Zeitschr. f. math. Logik und Grundlagen d. Math.
Bd. 23, S. 547 – 554 (1977)

SOME REMARKS ON GENERALIZED SPECTRA

by L. LOVÁSZ and P. GÁCS in Budapest (Hungary)

Introduction

A considerable part of combinatorics deals, even if it is not stated explicitly, with finite structures: the combinatorial objects considered can be described as finite structures, in such a way that the properties investigated are invariant under isomorphism. A property of structures is automatically isomorphism-invariant if it states the satisfaction of a given formula on the structure. The "simplicity" of the property can then be characterized by the simplicity of the defining formula. Another important characterisation of properties arises by considering the complexity of their computation. R. FAGIN [1–4] discovered that a class of structures can be defined by the satisfaction of an existential second order formula if and only if it is recognizable by a non-deterministic Turing machine in polynomial time. In this note we use his results to add some remarks on polynomially complete problems and diagonalization.

The simplest formulas are the first order ones. For a formula φ, $\mathrm{Mod}_\omega(\varphi)$ is the class of finite models of φ. A class $\boldsymbol{A}$ is called *elementary* if it is $\mathrm{Mod}_\omega(\varphi)$ for some φ. Many important classes of structures are elementary: finite groups, ordered sets, lattices etc. In dealing with structures we shall need not only classes of them but also their functions: let us call them operators. In analogy with elementary classes we shall define the notion of an elementary operator.

Definition 1. Let us have a similarity type $\mathscr{S} = \{Q_1, \ldots, Q_m\}$, where Q_j are relation or function symbols. By a *structure* $\mathfrak{A}$ *of similarity type* $\mathscr{S}$ we understand a set $|\mathfrak{A}|$ with an interpretation $Q_j^{\mathfrak{A}}$ of the relation and function symbols on it *and* a relation $E^{\mathfrak{A}}(x, y)$ written also as $x =^{\mathfrak{A}} y$ called equality which satisfies the usual equality axioms.

The pair of structures $\langle \mathfrak{A}_1, \mathfrak{A}_2 \rangle$ is a new structure given in the following way. If the similarity type of $\mathfrak{A}_i$ is $\mathscr{S}_i$ then the similarity type of $\mathfrak{A} = \langle \mathfrak{A}_1, \mathfrak{A}_2 \rangle$ is $\mathscr{S}_1 \cup \mathscr{S}_2 \cup \{U\}$ where U is a new unary predicate symbol. $|\mathfrak{A}| = |\mathfrak{A}_1| \cup |\mathfrak{A}_2|$. Equality and all old relations remain as they were on $|\mathfrak{A}_i|$. Elements from different $|\mathfrak{A}_i|$ are unequal. The old relations are false on mixed tuples of elements, $f(x_1, \ldots, x_k) = x_1$, if f is an old function symbol and $(x_1, \ldots, x_k)$ a mixed tuple of elements. $U^{\mathfrak{A}}(x) \leftrightarrow x \in |\mathfrak{A}_1|$. $\langle \mathfrak{A}_1, \mathfrak{A}_2, \mathfrak{A}_3 \rangle$ is defined as $\langle \langle \mathfrak{A}_1, \mathfrak{A}_2 \rangle, \mathfrak{A}_3 \rangle$ and so on.

Definition 2. Let $\boldsymbol{A}, \boldsymbol{B}$ be two elementary classes of similarity type $\mathscr{S}, \mathscr{T}$ respectively. $F: \boldsymbol{A} \to \boldsymbol{B}$ is called an *elementary operator* if it can be obtained in the following way:

1) We choose a natural number k and two formulas $\tau(x_1, \ldots, x_k)$, $\varepsilon(x_1, \ldots, x_{2k})$ in the language $\mathrm{L}(\mathscr{S})$.

2) With each predicate symbol R of degree d in $\mathscr{T}$ we associate a formula $\varrho_R(x_1, \ldots, x_{kd})$, with each function symbol R in $\mathscr{T}$ of degree d we associate a formula $\varrho_R(x_1, \ldots, x_{(d+1)k})$.

