, q_{2^n-1} >$. Let $Box_1(D_{i_1}), \ldots, Box_j(D_{i_j})$ be the list of complex boxed atoms occurring in α and let $C_1, \ldots, C_k$ be the complex negative atoms occurring in α. Then α is in the form

$\alpha(Box_1(D_{i_1}), \ldots, Box_j(D_{i_j}); C_1, \ldots, C_k)$.

Note that complex boxed atoms in α occur only positively, compound blocks in α occur only negatively and that β is a complex positive atom.

Now we shall make use of the substitution lemma 2. This lemma guarantees existence of a substitution S such that

$$(*) \qquad \left\| \begin{array}{lll} q_1 & \equiv & S(D_1) \\ q_1 \wedge q_2 & \equiv & S(D_1 \wedge D_2) \\ \ldots & & \ldots \\ q_1 \wedge q_2 \wedge \ldots \wedge q_{2^n-1} & \equiv & S(D_1 \wedge D_2 \wedge \ldots \wedge D_{2^n-1}), \end{array} \right.$$

$$(**)\quad \left\| \begin{array}{lcl} q_1 & \vdash & S(D_1) \\ q_2 & \vdash & S(D_2) \\ \ldots & & \ldots \\ q_{2^n-1} & \vdash & S(D_{2^n-1}). \end{array} \right.$$

It can be assumed that the substitution S acts only on the variables from p and leaves unchanged the other variables.

Using $(*)$ the following statement can be proved by an easy induction.

Statement. *If B is a complex positive or negative atom then the following equivalence is true: $S(B) \equiv \tau_p(B)$.*

Applying S to A we obtain:

(1) $S(A) = \alpha(Box_1(S(D_{i_1})), \ldots, Box_j(S(D_{i_j})); S(C_1), \ldots, S(C_k)) \Rightarrow S(\beta)$

Applying the above Statement to (1) we obtain

(2) $S(A) \equiv \alpha(Box_1(S(D_{i_1})), \ldots, Box_j(S(D_{i_j})); \tau_p(C_1), \ldots, \tau_p(C_k)) \Rightarrow \tau_p(\beta)$

Using the fact that complex boxed atoms occur in α only positively, then applying $(**)$ to (2) we obtain

(3) $S(A) \vdash \alpha(Box_1(q_{i_1}), \ldots, Box_j(q_{i_j})); \tau_p(C_1), \ldots, \tau_p(C_k)) \Rightarrow \tau_p(\beta)$.

But obviously we have

(4) $\quad \alpha(Box_1(q_{i_1}), \ldots, Box_j(q_{i_j})); \tau_p(C_1), \ldots, \tau_p(C_k)) = \tau_p(\alpha)$.

Now from (4) and (3) we obtain

(5) $\quad S(A) \vdash \tau_p(\alpha) \Rightarrow \tau_p(\beta)$

and consequently $S(A) \vdash \tau_p(A)$. This shows that starting from A we can derive in K $\tau_p(A)$. The proof is finished. ■

The following statement is an analog of the Sahlqvist's Theorem.

THEOREM 16 (The Main Theorem) *The following conditions are true for each complex Sahlqvist formula A:*

(i) A defines locally a first-order condition $F_A(x)$ which can effectively be computed from A.

(ii) A is canonical.

Proof. The theorem is a corollary of the Translation Lemma. ■

5 Extension of the Sahlqvist–van Benthem substitution method to the complex Sahlqvist formulas

The Translation Lemma gives an algorithm for finding the corresponding first-order condition for a given complex Sahlqvist formula A: first translate A to some original Sahlqvist formula $\tau(A)$ and then apply the standard

Sahlqvist-van Benthem substitution method [1]. Note that this method can not be directly applied to complex Sahlqvist formulas. The reason is that it is applicable only to formulas having the format of standard Sahlqvist formulas and strongly uses the fact that boxes in the boxed atoms are in front of propositional variables. If we look at the set-theoretical form of this method in [8] or in [3] we can see that the algorithm eliminates the second-order universal quantifiers on propositional variables (considered as ranging over subsets of the universe) *one by one in an arbitrary order*, making suitable "minimal" substitution of the corresponding variable. This feature characterizes the substitutions in Sahlqvist-van Benthem algorithm as *simple substitutions*, substitutions only for one given variable in each step. The fact that boxes in complex boxed atoms are in front of elementary disjunctions of several propositional variables suggests to consider "simultaneous substitutions" or substitutions of several variables simultaneously in a given step of the algorithm. This idea works and gives rise to a suitable extension of the Sahlqvist-van Benthem substitution method, now as a "method of simultaneous substitutions". This extended version of the substitution method works directly on complex Sahlqvist formulas for computing their first-order equivalents without going trough their translations into the original Sahlqvist's class.

