FORMULAS OVER MINIMAL STRUCTURES

Yerzhan Baissalov[1] and Jamalbek Tussupov[2]

Astana IT University[1] *and L. N. Gumilyov Eurasian National University*[2]

E-mail: yerzhan.baissalov@astanait.edu.kz and j.tussupov@gmail.com

Introduction. In this article, we present several syntactic results about formulas over minimal structures. These syntactic studies were inspired by the remarkable article [1], where, among others, the following Dichotomy Theorem on minimal structures with a definable generic type was proved: in such structures, either we can define a partial order with infinite chains, or every generic Morley sequence is an indiscernible set.

We follow the notation of [1]: the starting point is the minimal structure M (its universe is denoted by the same letter M, we hope this will not lead to any confusion). Let L be the language of M. We enrich this language with constants for the elements of M and denote this enriched language by L_M. In what follows, by a formula we mean a formula of L_M. Let $\varphi(M^n)$ denote the set of elements that satisfy the formula $\varphi(\overline{x})$ in the model M, where n is the length of the tuple $\overline{x}$; in this case, we say that the formula $\varphi(\overline{x})$ *defines* the set $\varphi(M^n)$. Two formulas are *equivalent* if they define the same set; this equivalence is denoted by $\sim$. The minimality of M is understood in the Eastern sense: every definable subset of M is finite or co-finite; in this case one may also say that the structure is *definably minimal*. The unique non-algebraic 1-type over M is called generic and is denoted by p.

As usual, we live in some homogeneous and saturated world, which is denoted in [1] by $\overline{M}$. But we will not denote it in any way, since in the initial sections we use only life in M, and in the last section this world will be clear from the context.

Definition 1. Let $\varphi(x, \overline{y})$ be a satisfiable formula with free variables x and $\overline{y} = (y_1, y_2, \dots, y_n)$. We say the variable x is

- *normal* if the set $\exists\overline{y}\varphi(M, \overline{y})$ is infinite
- *tame* if the set $\varphi\left(M, \overline{b}\right)$ is finite for any $\overline{b} \in M$
- *wild* if the set $\varphi\left(M, \overline{b}\right)$ is infinite or empty for any $\overline{b} \in M$

Definition 2. A formula is called *normal* (*tame*, *wild*) if every free variable in it is normal (tame, wild). It is called *faceted* if every free variable in it is either tame or wild.

Remark. A formula $\varphi(x)$ with a single free variable is abnormal and tame if $\varphi(M)$ is finite, otherwise it is normal and wild.

The definability of the generic type allows us to use the quantifiers $\exists^f$ and $\exists^c$.

Definition 3. The formula $\exists^f x\varphi(x,\overline{y})$ [$\exists^c x\varphi(x,\overline{y})$] defines all $\overline{y}$ such that the set $\varphi(M,\overline{y})$ is finite [infinite = co-finite].

Remark. In [1] the formula $\exists^c x\varphi(x,\overline{y})$ is denoted by $d_p x\varphi(x,\overline{y})$.

1. The definability of generic type and faceted formulae. In this section, we will prove that in a minimal structure with a definable generic type, every formula is equivalent to a disjunction of faceted formulas. Let M be such a structure.

Theorem 1. Every formula over M is equivalent to a disjunction of faceted formulas.

Proof is by induction on the number of free variables of the formula and is a direct consequence of the following lemma, since non-normal formulas are equivalent over M to a disjunction of formulas of the form $\theta \wedge x = a$, where $a \in M$ and θ does not contain a free variable x (therefore it contains a fewer number of free variables).

Lemma 1. Every formula over M is equivalent to a disjunction of faceted and non-normal formulas.

Proof of the lemma is carried out by induction on the number n of free variables of the formula. For $n = 1$ everything is clear. Let us consider in detail the case $n = 2$.

