

# Dependence logic with a majority quantifier\*

Arnaud Durand<sup>1</sup>, Johannes Ebbing<sup>2</sup>, Juha Kontinen<sup>3</sup>, and Heribert Vollmer<sup>2</sup>

- 1 Université Paris Diderot, IMJ, CNRS UMR 7586, Case 7012, 75205 Paris cedex 13, France, durand@logique.jussieu.fr
- 2 Leibniz Universität Hannover, Theoretical Computer Science, Appelstr. 4, 30167 Hannover, Germany, {ebbing,vollmer}@thi.uni-hannover.de
- 3 University of Helsinki, Department of Mathematics and Statistics, P.O. Box 68, 00014, Finland, juha.kontinen@helsinki.fi.

---

## Abstract

We study the extension of dependence logic  $\mathcal{D}$  by a majority quantifier  $M$  over finite structures. We show that the resulting logic is equi-expressive with the extension of second-order logic by second-order majority quantifiers of all arities. Our results imply that, from the point of view of descriptive complexity theory,  $\mathcal{D}(M)$  captures the complexity class counting hierarchy.

**1998 ACM Subject Classification** F.1.1 Computability theory, Relations between models; F.1.3 Complexity hierarchies; F.2.2 Computations on discrete structures; F.4.1 Computability theory, Computational logic, Model theory;

**Keywords and phrases** dependence logic, counting hierarchy, majority quantifier, second order logic, descriptive complexity, finite model theory

**Digital Object Identifier** 10.4230/LIPIcs.FSTTCS.2011.252

## 1 Introduction

We study the extension of dependence logic  $\mathcal{D}$  by a majority quantifier  $M$  over finite structures. Dependence logic [19] extends first-order logic by dependence atomic formulas

$$=(t_1, \dots, t_n)$$

the intuitive meaning of which is that the value of the term  $t_n$  is completely determined by the values of  $t_1, \dots, t_{n-1}$ . While in first-order logic the order of quantifiers solely determines the dependence relations between variables, in dependence logic more general dependencies between variables can be expressed. Historically dependence logic was preceded by partially ordered quantifiers (Henkin quantifiers) of Henkin [8] and Independence-Friendly (IF) logic of Hintikka and Sandu [9]. It is known that both IF logic and dependence logic are equivalent to existential second-order logic ESO in expressive power. From the point of view of descriptive complexity theory, this means that dependence logic captures the class NP.

The framework of dependence logic has turned out to be flexible to allow interesting generalizations. For example, the extensions of dependence logic in terms of so-called intuitionistic implication and linear implication was introduced in [1]. In [23] it was shown that extending  $\mathcal{D}$  by the intuitionistic implication makes the logic equivalent to full second-order logic SO.

---

\* The third author was supported by grants 127661 and 138163 of the Academy of Finland. The second and fourth author were supported by a grant from DAAD within the PPP programme. The fourth author was also supported by DFG grant VO 630/6-2.



© A. Durand, J. Ebbing, J. Kontinen, and H. Vollmer;  
licensed under Creative Commons License NC-ND

31<sup>st</sup> Int'l Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2011).

Editors: Supratik Chakraborty, Amit Kumar; pp. 252–263



Leibniz International Proceedings in Informatics

LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Recently, new variants of the dependence atomic formulas have been introduced in [7] and [6]. Also a modal version of dependence logic was introduced in [20] and has been studied in [14] and [15]. In this paper we are concerned with introducing a new quantifier to dependence logic: the majority quantifier. Adding majority and, more generally, counting capabilities to logical formalisms or computational devices has deserved a lot of attention in theoretical computer science. Understanding the power of counting is an important problem both in logic and in computational complexity:

- The circuit class  $TC^0$ , the class of problems solvable by polynomial-size constant-depth circuits with majority gates, is at the current frontier for lower bound techniques (see, e.g., [21]). We have strict separations of classes within  $TC^0$ , but above  $TC^0$  we have essentially no lower bounds. By a diagonalization it follows that  $TC^0$  is different from the second level of the exponential-time hierarchy and that uniform  $TC^0$  is strictly included in the class PP of probabilistic polynomial time [2], but a separation from a lower class seems to be far away. In particular, the question if  $TC^0$  equals  $NC^1$  (logarithmic-depth circuits with bounded fan-in gates) is considered the P-NP problem of circuit complexity.
- The counting-hierarchy (the oracle hierarchy built upon PP) can be characterized using majority quantifiers in just the same way as by Wrathall's theorem existential and universal quantifiers characterize the polynomial hierarchy [17].
- By Toda's theorem, one majority quantifier is as powerful as the whole polynomial hierarchy [16].

Here we suggest a definition of a majority quantifier for dependence logic. The proposed semantics mimics that of the existential and universal quantifiers in  $\mathcal{D}$ . The present paper is devoted to a first study of the resulting logic, denoted by  $\mathcal{D}(M)$ . We examine some of its basic properties, prove strong normal forms (some of our technically most involved proofs are found here), and show in our main result, that dependence logic with the majority quantifier leads to a new descriptive complexity characterization of the counting hierarchy:  $\mathcal{D}(M)$  captures CH.

Engström [5] has also studied generalized quantifiers in dependence logic. He considered different conservative extensions of  $\mathcal{D}$ —informally this means that he extends  $\mathcal{D}$  by generalized quantifiers in a first-order manner. From a descriptive complexity point of view, his logics do not lead out of NP, i.e., ESO, assuming the quantifier in question is ESO-definable (e.g., the majority quantifier). Our approach and results differ from that of Engström since we are in a sense extending dependence logic by a dependence majority quantifier, whose semantics is defined in close analogy with the semantics of  $\exists$  and  $\forall$  in dependence logic. The results of our paper show that our extension behaves like an extension of SO by second-order generalized quantifiers.

This article is organized as follows. In Sect. 2 we defined dependence logic and discuss some basic results on it. Then we introduce a majority quantifier for the dependence logic setting and discuss the basic properties of  $\mathcal{D}(M)$ . In Subsect. 2 we discuss the complexity class counting hierarchy and the second-order majority quantifiers  $\text{Most}^k$  that have been used to characterize it in [10]. In Sect. 3, we introduce second-order majority quantifiers  $\text{Most}_f^k$  ranging over functions and in Sect. 4 we show that, for sentences the logics  $\text{SO}(\text{Most}_f)$  (the extension of second-order logic SO by  $\text{Most}_f^k$  for  $k \geq 1$ ) and  $\mathcal{D}(M)$  are equivalent. Due to space restrictions, some proofs have to be omitted in this paper, but can be found in the full version at <http://arxiv.org/abs/1109.4750>.