We will demonstrate the new method by an example. Consider the formula from example 3:

$A = A(p,q) = \Diamond\Box(p \vee q) \wedge \Diamond\Box(p \vee \neg q) \wedge \Diamond\Box(\neg p \vee q) \Rightarrow \Box\Diamond(p \wedge q)$.

The corresponding first-order condition $F(x)$ is:

$(\forall y_1 y_2 y_3 y_4)(xRy_1 \& xRy_2 \& xRy_3 \& xRy_4 \rightarrow (\exists z)(y_1Rz \& y_2Rz \& y_3Rz \& y_4Rz))$

We consider $A(p,q)$ as a set-theoretical expression over the modal algebra of a given frame (W,R). The variables p,q range over subsets of W and the variables $x, y_1, \ldots, y_4$ – range over W. The (local) validity of $A(p,q)$ at a point $x \in W$ means the validity of the following second-order formula in (W,R):

$(\forall pq)(x \in A(p,q))$.

Rewriting it in an equivalent way using the known notations from Boolean modal set-algebras we obtain ($-p$ denotes the complement of p and $x \in \Box p$ is equivalent to $R(x) \subseteq p$,where $y \in R(x)$ iff xRy):

(1) $(\forall pq)(\forall y_1 y_2 y_3)(xRy_1 \& xRy_2 \& xRy_3 \& R(y_1) \subseteq (p \cup q) \& R(y_2) \subseteq (p \vee -q) \& R(y_3) \subseteq (-p \cup q) \rightarrow x \in \Box\Diamond(p \cap q))$.

By the Boolean Lemma there exist subsets $p, q \subseteq W$ such that the following equations are satisfied:

(2) $R(y_1) \subseteq (p \cup q)$, $R(y_2) \subseteq (p \cup -q)$, $R(y_3) \subseteq (-p \cup q)$, $R(y_1) \cap R(y_2) \cap R(y_3) = p \cap q$.

Substituting p and q from (2) in (1) we obtain:

(3) $(\forall y_1 y_2 y_3)(xRy_1 \& xRy_2 \& xRy_3 \rightarrow x \in \Box\Diamond(R(y_1) \cap R(y_2) \cap R(y_3))$.

Rewriting $x \in \Box\Diamond(R(y_1) \cap R(y_2) \cap R(y_3))$ and substituting the result in (3) we obtain just the required first-order condition $F(x)$ which is the local correspondent of $A(p,q)$.

This shows that if $x \in A(p,q)$ holds in (W,R) for all subsets p,q then the first-order formula $F(x)$ holds in (W,R) (for that given $x \in W$). Note that to obtain $F(x)$ we have substantially used the Boolean Lemma to find the appropriate substitutions for p and q. This "simultaneous" substitution is in a sense " minimal": the obtained condition $F(x)$ from this "minimal" assumption guarantees the converse implication: if $F(x)$ holds in (W,R) (for that given $x \in W$) then we have $(\forall pq)(x \in A(p,q))$.

Indeed, let p,q be arbitrary subsets of W. In order to show that $x \in A(p,q)$ suppose

(4) $\quad x \in \Diamond\Box(p \cup q) \cap \Diamond\Box(p \cup -q) \cap \Diamond\Box(-p \cup q)$

and proceed to show that

(5) $\quad x \in \Box\Diamond(p \cap q)$.

The condition (4) can be rewritten as follows:

(6) $\quad (\exists y_1 y_2 y_3)(xRy_1 \& xRy_2 \& xRy_3)$
such that

(7) $\quad R(y_1) \subseteq (p \cup q)$, $R(y_2) \subseteq (p \vee -q)$ and $R(y_3) \subseteq (-p \cup q)$.

