On the Expressive Power of Monotone Natural Language Quantifiers over Finite Models*

Jouko Väänänen†
University of Helsinki, Department of Mathematics
jouko.vaananen@helsinki.fi

Dag Westerståhl
Gothenburg University, Department of Philosophy
dag.westerstahl@phil.gu.se

April 10, 2002

Abstract

We study definability in terms of monotone generalized quantifiers satisfying Isomorphism closure, Conservativity and Extension. Among the quantifiers with the latter three properties — here called CE quantifiers — one finds the interpretations of determiner phrases in natural languages. The property of monotonicity is also linguistically ubiquitous, though some determiners like *an even number of* are highly non-monotone. They are nevertheless definable in terms of monotone CE quantifiers: we give a necessary and sufficient condition for such definability. We further identify a stronger form of monotonicity, called smoothness, which also has linguistic relevance, and we extend our considerations to smooth quantifiers. The results lead us to propose two tentative universals concerning monotonicity and natural language quantification. The notions involved as well as our proofs are presented using a graphical representation of quantifiers in the so-called number triangle.

Keywords: generalized quantifier, natural language semantics, finite model, monotone quantifier. MSC 2000 classification numbers 03B65, 03C80, 03C13.

*The final manuscript was prepared while the authors visited the Mittag-Leffler Institute in Djursholm, Sweden. We are grateful to the institute for its hospitality. We also thank an anonymous referee for several useful comments.

†Research partially supported by grant 40734 of the Academy of Finland.

1 Introduction, motivation, and summary

1.1 CE quantifiers and NL quantifiers

We shall study in the context of finite models the logical expressive power of monotone (generalized) quantifiers of a familiar kind: functions Q which with each universe M associate a binary relation Q_M between subsets of M,[1] and which satisfy the conditions of *Isomorphism closure*, *Conservativity* and *Extension*: For all M, M', all $A, B \subseteq M$, and $A', B' \subseteq M'$,

(Isom) If $(M, A, B) \cong (M', A', B')$, then $Q_M(A, B) \Leftrightarrow Q_{M'}(A', B')$.

(Cons) $Q_M(A, B) \Leftrightarrow Q_M(A, A \cap B)$.

(Ext) If $M \subseteq M'$, then $Q_M(A, B) \Leftrightarrow Q_{M'}(A, B)$.

Such quantifiers will be called *CE quantifiers* for short. *Monotonicity* is the extra property that

(Mon) If $A \subseteq M$ and $B \subseteq B' \subseteq M$, then $Q_M(A, B) \Rightarrow Q_M(A, B')$.

The class of CE quantifiers is a natural one in many respects. Isomorphism closure guarantees that these quantifiers really are relations between 'quantities', i.e., between numbers. It furthermore allows one to treat quantifiers as *logical constants*, which can be added to, say, first-order logic (FO). If we, as in this paper, restrict attention to finite universes, then each CE quantifier in fact corresponds to a *binary relation among natural numbers*, and vice versa (Section 2.1, Definition 2). The class of CE quantifiers is closed under Boolean operations. But perhaps the most familiar feature of CE quantifiers is their close connection to certain expressions in natural languages.

Many natural languages, among them English, contain a rich variety of (simple and complex) *determiner* expressions, and many of these expressions can be taken to denote, on each universe, binary relations between sets. Here are some English determiners:

(1) every, some, no, the, at least five, no more than ten, exactly seven, all but three, between five and eight, at most half of the, more than three quarters of the.

And here are some corresponding (second-order) relations on a universe M:[2]

[1] Or, equivalently, classes of structures of the form (M, A, B), where $A, B \subseteq M$. Then $(M, A, B) \in Q \Leftrightarrow Q_M(A, B)$.

[2] We use the convention of letting the English expressions in italics stand for the quantifier denoted by that expression. $|X|$ is the cardinality of the set X.

$every_M(A, B) \iff A \subseteq B$

$no_M(A, B) \iff A \cap B = \emptyset$

$the_M(A, B) \iff |A| = 1$ and $A \subseteq B$

$at\text{-}least\text{-}five_M(A, B) \iff |A \cap B| \geq 5$

$all\text{-}but\text{-}three_M(A, B) \iff |A - B| = 3$

$between\text{-}five\text{-}and\text{-}eight_M(A, B) \iff 5 \leq |A \cap B| \leq 8$

$at\text{-}most\text{-}half\text{-}of\text{-}the_M(A, B) \iff |A \cap B| \leq 1/2 \cdot |A|$

The plausibility of these interpretations is immediate from the truth conditions of sentences like

Every cat purrs.

No train was on time.

The boy cried.

At least five linguists have seen *Gone with the wind*.

All but three professors came to the meeting.

At most half of the students passed the exam.

Let us stipulate for this paper that a *natural language* (NL) quantifier is one which is denoted by some determiner in some natural language. It then appears to be generally true that

(U_1) *NL quantifiers are CE quantifiers.*[3]

(U_1) expresses a *linguistic universal*: a generalization across languages concerning, in this case, the interpretation of certain kinds of expressions. At first blush, it would appear quite possible that there could be a natural language which had a determiner which denoted, on each universe, a

[3]We hasten to add that (1) several other types of quantifiers can be associated with expressions in natural languages (for example, some determiner expressions denote *ternary* relations between sets, and some syntactic constructions appear to require *polyadic* quantifiers (relations between relations, not just between sets) for their interpretation; cf. Keenan and Westerståhl [7] for examples); (2) other expressions than determiners can involve CE quantification; (3) some determiners like *many* and *few* have a strongly context-dependent or intensional meaning — they are not considered here. For our purposes in this paper the present restricted notion of an NL quantifier suffices.

binary relation between sets, but which failed to satisfy Conservativity, or Extension, or Isomorphism closure. For example, there could be a quantifier *sovery*, which meant *some* on universes with less than 10 elements, and *every* on other universes; this would be Isom and Cons, but not Ext. Or, there could be a determiner denoting the Härtig quantifier I, one of the first to be studied in the theory of generalized quantifiers, defined by

$$I_M(A, B) \iff A, B \text{ have the same cardinality;}$$

this is Isom and Ext, but not Cons. There is no problem in constructing an artificial (formal) language with such quantifiers. But it seems that there are no such languages in nature.

Of course, even better than a true generalization is an account of *why* it holds, and indeed a lot has been written on this theme concerning (U_1), which seems fairly well explained.[4] We have nothing to add here, but we shall formulate in this introductory section two other universals, both concerned with monotonicity, that appear to be less familiar.

The notion of an NL quantifier is not intended to be precise. But some CE quantifiers are *clearly* NL ones, and others are *clearly not*. In the first category we have, to begin, the quantifiers exemplified in (1) above. In particular, we may take all quantifiers of the form *at-least-m* (as well as *more-than-m*, *at-most-m*, *exactly-m*, etc.), and likewise all proportional quantifiers of the form *more-than-m/n′ths-of-the* (and their variants with *at least*, *at most*, etc.), to be NL quantifiers (for $0 \leq m < n$). A prominent example here is the quantifier *most*, which (let us stipulate) is taken in the sense of *more than half of the*, i.e.,

$$most_M(A, B) \iff |A \cap B| > 1/2 \cdot |A|.$$

Furthermore, in view of English determiners like *some but not all*, *between five and eight*, etc., we may as well idealize a bit and take the class of NL quantifiers to be closed under Boolean operations. This already gives us a sizeable stock of NL quantifiers; some additional examples will appear below.

Since any binary relation on numbers corresponds to a CE quantifier, one may expect that many such relations belong to mathematics but have nothing to do with natural languages, at least not in the sense of being denoted by determiners. Here are two examples:

(2) $Div_M(A, B) \iff |A \cap B| \text{ divides } |A|.$

[4]See, for example, Barwise and Cooper [1], Keenan and Stavi [6], van Benthem [2], Westerståhl [11], Keenan and Westerståhl [7].

$$Sqrt_M(A, B) \iff |A \cap B| > \sqrt{|A|}.$$

These are mathematically natural CE quantifiers but apparently not NL quantifiers, and of course lots of similar examples can be given.

Now, it was noticed early on that most NL quantifiers have some sort of monotonicity properties. Barwise and Cooper [1] proposed a universal to the effect that all quantifiers denoted by non-complex determiners are monotone, or negations of monotone ones (like $no = \neg some$), and it might appear reasonable to conjecture that all NL quantifiers are Boolean combinations of monotone ones. Indeed, one easily checks that this is true of all the NL quantifiers mentioned so far. However, the following quantifier (and Boolean combinations from it) is a clear counterexample:

$$\textit{an-even-number-of}_M(A, B) \iff |A \cap B| \text{ is even.}$$

But could it not be that *an-even-number-of*, though not a Boolean combination of monotone quantifiers, is nevertheless definable in some other way from such quantifiers? In order to discuss this issue, we need to introduce one more class of quantifiers.

1.2 Simple unary quantifiers

A *simple unary* quantifier Q associates with each universe M a unary relation Q_M between subsets of M (i.e., a set of subsets of M), and satisfies the version of isomorphism closure appropriate for the unary case:

(Isom) If $(M, A) \cong (M', A')$, then $Q_M(A) \Leftrightarrow Q_{M'}(A')$.[5]

Examples are the quantifiers $\forall$ and $\exists$ from FO logic (where $\forall_M(A) \Leftrightarrow A = M$ and $\exists_M(A) \Leftrightarrow A \neq \emptyset$), as well as $\exists_{\geq n}$, and the *Rescher* quantifier Q^R defined by

$$Q^R_M(A) \iff |A| > 1/2 \cdot |M|.$$

There is a very close link between the class of simple unary quantifiers and the class of CE quantifiers. As will be explained in Section 2.1, the operation $\cdot^{\text{rel}}$ of *relativization* is a bijection from the first to the second which preserves various properties, in particular monotonicity. (A simple

[5] In the logical literature (where isomorphism closure is part of the concept of a quantifier), simple unary quantifiers are also called quantifiers of type $\langle 1 \rangle$ (or of type $(1; 1)$), and CE quantifiers are quantifiers of type $\langle 1, 1 \rangle$ (or $(1; 2)$) satisfying Conservativity and Extension.

unary Q is monotone if $Q_M(A)$ and $A \subseteq A' \subseteq M$ imply $Q_M(A')$.) For example, $\forall^{\mathrm{rel}} =$ *all*, $\exists^{\mathrm{rel}}_{\geq n} =$ *at-least-n*, and $(Q^R)^{\mathrm{rel}} =$ *most*. Moreover, a simple unary Q also corresponds to a binary relation between numbers, in fact, the *same* relation that corresponds to the CE quantifier Q^{rel}.