## 2 Preliminaries

In this section we first define dependence logic and discuss its basic properties. Then we define the counting hierarchy and the logic corresponding to it.

### 2.1 Dependence Logic

Dependence logic ( $\mathcal{D}$ ) extends the syntax of first-order logic by new dependence atomic formulas. In this article we consider only formulas of  $\mathcal{D}$  that are in negation normal form.

► **Definition 2.1** ([19]). Let  $\tau$  be a vocabulary. The  $\tau$ -formulas of dependence logic ( $\mathcal{D}[\tau]$ ) is defined by extending  $\text{FO}[\tau]$ , defined in terms of  $\vee, \wedge, \neg, \exists$  and  $\forall$ , by atomic dependence formulas

$$=(t_1, \dots, t_n), \quad (1)$$

where  $t_1, \dots, t_n$  are terms.

The meaning of the formula (1) is that the value of the term  $t_n$  is functionally determined by the values of the terms  $t_1, \dots, t_{n-1}$ . The formula  $=()$  is interpreted as  $\top$ . The semantics of  $\mathcal{D}$  will be formally presented shortly.

► **Definition 2.2.** Let  $\phi \in \mathcal{D}$ . The set  $\text{Fr}(\phi)$  of free variables of a formula  $\phi$  is defined as for first-order logic, except that we have the new case

$$\text{Fr}(=(t_1, \dots, t_n)) = \text{Var}(t_1) \cup \dots \cup \text{Var}(t_n),$$

where  $\text{Var}(t_i)$  is the set of variables occurring in term  $t_i$ . If  $\text{Fr}(\phi) = \emptyset$ , we call  $\phi$  a sentence.

The semantics of  $\mathcal{D}$  is formulated using the concept of a *Team*. Let  $\mathfrak{A}$  be a model with domain  $A$ . *Assignments* of  $\mathfrak{A}$  are finite mappings from variables into  $A$ . The value of a term  $t$  in an assignment  $s$  is denoted by  $t^{\mathfrak{A}}\langle s \rangle$ . If  $s$  is an assignment,  $x$  a variable, and  $a \in A$ , then  $s(a/x)$  denotes the assignment (with domain  $\text{dom}(s) \cup \{x\}$ ) that agrees with  $s$  everywhere except that it maps  $x$  to  $a$ .

► **Definition 2.3.** Let  $A$  be a set and  $\{x_1, \dots, x_k\}$  a finite (possibly empty) set of variables.

1. A *team*  $X$  of  $A$  with domain  $\text{dom}(X) = \{x_1, \dots, x_k\}$  (we call  $A$  the *co-domain* of  $X$ ) is any set of assignments  $s: \{x_1, \dots, x_k\} \rightarrow A$ .
2. The relation  $\text{rel}(X) \subseteq A^k$  corresponding to  $X$  is defined as

$$\text{rel}(X) = \{(s(x_1), \dots, s(x_k)) : s \in X\}.$$

3. For a function  $F: X \rightarrow A$ , we define

$$\begin{aligned} X(F/x) &= \{s(F(s)/x) : s \in X\} \\ X(A/x) &= \{s(a/x) : s \in X \text{ and } a \in A\}. \end{aligned}$$

We will next define the semantics of dependence logic. Below, atomic formulas and their negations are called literals.

► **Definition 2.4** ([19]). Let  $\mathfrak{A}$  be a model and  $X$  a team of  $A$ . The satisfaction relation  $\mathfrak{A} \models_X \phi$  is defined as follows:

1. If  $\phi$  is a first-order literal, then  $\mathfrak{A} \models_X \phi$  iff for all  $s \in X$  we have  $\mathfrak{A} \models_s \phi$ .

2.  $\mathfrak{A} \models_X (t_1, \dots, t_n)$  iff for all  $s, s' \in X$  such that  $t_1^{\mathfrak{A}}(s) = t_1^{\mathfrak{A}}(s'), \dots, t_{n-1}^{\mathfrak{A}}(s) = t_{n-1}^{\mathfrak{A}}(s')$ , we have  $t_n^{\mathfrak{A}}(s) = t_n^{\mathfrak{A}}(s')$ .
3.  $\mathfrak{A} \models_X \neg(t_1, \dots, t_n)$  iff  $X = \emptyset$ .
4.  $\mathfrak{A} \models_X \psi \wedge \phi$  iff  $\mathfrak{A} \models_X \psi$  and  $\mathfrak{A} \models_X \phi$ .
5.  $\mathfrak{A} \models_X \psi \vee \phi$  iff  $X = Y \cup Z$  such that  $\mathfrak{A} \models_Y \psi$  and  $\mathfrak{A} \models_Z \phi$ .
6.  $\mathfrak{A} \models_X \exists x \psi$  iff  $\mathfrak{A} \models_{X(F/x)} \psi$  for some  $F: X \rightarrow A$ .
7.  $\mathfrak{A} \models_X \forall x \psi$  iff  $\mathfrak{A} \models_{X(A/x)} \psi$ .

Above, we assume that the domain of  $X$  contains the variables free in  $\phi$ . Finally, a sentence  $\phi$  is true in a model  $\mathfrak{A}$  (in symbols:  $\mathfrak{A} \models \phi$ ) if  $\mathfrak{A} \models_{\{\emptyset\}} \phi$ . Above,  $A \models_s \phi$  denotes satisfaction in first-order logic.

Let us then recall some basic properties of dependence logic that will be needed later. The following lemma shows that the truth of a  $\mathcal{D}$ -formula depends only on the interpretations of variables occurring free in the formula. Below, for  $V \subseteq \text{dom}(X)$ ,  $X \upharpoonright V$  is defined by

$$X \upharpoonright V := \{s \upharpoonright V \mid s \in X\} \text{ and}$$

$$s \upharpoonright V := \{(a, s(a)) \mid a \in \text{dom}(s) \cap V\}.$$

► **Lemma 2.5** ([19]). *Suppose  $V \supseteq \text{Fr}(\phi)$ . Then  $\mathfrak{A} \models_X \phi$  if and only if  $\mathfrak{A} \models_{X \upharpoonright V} \phi$ .*

All formulas of dependence logic also satisfy the following strong monotonicity property called *Downward Closure*.