3) We define the underlying set of $F(\mathfrak{A}) = \mathfrak{B}$ ($\mathfrak{A} \in \boldsymbol{A}$) as $\{(a_1, \ldots, a_k) \mid a_i \in |\mathfrak{A}|$ and $\tau^{\mathfrak{A}}(a_1, \ldots, a_k)$ holds$\}$,

$$(a_1, \ldots, a_k) =^{\mathfrak{B}} (a_{k+1}, \ldots, a_{2k}) \leftrightarrow \varepsilon^{\mathfrak{A}}(a_1, \ldots, a_{2k})$$

and the relation R by

$$R^{\mathfrak{B}}((a_1, \ldots, a_k), \ldots, (a_{(d-1)k+1}, \ldots, a_{dk})) \leftrightarrow \varrho_R^{\mathfrak{A}}(a_1, \ldots, a_{dk}).$$

Function symbols are defined analogously. Moreover, we choose ε such that $=^{\mathfrak{B}}$ should satisfy the equality axioms.

An *elementary operator* of two arguments is an elementary operator on the pair of structures. The composition of elementary operators is again elementary. Elementary operators often arise when we reduce a combinatorial problem to another one, e.g. the problem of finding a maximum matching to a problem of integer programming.

There are very common non-elementary classes of structures: the simplest example is the class of finite sets with odd cardinality. To express this class we have to allow at least one quantifier on relations: we need a second order formula. The simplest second order formulas contain, in prenex form, only existential quantifiers on relations and functions. They are of the form $(\exists R_1, \ldots, \exists R_k)\, \varphi$ where φ is a first order formula. A set $\boldsymbol{A}$ of natural numbers is the *spectrum* of the first order formula $\varphi \in \mathrm{L}(\mathscr{S})$ ($\mathscr{S} = \{Q_1, \ldots, Q_m\}$) if $\boldsymbol{A} = \{n \mid (\exists Q_1, \ldots, Q_m)\, \varphi$ is true on $\{1, \ldots, n\}\}$ (see [1]).

A class $\boldsymbol{A}$ of structures of similarity type $\mathscr{S}$ is the *generalized spectrum* of the first order formula $\varphi \in \mathrm{L}(\mathscr{S} \cup \mathscr{T})$ ($\mathscr{T} = \{R_1, \ldots, R_n\}$) if

$$\boldsymbol{A} = \{\mathfrak{A} \mid (\exists R_1, \ldots, R_n)\, \varphi \text{ is true on } \mathfrak{A}\}.$$

Spectra and generalized spectra can be regarded as projections of elementary classes: $\boldsymbol{A}$ is the projection of $\boldsymbol{B} = \{\mathfrak{B} \in \mathrm{Fin}(\mathscr{S} \cup \mathscr{T}) \mid \varphi$ is true on $\mathfrak{B}\}$ to $\mathrm{Fin}(\mathscr{S})$. We shall denote the generalized spectrum of φ by $\mathrm{Mod}_\omega(\exists \mathscr{T} \varphi)$. Every generalized spectrum can be obtained as the projection of an elementary class of a very simple type: the formula φ can be required to have the form $(\forall x_1, \ldots, x_k)\, \sigma$ where σ is quantifier free. (The transformation process is known from logic: The essence is to replace parts of the form $\forall x\, \exists y$ by parts of the form "there is a function such that" . . .) On the other hand, projections of more general classes give also generalized spectra. FAGIN has shown in [1] that the projection of any class of structures which is recognizable in polynomial time on a (deterministic or non-deterministic) Turing machine is a generalized spectrum. It is relatively easy to decide the satisfaction of a first order formula φ on the structure $\mathfrak{A}$: on an appropriate *deterministic* Turing machine this does not take running time more than $(\mathrm{card}\, \mathfrak{A})^{g(\varphi)}$ where $g(\varphi)$ is some function of φ. Not all classes decidable in polynomial time are elementary: an example is the class of odd numbers. Generalized spectra are characterizable as classes recognizable in polynomial time on a *non-deterministic* Turing machine.

For elementary classes no such characterization is known. A hierarchy among elementary classes and the existence of non-elementary generalized spectra can be established by model-theoretic methods. A hierarchy of classes given by second order formulas (the "analytic hierarchy") is not yet proved to exist: if the famous hypothesis P = NP holds then all these classes are generalized spectra, i.e. can be given already by an existential second order formula. More generally: denote by CNP the class of all sets which are complements of NP sets. Then CNP = NP is a weaker assumption, but it also implies that the above classes are generalized spectra. We return to this statement and give some estimates in Section 2.