The latter fact should not lead one to believe, however, that simple unary quantifiers and CE quantifiers are equally expressive. In fact, much more can be expressed by CE quantifiers than by simple unary ones. A simple unary Q allows us to relate the size of *one* set to the size of the universe, whereas with the corresponding CE quantifier Q^{rel} we can express the same relation between (the sizes of) *two* arbitrary sets, provided the first is a subset of the latter. An instance of the added expressive power of CE quantifiers was proved already in Barwise and Cooper [1], namely, that *most* is not definable in $\mathrm{FO}(Q^R)$: FO logic with Q^R as an added generalized quantifier. (The notions of definability and of logics with generalized quantifiers are explained in Section 3.1 below.) This is an instance of the result, proved in Westerståhl [12] and Kolaitis and Väänänen [8], that for Q simple unary and monotone, Q^{rel} is definable in $\mathrm{FO}(Q)$ if and only if Q is already FO definable. The result about *most* was generalized further in Kolaitis and Väänänen [8] by showing that *most* is not definable in terms of any finite number of simple unary quantifiers.

Now, *an-even-number-of* is the relativization of the simple unary Q_{even}, where

$$(Q_{even})_M(A) \iff |A| \text{ is even.}$$

These two quantifiers are equally expressive, since

$$(Q_{even})_M(A) \iff \textit{an-even-number-of}_M(M, A)$$

and

$$\textit{an-even-number-of}_M(A, B) \iff (Q_{even})_M(A \cap B).$$

But are they definable from monotone quantifiers? The following is a consequence of the Bounded Oscillation Theorem (Theorem 14 below), which was proved in Väänänen [10]:

(3) *an-even-number-of* is not definable in terms of any (finite number of) monotone simple unary quantifiers.

1.3 Definability from monotone CE quantifiers

However, our starting-point in the present paper is the, perhaps surprising, observation (Section 4) that

(4) *an-even-number-of* is nevertheless definable in terms of a monotone CE quantifier.

(In fact, this is an instance of the more general observation that all *intersective* quantifiers are so definable, cf. Section 4.1.) Since the CE quantifiers are the proper ones to focus on in a natural language context, it becomes of interest to know precisely which quantifiers are FO definable from monotone CE quantifiers. Our main result in Section 4, Theorem 17, provides a characterization of these. It generalizes the Bounded Oscillation Theorem from simple unary to CE quantifiers. In the former case, the necessary and sufficient condition was for a quantifier to have *bounded oscillation*, a property which is easily visualizable when quantifiers are represented in the *number triangle* (as binary relations between natural numbers, cf. Section 2.2). In the CE case there is a corresponding property, which we call *bounded color oscillation*.

Our Theorem 17 gives a way, in principle, to find quantifiers that are *not* definable from monotone CE quantifiers. It turns out to be rather difficult, however, to provide examples. We conjecture, but we do not prove, that the divisibility quantifier *Div*, defined in (2) above, is such an example. But we do construct another quantifier, with a regular if somewhat complex pattern in the number triangle, and show that it does not have bounded color oscillation. This proof makes use of van der Waerden's Theorem, and it seems to us that any such example would hinge on similar mathematical facts, and in any case be rather far removed from natural languages. Therefore, we propose the following, as a linguistic universal concerning NL quantifiers and monotonicity:

(U_2) *All NL quantifiers are FO definable from monotone CE quantifiers.*

(U_2) is perhaps not a very bold statement, and one may wonder if it can be strengthened. One way would be to put constraints on which kind of defining sentences are allowed; we noted that Boolean combinations are not enough but will not pursue this further here. But another strengthening that might seem natural is to replace 'CE' by 'NL' in (U_2). That would single out monotone NL quantifiers as 'building blocks' from which all other NL quantifiers are constructed.

However, such a strengthened form of (U_2) does not seem to be true. Our main motivation for this claim does not take the form of a direct argument that some particular NL quantifier, say, *an-even-number-of*, is not definable from monotone NL quantifiers.[6] Rather, it follows from an analysis of the kind of monotonicity which occurs in natural languages. It seems that monotone NL quantifiers actually have a much stronger property than monotonicity, which we call *smoothness*. Think of the monotonicity of Q as saying that $Q_M(A, B)$ holds whenever B is sufficiently 'big', compared to A. So if you add an element to B, the result is still 'big', compared to A. Smoothness means, roughly, that the result is still 'big' if you *also* add an element to A and compare to that set. I.e., the standard of 'bigness' does not change drastically if we make a small change in the size of the comparison set.

In Section 5 we look at some properties of smooth quantifiers. In particular, we give a necessary condition for a quantifier to be FO definable from smooth CE quantifiers. It follows from this that *an-even-number-of* is not so definable. Now, as noted, it does seem probable that the following universal is true:

(U_3) *All monotone NL quantifiers are smooth.*

It follows that the suggested strengthening of (U_2) is not valid. If *an-even-number-of* is an NL quantifier — and we see no reason to doubt that it is — then (U_2) is best possible in this respect.

1.4 Plan

The results of the paper are presented and proved in Sections 4 and 5, as outlined above. In Section 2 we state precisely the properties and facts about simple unary and CE quantifiers that are used. In particular, we recall how to think about such quantifiers visually in the number triangle. Section 3 presents the machinery we use to prove facts of definability and undefinability in this paper. Some conclusions and questions are listed in section 6.

The paper is essentially self-contained, with all definitions and proofs in place, except that we in Section 3 state without proof a characterization of FO($\mathbf{Q}$)-equivalence up to quantifier rank r (Proposition 9), that is quite

[6] Although this seems indeed to be the case. *an-even-number-of* is definable from a monotone CE quantifier (Proposition 18), but that quantifier appears not to be NL. We discuss this issue further at the end of Section 4.1.

familiar from the literature. Also, we use van der Waerden's Theorem at one point in section 4.

Finally, why *finite* models? One reason is that some NL quantifiers, like the proportional ones, only make sense for finite sets. On the other hand, others work for all sets, and a few (e.g. *a-finite-number-of*) even presuppose infinite models. Another reason could be that finite structures is the appropriate context for the study of *computational complexity*, for which the theory of generalized quantifiers on finite models is quite relevant. However, that is not our topic here. Rather, our main motivation for restricting attention to finite sets is that this allows a simple and visual representation of CE and simple unary quantifiers in the 'number triangle', a representation which goes back to van Benthem [2], and which suggested several of the notions and results in this paper.

2 Representation and properties of simple unary and CE quantifiers

2.1 Quantifiers as binary relations in $\mathbb{N}$

CE quantifiers are relativizations of simple unary quantifiers, and can be viewed as the same relations between numbers. We now recapitulate these facts.

Definition 1 If Q is simple unary, its *relativization* Q^{rel} is defined by

$$(Q^{\mathrm{rel}})_M(A,B) \iff Q_A(A\cap B),$$

for all M and all $A,B\subseteq M$. (The operation of relativization can be applied to quantifiers of any type: one adds a new set argument and considers the behavior of the given quantifier with that set as universe and the arguments restricted to it. In particular, if Q is CE, Q^{rel} is defined by $(Q^{\mathrm{rel}})_M(A,B,C)\Leftrightarrow Q_A(A\cap B, A\cap C)$.)

We allow that $A=\emptyset$ here, and in general that $M=\emptyset$ for quantifiers Q_M over M. So, for example, $\forall_\emptyset(\emptyset)$ holds, but not $\exists_\emptyset(\emptyset)$ or $Q^R_\emptyset(\emptyset)$.

Definition 2 (a) Let Q be simple unary. We define a binary relation, also denoted Q, over the set $\mathbb{N}$ of natural numbers as follows:

$$Q(k,m) \iff \text{there is } M \text{ and } A\subseteq M \text{ such that } |M-A|=k, |A|=m, \text{ and } Q_M(A).$$

(b) Let Q be CE. Define a binary relation, also denoted Q, over $\mathbb{N}$ by:

$$Q(k,m) \iff \text{there is } M \text{ and } A, B \subseteq M \text{ such that } |A-B| = k, |A \cap B| = m, \text{ and } Q_M(A,B).$$

Fact 3 (a) *If Q is simple unary, then for all M and all $A \subseteq M$,*

$$Q_M(A) \iff Q(|M-A|, |A|).$$

(b) *If Q is CE, then for all M and all $A, B \subseteq M$,*

$$Q_M(A,B) \iff Q(|A-B|, |A \cap B|).$$

(c) *Q is CE iff $Q = Q_0^{rel}$ for some simple unary Q_0.*

(d) *If Q is simple unary, then, for all $k, m \in \mathbb{N}$,*

$$Q(k,m) \iff Q^{rel}(k,m).$$

Proof. (a): The left to right direction is immediate by Definition 2 (a). The other direction follows from isomorphism closure, since if $|M-A| = |M'-A'|$ and $|A| = |A'|$, then $(M,A) \cong (M',A')$.

(b): Similar to (a), but here ones uses Conservativity, Extension, and isomorphism closure to verify that if $|A-B| = |A'-B'|$ and $|A \cap B| = |A' \cap B'|$, then $Q_M(A,B) \Leftrightarrow Q_{M'}(A',B')$.

(c): One checks directly from Definition 1 that Q^{rel} is always a CE quantifier. In the other direction, suppose Q is CE, and define Q_0 by $(Q_0)_M(A) \Leftrightarrow Q_M(M,A)$. By Conservativity and Extension it follows that $Q = Q_0^{\mathrm{rel}}$.