► **Proposition 2.6** ([19]). *Let  $\phi$  be a formula of dependence logic,  $\mathfrak{A}$  a model, and  $Y \subseteq X$  teams. Then  $\mathfrak{A} \models_X \phi$  implies  $\mathfrak{A} \models_Y \phi$ .*

On the other hand, the expressive power of sentences of  $\mathcal{D}$  coincides with that of existential second-order sentences:

► **Theorem 2.7** ([19]).  $\mathcal{D} = \text{ESO}$ .

Finally, we note that dependence logic is a conservative extension of first-order logic.

► **Definition 2.8.** A formula  $\phi$  of  $\mathcal{D}$  is called a first-order formula if it does not contain dependence atomic formulas as subformulas.

First-order formulas of dependence logic satisfy the so-called *flatness* property:

► **Theorem 2.9** ([19]). *Let  $\phi$  be a first-order formula of dependence logic. Then for all  $\mathfrak{A}$  and  $X$ :*

$$\mathfrak{A} \models_X \phi \text{ if and only if for all } s \in X \text{ we have } \mathfrak{A} \models_s \phi.$$

## 2.2 Dependence logic with a majority quantifier

The main topic of the present paper is the study of a logic obtained from  $\mathcal{D}$  by the introduction of a majority quantifier  $M$ . We denote this extended logic by  $\mathcal{D}(M)$ . It is formally defined by extending the syntax and semantics of dependence logic by the following clause:

$$\mathfrak{A} \models_X Mx\phi(x) \text{ iff for at least } |A|^{|X|}/2 \text{ many functions } F: X \rightarrow A \text{ we have } \mathfrak{A} \models_{X(F/x)} \phi(x).$$

Analogously to  $\mathcal{D}$  the logic  $\mathcal{D}(M)$  has the so-called empty team property:

► **Proposition 2.10.** *For all models  $\mathfrak{A}$  and formulas  $\phi$  of  $\mathcal{D}(M)$ , it holds that  $\mathfrak{A} \models_{\emptyset} \phi$ .*

**Proof.** The claim is proved using induction on  $\phi$ . ◀

We also observe that  $\mathcal{D}(\mathbf{M})$  satisfies the downward closure property (compare to Proposition 2.6).

► **Proposition 2.11.** *Let  $\phi$  be a formula of  $\mathcal{D}(\mathbf{M})$ ,  $\mathfrak{A}$  a model, and  $Y \subseteq X$  teams. Then  $\mathfrak{A} \models_X \phi$  implies  $\mathfrak{A} \models_Y \phi$ .*

**Proof.** The claim is proved using induction on  $\phi$ . We consider the case where  $\phi$  is  $\mathbf{M}x\psi$ . The other cases are proved exactly as for dependence logic (see Proposition 3.10 in [19]). By the induction assumption,  $\psi$  satisfies the claim. Let  $\mathfrak{A}$ ,  $X$  and  $Y$  be as above and suppose that  $|A| = n$ ,  $|X| = m$ , and  $|Y| = m - 1$ . Let us assume  $\mathfrak{A} \models_X \phi$ . Then for at least  $(n^m)/2$  many functions  $F: X \rightarrow A$  it holds that  $\mathfrak{A} \models_{X(F/x)} \psi$ . Since  $\psi$  satisfies the claim, it holds that if  $\mathfrak{A} \models_{X(F/x)} \psi$ , then  $\mathfrak{A} \models_{Y(F'/x)} \psi$ , where

$$F' = F \upharpoonright Y. \quad (2)$$

Note that, in the worst case, at most  $n$  different functions  $F$  gives rise to the same reduct  $F'$  in (2). Therefore, the number of functions  $F: Y \rightarrow A$  satisfying  $\mathfrak{A} \models_{Y(F'/x)} \psi$  is at least  $(n^m)/2n = n^{m-1}/2$  and hence  $\mathfrak{A} \models_Y \phi$ . It is easy to see that the analogous argument can be used with any  $Y \subseteq X$ . ◀

A well-studied property in the context of dependence logic is that of *coherence*, defined as follows. A formula  $\phi$  is called *k-coherent* if and only if for all structures  $\mathfrak{A}$  and teams  $X$  it holds that

$$\mathfrak{A} \models_X \phi \Leftrightarrow \text{for every } k\text{-element subteam } X' \subseteq X \text{ it holds that } \mathfrak{A} \models_{X'} \phi.$$

1-coherent formulas are also called *flat*.

► **Proposition 2.12.** *There is a formula  $\phi \in \mathcal{D}(\mathbf{M})$  without dependence atoms such that  $\phi$  is not *k-coherent* for any  $k \in \mathbb{N}$ .*

We also note that the analogue of Proposition 2.5 does not hold for  $\mathcal{D}(\mathbf{M})$ .

► **Proposition 2.13.** *The truth of a  $\mathcal{D}(\mathbf{M})$ -formula  $\phi$  may depend on the interpretations of variables that do not occur free in  $\phi$ .*

### 2.3 Second-order Majority Quantifiers and the Counting Hierarchy

In this section we define the counting hierarchy and the relevant generalized quantifiers.

► **Definition 2.14.** Let  $k \geq 1$ . We define the *k-ary second-order generalized quantifier*  $\text{Most}^k$  binding a *k-ary relation symbol*  $X$  in a formula  $\phi$ . Assume  $\mathfrak{A}$  is a structure with domain  $A$  such that  $|A| = n$ . Then the semantics of this quantifier is defined as follows:

$$\mathfrak{A} \models \text{Most}^k X \phi(X) \Leftrightarrow |\{B \subseteq A^k \mid \mathfrak{A} \models \phi(B)\}| \geq 2^{n^k}/2.$$

We will also make use of the so-called *k-ary second-order Rescher quantifier*, defined as follows:

$$\mathfrak{A} \models \mathbf{R}^k X, Y(\phi(X), \psi(Y)) \Leftrightarrow |\{B \subseteq A^k \mid \mathfrak{A} \models \phi(B)\}| \geq |\{B \subseteq A^k \mid \mathfrak{A} \models \psi(B)\}|.$$

It is quite easy to see that the  $\text{Most}^k$ -quantifier can be defined in terms of the quantifier  $\text{R}^k$ . In [10] it was shown that the  $k$ -ary Rescher quantifier  $\text{R}^k$  can be defined in first order logic with  $\text{Most}^{k+1}$ , and, for  $k \geq 2$ , already with  $\text{Most}^k$ . It is worth noting that in [10] the quantifiers  $\text{Most}^k$  and  $\text{R}^k$  are interpreted as strict majority and strict inequality, respectively. All the results of [10] that we use also hold under the "non-strict" interpretation adopted in this article.