Let $\varphi(x,y)$ be a formula over M with two free variables x and y. If the formula φ is not normal, then everything follows from the case $n = 1$, so suppose that φ is normal.

Obviously, we have the equivalence of formulas

$$\exists x\varphi(x,y) \sim \exists^f x\varphi(x,y) \vee \exists^c x\varphi(x,y),$$

where two disjunctive terms of a right-hand formula define disjoint subsets, whose union is co-finite in M. Hence, then exactly one of the formulas $\exists^f x\varphi(x,y)$ and $\exists^c x\varphi(x,y)$ is finite and the other is cofinite. So, for some $\alpha \in \{f,c\}$, the formula $\exists^c y\exists^\alpha x$ is satisfied in the model M. Similarly, for some $\beta \in \{f,c\}$, the model M satisfies $\exists^c x\exists^\beta y\,\varphi$.

Now we use the fact that the formula φ is equivalent to the disjunction of the following three formulas:

$$\varphi \sim \left(\varphi \wedge \exists^\alpha x\,\varphi \wedge \exists^\beta y\,\varphi\right) \vee (\varphi \wedge \exists^{\overline{\alpha}} x\,\varphi) \vee \left(\varphi \wedge \exists^{\overline{\beta}} y\,\varphi\right),$$

where $\bar{\gamma}$ denotes the "negation" of $\gamma \in \{f,c\}$, i.e., $\bar{\gamma} \in \{f,c\}\backslash\{\gamma\}$. Since the formulas $\exists^{\overline{\alpha}} x\,\varphi$, $\exists^{\overline{\beta}} y\,\varphi$ are finite, the last two disjunctive terms are non-normal, and the first term $\psi(x,y) \coloneqq \varphi \wedge \exists^\alpha x\,\varphi \wedge \exists^\beta y\,\varphi$, as it is easy to check, is a faceted formula: indeed, if, for example, for $a \in M$ the formula $\psi(a,y)$ is satisfiable, then it is always a β-formula, as the intersection of the β-formula $\varphi(a, y)$ with the cofinite formula $\exists^{\alpha} x\varphi$.

Thus, the case $n = 2$ is completely dismantled.

Step of induction in the proof of the lemma. Assume that the lemma has been proved for $n \leq k$ and consider the formula $\varphi(x, \bar{x})$ with $l(\bar{x}) = k$.

Let us divide φ into two parts: $\varphi \wedge \exists^{f} x\, \varphi$ and $\varphi \wedge \exists^{c} x\, \varphi$. Each of these parts is a faceted formula in x. Therefore, we can assume that the formula $\varphi(x, \bar{x})$ is faceted in x.

Consider the formula $\exists x\varphi$. By the induction hypothesis, it is equivalent to the disjunction of faceted and non-normal formulas:

$$\exists x\varphi \sim \bigvee_{i=1}^{s} \delta_i \vee \bigvee_{j=1}^{t} \varepsilon_j \,,$$

where the formulas δ_i are faceted and ε_j are not normal. Then it is easy to see that

$$\varphi \sim \bigvee_{i=1}^{s} (\varphi \wedge \delta_i) \vee \bigvee_{j=1}^{t} (\varphi \wedge \varepsilon_j)$$

and each $\varphi \wedge \delta_i$ is faceted, each $\varphi \wedge \varepsilon_j$ is not normal. The lemma and thus the theorem are proved.

Example 1. The generic type of the minimal structure $(\omega + \omega^*, <)$ is not definable and the formula $x < y$ cannot be represented as a disjunction of faceted formulas.

2. Unordered minimal structures with definable generic types and tame formulae. This section devoted to the following theorem.

Theorem 2. Let M be a minimal structure with the definable generic type. If there does not exist a definable partial order with infinite chains in M, then each formula over M is equivalent to a Boolean combination of tame formulas.