(d): Immediate from Definitions 1 and 2. □

So, for example,

$$\forall(k,m) \Leftrightarrow k = 0 \Leftrightarrow \textit{all}(k,m),$$

$$\exists_{\geq 3}(k,m) \Leftrightarrow m \geq 3 \Leftrightarrow \textit{at-least-3}(k,m),$$

$$Q^R(k,m) \Leftrightarrow m > 1/2 \cdot (k+m) \Leftrightarrow m > k \Leftrightarrow \textit{most}(k,m),$$

$$Q_{even}(k,m) \Leftrightarrow m \text{ is even} \Leftrightarrow \textit{an-even-number-of}(k,m).$$

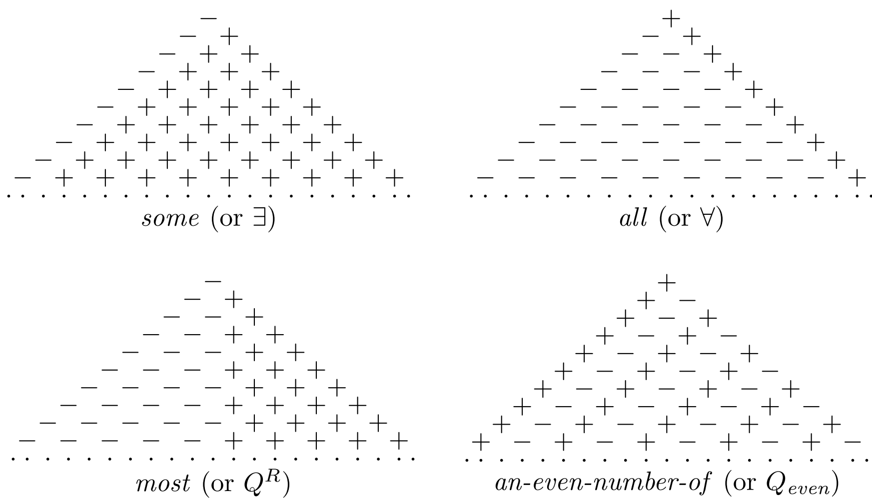

Figure 1: Some quantifiers in the number triangle.

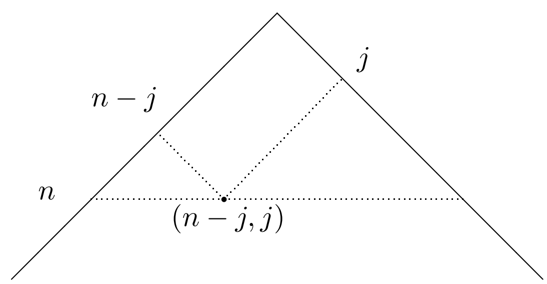

Figure 2: A point at level n

2.2 Quantifiers in the number triangle

The *number triangle* is simply $\mathbb{N} \times \mathbb{N}$, placed as a triangle with $(0,0)$ at the top. So a simple unary or CE quantifier Q is a subset of $\mathbb{N} \times \mathbb{N}$, which we represent by putting a "+" at those (k,m) which belong to Q, and a "−" elsewhere, as in Figure 1.

The *$n'th$ level* of the number triangle (Figure 2) is the 'diagonal'

$$(n,0),(n-1,1),\ldots,(n-j,j),\ldots,(0,n).$$

Here n is the size of the universe for simple unary quantifiers, and of the first (set) argument for CE quantifiers.

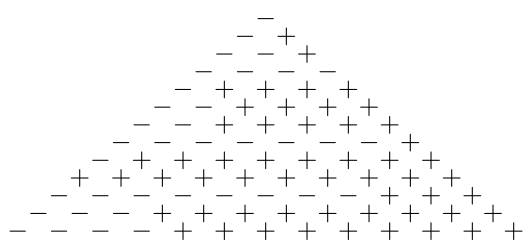

Figure 3: A monotone quantifier.

2.3 Monotonicity

Consider first a simple unary quantifier Q. Clearly, Q is monotone iff on each level n, there is a smallest size, say $f(n)$, such that if $|M| = n$ and $B \subseteq M$,

(5) $\quad Q_M(B) \iff |B| \geq f(n).$

If $Q_M(B)$ happens to be false for all $B \subseteq M$ we stipulate that $f(n) = n+1$; then (5) still holds. This motivates the following

Definition 4 If $f : \mathbb{N} \rightarrow \mathbb{N}$ is a function such that for all $n \in \mathbb{N}$, $f(n) \leq n+1$, define the simple unary quantifier Q_f by

$$(Q_f)_M(B) \iff |B| \geq f(|M|),$$

or, in other words,

$$Q_f(n-j, j) \iff j \geq f(n).$$

Fact 5 (a) *A simple unary quantifier is monotone iff it is equal to Q_f for some f as in Definition 4.*

(b) *A CE quantifier is monotone iff it is of the form Q_f^{rel}. We have*

$$(Q_f^{rel})_M(A, B) \iff |A \cap B| \geq f(|A|).$$

Thus, a monotone quantifier (simple unary or CE) has the characteristic pattern that if there is a + on some level, then all points to the right on the same level have a + too (Figure 3).

2.4 Duals

The dual Q^{d} of a simple unary quantifier Q is defined by

$$Q^{\mathrm{d}}_M(A) \iff \neg Q_M(M - A).$$

Likewise, if Q is CE,

$$Q^{\mathrm{d}}_M(A, B) \iff \neg Q_M(A, A - B).$$

Then

$$(Q^{\mathrm{d}})^{\mathrm{rel}} = (Q^{\mathrm{rel}})^{\mathrm{d}},$$

and we have examples like the following: $\textit{all}^{\mathrm{d}} = \textit{some}$, $\textit{more-than-n}^{\mathrm{d}} = $ *all-but-at-most-n*, $\textit{most}^{\mathrm{d}} = $ *at-least-half.* In the number triangle, we obtain the Q^{d} from Q by first changing all $+'$s to $-'$s and vice versa, and then rotating the triangle 180° along the vertical axis through $(0, 0)$:

(6) $\quad Q^{\mathrm{d}}(k, j) \iff \neg Q(j, k).$

Furthermore, we have the following

Fact 6 *A quantifier is monotone iff its dual is monotone. In fact,*

$$(Q_f)^d = Q_{f^d},$$

where

$$f^d(n) \;=\; n - f(n) + 1.$$

Likewise,

$$(Q^{rel}_f)^d = (Q^d_f)^{rel} = (Q_{f^d})^{rel}.$$

3 Definability, logics, and FO(Q)-equivalence

3.1 Expressive power

To formalize the notion of a property (of models) expressible by means of Q and standard first-order machinery, one constructs the logical language, or simply the *logic*, FO(Q), as follows: Add a variable-binding operator, also denoted 'Q', to FO logic with the new formation rule

(Q-syn) If φ and ψ are formulas, so is

$$Qx(\varphi, \psi)$$

(or '$Qx\varphi$' in the simple unary case),

and with the corresponding semantic rule

(Q-sem) If $\mathbf{M}$ is a model with universe M,

$$\mathbf{M} \models Qx(\varphi, \psi)[\mathbf{a}] \iff Q_M(\varphi^{x,\mathbf{M},\mathbf{a}}, \psi^{x,\mathbf{M},\mathbf{a}})$$

(and similarly in the simple unary case),

where, if $\varphi = \varphi(x, \mathbf{y}), \psi = \psi(x, \mathbf{y})$ have free variables among $x, \mathbf{y}$ and $\mathbf{a}$ is a sequence of elements of M corresponding to $\mathbf{y}$,

$$\varphi(x, \mathbf{y})^{x,\mathbf{M},\mathbf{a}} = \{a \in M : \mathbf{M} \models \varphi(a, \mathbf{a})\}.$$

Similarly for logics FO($\mathbf{Q}$), where $\mathbf{Q} = \{Q_0, \ldots, Q_{u-1}\}$ is a set of (simple unary or CE) quantifiers. If FO($\mathbf{Q}$) and FO($\mathbf{Q'}$) are two such logics we say that FO($\mathbf{Q}$) $\leq$ FO($\mathbf{Q'}$) if every FO($\mathbf{Q}$)-sentence is logically equivalent to a FO($\mathbf{Q'}$)-sentence, and that FO($\mathbf{Q}$) $\equiv$ FO($\mathbf{Q'}$) if FO($\mathbf{Q}$) $\leq$ FO($\mathbf{Q'}$) and FO($\mathbf{Q'}$) $\leq$ FO($\mathbf{Q}$), i.e., if the two logics have the same expressive power.

Expressive power can also be formulated in terms of definability of quantifiers. Q is *definable* in FO($\mathbf{Q}$) if there is a sentence φ in this logic, with two unary predicates as the only non-logical symbols, such that

$$Q_M(A, B) \iff (M, A, B) \models \varphi.$$

Similarly for simple unary Q. Then it holds that

(7) FO($\mathbf{Q}$) $\leq$ FO($\mathbf{Q'}$) iff each quantifier in $\mathbf{Q}$ is definable in FO($\mathbf{Q'}$).

In the rest of this paper we will be concerned with the question

(*) When is Q definable in FO($\mathbf{Q}$)?

for certain choices of Q and $\mathbf{Q}$. Since Q^{d} is definable from Q, by

$$Q^{\mathrm{d}}_M(A, B) \iff (M, A, B) \models \neg Qx(Ax, \neg Bx),$$

we can always assume, when asking (*), that $\mathbf{Q}$ is closed under duals, and also that it contains *some* (or $\exists$). This in fact simplifies certain things that follow.

Next, a simple unary Q is always definable from Q^{rel}:

$$(8) \quad Q_M(A) \iff (M, A) \models Q^{\mathrm{rel}}x(x = x, Ax).$$

So if Q is definable in $\mathrm{FO}(\mathbf{Q})$, it is definable in $\mathrm{FO}(\mathbf{Q}^{\mathrm{rel}})$, where $\mathbf{Q}^{\mathrm{rel}} = \{Q_0^{\mathrm{rel}}, \ldots, Q_{u-1}^{\mathrm{rel}}\}$ and the Q_t are simple unary. But (as we noted in Section 1.2) a quantifier might be definable in $\mathrm{FO}(\mathbf{Q}^{\mathrm{rel}})$ without being definable in $\mathrm{FO}(\mathbf{Q})$.