The counting hierarchy (CH) is the analogue of the polynomial hierarchy, defined as the oracle hierarchy using as building block probabilistic polynomial time (the class PP) instead of NP:

1.  $\text{C}_0\text{P} = \text{P}$ ,
2.  $\text{C}_{k+1}\text{P} = \text{PP}^{\text{C}_k\text{P}}$ ,
3.  $\text{CH} = \bigcup_{k \in \mathbb{N}} \text{C}_k\text{P}$ .

The counting hierarchy was first defined by Wagner [22] but the above equivalent formulation is due to Torán [17]. The class PP consists of languages  $L$  for which there is a polynomial time-bounded nondeterministic Turing machine  $N$  such that, for all inputs  $x$ ,  $x \in L$  iff more than half of the computations of  $N$  on input  $x$  accept.

In [10] it was shown that the extension  $\text{FO}(\text{Most})$  of FO by the quantifiers  $\text{Most}^k$ , for  $k \in \mathbb{N}$ , describes exactly the problems in the counting hierarchy. The proof therein used the fact that the second-order existential quantifier can be simulated by  $\text{Most}^k$  and first-order logic.

► **Theorem 2.15.**  $\text{FO}(\text{Most}) = \text{SO}(\text{Most}) = \text{CH}$ .

By the above remark we see that in the previous theorem the  $\text{Most}$  quantifiers can be replaced by Rescher quantifiers.

### 3 Majority over Functions

For our main result that compares second-order logic and dependence logic with majority-quantifiers, it turns out to be helpful to consider a version of the  $\text{Most}$ -quantifier that ranges over functions instead of relations.

► **Definition 3.1.** Let  $k \geq 1$ . We define the  $k$ -ary second-order generalized quantifier  $\text{Most}_f^k$  binding a  $k$ -ary function symbol  $g$  in a formula  $\phi$ . Assume  $\mathfrak{A}$  is a structure with domain  $A$  such that  $|A| = n$ . Then

$$\mathfrak{A} \models \text{Most}_f^k g \phi(g) \iff |\{f: A^k \rightarrow A \mid \mathfrak{A} \models \phi(f)\}| \geq n^{n^k}/2.$$

We denote by  $\text{SO}(\text{Most}_f)$  the extension of SO by the quantifiers  $\text{Most}_f^k$  for all  $k \geq 1$ . The following elementary properties of  $\text{SO}(\text{Most}_f)$  will be useful.

► **Proposition 3.2.** *The following equivalences hold:*

1.  $(\phi \vee \text{Most}_f^k g \psi) \equiv \text{Most}_f^k g (\phi \vee \psi)$ , if  $g$  does not appear free in  $\phi$ ,
2.  $(\phi \wedge \text{Most}_f^k g \psi) \equiv \text{Most}_f^k g (\phi \wedge \psi)$ , if  $g$  does not appear free in  $\phi$ .

Note that since also  $\neg \text{Most}_f^k g \psi$  is equivalent to  $\text{Most}_f^k g \neg \psi$ , Proposition 3.2 allows us to transform formulas of  $\text{SO}(\text{Most}_f)$  to prenex normal form. The equivalences of Proposition 3.2 obviously hold also for the relational majority quantifiers  $\text{Most}^k$ .

The next proposition states the intuitively obvious fact that the extensions of SO by the quantifiers  $\text{Most}^k$  or alternatively by  $\text{Most}_f^k$ , for  $k \in \mathbb{N}$ , are equal in expressive power.

► **Proposition 3.3.**  $\text{SO}(\text{Most}) = \text{SO}(\text{Most}_f^k)$ .

**Proof.** We prove the claim by an argument analogous to Theorem 3.4 in [10]. We will show how to express the quantifier  $\text{Most}_f^k$  in the logic  $\text{SO}(\text{Most})$  implying  $\text{SO}(\text{Most}_f^k) \leq \text{SO}(\text{Most})$ . The converse inclusion is proved analogously.

Let us consider a formula of the form  $\text{Most}_f^k g \phi(g) \in \text{SO}(\text{Most}_f^k)$ . Let  $\mathfrak{A}$  be a structure. We may assume that  $\mathfrak{A}$  is ordered (we can existentially quantify it) and hence there is a FO-formula  $\delta(\bar{x}, \bar{y})$  defining the lexicographic ordering of the set  $A^{k+1}$ . We can construct a formula  $\chi(X, Y)$  which, for  $A_1, A_2 \subseteq A^{k+1}$ , defines the lexicographic ordering ( $A_1 \leq_l A_2$ ) of  $k+1$ -ary relations induced by  $\delta(\bar{x}, \bar{y})$ .

It is now fairly straightforward to express  $\text{Most}_f^k g \phi(g)$  in the logic  $\text{SO}(\text{Most})$ . Let

$$\begin{aligned} G &= \{B \subseteq A^{k+1} \mid B \text{ is the graph of some } g \text{ and } \mathfrak{A} \models \phi(g)\}, \\ G^c &= \{B \subseteq A^{k+1} \mid B \text{ is the graph of some } g \text{ and } \mathfrak{A} \not\models \phi(g)\}. \end{aligned}$$

It now suffices to express  $|G| \geq |G^c|$  in the logic  $\text{SO}(\text{Most})$ . For a  $D \subseteq A^{k+1}$ , define the set  $\text{IS}(D)$  (the “initial segment” determined by  $D$ ) by

$$\text{IS}(D) = \{D' \subseteq A^{k+1} \mid D' \not\subseteq G \cup G^c \text{ and } D' \leq_l D\}.$$

The condition  $|G| \geq |G^c|$  can be now expressed by

$$\forall D (|G^c \cup \text{IS}(D)| \geq 2^{n^{k+1}}/2 \Rightarrow |G \cup \text{IS}(D)| \geq 2^{n^{k+1}}/2).$$

It is straightforward to express this in the logic  $\text{SO}(\text{Most})$ . ◀

The following lemma will be needed in the proof of the next proposition.