From (7) by simple Boolean transformations we obtain

(8) $\quad R(y_1) \cap R(y_2) \cap R(y_3) \subseteq (p \cap q)$.

By the monotonicity of $\Box$ and $\Diamond$ we obtain from (8)

(9) $\quad \Box\Diamond(R(y_1) \cap R(y_2) \cap R(y_3)) \subseteq \Box\Diamond(p \cap q)$.

From (6) and $F(x)$ (in the form of (3)) we obtain
(10) $\quad x \in \Box\Diamond(R(y_1) \cap R(y_2) \cap R(y_3))$.

Then by (9) we obtain $x \in \Box\Diamond(p,q)$, i.e. (5), which had to be proved. Note that in this direction there were no need of the Boolean Lemma – the only thing from which the proof depends essentially is the monotonicity of the formula $\Box\Diamond(p \wedge q)$.

Let us note that the procedure demonstrated in the above example can be extended to an algorithm for obtaining the first-order equivalents for all complex Sahlqvist formulas which extends the Sahlqvist-van Benthem substitution method. The extension is based on the Boolean Lemma which guarantees the existence of a suitable simultaneous substitution of the complex variables.

REMARK 17 The method of simultaneous substitutions gives another motivation for complex Sahlqvist formulas: if the standard Sahlqvist-van Benthem method considered as a method of "simple substitutions" corresponds to the original Sahlqvist's class, then the method of "simultaneous substitutions" naturally corresponds to the class of complex Sahlqvist formulas. There is however a third kind of substitutions – compound or composed: the result of one substitution is used to define another substitution. Is there a class of "extended" Sahlqvist formulas corresponding to such a kind of substitutions? Yes, this is just the extension, given in [5]. In this sense the class of complex Sahlqvist formulas completes the picture: for each of the three kinds of substitutions – simple, compound and simultaneous – we have the most natural class of modal formulas for which the corresponding method for obtaining first-order equivalents works.

6 Concluding remarks

This paper has to be considered as an initial step to modal definability in languages with finitely many propositional variables with some applications to Sahlqvist theory. The notion of minimal modal equivalent of a first-order condition has been introduced and studied. Here an interesting phenomenon has been discovered: there are first-order conditions corresponding to Sahlqvist formulas with minimal modal equivalents not in the Sahlqvist's class. This fact suggests the idea of introducing a new extension of the class of Sahlqvist formulas, called "complex Sahlqvist formulas". For reasons of simplicity this extension is given for the simplest modal language but the definition and the obtained results can be easily extended to the case of Sahqvist formulas in the polyadic modal languages introduced in [5].

Complex Sahlqvist formulas violate one of the forbidden conditions for ordinary Sahlqvist formulas – disjunctions preceded by boxes in the Sahlqvist antecedents. Nevertheless complex Sahlqvist formulas possess two nice properties of the ordinary Sahlqvist formulas: they define locally first-order conditions and are canonical. A suitable extension of the Sahlqvist-van Benthem substitution method for obtaining (local) first-order equivalents is adapted for the class of complex Sahlqvist formulas called here the method of "simultaneous substitutions". The method of simultaneous substitutions can be considered also as a method of eliminating second-order quantifiers. In the literature there are other methods for eliminating second-order quantifiers, which have also applications in definability theory in modal logic. One for instance is the method SCAN described in [4]. It can be said that this paper is in a sense a result of experiments with SCAN: SCAN presented me first-order conditions for various non-Sahlqvist formulas for which there were no existing theory and methods to deal with. The present paper is an attempt for building of such a theory.