A final observation in connection with (*) is that, when the logic is of the form $\mathrm{FO}(\mathbf{Q}^{\mathrm{rel}})$, we can restrict attention to the case when Q, the quantifier about which we ask whether it is definable, is simple unary. This follows from

Lemma 7 *The following are equivalent, for a simple unary quantifier Q:*[7]

(a) *Q is definable in $\mathrm{FO}(\mathbf{Q}^{rel})$.*

(b) *Q^{rel} is definable in $\mathrm{FO}(\mathbf{Q}^{rel})$.*

Proof. That (b) implies (a) is clear from (8). In the other direction, assume (a). It is a general (and not hard to prove, using induction on the defining formula) fact about relativization that

(9) If Q is definable in $\mathrm{FO}(\mathbf{Q})$ then Q^{rel} is definable in $\mathrm{FO}(\mathbf{Q}^{\mathrm{rel}})$.

But also,

(10) $(Q^{\mathrm{rel}})^{\mathrm{rel}}$ is definable from Q^{rel}.

This is because relativizing first to B and then to A is the same as relativizing once to $A \cap B$. From these observations, (b) follows. □

3.2 Criteria for $\mathrm{FO}^r(\mathbf{Q})$-equivalence

In proving results about expressive power, it is normally *un*definability that requires most work, since, in principle, you have to go through the whole language and verify that no definition does the job. Various methods have been employed, for example, quantifier elimination (starting with Mostowski), or counting arguments (cf. Hella, Luosto and Väänänen [5]). A straightforward yet powerful and very general method (which goes back to R. Fraïssé and A. Ehrenfeucht) is based on the following idea: To prove that one logic

[7] In fact, it holds for quantifiers of arbitrary types.

is more powerful than another, find two models which are *close* in the sense that they cannot be distinguished in the second logic (at least by sentences of a certain complexity), but so that they can be distinguished in the first. This is the method we shall employ here, so we have to begin by defining a suitable notion of closeness.

3.2.1 Closeness in terms of partial isomorphisms

By the remarks in the previous subsection, it will be enough for our purposes to consider models of the form $\mathbf{M} = (M, A)$ with $A \subseteq M$, and logics of the form $\mathrm{FO}(\mathbf{Q})$, where $\mathbf{Q}$ contains only CE quantifiers, among them *some*, and is closed under duals.[8] $\mathbf{M}$ partitions M into the 2 sets $M - A$ and A, from which we can form the 4 'unions'

$$U_{1,\mathbf{M}} = \emptyset,\ U_{2,\mathbf{M}} = A,\ U_{3,\mathbf{M}} = M - A,\ U_{4,\mathbf{M}} = M.$$

We let $B \triangle C = (B - C) \cup (C - B)$ be the symmetric difference between B and C, and we say that C is an X-*variant* of B if $B \triangle C \subseteq X$. If X is small, this means that B and C are 'almost' the same.

If $\mathbf{M} = (M, A)$, $\mathbf{M}' = (M', A')$, $a_1, \ldots, a_n \in M$, and $a'_1, \ldots, a'_n \in M$, then

$$(a_1, \ldots, a_n) \simeq (a'_1, \ldots, a'_n)$$

means that $\{(a_i, b_i) : 1 \leq i \leq n\}$ is a *partial isomorphism*, i.e., that

$$a_i = a_j \iff a'_i = a'_j,$$

and

$$a_i \in A \iff a'_i \in A'.$$

Now we can give one formulation of the notion of closeness. Let $r > 0$.

Definition 8

$$\mathbf{M} \approx_{r,\mathbf{Q}} \mathbf{M}'$$

holds iff whenever $Q_t \in \mathbf{Q}$, $(a_1, \ldots, a_{r-1}) \simeq (a'_1, \ldots, a'_{r-1})$, X_1, X_2 are $\{a_1, \ldots, a_{r-1}\}$-variants of unions $U_{i_1,\mathbf{M}}, U_{i_2,\mathbf{M}}$, respectively, and X'_1, X'_2 are

[8] But what follows in this subsection generalizes easily to the case of models of the form $(M, A_1, \ldots, A_n)$, and arbitrary monadic quantifiers.

the corresponding[9] $\{a'_1, \ldots, a'_{r-1}\}$-variants of $U_{i_1,\mathbf{M}'}, U_{i_2,\mathbf{M}'}$, it holds that

(11) $(Q_t)_M(X_1, X_2) \iff (Q_t)_{M'}(X'_1, X'_2)$,

or, equivalently, that

$$Q_t(|X_1 - X_2|, |X_1 \cap X_2|) \iff Q_t(|X'_1 - X'_2|, |X'_1 \cap X'_2|).$$

Roughly, $\mathbf{M} \approx_{r,\mathbf{Q}} \mathbf{M}'$ iff whenever a $Q_t \in \mathbf{Q}$ relates two unions of partition sets in $\mathbf{M}$, it also relates the corresponding unions in $\mathbf{M}'$, and vice versa, even after at most $r-1$ elements have been 'moved around' in $\mathbf{M}$, and at most $r-1$ elements in $\mathbf{M}'$, provided they were 'moved around' in the same way.

When r increases, with fixed (finite) $\mathbf{M}$ and $\mathbf{M}'$, $\approx_{r,\mathbf{Q}}$ eventually becomes the relation of *isomorphism*, since

(12) If $\mathbf{M} \approx_{r,\mathbf{Q}} \mathbf{M}'$ then either $|A|, |A'| \geq r$ or $|A| = |A'|$, and similarly for $M - A$ and $M' - A'$.

(12) follows from our assumption that *some* (or $\exists$) is in $\mathbf{Q}$: If, for example, $|A| < r$ but $|A'| > |A|$, then removing $|A|$ elements from A, and the same number of elements from A', we obtain $X_1 = X_2 = \emptyset$ in $\mathbf{M}$, but $X'_1 = X'_2 \neq \emptyset$ in $\mathbf{M}'$, contradicting (11) when $Q_t =$ *some*.

Now to connect this to the truth or falsity of sentences in models, we need to recall the notion of *quantifier rank* of an FO($\mathbf{Q}$)-formula φ; i.e, the maximum number of nestings of quantifiers from $\mathbf{Q}$ that occur in φ. We use FO^r ro denote the fragment of FO consisting of formulas with quantifier rank $\leq r$. $\mathbf{M}$ and $\mathbf{M}'$ are FO($\mathbf{Q}$)-*equivalent*, in symbols,

$$\mathbf{M} \equiv_{\mathrm{FO}(\mathbf{Q})} \mathbf{M}',$$

if the same FO($\mathbf{Q}$)-sentences are true in each.

$$\mathbf{M} \equiv_{\mathrm{FO}^r(\mathbf{Q})} \mathbf{M}'$$

means that this holds for sentences of quantifier rank at most r.

A proof of the following proposition can be found in Westerståhl [11], where some easy applications to undefinability results are also made. A thoroughgoing overview of definability among monadic quantifiers, in particular simple unary ones, is given in Väänänen [10].

[9] If C is an $\{a_1, \ldots, a_{r-1}\}$-variant of B, then the *corresponding* $\{a'_1, \ldots, a'_{r-1}\}$-variant of B' is $(B' - \{a'_i : a_i \notin C\}) \cup \{a'_i : a_i \in C\}$.

Proposition 9 *Let* $\mathbf{Q}$ *be a finite set of CE quantifiers containing* some.[10] *Then the following are equivalent:*

(a) $\mathbf{M} \approx_{r,\mathbf{Q}} \mathbf{M}'$

(b) $\mathbf{M} \equiv_{\mathrm{FO}^r(\mathbf{Q})} \mathbf{M}'$

3.2.2 Closeness in terms of colors

It will be convenient to give a slightly different formulation of closeness, this time in terms of colors. By a *coloring* of a set X we mean simply a function c from X to some finite set C. The elements of C are called colors but can be any objects, as long as there are a finite number of them.

We shall color structures of the form $\mathbf{M} = (M, A)$, relative to r and $\mathbf{Q}$. The coloring might seem slightly unintuitive at first blush, but the sole point is that the color of $\mathbf{M}$ must contain information about how each $Q_t \in \mathbf{Q}$ behaves on pairs (X_1, X_2), where the X_i come from $\emptyset, A, M-A, M$ by adding or deleting at most $r-1$ elements (in all) from M. In general there would be 16 forms of such pairs to consider, but since $\mathbf{Q}$ is closed under duals and contains only CE quantifiers, it turns out we can restrict attention to the following 5 (modulo addition or removal of at most $r-1$ elements):

$$(M, A),\ (M, \emptyset),\ (A, \emptyset),\ (M-A, \emptyset),\ (\emptyset, \emptyset)$$

and only the first 4 of these depend on $\mathbf{M}$. Thus, there are 4 cases to consider ($s = 0, \ldots, 3$ below). We give the definition and then illustrate how it works by means of an example.

Definition 10 Given $\mathbf{M}$, r, and $\mathbf{Q} = \{Q_0, \ldots, Q_{u-1}\}$ as above, let[11]

(13) $c_{r,\mathbf{Q}}(\mathbf{M}) =$

$$\begin{aligned}\{(s,t,j,m) : t \in [0,u)\ \&\ j, m \in (-r, r)\ \&\ |j+m| < r\ \&\\ ((s = 0\ \&\ Q_t(|M-A|+m, |A|+j)\ \&\ j+m \leq 0)\\ \text{or } (s = 1\ \&\ Q_t(|M|+m, j)\ \&\ j+m \leq 0)\\ \text{or } (s = 2\ \&\ Q_t(|A|+m, j)\ \&\ j+m \leq |M-A|)\\ \text{or } (s = 3\ \&\ Q_t(|M-A|+m, j)\ \&\ j+m \leq |A|))\}.\end{aligned}$$

$C_{r,\mathbf{Q}}$ is the set of colors of the form $c_{r,\mathbf{Q}}(\mathbf{M})$. Clearly $C_{r,\mathbf{Q}}$ is finite.

[10] Actually, type $\langle 1, 1\rangle$ quantifiers satisfying Isom will do. Closure under duals is not used in this result.

[11] $[0, u)$ is the interval $0, 1, \ldots, u-1$, and $(-r, r)$ is $-(r-1), \ldots, -1, 0, 1, \ldots, r-1$, etc. If j is an integer, $|j|$ is its absolute value.