► **Lemma 3.4.** *Let  $k \geq 1$ . There exists an ESO sentence  $\chi(g)$ , where  $g$  is  $k$ -ary, such that for all  $\mathfrak{A}$  with domain  $|A| = n$ ,  $\chi(g)$  is satisfied by exactly  $\lceil n^{n^k}/2 \rceil - 2^{n^k-1}$  many  $k$ -ary functions  $g$  none of which is a characteristic function of some  $k$ -ary relation.*

The next proposition gives a useful normal form for sentences of the logic  $\text{SO}(\text{Most}_f^k)$ .

► **Proposition 3.5.** *Every sentence of  $\text{SO}(\text{Most}_f^k)$  is equivalent to a sentence of the form*

$$\exists \bar{h}^1 \text{Most}_f^k g_1 \cdots \text{Most}_f^k g_l \exists \bar{h}^2 \theta,$$

where the function symbols in  $\bar{h}^1$ , and  $g_i$  for  $1 \leq i \leq l$ , are  $k$ -ary ( $k \geq 3$ ), and  $\theta$  is a universal first-order sentence.

**Proof.** Note that by Proposition 3.3 it suffices to show that every sentence of the logic  $\text{SO}(\text{Most})$  can be transformed to this form. The result in [10] shows (as pointed out in Lemma 10.5 in [11]) that, in the presence of built-in relations  $\{<, +, \times\}$ , sentences of  $\text{SO}(\text{Most})$  can be assumed to have the form

$$\text{Most}^{i_1} Y_1 \cdots \text{Most}^{i_l} Y_l \psi, \tag{3}$$

where  $\psi$  is first-order. Furthermore, when  $l$  in (3) is fixed, we get a fragment of  $\text{SO}(\text{Most})$  characterizing the  $l$ th level of CH, i. e., the class  $C_l\text{P}$ .

We will next show how to transform any sentence of the form (3) to the required form. The first step is to quantify out the built-in relations  $\{<, +, \times\}$  to get a sentence of the form

$$\exists X_{<} \exists X_{+} \exists X_{\times} \text{Most}^{i_1} Y_1 \cdots \text{Most}^{i_l} Y_l \psi^*. \tag{4}$$

The relations  $X_{<}$ ,  $X_{+}$ , and  $X_{\times}$  can be axiomatized as part of  $\psi^*$  (compare to case 2 of Proposition 3.2). Then we modify the sentence (4) to change the arities of all the quantified relations to some big enough  $k$ . We need only to replace all occurrences, say  $Y_i(t_1, \dots, t_{i_j})$ , of the quantified relation symbols in  $\psi^*$  by  $Y_i(t_1, \dots, t_{i_j}, 0, \dots, 0)$ . (Note that the needed constant 0 can be defined using the linear order.) Increasing the arity of the second-order existential quantifiers in (4) is clearly unproblematic. For the majority quantifiers  $\text{Most}^{i_j}$ , we note that for any structure  $\mathfrak{A}$  of cardinality  $n$  and  $B \subseteq A^v$ , the number of  $k$ -ary relations  $D \subseteq A^k$  such that

$$\{\bar{a} \in A^v \mid (\bar{a}, 0, \dots, 0) \in D\} = B \quad (5)$$

is  $2^{n^k - n^v}$ , which is independent of  $B$ . Furthermore, obviously the truth of  $\psi^*$  with respect to a tuple of  $k$ -ary relations  $D_1, \dots, D_{l+3}$  only depends on whether  $\psi^*(B_1, \dots, B_{l+3})$  holds, where  $B_i$  is the restriction of  $D_i$  defined analogously to (5). This fact allows us to increase also the arity of the majority quantifiers without changing the meaning of the sentence (4).

Let us then show how to transform the relational quantifiers in (4) into function quantifiers. We claim that it is possible to replace  $\psi^*(X_{<}, X_{+}, X_{\times}, Y_1, \dots, Y_l)$  by a formula of the form

$$\theta(\bar{g}) \vee (\forall \bar{x} (\bigwedge_{1 \leq i \leq l} g_i(\bar{x}) \in \{0, 1\}) \wedge \psi'(g_{<}/X_{<}, g_{+}/X_{+}, g_{\times}/X_{\times}, g_1/Y_1, \dots, g_l/Y_l)), \quad (6)$$

where  $\bar{g} = (g_{<}, g_{+}, g_{\times}, g_1, \dots, g_l)$ , the new function symbols are all  $k$ -ary and  $\psi'$  is obtained from  $\psi^*$  by substituting subformulas  $Z(t_1, \dots, t_k)$  by the corresponding  $g_{(\cdot)}(1_1, \dots, 1_k) = 1$ , where  $Z \in \{Y_1, \dots, Y_l, X_{<}, X_{+}, X_{\times}\}$ .

The formula  $\theta(\bar{g})$  is a ESO-formula that accepts certain dummy functions in order to shift the border of acceptance from  $(2^{|\mathfrak{A}|^k})/2$  (half of  $k$ -ary relations) to  $|\mathfrak{A}|^{|\mathfrak{A}|^k}/2$  (half of  $k$ -ary functions). The logical form of  $\theta$  is

$$\chi(g_1) \vee \chi(g_2) \vee \dots \vee \chi(g_l),$$

where  $\chi(g)$  is defined in Lemma 3.4. Note that we repeatedly use case 1 of Lemma 3.2 to gather all the formulas  $\chi(g_i)$  into  $\theta$  which is placed after the block of all majority quantifiers.