Proof. Let $\varphi(\overline{z})$ be a satisfiable formula with n free variables $\overline{z} = (z_1, \dots, z_n)$. We use induction on n. The theorem is obviously true for $n = 1$, so it suffices to prove the induction step. Assume that the theorem is true for $n < k$ and let's prove it for $n = k$.

By Theorem 1 we may assume that φ is a faceted formula.

If φ is tame, then the job is done, and if it is wild, then the following lemma reduces the proof to the induction hypothesis.

Lemma 2. The negation of a wild formula with k free variables is equivalent to the disjunction of a tame formula and formulas with less than k free variables.

Proof of the lemma. Let $\varphi(\overline{z})$ be a wild formula with k free variables $\overline{z} = (z_1, \dots, z_k)$. Then

$$\neg\varphi \sim \bigvee_{i=1}^{k} (\neg\exists z_i \varphi \wedge z_i = z_i) \vee \left(\neg\varphi \wedge \bigwedge_{i=1}^{k} \exists z_i \varphi\right),$$

and in the formula on the right side, the last disjunctive term is tame, and the remaining disjunctive terms are Boolean combinations of formulas with less than k free variables (we leave the verification of these facts to the reader).

Thus, it remains to study the case when the formula φ is of mixed form, i.e., it has both tame and wild variables. Let's also assume that the formula is normal, otherwise it would be equivalent to a Boolean combination of formulas with fewer free variables.

Lemma 3. Let $\varphi(x, \overline{x}, \overline{y})$ be a normal mixed formula with wild variables x, $\overline{x} = (x_1, \dots, x_l)$ and tame variables $\overline{y} = (y_1, \dots, y_m)$. Then $m > 1$.
Proof of the lemma. Assume the contrary $m = 1$ and $\overline{y} = y$. We'll show by induction on l that then an order with infinite chains is definable.
Base of Induction. Let $l = 0$ and $\overline{x} = \emptyset$. We construct by induction ω-sequence $\{c_i\} \subset M$. Let c_0 be an element with the smallest $\varphi(c_0, M)$ with respect to $\subseteq$. If c_i is already defined and $\varphi(c_i, M) = \{a_1, \dots, a_s\}$, then we choose any $a_{s+1} \in \exists x \varphi(x, M) \setminus \varphi(c_i, M)$ and then any $c_{i+1} \in \cap_{j=1}^{s+1} \varphi(M, a_j)$. Obviously, $\{c_i\}$ will be an ω-chain with respect to the partial order $u \leq v$ defined by $\forall y[\varphi(u, y) \to \varphi(v, y)]$.
Step of Induction. If for some $a \in \exists \overline{x} y \varphi(M, \overline{x}, y)$ the formula $\varphi(a, \overline{x}, y)$ is normal, then the induction hypothesis is already applicable to this formula. Otherwise, we have the following: the set $\exists \overline{x} \varphi(a, \overline{x}, M)$ is finite for each $a \in M$, so in this case the normal formula $\exists \overline{x} \varphi(x, \overline{x}, y)$ with wild x and tame y applies.

Let $\varphi(\overline{x}, \overline{y})$ be a normal mixed formula with wild variables $\overline{x} = (x_1, \dots, x_l)$ and tame variables $\overline{y} = (y_1, \dots, y_m)$, $m > 1$.
Lemma 4. The formula $\psi(\overline{y}) \coloneqq \exists \overline{x} \varphi(\overline{x}, \overline{y})$ is tame (and obviously normal).
Proof of the lemma. Let us change $\overline{y}$ by $y\overline{z}$ and see that y is tame in $\exists \overline{x} \varphi(\overline{x}, y\overline{z})$. If it wouldn't be so then for some $\overline{b} \in M$ the formula $\exists \overline{x} \varphi(\overline{x}, y\overline{b})$ would be infinite, giving the normal formula $\varphi(\overline{x}, y\overline{b})$ with wild $\overline{x}$ and tame y, which contradicts Lemma 3.