PUDLÁK [10] proved that for all k, there is a generalized spectrum $\mathrm{Mod}_\omega(\exists\mathscr{R}\varphi)$, where φ contains k quantifiers, which is not representable as $\mathrm{Mod}_\omega(\exists\mathscr{T}\psi)$ with ψ having less than k quantifiers. His result relies on COOK's result on the hierarchy of non-deterministic Turing machine computations [9].

Among the generalized spectra there is a supposed hierarchy according to the maximum degree d of the predicates $R \in \mathscr{T}$ quantified. If $d = 1$, the generalized spectrum is called monadic. The class of generalized spectra with $d \leqq k$ is denoted by $\mathrm{J}_k(\mathscr{S})$. FAGIN proved in [2] that the class of connected graphs, which is of the form $\mathrm{Mod}_\omega(\exists\mathscr{R}\varphi)$ with $\mathscr{R}$ of degree two, is not monadic. (He uses the method of Fraissé-games. The same result was obtained somewhat later by HÁJEK, who used semiset technique. Ultraproducts also yield this result.) Whether non-monadic spectra can be further classified is unknown. [3–4] contains much information about this supposed hierarchy. The results of Section 2 imply that if CNP = NP holds then a hierarchy related to the degree hierarchy exists among the generalized spectra, thus compensating for the collapse of the analytic hierarchy. In fact, let us bring the second order formula φ to an equivalent form

$$\exists\mathscr{R}_1 \neg \exists\mathscr{R}_2 \ldots \neg\exists\mathscr{R}_m \neg (\exists x_1, \ldots, x_k)\,\psi \tag{1}$$

where ψ is quantifier-free and the $\mathscr{R}_i$ are sets of relation symbols. We define the *depth* $d(\varphi)$ of φ as the maximum degree of the relation symbols present in ψ and the *height* $h(\varphi)$ of φ as the m in (1). (The degree of a function symbol $f(x_1, \ldots, x_r)$ is $r + 1$.) Then CNP = NP implies that for all d, h there are generalized spectra which are not $\mathrm{Mod}_\omega(\varphi)$ for any sentence φ with $d(\varphi) \leqq d$ and $h(\varphi) \leqq h$.

The results of COOK and LEVIN imply that among the generalized spectra there are universal ones: the decision problem for any other generalized spectrum can be reduced to the decision problem of a universal one. We strengthen these results slightly in Section 1 by showing that the functions accomplishing the reduction can be required to be elementary operators. See JONES [12] and SCHNORR [11] for an analogous result. This also proves the existence of "complement-complete" generalized spectra without reference to automata theory.

Classical diagonalization fails to work for non-deterministic Turing machines: this makes the P = NP problem so hard. Yet the more modest results of Section 2 are obtained by this procedure.

1.

The results of COOK, KARP and LEVIN on polynomially complete combinatorial problems were reformulated by FAGIN for spectra not in their full power. He proved that there are certain generalized spectra – called complement-complete ones – the complement of which is a generalized spectrum if and only if the complement of every generalized spectrum is a generalized spectrum. The reason for this weakening was perhaps that until now there is no appropriate translation of the notion of a function computable in polynomial time on a *deterministic* Turing machine. Now we introduce a stricter reducibility notion replacing polynomially computable functions by elementary operators. For convenience, we shall treat generalized spectra as subsets of elementary classes:

Definition 3. Let $\mathscr{S} = \{Q_1, \ldots, Q_m\}$, $\mathscr{T} = \{R_1, \ldots, R_n\}$ be two similarity types, $\varphi \in \mathrm{L}(\mathscr{S})$, $\psi \in \mathrm{L}(\mathscr{S} \cup \mathscr{T})$ two first order formulas. The set of structures $\{\mathfrak{A} \in \mathrm{Fin}(\mathscr{S}) \mid \varphi$ holds on $\mathfrak{A}$ and $(\exists R_1, \ldots, R_n)\, \psi$ holds on $\mathfrak{A}\}$ is called the *generalized spectrum of the pair* of formulas (φ, ψ). It is denoted by $\mathrm{Mod}_\omega(\varphi \mathbin{\&} \exists \mathscr{T} \psi)$. The pair of formulas $M = (\varphi, \psi)$ will be called an (elementary) *search problem*. (On a structure satisfying φ we are searching for the appropriate relations $\mathscr{T}_i$.)