To prove the claim we finally transform the formula (6) into Skolem normal form to get a sentence of the form

$$\exists g_{<} \exists g_{+} \exists g_{\times} \text{Most}_f^k g_1 \dots \text{Most}_f^k g_l \exists \bar{g} \psi', \quad (7)$$

where  $\psi'$  is a universal FO-sentence. ◀

#### 4 SO(Most) = $\mathcal{D}(\text{M})$

In this section we show that the logics  $\text{SO}(\text{Most}_f)$  (and thus, by the previous section,  $\text{SO}(\text{Most})$ ) and  $\mathcal{D}(\text{M})$  are equivalent with respect to sentences.

We will first show a compositional translation mapping formulas of  $\mathcal{D}(\text{M})$  into sentences of  $\text{SO}(\text{Most}_f)$ . This translation is analogous to the translation from  $\mathcal{D}$  into ESO of [19].

► **Lemma 4.1.** *Let  $\tau$  be a vocabulary. For every  $\mathcal{D}(\text{M})[\tau]$ -formula  $\phi$  there is a  $\tau \cup \{S\}$ -sentence  $\psi$  of  $\text{SO}(\text{Most}_f)$  such that for all models  $\mathfrak{A}$  and teams  $X$  with  $\text{dom}(X) = \text{Fr}(\phi)$  it holds that*

$$\mathfrak{A} \models_X \phi \iff (\mathfrak{A}, \text{rel}(X)) \models \psi.$$



**Proof.** For technical reasons to be motivated shortly, we will actually prove a slightly more general result showing that for every  $\mathcal{D}(\mathbf{M})[\tau]$ -formula  $\phi$  and every finite set of variables  $\{y_1, \dots, y_n\} \supseteq \text{Fr}(\phi)$  there is a  $\text{SO}(\text{Most}_f)[\tau \cup S]$ -sentence  $\psi$  such that for all  $\mathfrak{A}$  and teams  $X$  with  $\text{dom}(X) = \{y_1, \dots, y_n\}$  it holds that

$$\mathfrak{A} \models_X \phi \iff (\mathfrak{A}, \text{rel}(X)) \models \psi.$$

We will prove the claim using induction on the structure of  $\mathcal{D}(\mathbf{M})$ -formulas. In the following we write  $\phi(y_1, \dots, y_n)$  to mean that  $\text{Fr}(\phi) \subseteq \{y_1, \dots, y_n\}$ . The quantifiers  $R^k$  can be uniformly defined in the logic  $\text{SO}(\text{Most})$ , hence by the results of the previous section, also in  $\text{SO}(\text{Most}_f)$ . Therefore, we may freely use the quantifiers  $R^k$  in the translation.

Atomic formulas and their negations are translated exactly in the same way as in the analogous translation from  $\mathcal{D}$  into ESO in [19]. The cases  $\gamma := \exists y_n \phi(y_1, \dots, y_n)$  and  $\gamma := \forall y_n \phi(y_1, \dots, y_n)$  are also translated as in [19]. Suppose then that  $\gamma := (\phi \vee \psi)(y_1, \dots, y_n)$  and that  $\phi^*(S)$  and  $\psi^*(S)$  already exist by induction hypothesis. We translate  $\gamma$  as follows:

$$\gamma^*(S) := \exists Y \exists Z (\phi^*(Y/S) \wedge \psi^*(Z/S) \wedge \forall y_1 \dots \forall y_n (S(\bar{y}) \rightarrow R(\bar{y}) \vee T(\bar{y}))). \quad (8)$$

Note that  $\gamma^*(S)$  is defined as in [19]. The only difference is that in the case of dependence logic the sentence (8) can be written using a single sentence  $\phi^*(S)$  (and  $\psi^*(S)$ ) that translates  $\phi$  over teams with domain  $\text{Fr}(\phi)$  (see Proposition 2.5). In the case of  $\mathcal{D}(\mathbf{M})$  the behavior of  $\phi$  and  $\psi$  over teams  $X$  with  $\text{dom}(X) = \{y_1, \dots, y_n\}$  does not in general reduce to their behavior over  $X \upharpoonright \text{Fr}(\phi)$  and  $X \upharpoonright \text{Fr}(\psi)$  (see Proposition 2.13). Therefore, to formulate the sentence (8), we need sentences  $\phi^*(S)$  and  $\psi^*(S)$  that are correct translations of  $\phi$  and  $\psi$  with respect to teams with domain  $\{y_1, \dots, y_n\}$ .

The case  $\gamma := (\phi \wedge \psi)(y_1, \dots, y_n)$  is also analogous to [19]. It remains to consider the case where our formula  $\gamma$  is of the form

$$\gamma := \text{M}y_n \phi(y_1, \dots, y_n) \quad (9)$$

and  $\phi$  is a formula for which we have already a translation into an  $\text{SO}(\text{Most}_f)[\tau \cup S]$  sentence  $\phi^*(S)$ . We claim that  $\gamma$  can be translated as follows:

$$\gamma^*(S) := R^n Y, Z(\theta_1(Y), \theta_2(Z)) \quad (10)$$

where

$$\begin{aligned} \theta_1(Y) &:= \phi^*(Y/S) \wedge \forall y_1 \dots \forall y_{n-1} \exists^=1 y_n Y(\bar{y}) \wedge \forall y_1 \dots \forall y_{n-1} (\exists y_n Y(\bar{y}) \leftrightarrow S(\bar{y}')) \\ \theta_2(Z) &:= \neg \phi^*(Z/S) \wedge \forall y_1 \dots \forall y_{n-1} \exists^=1 y_n Z(\bar{y}) \wedge \forall y_1 \dots \forall y_{n-1} (\exists y_n Z(\bar{y}) \leftrightarrow S(\bar{y}')) \\ \text{and } \bar{y}' &:= y_1, \dots, y_{n-1} \end{aligned}$$

The following equivalence is now obvious for all  $\mathfrak{A}$  and  $X$ :

$$\mathfrak{A} \models_X \gamma \iff (\mathfrak{A}, \text{rel}(X)) \models \gamma^*(S).$$

◀

Next we will show that, for sentences, Lemma 4.1 can be reversed.