The side conditions on $j+m$ in each case reflect the fact that we are only allowed to 'move around' elements inside the universe M, not to add elements to it. In the case $s=0$, for example, we add (if $m>0$) or delete (if $m<0$) m elements to $M-A$, and we add or delete j elements to A. As long as $j+m \leq 0$, this can be done by moving around elements inside M.

To look at another case in more detail, consider the 4-tuple $(2,3,-2,4)$ (assuming $r>4$). It concerns the behavior of Q_3 on a pair of sets, where the first is obtained by removing 2 elements, say a_1, a_2, from A, and the second by adding 4 elements to $\emptyset$. Of these, two can be the ones removed from A, and two, say b_1, b_2, in general would have to come from $M-A$. Thus, it is necessary that $-2+4=2 \leq |M-A|$. Letting $X_1 = A \cup \{b_1, b_2\}$ and $X_2 = \{a_1, a_2, b_1, b_2\}$, we then have

$$\begin{aligned} (2,3,-2,4) \in c_{r,\mathbf{Q}}(\mathbf{M}) &\iff Q_3(|A|-2,4) \\ &\iff Q_3(|X_1-X_2|, |X_1 \cap X_2|) \\ &\iff (Q_3)_M(X_1, X_2). \end{aligned}$$

Next, we observe that all these colors are describable by $\mathrm{FO}^r(\mathbf{Q})$-sentences.

Lemma 11 *For each $c \in C_{r,\mathbf{Q}}$ there is a sentence φ_c in $\mathrm{FO}^r(\mathbf{Q})$ such that for all $\mathbf{M}$,*

$$\mathbf{M} \models \varphi_c \iff c_{r,\mathbf{Q}}(\mathbf{M}) = c.$$

Proof. It will suffice to find $\mathrm{FO}^r(\mathbf{Q})$-sentences ψ_{stjm} satisfying

(14) $\mathbf{M} \models \psi_{0tjm} \iff Q_t(|M-A|+m, |A|+j) \;\&\; j+m \leq 0$

(15) $\mathbf{M} \models \psi_{1tjm} \iff Q_t(|M|+m, j) \;\&\; j+m \leq 0$

(16) $\mathbf{M} \models \psi_{2tjm} \iff Q_t(|A|+m, j) \;\&\; j+m \leq |M-A|$

(17) $\mathbf{M} \models \psi_{3tjm} \iff Q_t(|M-A|+m, j) \;\&\; j+m \leq |A|$.

For then, if Z is the (finite) set $\{(s,t,j,m) : s \in [0,3] \;\&\; t \in [0,u) \;\&\; j,m \in (-r,r) \;\&\; |j+m| < r\}$, we can let

$$\varphi_c = \bigwedge_{(s,t,j,m)\in c} \psi_{stjm} \;\wedge \bigwedge_{(s,t,j,m)\in Z-c} \neg\psi_{stjm}.$$

Let us look at (16) as an example; the others are similar. In this case we may assume $j \geq 0$.

Case 1. $m \geq 0$. Let ψ_{2tjm} say

$\exists$ distinct $x_1, \ldots, x_{j+m} \in M-A$ such that
$(Q_t)_M(A \cup \{x_1, \ldots, x_{j+m}\}, \{x_1, \ldots, x_j\})$.

Then (16) holds. Since $j + m < r$, this can be expressed in $\mathrm{FO}^r(\mathbf{Q})$.

Case 2. $m < 0$. Let $m' = -m$.

Case 2.1. $j \leq m'$. Now $|M - A| \geq j$ holds trivially. Let ψ_{2tjm} say

$$\begin{array}{l}\exists \text{ distinct } x_1, \ldots, x_{m'} \in A \text{ such that} \\ (Q_t)_M(A - \{x_{j+1}, \ldots, x_{m'}\}, \{x_1, \ldots, x_j\}).\end{array}$$

Then (16) holds.

Case 2.2. $j > m'$. (Cf. the example discussed after Definition 10.) This time, let ψ_{2tjm} say

$$\begin{array}{c}\exists \text{ distinct } x_1, \ldots, x_{j-m'} \in M - A \text{ and distinct } y_1, \ldots, y_{m'} \in A \text{ s.t.} \\ (Q_t)_M(A \cup \{x_1, \ldots, x_{j-m'}\}, \{x_1, \ldots, x_{j-m'}, y_1, \ldots, y_{m'}\}).\end{array}$$

Again it is clear that (16) holds. □

Now we get our second characterization of the 'closeness' relation $\equiv_{\mathrm{FO}^r(\mathbf{Q})}$:

Proposition 12 *Let* $\mathbf{Q}$ *be a finite set of CE quantifiers containing* some *and closed under duals. The following are equivalent, for* $r > 0$*:*

(a) $c_{r,\mathbf{Q}}(\mathbf{M}) = c_{r,\mathbf{Q}}(\mathbf{M}')$

(b) $\mathbf{M} \equiv_{\mathrm{FO}^r(\mathbf{Q})} \mathbf{M}'$

Proof. (b) ⇒ (a): This is a consequence of Lemma 11.

(a) ⇒ (b): Assume $c_{r,\mathbf{Q}}(\mathbf{M}) = c_{r,\mathbf{Q}}(\mathbf{M}')$, with $\mathbf{M} = (M, A)$ and $\mathbf{M}' = (M', A')$. Suppose $(a_1, \ldots, a_{r-1}) \simeq (a'_1, \ldots, a'_{r-1})$, let X_1, X_2 be $\{a_1, \ldots, a_{r-1}\}$-variants of two of $\emptyset, A, M - A, M$, and let X'_1, X'_2 be the corresponding $\{a'_1, \ldots, a'_{r-1}\}$-variants of the same two sets. As an example, suppose X_1 and X_2 are both $\{a_1, \ldots, a_{r-1}\}$-variants of A; the other cases are similar. It is easy to see that there are $j, m \in (-r, r)$ such that $|j + m| < r$, $j + m \leq |M - A|, |M' - A'|$, and $|X_1 \cap X_2| = |A| + m$, $|X'_1 \cap X'_2| = |A'| + m$, $|X_1 - X_2| = |X'_1 - X'_2| = j$. Take any $t \in [0, u)$ and let $Q_t^{\mathrm{d}} = Q_{t'}$. We now

$$
\begin{array}{c}
- \\
- \; + \\
- \; - \; + \\
- \; + \; + \; + \\
- \; + \; + \; - \; + \\
- \; - \; + \; + \; - \; + \\
- \; - \; + \; + \; + \; - \; + \\
- \; - \; + \; + \; + \; - \; - \; + \\
- \; - \; - \; + \; + \; + \; - \; - \; + \\
- \; - \; - \; + \; + \; + \; - \; - \; - \; + \\
- \; - \; - \; + \; + \; + \; + \; - \; - \; - \; + \\
- \; - \; - \; - \; + \; + \; + \; + \; - \; - \; - \; + \\
- \; - \; - \; - \; - \; + \; + \; + \; + \; - \; - \; - \; +
\end{array}
$$

Figure 4: A boundedly oscillating quantifier.

have (cf. (6) in Section 2.4)

$$
\begin{aligned}
(Q_t)_M(X_1, X_2) &\iff Q_t(j, |A| + m) \\
&\iff \neg Q_{t'}(|A| + m, j) \\
&\iff (2, t', j, m) \notin c_{r,\mathbf{Q}}(\mathbf{M}) \\
&\iff (2, t', j, m) \notin c_{r,\mathbf{Q}}(\mathbf{M}') \\
&\iff \neg Q_{t'}(|A'| + m, j) \\
&\iff Q_t(j, |A'| + m) \\
&\iff (Q_t)_{M'}(X'_1, X'_2).
\end{aligned}
$$

By definition, this means that $\mathbf{M} \approx_{r,\mathbf{Q}} \mathbf{M}'$, so (b) follows from Proposition 9. □

4 Definability from monotone quantifiers

4.1 Definability from Q_f vs. from Q_f^{rel}

Consider first the question of when (a simple unary) Q is definable from monotone simple unary quantifiers, i.e., from quantifiers of the form Q_f. This question has an answer which is easily representable in the number triangle. The following definition and result are from Väänänen [10].

Definition 13 The *oscillation* of a simple unary or CE quantifier Q at level n in the number triangle is the number of times Q switches from $+$ to $-$ or vice versa at that level. Q has *bounded oscillation* if there is a finite bound m on the oscillation of Q (at any level). See Figure 4.

For example, any monotone quantifier has bounded oscillation (with $m = 1$). The quantifier *either-between-3-and-5-or-more-than-8* has bounded oscillation (with $m = 3$). *an-even-number-of* is a typical quantifier with unbounded oscillation.

Theorem 14 (**Bounded Oscillation Theorem** [10]) *Q is definable in terms of quantifiers of the form Q_f if and only if Q has bounded oscillation.*

This is no longer true if we consider definability in terms of quantifiers of the form Q_f^{rel} instead. Indeed, *an-even-number-of* is definable from one such quantifier. This is a consequence of the next observation.

In the linguistic literature a CE quantifier Q is often called *intersective* if the relation $Q_M(A, B)$ only depends on (the size of) $A \cap B$, i.e., if

$$|A \cap B| = |A' \cap B'| \text{ implies } (Q_M(A, B) \Leftrightarrow Q_{M'}(A', B'))$$

(for $A, B \subseteq M$ and $A', B' \subseteq M'$).[12] When Q is CE, this condition is easily seen to be equivalent to the *symmetry* of Q, i.e.,

$$Q_M(A, B) \implies Q_M(B, A).$$

Examples are *some*, *between-three-and-six*, and *an-even-number-of*, but not e.g. *most* or *all*.[13]

Another equivalent characterization of intersectivity is the following:

There is a set $S \subseteq \mathbb{N}$ such that for all M and all $A, B \subseteq M$,
$Q_M(A, B) \iff |A \cap B| \in S$.