Definition 4. For two classes of structures $\boldsymbol{B}_i \subseteqq \mathrm{Fin}(\mathscr{S}_i)$ $(i = 0, 1)$ we say that $\boldsymbol{B}_0$ is (*elementarily*) *reducible* to $\boldsymbol{B}_1$ and write $\boldsymbol{B}_0 \leqq \boldsymbol{B}_1$ if an elementary operator $F\colon \mathrm{Fin}(\mathscr{S}_0) \to \mathrm{Fin}(\mathscr{S}_1)$ exists with $\mathfrak{A} \in \boldsymbol{B}_0$ iff $F(\mathfrak{A}) \in \boldsymbol{B}_1$. Clearly if $\boldsymbol{B}_1$ is a generalized spectrum and $\boldsymbol{B}_0 \leqq \boldsymbol{B}_1$ then so is $\boldsymbol{B}_0$ and, if $\boldsymbol{B}_1 \in \mathsf{J}_k(\mathscr{S}_1)$ then $\boldsymbol{B}_0 \in \mathsf{J}_{kk_1}(\mathscr{S}_0)$ where k_1 is the k in the definition of an elementary operator. Let us have two pairs of similarity types $(\mathscr{S}_i, \mathscr{T}_i)$ $(i = 0, 1)$, and two pairs of formulas (φ_i, ψ_i) determining the search problems $\boldsymbol{M}_0, \boldsymbol{M}_1$. Denote $\boldsymbol{A}_i = \mathrm{Mod}_\omega(\varphi_i)$, $\boldsymbol{B}_i = \mathrm{Mod}_\omega(\varphi_i \mathbin{\&} \exists \mathscr{T}_i \psi_i)$. We say that M_0 is *elementarily reducible* to M_1 (we write $M_0 \leqq M_1$) if we have the following:

1. An elementary operator $F\colon \boldsymbol{A}_0 \to \boldsymbol{A}_1$.
2. An elementary operator $G\colon \boldsymbol{A}_0 \times \mathrm{Mod}_\omega(\varphi_1 \mathbin{\&} \psi_1) \to \mathrm{Mod}_\omega(\varphi_0 \mathbin{\&} \psi_0)$.
3. For every $\mathfrak{A} \in \boldsymbol{A}_0$, $\mathfrak{A} \in \boldsymbol{B}_0$ iff $F(\mathfrak{A}) \in \boldsymbol{B}_1$.
4. If $F(\mathfrak{A}) \in \boldsymbol{B}_1$, then the corresponding extension of $\mathfrak{A}$ can be found effectively by G: For a structure $\mathfrak{C}$ over $\mathscr{S} \cup \mathscr{T}$ denote by $\mathfrak{C} \restriction \mathscr{S}$ its restriction to $\mathscr{S}$. We require that if $\mathfrak{B} \restriction \mathscr{S}_1 = F(\mathfrak{A})$, $\mathfrak{B} \in \mathrm{Mod}_\omega(\varphi_1 \mathbin{\&} \psi_1)$ then $G(\mathfrak{A}, \mathfrak{B}) \restriction \mathscr{S}_0 = \mathfrak{A}$.

Since composition preserves elementarity $\leqq$ is a partial order.

Now we describe a variant of the satisfiability problem in propositional logic as a search problem and then show that it is universal in some sense.

Definition 5. The *satisfiability problem* $\mathrm{SAT} = (\varrho, \sigma)$ is defined as follows. The similarity type $\mathscr{S}_0$ consists of one binary relation symbol $\prec$ and three unary predicate symbols P, U_0, U_1. Intuitively, $\prec$ defines a logical network, $P(x)$ tells whether at the point x of the network stands a sign $\neg$ or $\vee$, U_0 tells in which nodes is the truth value prescribed and U_1 gives the prescribed values in these nodes. The similarity type $\mathscr{T}_0$ consists of a unary predicate symbol T which, intuitively, orders truth values to the nodes. ϱ says that $\prec$ is a strict partial order. We say that y is an immediate predecessor of x if $y \prec x$ and there is no z with $y \prec z \prec x$. σ says that $U_0(x) \to (U_1(x) \leftrightarrow T(x))$ and tells how to compute the truth values $T(x)$: if $P(x)$ holds and x has an immediate predecessor y then $T(x) = \neg T(y)$. If $\neg P(x)$ and x has some immediate predecessors then $T(x)$ iff for at least one of them, y, $T(y)$ holds.