There are of course many natural open problems and questions which can be formulated by everyone interested in the field and in the present subject. I will mention only one: find enough invariants of the original Sahlqvist's class making possible to prove statements in the following form: "there is no Sahlqvist formula having such and such property". Example. We proved in Proposition 13 that the non-Sahlqvist formula $\mathbf{CR}_{k-1}^{min}$ having $n+1$ propositional variables is a minimal modal equivalent of the first-order condition $CR_{k-1}(x)$. It will be interesting to know if there exists a (classical) *Sahlqvist* formula with $n+1$ variables which is a minimal modal equivalent of $CR_{k-1}(x)$. For some cases such a formula exists. For instance for $k=3$ the Sahlqvist formula $\Diamond\Box q_1 \wedge \Diamond\Box q_2 \Rightarrow \Box\Diamond(q_1 \wedge q_2)$ of two variables is a minimal equivalent of $CR_{3-1}(x)$. But for $k=4$ I conjecture that such a Sahlqvist formula of two variables does not exist. At this stage however our knowledge about the structure of the Sahlqvist's class is too limited for proving such statements.

And one more question before the end: can we extend further the class of complex Sahlqvist formulas? Probably yes. Just before preparing the final version of this paper I have found several examples of formulas which are candidates for a such an extension. Here is one which defines (locally) a first-order condition and is canonical:

$\Diamond\Box(p_1 \vee \Box(q \vee r)) \wedge \Diamond\Box(\neg p_1 \vee \Box(q \vee r)) \wedge \Diamond\Box(p_2 \vee \Box(q \vee \neg r)) \wedge \Diamond\Box(\neg p_2 \vee \Box(q \vee \neg r)) \wedge \Diamond\Box(p_3 \vee \Box(\neg q \vee r)) \wedge \Diamond\Box(\neg p_3 \vee \Box(\neg q \vee r)) \wedge \Diamond\Box(p_4 \vee \Box(\neg q \vee \neg r)) \wedge \Diamond\Box(\neg p_4 \vee \Box(\neg q \vee \neg r)) \Rightarrow \bot$.

In this formula "boxed atoms" are not boxed pure elementary disjunctions but contain also boxes inside the elementary disjunctions . The proof of first-order definability and canonicity is based on some theory of "solving equations" in modal algebras similar to those in the Boolean Lemma.

Acknowledgements

I am grateful to Andreas Herzig for introducing me to SCAN and to Renate Schmidt and Valentin Goranko for discussing some problems related to SCAN and Sahlqvist's class.

BIBLIOGRAPHY

[1] Benthem, J. van 1983. *Modal Logic and Classical logic.* Naples: Bibliopolis.

[2] Blackburn, P., M. de Rijke, and Y. Venema. 2000. *Modal Logic.* Cambridg Tracts in Theoretical Computer Science.

[3] Chagrov, A., and M. Zakharyaschev. 1997. *Modal Logic.* Oxford Logic Guides, Vol. 35. Oxford: Oxford University Press.

[4] Gabbay, D., and H. J. Ohlbach. 1992. Quantifier elimination in second-order predicate logic. In: Proceedings of the KR'92, *Princiles of Knowledge Representation and Reasoning*, Morgan Kaufmann, 425-436.
[5] Goranko, V., and D. Vakarelov.2002. Sahlqvist Formulas Unleashed in Polyadic Modal Languages. In F. Wolter, H. Wansing, M. de Rijke, and M. Zakharyaschev, eds. *Advances in Modal Logic, vol. 3*, World Scientific, New Jersey, London, Singapore, Hong Kong, 221–240.
[6] de Rijke, M., and Y. Venema. 1995. Sahlqvist's Theorem for Boolean algebras with Operators with an Application to Cylindric Algebras. *Studia Logica* 54:61-78.
[7] Sahlqvist, H. 1975. Completeness and Correspondence in the First and Second Order Semantics for Modal Logic. In: Kanger, S. (ed.) *Proceedings of the Third Scandinavian Logic Symposium. Uppsala 1973.* Amsterdam, North-Holland.
[8] Sambin, G., and V. Vaccaro. 1989. A new Proof of Sahqvist's Theorem on Modal definability and Completeness. *Journal of Symbolic Logic* 54:992-999.
[9] Segerberg, K. 1971. *An Essay in Classical Modal Logic*, Filosofiska Studier 13, Department of Philosophy, University of Uppsala.

Dimiter Vakarelov
Department of Mathematical Logic
Faculty of Mathematics and Computer Science
Sofia University
blvd James Bouchier 5, 1126 Sofia, Bulgaria
E-mail: dvak@fmi.uni-sofia.bg