From this we also see that if Q^{rel} is intersective then Q^{rel} is definable in terms of Q, and that

(18) $Q_M(A) \iff |A| \in S.$

Proposition 15 *If a CE quantifier is intersective, then it is definable in* FO(Q_f^{rel}) *for some function f.*

Proof. We may assume (cf. the remarks above, or Lemma 7) that the quantifier Q in question is simple unary and as in (18). Define

$$f(n) = \begin{cases} 1 & \text{if } n \in S \\ 2 & \text{if } n \notin S. \end{cases}$$

[12] Intersectivity has been invoked in the context of various natural language phenomena; cf. Keenan and Westerståhl [7].

[13] On the other hand, *all* and *all-but-four* (but not *most*) are *co-intersective*, in that $Q_M(A, B)$ only depends on $A - B$, or in other words that the dual is intersective. So Proposition 15 below holds for co-intersective quantifiers as well.

Then

$$
\begin{aligned}
Q_M(A) &\iff 1 \geq f(|A|) \\
&\iff A = \emptyset \text{ or } \exists a \in A(|\{a\}| \geq f(|A|)) \\
&\iff (M, A) \models \neg\exists x Ax \ \vee \ \exists x(Ax \wedge Q_f^{\mathrm{rel}} y(Ay, y = x)) \qquad (19)
\end{aligned}
$$

(assuming $0 \in S$; otherwise delete the first disjunct). □

Looking at the proof of this proposition, can we conclude anything about the expressibility in English of *an-even-number-of* in terms of monotone quantifiers? The question is vague but appears to split into two parts: whether Q_f^{rel} in the proof (when S is the set of even numbers) is an NL quantifier, and whether the form of the definition itself (the sentence used in (19)) is 'natural' in English.

As to the first issue, note that Q_f^{rel} is defined by

(20) $(\textit{some}(A, B) \wedge \textit{an-even-number-of}(A, A))$
$\vee\ (\textit{at-least-two}(A, B) \wedge \textit{an-odd-number-of}(A, A))$.

This condition is readily expressed in English as

> Either some A's are B and there is an even number of A's, or at least two A's are B and there is an odd number of A's.

However, the issue at hand is not whether the condition is easily expressible in English, but whether it can be seen as the denotation of an English *determiner*. And it seems to us that this is not the case — note that (20) is not a Boolean combination of determiner denotations.

Irrespective of this, it also seems to us that the defining sentence used in (19), though quite natural in FO(Q_f^{rel}), is not so easily expressed in English, the main problem being the unit set occurring as the second argument of Q_f^{rel}.[14] We conclude, then, that although *an-even-number-of* is logically definable from a monotone CE quantifier, it does not seem to be expressible by an English determiner phrase involving a monotone NL quantifier.

4.2 Color oscillation

To generalize Theorem 14 to the case of definability in terms of monotone CE quantifiers we first need to generalize the notion of bounded oscillation.

[14] This second argument should correspond to an English VP. One might try, relative to *some A* introduced in the first part of the sentence, the VP "is identical to *it*". But this seems strained. The use of the identity relation in English appears to be much more restricted than its use in FO languages — it would be interesting to know exactly how.

Let α be any coloring of the set $\mathbb{N}$ of natural numbers. α *induces* a coloring of the number triangle $\mathbb{N}^2$ as follows:

$$\alpha(i,j) \;=\; (\alpha(i),\alpha(j)).$$

The *color class* of a point in the number triangle is the equivalence class of that point under the relation of having the same (induced) color. On each level of the triangle, the points are partitioned into the color classes restricted to that level.

Definition 16 A (simple unary or CE) quantifier Q has *bounded color oscillation* if there is a coloring α of $\mathbb{N}$ and a finite bound m such that, at each level n, Q oscillates (i.e., changes from + to − or vice versa) at most m times *inside each color class.*

Note first that it is crucial that we consider only induced colorings of $\mathbb{N}^2$ — otherwise *every* CE quantifier would have bounded color oscillation (with 2 colors: one for the $+'$s and one for the $-'$s).

If Q has bounded oscillation — in particular, if it is monotone — then Q trivially has bounded color oscillation, with just one color. Now let $\alpha(n)$ = red, if n is even, and $\alpha(n)$ = green otherwise. There are 4 induced colors (ordered pairs of red and green). If Q is the quantifier *an-even-number-of* then, at each level, and inside each color class, Q does not oscillate at all. Thus Q has bounded color oscillation.

Here is a slightly more complex example of a CE quantifier with bounded color oscillation (same colors as above, $m = 2$):

$$\begin{aligned} Q_M(A,B) \;\;\Leftrightarrow\;\; & |A\cap B| \text{ is even and } 1/2\cdot|A| \le |A\cap B| \le 3/4\cdot|A|, \\ & \text{or } |A\cap B| \text{ is odd and } 1/3\cdot|A| \le |A\cap B| \le 2/3\cdot|A|. \end{aligned}$$

One might think this is far removed from things we say in natural languages. However, according to our somewhat idealized notion of an NL quantifier in Section 1.1, this *is* an NL quantifier, since it is a Boolean combination of some simple proportional quantifiers and *an-even-number-of.*

The next theorem is our main result, of which Theorem 14 is a special case.

Theorem 17 *Q is definable in terms of quantifiers of the form Q^{rel}_f if and only if Q has bounded color oscillation.*

Proof. From right to left: This is the easier direction. Suppose Q has bounded color oscillation, relative to a coloring α of $\mathbb{N}$ and with the bound

m. Inside each color class, Q divides the number triangle into at most $m+1$ sectors, containing alternately only $+'$s or only $-'$s. These sectors can be described using certain functions g_i such that each interval $[g_i(n), g_{i+1}(n))$, restricted to points in the color class, constitutes one sector at level n. Moreover, each color c can be represented by its characteristic function f_c. Then $(n-j, j)$ is in Q iff there are colors c, d such that $(n-j, j)$ is in one of the $+$ sectors with the induced color (c, d). The number of colors is finite, and the number of sectors is bounded by $m+1$. This means that Q can be defined in terms of quantifiers of the form $Q^{\text{rel}}_{f_c}$ and $Q^{\text{rel}}_{g_i}$, as follows.

For each α-color c, define

$$f_c(n) = \begin{cases} 0 & \text{if } \alpha(n) = c \\ 1 & \text{otherwise.} \end{cases}$$

Next, given n and α-colors c, d, let $C = \{j \in [0, n] : \alpha(n-j) = c \text{ and } \alpha(j) = d\}$. We select numbers $J_n(s, c, d)$ for $0 \leq s \leq m+2$ which enumerate in order of magnitude the oscillation points of Q at level n inside C:

If $C = \emptyset$, let each $J_n(s, c, d) = 0$. Otherwise, if

$$C = \{x_0, \ldots, x_p\} \text{ with } x_0 < \ldots < x_p,$$

call x_i a $+$ point if $(n - x_i, x_i) \in Q$, and a $-$ point otherwise. Define

- $J_n(0, c, d) = x_0$ and $J_n(m+2, c, d) = x_p + 1$.
- $J_n(1, c, d)$ is the first $-$ point in C if it exists; otherwise $J_n(1, c, d) = x_p + 1$.
- For $s \in [1, m]$, let $J_n(s+1, c, d)$ be the first point in C whose sign is different from $J_n(s, c, d)$ if there is such a point; otherwise we let $J_n(s+1, c, d) = x_p + 1$.

From this we see that, for $j \in C$,

$$Q(n-j, j) \iff \exists t\, (J_n(2t, c, d) \leq j < J_n(2t+1, c, d)).$$

Now let $g^s_{c,d}(n) = J_n(s, c, d)$. Then, $Q(n-j, j)$ iff there are α-colors c, d and a number t such that $0 \leq 2t \leq m+1$ and

(i) $f_c(n-j) = 0$

(ii) $f_d(j) = 0$

(iii) $g^{2t}_{c,d}(n) \leq j < g^{2t+1}_{c,d}(n)$.

Now (i) is expressed by the sentence

$$\psi_c \ = \ Q^{\text{rel}}_{f_c} x(Ax \wedge \neg Bx, x \neq x),$$

(ii) by

$$\theta_d \ = \ Q^{\text{rel}}_{f_d} x(Ax \wedge Bx, x \neq x),$$

and (iii) by

$$\varphi_{cdt} \ = \ Q^{\text{rel}}_{g^{2t}_{c,d}} x(Ax, Bx) \wedge \neg Q^{\text{rel}}_{g^{2t+1}_{c,d}} x(Ax, Bx).$$

That is, Q is definable as the finite disjunction

$$\bigvee_{\substack{c,d \in \text{range}(\alpha) \\ 0 \leq 2t \leq m+1}} (\psi_c \ \wedge \ \theta_d \ \wedge \ \varphi_{cdt}).$$

Now, for the other direction, suppose Q is $\text{FO}^r(\mathbf{Q})$-definable, where $\mathbf{Q} = \{Q_0, \ldots, Q_{u-1}\}$ is closed under duals, contains *some*, and $Q_t = Q^{\text{rel}}_{f_t}$ for $t \in [0, u)$. Define a coloring α by

$$\alpha(j) = \{(t, z, s) : f_t(j + z) = s, \ t \in [0, u), \ z \in (-r, r), \ s \in [0, r)\},$$

and let $m = u(2r-1)^2$. That is, j_1 and j_2 have the same α-color iff $f_t(j_1+z)$ and $f_t(j_2+z)$ ($t \in [0, u)$, $z \in (-r, r)$) agree on *small* ($< r$) values. We shall use Proposition 12 to show that within each of the induced color classes there are at most m oscillations at every level.

First note that since *some* $\in \mathbf{Q}$, we may assume f_0 is the constant function defined by $f_0(n) = 1$ for all $n \geq 0$. This means that for $z \in [0, r)$, $(0, -z, 1) \in \alpha(n) \Leftrightarrow n \geq z$, and so

(21) For $z \in [0, r)$, $\alpha(n_1) = \alpha(n_2)$ implies $n_1 \geq z \Leftrightarrow n_2 \geq z$.

Now we make the following

Claim: For any $(n - j_1, j_1)$ and $(n - j_2, j_2)$ in the same color class, if

$f_t(n+y)+z \leq j_1 \Leftrightarrow f_t(n+y)+z \leq j_2$ for $t \in [0, u)$ and $y, z \in (-r, r)$,

then $Q(n - j_1, j_1) \Leftrightarrow Q(n - j_2, j_2)$.