Theorem 1. *Let us have a search problem* $M = (\varphi, \psi)$ *with* $\boldsymbol{A} = \mathrm{Mod}_\omega(\varphi)$ *where* $\varphi \in \mathrm{L}(\mathscr{S})$, $\psi \in \mathrm{L}(\mathscr{S} \cup \mathscr{T})$. *Suppose that a formula* $\delta(x) \in \mathrm{L}(\mathscr{S})$ *can be given such that for all* $\mathfrak{A} \in \boldsymbol{A}$ *there are two elements* x, y *of* $\mathfrak{A}$ *with* $\delta(x)$ *true and* $\delta(y)$ *false. Then* $M \leqq \mathrm{SAT}$.

Note. The assumption of the theorem requires some minimal expressing power of the structures of $\boldsymbol{A}$.

Proof. For simplicity we suppose that in our formulas all terms are variables, i.e. we have minimal subformulas only of the forms $Q(x_1, \ldots, x_k)$ and $f(x_1, \ldots, x_k) = y$. This can always be achieved. We confine ourselves to formulas using only $\neg$, $\vee$ and $\exists$.

Definition 6. Let us fix the similarity type $\mathscr{L} = \{\subseteqq, \trianglelefteq, \sim, V, P_1, P_2, f, g\}$ where $\subseteqq$, $\trianglelefteq$, $\sim$ are binary, V, P_1, P_2 are unary predicate symbols, f and g are unary function symbols. Having a similarity type $\mathscr{S} = \{Q_1, \ldots, Q_m\}$ and the second order formula $\varphi \in \mathrm{L}_2(\mathscr{S})$ we define the structure $\mathfrak{F}(\varphi) \in \mathrm{Fin}(\mathscr{L})$ (it is intended as an encoding, a "numbering" of φ) as follows. An *occurrence* of a variable x in φ is every written occurrence of it not counted when it occurs after a quantor.

$$|\mathfrak{F}| = \{\psi \in \mathrm{L}(\mathscr{S}) \mid \psi \text{ is a subformula of } \varphi\} \cup \cup \{t \mid t \text{ is the occurrence of a variable in } \varphi\}.$$

$V^{\mathfrak{F}}(t) \leftrightarrow t$ is the occurrence of a variable in φ.

On a subformula t, P_1 and P_2 tell if it is a negation or a disjunction. (If both are false, then t is a quantified subformula.) $\subseteqq$ is the ordering of subformulas according to inclusion.

On the minimal subformulas, $\trianglelefteq$ establishes an ordering according to the order of the corresponding predicate or function symbols in $\mathscr{S}$ (e.g. $Q_i(x_1, \ldots, x_k) \trianglelefteq Q_j(y_1, \ldots, y_m)$ iff $i \leqq j$). On occurrences of variables, $\trianglelefteq$ establishes the order of their first occurrence. For occurrences of variables $t_1 \sim t_2$ holds iff t_1 and t_2 occur on the same place of two occurrences of the same predicate symbol. f orders to every occurrence of a variable the minimal subformula in which it occurs. g orders to every variable the subformula where it is being quantified and to every minimal subformula the subformula where the corresponding relation symbol is quantified. These predicates and functions can be defined in an arbitrary way on any place where they are not given by the above definition.

Clearly, $\mathfrak{F}(\varphi)$ is a one-to-one coding of formulas by structures. It is easy to give the elementary class $\boldsymbol{L} \subseteqq \mathrm{Fin}(\mathscr{L})$ of codes of formulas.

Having this definition and the formula $\delta(x)$ we can easily give an elementary operator $D\colon \boldsymbol{A} \to \{\mathfrak{F}(\exists \mathscr{T} \psi)\}$ i.e. to define $\mathfrak{F}$ from any element of $\boldsymbol{A}$. We now define $F(\mathfrak{A}() = \mathfrak{B} \in \mathrm{Mod}_\omega(\varrho)$ as follows. Denote by k the number of different variables in ψ. $|\mathfrak{B}| = \{(c, a_1, \ldots, a_k) \mid c \in |\mathfrak{F}|,\ c \text{ is a subformula of } \psi,\ a_i \in |\mathfrak{A}|\ (i = 1, \ldots, k)\}$.