The formula $\neg\varphi \wedge \psi$ is equivalent to the disjunction of a tame formula and formulas with fewer free variables:

$$\neg\varphi \wedge \psi \sim \bigvee_{i=1}^{l} (\neg\exists x_i \varphi \wedge x_i = x_i \wedge \psi) \vee \left(\neg\varphi \wedge \psi \wedge \bigwedge_{i=1}^{l} \exists x_i \varphi \right).$$

Here, the last disjunctive term of the right-side formula is tame, and its remaining disjunctive terms are Boolean combinations of formulas with fewer free variables (we again leave the verification to the reader).

Now the equivalences

$$\varphi \sim \varphi \wedge \psi \sim \neg(\neg\varphi \wedge \psi) \wedge \psi$$

complete the induction and thus the proof of the theorem.

Example 2. The generic type of the minimal structure $(\omega, <)$ is definable and the faceted formula $x < y$ cannot be represented as a Boolean combination of tame formulas.

3. Asymmetric minimal structures and Skolem predicates. In this section, we observe some simple facts about asymmetric minimal structures that are implicitly present in [1].

Let M be an asymmetric minimal structure with definable partial order $<$ with infinite chains. As noted in [1], in a such structure there is a definable partial order with ω-chain or ω^*-chain. In fact, only one of these types of chains can be present in M. If, for example, $a_0 < a_1 < \cdots$ is an ω-chain and $b_0 > b_1 > \cdots$ is a ω^*-chain, then the sets $\{a_i\}$ and $\{b_j\}$ do not intersect (a common point would divide M into two infinite parts). The formulas $a_i < x < b_j$, $i, j \in \omega$, are all generic, so the generic $(x < y)$-type (here y is a parameter) is not definable, a contradiction.

Up to anti-isomorphism, ω and ω^* are the same, so w.l.o.g. we can suppose that M contains a ω-chain for some definable partial order.

The generic type contains a finite set of formulas of the form $\neg(c < x)$, $c \in M$. Therefore, we can definably modify the relation $<$ so that p contains all the formulas $c < x$, $c \in M$. This modified relation will be a directed, well partial order with no chains of type $\omega + 1$.

Observation 1. Up to isomorphism, $M = acl(\emptyset)$ is the only minimal model of its theory. It is also minimal in the Western sense: it does not contain a proper elementary submodel.

Question 1. Is it possible to prove that Th(M) has, up to isomorphism, an infinite number of countable models?

For each formula $\varphi(x,\overline{y})$, one can define a formula $\sigma_\varphi(x,\overline{y})$ such that the following propositions hold in the model M:

$$\forall x\overline{y}(\sigma_\varphi(x,\overline{y}) \to \varphi(x,\overline{y}))$$

$$\forall\overline{y}(\exists x\varphi(x,\overline{y}) \to \exists^f x\sigma_\varphi(x,\overline{y}))$$

We will call the formula σ_φ the *Skolem predicate* for φ. Semantically, $\sigma_\varphi(x,\overline{y})$ says: "for a given $\overline{y}$, x is a minimal element with respect to the order $<$ that satisfies the formula $(x,\overline{y})$". In [1] it was proved that the number of minimal (relative to $<$) elements of any subset of the structure M is finite.

In [1], a closure operation cl was defined for minimal structures with the definable generic type. Skolem's predicates allow us to notice the following.

Observation 2. Let A be a subset in some model N of the theory Th(M). Then $cl(A)$ is an elementary submodel of N.

Question 2. When is the closure $cl(A)$ of a subset A of some model of the theory Th(M) an elementary submodel if M is a symmetric minimal model in the sense of [1]?

Acknowledgments. We would like to express our deep gratitude to the group of logicians of University Lyon-1 for organizing and inviting to the 7th Franco-Kazakh Colloquium on Model Theory.

REFERENCES

1. K. Krupiński, P. Tanović and F. O. Wagner, Around Podewski's conjecture, arXiv:1201.5709v2.