Note that $m = u(2r-1)^2$ is the variation of (t, y, z) in the Claim. So if Q changes sign at some level inside some color class, there must be t, y, z violating the condition of the Claim, and this can happen at most m times.

To prove the Claim, it suffices by Proposition 12 to show that, under the given assumptions,

$$c_1 = c_{r,\mathbf{Q}}((n-j_1, j_1)) \;=\; c_{r,\mathbf{Q}}((n-j_2, j_2)) = c_2$$

(where $(n-j_1, j_1)$ is a structure (M, A) such that $|M| = n$ and $|A| = j_1$, etc.). We have

$$\begin{aligned}
(0,t,y,z) \in c_1 \quad &\Rightarrow \quad y, z \in (-r, r),\ |y+z| < r,\ Q_t(n-j_1+z, j_1+y) \\
&\Rightarrow \quad j_1 \geq f_t(n+y+z) - y \text{ (since } Q_t = Q^{\mathrm{rel}}_{f_t}) \\
&\Rightarrow \quad j_2 \geq f_t(n+y+z) - y \text{ (by assumption in Claim)} \\
&\Rightarrow \quad Q_t(n-j_2+z, j_2+y) \\
&\Rightarrow \quad (0,t,y,z) \in c_2.
\end{aligned}$$

The next case uses only the fact that the two points are at the same level:

$$\begin{aligned}
(1,t,y,z) \in c_1 \quad &\Rightarrow \quad y, z \in (-r, r),\ |y+z| < r,\ Q_t(n+z, y) \\
&\Rightarrow \quad (1,t,y,z) \in c_2.
\end{aligned}$$

Next,

$$\begin{aligned}
(2,t,y,z) \in c_1 \quad &\Rightarrow \quad y, z \in (-r, r),\ |y+z| < r,\ y+z \leq n-j_1, \\
&\qquad \text{and } Q_t(j_1+z, y) \\
&\Rightarrow \quad y \geq s = f_t(j_1+y+z) \text{ and } y+z \leq n-j_1 \\
&\Rightarrow \quad y \geq s = f_t(j_2+y+z) \ \text{ (since } \alpha(j_1) = \alpha(j_2)) \text{ and} \\
&\qquad y+z \leq n-j_2 \ \text{ (by (21), since } \alpha(n-j_1) = \alpha(n-j_2)) \\
&\Rightarrow \quad Q_t(j_2+z, y) \text{ and } y+z \leq n-j_2 \\
&\Rightarrow \quad (2,t,y,z) \in c_2.
\end{aligned}$$

The case when $(3,t,y,z) \in c_1$ is similar. By symmetry, we can conclude that $c_1 = c_2$. □

4.3 An example

We now give an example of a CE quantifier Q_u which is not of bounded color oscillation, hence not definable in terms of monotone CE quantifiers. Q_u is best described by its pattern in the number triangle (Figure 5).

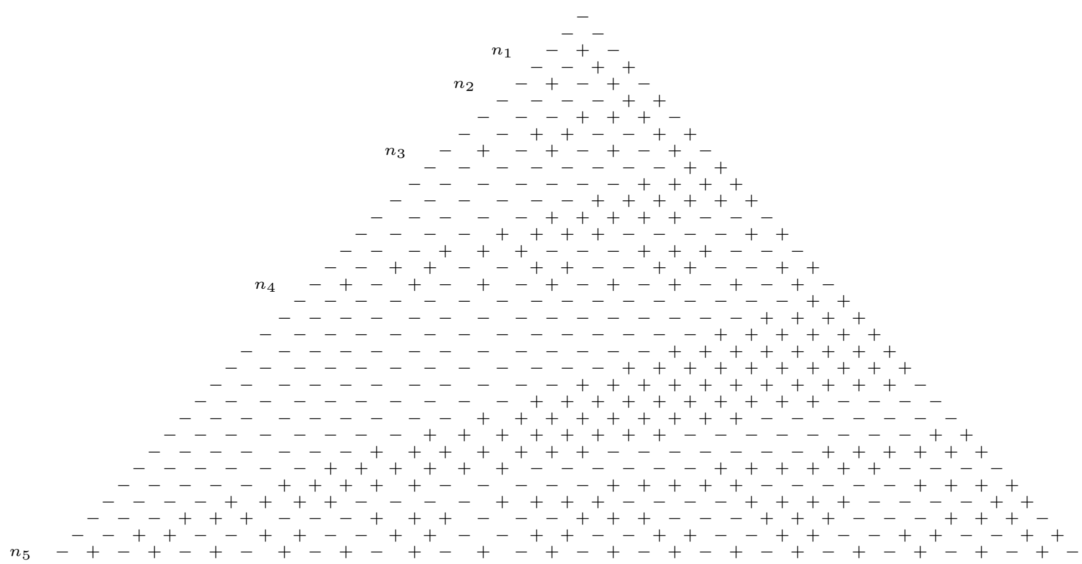

Figure 5: The quantifier Q_u.

Let $n_j = 2^j$, $j = 1, 2, \ldots$. At level n_j, Q_u switches between $-$ and $+$ at every step. One level up, at $n_j - 1$, there are two $-'$s, then two $+'$s, then two $-'$s, etc. At level $n_j - 2$, there are three $-'$s, then three $+'$s, etc. And so on, until we reach level n_{j-1}, where the same pattern begins all over again.

To show that Q_u does not have bounded color oscillation we use van der Waerden's Theorem. This theorem is a generalization of the Pigeon Hole Principle: Suppose you are putting objects in boxes. If there are more objects than boxes, at least one box is going to contain more than one object. If there are *many* more objects than boxes, chances are at least one box will contain a lot of objects. Instead of boxes we use colors. Suppose the objects are the numbers $0, 1, \ldots, n$, for some n. If there are many more numbers than colors, chances are many numbers will get the same color, and you can ask questions about what *kinds* of numbers get the same color. Van der Waerden's Theorem is about when numbers of the form

$$a + d, a + 2d, \ldots, a + md,$$

i.e., arithmetic progressions of finite length, all get the same color. It says that for any finite number of colors, if you take enough numbers you will get arithmetic progressions of arbitrary length in some color class (box). More precisely:

Van der Waerden's Theorem (cf. Graham, Rotschild and Spencer [4]) For any k and any m there is a number $W_k(m)$ such that if the numbers $0, 1, \ldots, W_k(m)$ are colored with k colors, then one color class contains an arithmetic progression of length m.

Now suppose Q_u had bounded color oscillation, relative to some coloring $\chi : \mathbb{N} \longrightarrow \{0, 1, \ldots, k-1\}$, and with bound $m-1$. To contradict this we need to find a large enough level n such that in some induced color class at level n, Q_u changes sign at least m times.

Take $n_j \geq W_{k^2}(m+2)$. Color each number i in the set

$$I = \{0, 1, \ldots, n_j\}$$

with the color $\beta(i) = (\chi(n_j + 1 - i), \chi(i))$. By van der Waerden's Theorem there are a and d such that the numbers

$$a + d, a + 2d, \ldots, a + (m+2)d$$

are in I and have the same β-color. Now look at the level $n = n_j - (d-1)$ of Q_u (Figure 6).[15] The pairs

$$(n_j - (d-1) - (a + ld), a + ld), \quad l = 1, \ldots, m+1,$$

at that level all have the same induced χ-color, since it follows from the fact that the elements of the arithmetic progression have the same β-color that

$$\chi(a + ld) = \chi(a + l'd)$$

and

$$\chi(n_j - (d-1) - (a + ld)) = \chi(n_j - (d-1) - (a + l'd))$$

for $1 \leq l, l' \leq m+1$. Thus, Q_u changes sign at least m times inside one color class at level n, a contradiction. We have proved

Proposition 18 *The CE quantifier Q_u does not have bounded color oscillation.*

[15] Note that $2d < n_j$, whence $n_j - (d-1) > n_j/2 = n_{j-1}$.

$$\cdots \underbrace{- \; - \; \overset{\overset{a+d}{\downarrow}}{-} \; \cdots \; -}_{d} \; \underbrace{+ \; + \; \overset{\overset{a+2d}{\downarrow}}{+} \; \cdots \; +}_{d} \; \underbrace{- \; - \; \overset{\overset{a+3d}{\downarrow}}{-} \; \cdots \; -}_{d} \; + \; + \; \cdots$$

Figure 6: Level $n_j - (d-1)$ of Q_u.

5 Smooth quantifiers

Mere monotonicity of a CE quantifier is a rather weak constraint: it is easy to find monotone CE quantifiers which are not NL quantifiers, for example,

$$Q_M(A,B) \iff |A| \text{ is even.}$$

From a natural language perspective one would like to find a strengthening of monotonicity. Requiring monotonicity in the left argument is too strong; and in any case all such left monotone quantifiers are FO definable (cf. Westerståhl [11]). Rather, what one would like is a property, say **P**, which implies monotonicity and is such that

(+) If Q is a monotone NL quantifier, then Q has **P**.

We now propose such a property; it was introduced by van Benthem and applied by him to various aspects of the computational behavior of quantifiers (cf. van Benthem [3]), but not as far as we know in the present context.[16]

Definition 19 A function $f : \mathbb{N} \longrightarrow \mathbb{N}$ such that $f(n) \leq n+1$ is *smooth* if, for all n,

$$f(n) \leq f(n+1) \leq f(n) + 1.$$

Also, Q_f and Q_f^{rel} are called *smooth* if f is smooth.

As we stated in the universal (U_3) (Section 1.3), it appears that all monotone NL quantifiers are indeed smooth. A typical example would be *at-least-m/n′ths-of*. Smooth quantifiers have a characteristic pattern in the number triangle. Since they are monotone, each level consists of a sequence of $-$′s followed by a sequence of $+$′s. But smoothness also means that if

[16]Van Benthem called the property *continuity*. This term has, however, been used in various senses in the literature on NL quantifiers; cf. Westerståhl [11], sections 3.6 and 4.2, for an overview (there the present property was called ‘SUPER CONT’).

$(n-j, j)$ is the leftmost + at level n (i.e., if $f(n) = j$), then the leftmost + at the next level $n+1$ is one of the two 'successors' $(n+1-j, j)$ and $(n-j, j+1)$.