Definition 7. For a structure $\mathfrak{A}$ of similarity type $\mathscr{S}$ an $\mathfrak{A}$-*formula* is an arbitrary first order formula $\varphi \in \mathrm{L}(\mathscr{S})$ in which every free variable is replaced by an element of $|\mathfrak{A}|$.

The elements of $|\mathfrak{B}|$ define in a natural way $\mathfrak{A}$-subformulas of ψ. $(c, a_1, \ldots, a_k)$ gives us namely a subformula c of ψ. Substituting a_i if the i-th variable of ψ is free in c gives us the desired $\mathfrak{A}$-subformula. Let us have two elements of $|\mathfrak{B}|$, $b_i = (c_i, \ldots)$, $i = 0, 1$. If c_i are not both minimal subformulas of ψ, then $b_0 =^{\mathfrak{B}} b_1$ iff $c_0 = c_1$ and for all free variables of c_i, the corresponding substituted elements of $|\mathfrak{A}|$ are equal. If c_i are minimal, then $b_0 =^{\mathfrak{B}} b_1$ iff a) $c_0 \trianglelefteq c_1$ and $c_1 \trianglelefteq c_0$ and b) for every pair of occurrences of variables t_0, t_1 such that $f(t_i) = c_i$ and $t_0 \sim t_1$ the corresponding substituted elements of $|\mathfrak{A}|$ are equal. This defines $E^{\mathfrak{B}}$. $b_0 \prec b_1$ holds iff for the corresponding $\mathfrak{A}$-formulas the first is equal in $\mathfrak{B}$ to a subformula of the second. $P(b)$ is true if the corresponding subformula is a negation. $U_0(c, \ldots)$ holds iff 1. c is the whole formula ψ, or 2. c is a minimal subformula and the corresponding relation symbol is not quantified in $\exists\mathcal{T}\psi$. U_1 holds on $(c, \ldots)$ in case 1. of the above definition and in case that U_1 has the same truth value as the corresponding relation of $\mathfrak{A}$ on the corresponding elements of $|\mathfrak{A}|$.

This completes the definition of F. The definition of G as well as the verification that F and G satisfy the conditions of a reduction is trivial.

Definition 8. The search problem $M = (\varphi, \psi)$, is called a *search problem with order* if a formula $\lambda(x, y) \in \mathrm{L}(\mathcal{S})$ can be given such that φ implies that 1. the structure has at least two different elements, 2. $\lambda(x, y)$ defines a linear order on the structure.

Consider the combinatorial problems which were proved to be complete in [6]. Exclude of them the last four, i.e. KNAPSACK, PARTITION, SEQUENCING and MAX CUT. The remaining ones are easily reformulable as search problems.

Theorem 2. *Let M_0 be one of the aforementioned search problems. Every search problem with order is elementarily reducible to M_0.*

Sketch of the proof. By Theorem 1, every search problem with order reduces to SAT. The order can be transferred naturally, thus transforming SAT into a search problem with order. SAT with an order reduces easily to SATISFIABILITY of KARP, with order. Further every reduction which can be found in [6] (not counting the last four) can be easily transformed into an elementary reduction using an additional order relation. On the other hand, the order always generates an order on the defined structure.

2.

We consider the classes

$$\boldsymbol{E}(\mathcal{S}) = \{\langle \mathfrak{F}(\varphi), \mathfrak{A}\rangle \mid \varphi \in \mathrm{L}_2(\mathcal{S}), \mathfrak{A} \in \mathrm{Fin}(\mathcal{S})\},$$

$$\boldsymbol{E}_k(\mathcal{S}) = \{\langle \mathfrak{F}(\varphi), \mathfrak{A}\rangle \in \boldsymbol{E}(\mathcal{S}) \mid \varphi \text{ is an existential formula, the number of different variables in } \varphi \text{ is not more than } k\}.$$