► **Lemma 4.2.** *Let  $\tau$  be a vocabulary and  $\phi \in \text{SO}(\text{Most}_f)[\tau]$ . Then there is a sentence  $\psi \in \mathcal{D}(\mathbf{M})[\tau]$  such that for all models  $\mathfrak{A}$ :*

$$\mathfrak{A} \models \phi \iff \mathfrak{A} \models \psi.$$

**Proof.** By Proposition 3.5 we may assume that  $\phi$  is of the form:

$$\exists \bar{h}^1 \text{Most}_f^k g_1 \cdots \text{Most}_f^k g_n \exists \bar{h}^2 \forall x_1 \cdots \forall x_m \psi, \quad (11)$$

where the function symbols in  $\bar{h}^1$  and  $g_1, \dots, g_n$  are  $k$ -ary, and  $\psi$  is quantifier free. Before translating this sentence into  $\mathcal{D}(\mathbf{M})$ , we will first apply certain reductions to it. First of all, we make sure that the functions  $g_i$  have only occurrences of the form  $g_i(\bar{x})$  in  $\psi$  where  $\bar{x} = (x_1, \dots, x_k)$ . We can achieve this by existentially quantifying new names  $f_i$  for these symbols and passing on to the sentence

$$\exists \bar{h}^1 \text{Most}_f^k g_1 \cdots \text{Most}_f^k g_n \exists \bar{h}^2 \exists f_1 \cdots \exists f_n \forall x_1 \cdots \forall x_m \left( \bigwedge_{1 \leq j \leq n} g_j(\bar{x}) = f_j(\bar{x}) \wedge \psi^* \right), \quad (12)$$

where  $\psi^*$  is obtained from  $\psi$  by replacing all occurrences of  $g_i$  by  $f_i$  for  $1 \leq j \leq n$ . Analogously, we may also assume that the functions  $h$  in  $\bar{h}^1$  have only occurrences  $h(x_1, \dots, x_k)$  in  $\psi$ . Here  $m$  can always be made at least  $k$ .

The next step is to transform the quantifier-free part  $\psi^*$  to satisfy the condition that for each of the function symbols  $h$  in  $\bar{h}^2$  (also  $f_i$ ) there is a unique tuple  $\bar{x}$  of pairwise distinct variables such that all occurrences of it in  $\psi^*$  are of the form  $h(\bar{x})$  ( $f_i(\bar{x})$ ). In order to achieve this, we might have to introduce new existentially quantified functions and also universal first-order quantifiers (see Theorem 6.15 in [19]), but the quantifier structure of the sentence (11) does not change.

We will now assume that the sentence (11) has the properties discussed above:

1. The function symbols  $h \in \bar{h}^1$  and  $g_i$  have only occurrences of the form  $h(x_1, \dots, x_k)$  and  $g_i(x_1, \dots, x_k)$  in  $\psi$ , respectively.
2. For each  $h$  in  $\bar{h}^2$  ( $f_i$ , for  $1 \leq i \leq n$ ) there is a unique tuple  $\bar{x}$  of pairwise distinct variables such that all occurrences of  $h$  in  $\psi^*$  are of the form  $h(\bar{x})$  ( $f_i(\bar{x})$ ).

We will now show how the sentence (11) can be translated into  $\mathcal{D}(\mathbf{M})$ . For the sake of bookkeeping, we assume that  $\bar{h}^1 = h_1 \dots h_p$ ,  $\bar{h}^2 = h_{p+1} \dots h_r$ , and that  $h_i$  appears in  $\psi$  only as  $h_i(\bar{x}^i)$ . We claim now that the following sentence of  $\mathcal{D}(\mathbf{M})$  is a correct translation for (11):

$$\forall x_1 \cdots \forall x_k \exists y_1 \cdots \exists y_p \mathbf{M} z_1 \cdots \mathbf{M} z_n \forall x_{k+1} \cdots \forall x_m \exists y_{p+1} \cdots \exists y_r \left( \bigwedge_{p+1 \leq j \leq r} =(\bar{x}^j, y_j) \wedge \theta \right), \quad (13)$$

where  $\theta$  is obtained from  $\psi$  by replacing all occurrences of the term  $g_i(x_1, \dots, x_k)$  by the variable  $z_i$  and, similarly, each occurrence of  $h_i(\bar{x}^i)$  by  $y_i$ .

Let us then show that the sentence  $\phi$  (see (11)) and sentence (13) are logically equivalent. Let  $\mathfrak{A}$  be a structure and let  $\mathbf{h}_1, \dots, \mathbf{h}_r$  and  $\mathbf{g}_1, \dots, \mathbf{g}_n$  interpret the corresponding function symbols. We will show that the following holds:

$$(\mathfrak{A}, \bar{\mathbf{h}}, \bar{\mathbf{g}}) \models_X \psi \Leftrightarrow \mathfrak{A} \models_{X^*} \theta, \quad (14)$$

where  $X = \{\emptyset\}(A/x_1) \cdots (A/x_m)$  and

$$X^* = \begin{array}{ccccccc} \{\emptyset\}(A/x_1) & \cdots & (A/x_k) & (H_1/y_1) & \cdots & (H_p/y_p) & \\ & & (G_1/z_1) & \cdots & (G_n/z_n) & (A/x_{k+1}) & \cdots & (A/x_m) \\ & & (H_{p+1}/y_1) & \cdots & (H_r/y_r), & & & \end{array}$$

where the supplement functions  $H_i$  and  $G_i$  are defined using the functions  $\mathbf{h}_i$  and  $\mathbf{g}_i$  as follows:

$$\begin{aligned} H_i(s) &= \mathbf{h}_i(s(x_1), \dots, s(x_k)) \text{ for } 1 \leq i \leq p, \\ H_i(s) &= \mathbf{h}_i(s(\bar{x}^i)) \text{ for } p+1 \leq i \leq r, \\ G_i(s) &= \mathbf{g}_i(s(x_1), \dots, s(x_k)) \text{ for } 1 \leq i \leq n, \end{aligned}$$

and where  $s(\bar{x}^i)$  is the tuple obtained by pointwise application of  $s$ . The claim in (14) is now proved using induction on the structure of the quantifier-free formula  $\psi$ . Note that  $\psi$  is a first-order formula of dependence logic; hence, by Theorem 2.9, (14) holds iff the equivalence holds for each  $s \in X$  (equivalently  $s \in X^*$  since the values of the universally quantified variables functionally determine the values of all the other variables) individually. We can now show, using induction on the construction of  $\psi$ , that for all  $s \in X^*$  it holds that

$$\mathfrak{A} \models_s \theta \iff (\mathfrak{A}, \bar{\mathbf{h}}, \bar{\mathbf{g}}) \models_{s'} \psi, \quad (15)$$

where  $s' = s \upharpoonright \{x_1, \dots, x_m\}$ . The key to this result is the fact that, for every  $s$ , the interpretation of the variables  $z_i$  and  $y_i$  agree with the interpretation of the terms  $h_i(\bar{x}^i)$  and  $g(x_1, \dots, x_k)$ , respectively.