The next fact, which can easily be established by looking in the number triangle, gives alternative characterizations of smoothness.

Fact 20 *If Q is a CE quantifier, the following are equivalent:*

(a) *Q is smooth, i.e. $Q = Q_f^{rel}$ for some smooth f.*

(b) *$Q = Q_f^{rel}$ for some f such that both f and f^d are non-decreasing.*[17]

(c) *For all M and all $A, A', B \subseteq M$,*

  (c1) $QAB, A' \subseteq A, A' \cap B = A \cap B \Rightarrow QA'B$

  (c2) $QAB, A \subseteq A', A' - B = A - B \Rightarrow QA'B.$

Smoothness of f also means that, except in trivial cases, it eventually leaves the edges of the number triangle. To make this precise, let us call f *trivial* if either f or f^{d} is eventually constant. It was noted in Westerståhl [12] and Kolaitis and Väänänen [8] that f is trivial if and only if Q_f is FO definable.

Lemma 21 *If f is non-trivial and smooth, then*

$$\forall r\, \exists N\, \forall n > N\; (r < f(n) < n - r).$$

Proof. Suppose f is smooth and non-trivial but that for some r,

$$\forall N\, \exists n > N (f(n) \leq r \text{ or } n - f(n) \leq r).$$

Case 1. $\exists n_0 < n_1 < \ldots$ such that $f(n_i) \leq r$ for all i. Since f is non-decreasing there must be an I such that $f(n_i) = c \leq r$ for $i > I$. Then $f(n)$ is eventually c, contradicting our assumption.

Case 2. Not Case 1. Then, by assumption, $\exists n_0 < n_1 < \ldots$ such that $n_i - f(n_i) \leq r$ for all i. Since $n - f(n)$ is also non-decreasing, the same argument as in Case 1 shows that $n - f(n)$ is eventually constant, which also contradicts the assumption. □

[17] I.e. $f(n) \leq f(n+1)$ for all n.

We showed in Section 4.1 that *an-even-number-of* is definable in terms of monotone CE quantifiers. In fact, it is definable in terms of a quantifier of the form Q_g^{rel} with an non-decreasing g: Let

$$g(n) = \begin{cases} n & \text{if } n \text{ is even} \\ n-1 & \text{if } n \text{ is odd.} \end{cases}$$

g is a function which is non-decreasing but not smooth. We have

$$\begin{aligned} |A| \text{ is even} &\iff g(|A|) \not\leq |A|-1 \\ &\iff \neg[\exists a \in A\ (Q_g^{\mathrm{rel}})_M(A, A-\{a\})] \end{aligned}$$

(as before, it suffices to show that Q_{even}, i.e., the simple unary quantifier '$|A|$ is even', is definable in this way).

However, *an-even-number-of* is *not* definable in terms of smooth quantifiers. This is a consequence of our next result.

Theorem 22 *Suppose Q is definable in terms of smooth quantifiers. Then Q has bounded oscillation (without colors).*

Proof. Start as in the second half of the proof of Theorem 17, with Q definable in $\mathrm{FO}^r(\{Q_{f_0}^{\mathrm{rel}}, \ldots, Q_{f_{u-1}}^{\mathrm{rel}}\})$, and the coloring α defined there, which was shown to be such that $m = u(2r-1)^2$ bounds the oscillation of Q inside each induced color class at any level. This time we may assume that the f_t are smooth, and clearly also that they are non-trivial. For $t \in [0, u)$, let N_t be as in Lemma 21, and let $N = \max\{N_0, \ldots, N_{u-1}\}$. Thus,

$$\forall t \in [0, u)\, \forall n > N\ (r < f_t(n) < n - r).$$

But this means that, for large enough n, with $z \in (-r, r)$ and $s \in [0, r)$, $f_t(n+z) \neq s$. That is, if the functions leave the edges, no values are *small*, so trivially the functions eventually agree for small values. Hence,

$$\exists n_0\, \forall n, n' \geq n_0 (\alpha(n) = \alpha(n')).$$

It follows that there is n_1 such that at any level $n \geq n_1$, all the points between $(n - n_0, n_0)$ and $(n_0, n - n_0)$ are in the same induced color class. Hence on those points, Q oscillates at most m times. It may oscillate on the other points at level n, but at most $2n_0$ times, and it may oscillate differently at levels before n_1, but at most n_1 times. Thus, if $m' = \max\{m + 2n_0, n_1\}$, then m' bounds the total oscillation of Q. $\square$

Finally, we observe that the converse of Theorem 22 fails: there is a quantifier Q which has bounded oscillation but is not definable in terms of smooth quantifiers. Such a quantifier has already been mentioned, namely,

$$Q_M(A, B) \iff |A| \text{ is even},$$

or, since we might as well take the corresponding simple unary quantifier,

$$Q_M(A) \iff |M| \text{ is even}.$$

Clearly this quantifier has bounded oscillation: each level consists either entirely of $+'$s or entirely of $-'$s, so it doesn't oscillate at all (indeed it is monotone).

Proposition 23 '$|M|$ is even' *is not definable in terms of smooth quantifiers.*

Proof. We use Proposition 9. To simplify, think of the quantifier in question as having the empty vocabulary, noting that Proposition 9 holds for the empty vocabulary as well. Thus it suffices to find, for each r, two sets M and M', where one has even and the other has odd cardinality, such that $M \approx_{r,\mathbf{Q}} M'$. Here we may again assume that $\mathbf{Q} = \{Q^{\text{rel}}_{f_1}, \ldots, Q^{\text{rel}}_{f_k}\}$ is as in the proof of Theorem 22, and that N is as in that proof, so that

$$\forall t \in [0, u)\, \forall n > N\, (r < f_t(n) < n - r).$$

Now take M and M' such that $|M| > N + r$ and $|M'| = |M| + 1$. Let X_1, X_2 be variants of $\emptyset$ or M obtained by 'moving' at most $r - 1$ elements of M, and let X'_1, X'_2 be the corresponding variants in M'. We must check that

$$\begin{array}{llcll}
(22) & X_j \neq \emptyset & \iff & X'_j \neq \emptyset & j = 1, 2, \\
(23) & X_j = M & \iff & X'_j = M' & j = 1, 2, \\
(24) & |X_1 \cap X_2| \geq f_t(|X_1|) & \iff & |X'_1 \cap X'_2| \geq f_t(|X'_1|) & t \in [0, u).
\end{array}$$

Call a subset X of M (M') *small* if $|X| < r$, and *big* if $|M - X| < r$ ($|M' - X| < r$). The X_j and X'_j are either big or small, so (22) and (23) are immediate. As to (24), if X_1 is small, so is $X_1 \cap X_2$, so $|X_1| = |X'_1|$ and $|X_1 \cap X_2| = |X'_1 \cap X'_2|$, and thus (24) holds. If X_1 is big, so is X'_1, and

$$r < f_t(|X_1|), f_t(|X'_1|) < n - r.$$

Thus, (24) holds whether $X_1 \cap X_2$ is big or small. □

6 Concluding remarks

Hopefully, some light has now been shed on the question of when a quantifier is definable in terms of monotone CE quantifiers. We end with a list of questions, some logical and some linguistic, that arose during the course of this paper:

(1) Does the divisibility quantifier *Div* (Section 1.1) have bounded color oscillation? We conjecture that it doesn't.

(2) Theorem 22 gave a necessary condition for definability in terms of smooth quantifiers. What would a necessary *and* sufficient condition look like?

(3) Is universal (U_2) true, i.e., is it really the case that all NL quantifiers are definable in terms of monotone CE ones? Also, is it true that *an-even-number-of* is not definable in terms of monotone NL quantifiers?

(4) Our example of a CE quantifier with unbounded color oscillation used van der Waerden's Theorem; the size of the model needed in the proof of Proposition 18 grows as fast as the van der Waerden function $W_k(m)$. Is this necessary? If it is, one might even be able to prove that the existence of quantifiers with unbounded color oscillation implies van der Waerden's Theorem, along the lines of Luosto [9].

(5) Is universal (U_3) true?

References

[1] Barwise, J. and Cooper, R, 1981. Generalized quantifiers and natural language, *Linguistics & Philosophy* **4**, 159–219.

[2] van Benthem, J., 1984. Questions about quantifiers, *Journal of Symbolic Logic* **49**, 443–466.

[3] van Benthem, J., 1986. *Essays in Logical Semantics*, D. Reidel, Dordrecht.

[4] Graham, R., Rotschild, B., and Spencer, J., 1990. *Ramsey Theory*, John Wiley & Sons, New York.

[5] Hella, L., Luosto, K., and Väänänen, J., 1996. The hierarchy theorem for generalized quantifiers, *Journal of Symbolic Logic* **61**, 802–819.

[6] Keenan, E. and Stavi, J., 1986. A semantic characterization of natural language determiners, *Linguistics & Philosophy* **9**, 253–326.

[7] Keenan, E. and Westerståhl, D, 1997. Generalized quantifiers in linguistics and logic, in J. van Benthem and A. ter Meulen (eds), *Handbook of Logic and Language*, Elsevier, Amsterdam, 837–893.

[8] Kolaitis, Ph. and Väänänen, J., 1995. Generalized quantifiers and pebble games on finite structures, *Annals of Pure and Applied Logic* **74**, 23–75.

[9] Luosto, K., 1999. Ramsey theory is needed for solving definability problems of generalized quantifiers, *Lecture Notes in Computer Science* 1754, Springer, Berlin, 121–134.

[10] Väänänen, J., 1997. Unary quantifiers on finite models, *Journal of Logic, Language and Information* **6**, 275–304.

[11] Westerståhl, D., 1989. Quantifiers in formal and natural languages, in D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic*, vol. IV, D. Reidel, Dordrecht, 1–131.

[12] Westerståhl, D., 1991. Relativization of quantifiers in finite models, in J. van der Does and J. van Eijck (eds), *Generalized Quantifier Theory and Applications*, ILLC, Amsterdam, 187–205. Also in *Quantifiers: Logic, Models and Computation* (same editors), CSLI Publications, Stanford, 1996, 375–383.