We remind that the similarity type of $\mathfrak{F}(\varphi)$ was denoted by $\mathcal{L}$. Denote by $\mathcal{L}_d$ the similarity type $\mathcal{L} \cup \{R^{d+1}, U\}$ where R^{d+1} is a relation symbol of degree $d + 1$, U is a unary relation symbol. For any similarity type $\mathcal{S}$ with maximum degree of relations $\leqq d$, $\varphi \in \mathrm{L}_2(\mathcal{S})$, $\mathfrak{A} \in \mathrm{Fin}(\mathcal{S})$ we define the structure $\mathfrak{B} = [\mathfrak{F}(\varphi), \mathfrak{A}] \in \mathrm{Fin}(\mathcal{L}_d)$ as follows. $|\mathfrak{B}|$, the interpretation of U and the relations of $\mathfrak{F}(\varphi)$ are the same as for $\langle \mathfrak{F}(\varphi), \mathfrak{A}\rangle$. $R^{d+1}(t, t_1, \ldots, t_d)$ is true if and only if t is an element of $|\mathfrak{F}(\varphi)|$, corresponding to a minimal subformula $Q(x_1, \ldots, x_k)$ with Q free in φ, $t_i \in |\mathfrak{A}|$ and $Q(t_1, \ldots, t_k)$ holds on $\mathfrak{A}$. We define the classes

$$\boldsymbol{G}(d) = \{[\mathfrak{F}(\varphi), \mathfrak{A}] \mid \text{for some } \mathcal{S} \text{ with maximum degree} \leqq d,\ \varphi \in \mathrm{L}_2(\mathcal{S}), \mathfrak{A} \in \mathrm{Fin}(\mathcal{S})\}.$$

$$\boldsymbol{G}(d, h) = \{[\mathfrak{F}(\varphi), \mathfrak{A}] \in \boldsymbol{G}(d) \mid \varphi \text{ is in the form (1) with } d(\varphi) \leqq d,\ h(\varphi) \leqq h\}.$$

All these classes are easily seen to be elementary.

Theorem 3. a) *The class*

$$\boldsymbol{F}_k(\mathcal{S}) = \{\langle \mathfrak{F}(\varphi), \mathfrak{A}\rangle \in \boldsymbol{E}_k(\mathcal{S}) \mid \varphi \textit{ is true on } \mathfrak{A}\}$$

is a generalized spectrum, moreover, $\boldsymbol{F}_k \in \mathrm{J}_k$ (J_k was defined in the Introduction).

b) *Let* $d \geqq 1$. *Suppose* $\mathrm{P} = \mathrm{NP}$ *and, in particular, that* $\mathrm{Mod}_\omega(\neg\exists\mathscr{T}_0\sigma) \in \mathfrak{J}_s$ *where* $\mathrm{SAT} = (\varrho, \sigma)$. *Then the class*

$$\boldsymbol{H}(d, h) = \{[\mathfrak{F}(\varphi), \mathfrak{A}] \in \boldsymbol{G}(d, h) \mid \varphi \text{ is true on } \mathfrak{A}\}$$

is a generalized spectrum, and, moreover, $\boldsymbol{H}(d, h) \in \mathfrak{J}_l$ *where* $l = (d + 1)\, s^h$.

Proof a) The reduction to SAT given in Theorem 1 is uniform if we can use $\mathfrak{F}(\varphi)$ and the number of variables is bounded by k. Hence the statement easily follows.

b) By induction on h. Let $h = 0$, i.e. φ of the form $(\exists x_1, \ldots, x_k)\, \psi$ where ψ is quantifier free. In this case, on $[\mathfrak{F}(\varphi), \mathfrak{A}]$ we can define the truth of φ in the following way:

"There is a function $u(t)$ from the occurrences of variables to elements of $|\mathfrak{A}|$ (having the obvious properties of an assignment of values to $x_1, \ldots, x_k$), and an assignment of truth values to subformulas of ψ which is in accordance with the assignment u and makes ψ true."

This sentence can clearly by written in an existential second order formula of degree two. Suppose that the statement is true for h, we prove it for $h + 1$. If for any formula φ $\mathrm{Mod}_\omega(\varphi) \in \mathfrak{J}_n$ then by the assumption about SAT and the reduction to it, $\mathrm{Mod}_\omega(\neg\varphi) \in \mathfrak{J}_{sn}$. This is true especially of the formula φ with $\mathrm{Mod}_\omega(\varphi) = \boldsymbol{H}(d, h)$.

Now, to define $\boldsymbol{H}(d, h + 1)$ only one more quantor of the form $\exists T^{d+1}$ is needed.

Let us now turn to the actual aim of this preliminary theorem, the diagonalization result.