Finally, we note that there is a one-to-one correspondence between all possible interpretations  $\mathbf{h}_1, \dots, \mathbf{h}_r$  and  $\mathbf{g}_1, \dots, \mathbf{g}_n$  for the function symbols and teams  $X^*$  satisfying the dependence atomic formulas in (13). Therefore, sentence  $\phi$  (see (11)) and sentence (13) are logically equivalent.  $\blacktriangleleft$

## 5 Conclusion and Open Questions

We have seen that extending dependence logic by a majority quantifier increases the expressive power of dependence logic considerably. One particular consequence of our result is that  $\mathcal{D}(\mathbf{M})$  is closed under classical negation on the level of sentences. Note further that, for open formulas, this does not hold because of the downward closure property of formulas.

Several open questions remain and we now discuss some of them. Firstly, Proposition 2.12 shows that the fragment of  $\mathcal{D}(\mathbf{M})$  without dependence atoms does not satisfy the flatness property. It would be interesting to pin down the exact expressive power of sentences of  $\mathcal{D}(\mathbf{M})$  without dependence atoms.

The second open question concerns the open formulas of  $\mathcal{D}(\mathbf{M})$ . In [12] it was shown that the open formulas of  $\mathcal{D}$  correspond to the downwards monotone properties of NP (see [12] for the exact formulation). We conjecture that the open formulas of  $\mathcal{D}(\mathbf{M})$  correspond in an analogous manner to the downwards monotone properties of CH.

The majority quantifier is only one particular example of so-called *generalized quantifiers* (or, *Lindström quantifiers*), introduced in [13] and studied extensively in the context of descriptive complexity theory (see surveys [18] and [4]). In [3], second-order Lindström quantifiers were introduced and some results concerning their expressive power were obtained. We consider it an interesting study to enrich in a similar way dependence logic by further generalized quantifiers and relate the obtained logics to those studied in [3].

---

## References

- 1 S. Abramsky and J. Väänänen. From IF to BI. *Synthese*, 167(2):207–230, 2009.

- 2 E. Allender. The permanent requires large uniform threshold circuits. *Chicago J. Theoret. Comput. Sci.*, pages Article 7, 19 pp. (electronic), 1999.
- 3 H.-J. Burtschick and H. Vollmer. Lindström quantifiers and leaf language definability. *Int. J. Found. Comput. Sci.*, 9(3):277–294, 1998.
- 4 H.-D. Ebbinghaus and J. Flum. *Finite model theory, 2nd edition*. Perspectives in Mathematical Logic. Springer-Verlag, 1999.
- 5 F. Engström. Generalized quantifiers in dependence logic. arXiv:1103.0396.
- 6 P. Galliani. Inclusion and exclusion dependencies in team semantics: On some logics of imperfect information. arXiv:1106.1323.
- 7 E. Grädel and J. Väänänen. Dependence and independence. To appear in *Studia Logica*.
- 8 L. Henkin. Some remarks on infinitely long formulas. In *Infinitistic Methods (Proc. Sympos. Foundations of Math., Warsaw, 1959)*, pages 167–183. Pergamon, Oxford, 1961.
- 9 J. Hintikka and G. Sandu. Informational independence as a semantical phenomenon. In *Logic, methodology and philosophy of science, VIII (Moscow, 1987)*, volume 126 of *Stud. Logic Found. Math.*, pages 571–589. North-Holland, Amsterdam, 1989.
- 10 J. Kontinen. A logical characterization of the counting hierarchy. *ACM Trans. Comput. Log.*, 10(1), 2009.
- 11 J. Kontinen and H. Niemistö. Extensions of MSO and the monadic counting hierarchy. *Inf. Comput.*, 209(1):1–19, 2011.
- 12 J. Kontinen and J. Väänänen. On definability in dependence logic. *J. Log. Lang. Inf.*, 18(3):317–332, 2009.
- 13 P. Lindström. First order predicate logic with generalized quantifiers. *Theoria*, 32:186–195, 1966.
- 14 P. Lohmann and H. Vollmer. Complexity results for modal dependence logic. In A. Dawar and H. Veith, editors, *Computer Science Logic, 24th International Workshop, CSL 2010, 19th Annual Conference of the EACSL, Brno, Czech*, volume 6247 of *Lecture Notes in Computer Science*, pages 411–425. Springer, 2010.
- 15 M. Sevenster. Model-theoretic and computational properties of modal dependence logic. *J. Log. Comput.*, 19(6):1157–1173, 2009.
- 16 S. Toda. PP is as hard as the polynomial-time hierarchy. *SIAM J. Comput.*, 20(5):865–877, 1991.
- 17 J. Torán. Complexity classes defined by counting quantifiers. *J. Assoc. Comput. Mach.*, 38(3):753–774, 1991.
- 18 J. Väänänen. Generalized quantifiers, an introduction. In *Generalized quantifiers and computation (Aix-en-Provence, 1997)*, volume 1754 of *Lecture Notes in Comput. Sci.*, pages 1–17. Springer, Berlin, 1999.
- 19 J. Väänänen. *Dependence logic: A New Approach to Independence Friendly Logic*, volume 70 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2007.
- 20 J. Väänänen. Modal dependence logic. In K. Apt and R. van Rooij, editors, *New Perspectives on Games and Interaction*, volume 5 of *Texts in Logic and Games*, pages 237–254. Amsterdam University Press, 2008.
- 21 H. Vollmer. *Introduction to Circuit Complexity – A Uniform Approach*. Texts in Theoretical Computer Science. Springer Verlag, Berlin Heidelberg, 1999.
- 22 K. Wagner. The complexity of combinatorial problems with succinct input representation. *Acta Informatica*, 23:325–356, 1986.
- 23 F. Yang. Expressing second-order sentences in intuitionistic dependence logic. To appear in *Studia Logica*.