Theorem 4. a) *The complement of* $\boldsymbol{F}_{2k}(\mathscr{L})$ *is not the generalized spectrum of any formula with number of variables at most* k.

b) *The complement of* $\boldsymbol{H}(d, h)$ *is not* $\mathrm{Mod}_\omega(\varphi)$ *for any second order formula* φ *with* $d(\varphi) \leqq d$, $h(\varphi) \leqq h$.

Proof. Consider the class of structures

$$\boldsymbol{A}_k = \{\mathfrak{F}(\varphi) \mid \varphi \in \mathrm{L}_2(\mathscr{L}) \text{ is an existential formula with number of variables not more than } k \text{ and } \varphi \text{ is false on } \mathfrak{F}(\varphi)\}.$$

$\boldsymbol{A}_k$ is easily seen to be reducible to the complement of $\boldsymbol{F}_{2k}$ by an elementary operator F such that all formulas needed in the definition of F are quantifier free and the definition is in $\mathfrak{F}(\varphi) \times \mathfrak{F}(\varphi)$. Suppose that the complement of $\boldsymbol{F}_{2k}$ is $\mathrm{Mod}_\omega(\exists\mathscr{T}\varphi_0)$ for some φ_0 having number of variables not more than k. Then $\boldsymbol{A}_k = \mathrm{Mod}_\omega(\exists\mathscr{T}\varphi_1)$ for some $\varphi_1 \in \mathrm{L}(\mathscr{L})$ having number of variables not more than $2k$. In this case the question $\mathfrak{F}(\exists\mathscr{T}\varphi_1) \in \boldsymbol{A}_k$? leads us to a contradicion. The proof of b) is analogous.

Corollary. *Suppose* $\mathrm{NP} = \mathrm{CNP}$. *Then*

a) *for all* k *there exists a generalized spectrum which is not* $\mathrm{Mod}_\omega(\exists\mathscr{T}\varphi)$ *for any first-order formula* φ *with number of variables less than* k.

b) *For all* k *there exists a generalized spectrum which is not* $\mathrm{Mod}_\omega(\varphi)$ *for any second-order formula* φ *with* $d(\varphi) \leqq k$, $h(\varphi) \leqq k$.

Note. Corollary a) is true without the condition $\mathrm{NP} = \mathrm{CNP}$: it follows from Cook [9] as in Pudlák [10].

Aknowledgement. We are thankful to D. S. Johnson for many useful discussions during the preparation of this paper, R. Fagin for his valuable remarks and A. Slisenko and B. A. Trachtenbrot for their attention and making us aware of Fagin's work.

Literature

[1] FAGIN, R., Generalized first order spectra and polynomial time recognizable sets. In: Complexity of Computation (ed. R. KARP), SIAM-AMS Proceedings **7** (1974).
[2] FAGIN, R., Monadic generalized spectra. This Zeitschr. **21** (1975), 88–96.
[3] FAGIN, R., A two-cardinal characterization of double spectra. This Zeitschr. **21** (1975), 121–122.
[4] FAGIN, R., A spectrum hierarchy. This Zeitschr. **21** (1975), 123–124.
[5] COOK, S., The complexity of theorem-proving precedures. Conf. Record of 3rd ACM Symp. on Theory of Computing (1970), 151–158.
[6] KARP, R., Reducibility among combinatorial problems. In: Complexity of Computer Computations (R. E. MÜLLER, J. W. THATCHER, ed.) Plenum Press, New York 1972, pp. 85–104.
[7] LEVIN, L. A., Universal search problems. Problemy Peredači Informatsii **9** (1973), 115–116.
[8] HÁJEK, P., Projective classes of models in observational predicate calculi. In: Lecture Notes on Computer Science **32**, Springer-Verlag 1975, Proceedings of the 4th MFCS conference, Marianské Lázné 1975.
[9] COOK, S., A hierarchy for non-deterministic time complexity. J. C. S. S. **7** (1973), 343–353.
[10] PUDLÁK, P., The observational predicate calculus and complexity of computations (preliminary communication).
[11] SCHNORR, C. P., Satisfiability is quasi-linear complete. Preprint.
[12] JONES, N. D., Space-bounded reducibility among combinatorial problems. J. C. S. S. **11** (1975), 68–85.

(Eingegangen am 4. Mai 1976)