Definable sets in tree-like structures

Géraud Sénizergues

LaBRI, Bordeaux University,

Thursday June 29th 2017

Dedicated to Paul Schupp on his 80th birthday.

INTRODUCTION

Three fundamental works

Three fundamental works by D.E. Muller and P.E.Schupp :

[Groups, the theory of ends, and context-free languages. Journal of Computer and System Sciences, 1983]

[The theory of ends, pushdown automata, and second-order logic. Theoretical Computer Science, 1985]

[Alternating automata on infinite trees. Theoretical Computer Science, 1987]

Decidability/Definability/Selection

Decidability : Given a formula Φ, can we decide whether

$$\mathcal{M} \models \Phi$$

Definability :
A subset R is definable iff there exists a formula $\Phi(X)$ such that

$$\mathcal{M} \models \exists! X \Phi(X) \text{ and } \mathcal{M}, R \models \Phi(X).$$

Selection :
Given a formula $\Phi(X)$, a selector (for the formula and the structure $\mathcal{M}$) is a formula $\hat{\Phi}(X)$ such that

$$\begin{aligned} \mathcal{M} &\models (\exists X \cdot \Phi(X)) \to (\exists X \cdot \hat{\Phi}(X)) \\ \mathcal{M} &\models \forall X \cdot (\hat{\Phi}(X) \to \Phi(X)) \\ \mathcal{M} &\models \forall X \cdot \forall Y \cdot (\hat{\Phi}(X) \wedge \hat{\Phi}(Y)) \to (X = Y). \end{aligned}$$

contents

MSO logics

MSO :syntax

Let *Sig* $= \{r_1, \dots, r_n\}$ be a signature containing relational symbols only, where $\rho_i \in \mathbb{N}$ is the arity of symbol r_i.
Let *Var* $= \{x, y, z, \dots, X, Y, Z \dots\}$ be a set of variables, where $x, y, \dots$ denote first order variables and $X, Y, \dots$ second order variables.
The set of MSO-formulas over Sig, Var is the smallest set such that :
for every $x, x_1, \dots, x_\rho, X, Y, X_1 \dots X_\tau$ in Var and MSO formula Φ, Ψ

$$x \in X, \quad Y \subseteq X$$

$$\neg\Phi, \quad \Phi \vee \Psi, \quad \Phi \wedge \Psi, \quad \Phi \rightarrow \Psi, \quad \exists x.\Phi, \quad \exists X.\Phi, \quad \forall x.\Phi, \quad \forall X.\Phi,$$

are MSO-formulas.

MSO :semantics

Let $\mathcal{M} = \langle D_{\mathcal{M}}, r_1, \ldots, r_n \rangle$ be a structure over the signature *Sig*, and $\nu : Var \rightarrow D_{\mathcal{M}} \cup \mathcal{P}(D_{\mathcal{M}})$ a valuation

The validity of a MSO-formula in the structure $\mathcal{M}$ with valuation ν is then defined by induction on the structure of the formula.
Notation :

$$\mathcal{M}, \nu \models \Phi.$$

Tools :automata

Automata :
-*General* automata :
Right-action of the monoid A^* over a set C (set of configurations) :

$$(c, u) \mapsto c \odot u.$$

Initial configuration c_0 and set of final configurations C_f.
Right-equivalence over C :

$$c \equiv_r d \quad \Leftrightarrow \quad \{u \in A^* \mid c \odot u \in C_f\} = \{u \in A^* \mid d \odot u \in C_f\}$$

- *Pushdown* automata : case where $C = Q \cdot Z^*$ and $\odot$ consists of "small changes" on the right-end of the configuration.

Tools : alternating automata

Definition from [Muller-Schupp 87]
An *alternating* automata over binary trees, labelled on alphabet Σ is a tuple :

$$\langle Q, \Sigma, q_0, \delta, \Omega \rangle$$

where

$$\delta : Q \times \Sigma \to \mathcal{B}^+(Q \times \{\ell, r\}).$$

The tree $t : \{\ell, r\}^* \to \Sigma$ is recognized by the automaton iff player J0 is winning the following game :

Tools : alternating automata

$$V_0 \quad := \quad Q \times \{\ell, r\}^*$$
$$V_1 \quad := \quad \text{conjunctive monomials over } Q \times \{\ell, r\}$$

Edges :
- J0 (the "prover") chooses a conjunctive monomial that is allowed by t and δ
- J1 (the "attacker") chooses one atom $(q, d) \in Q \times \{\ell, r\}$ of that monomial.

Tools : alternating automata

The automaton is non-deterministic when every position of J1 accessible in the game has the form

$$(p, \ell) \wedge (q, r).$$

Theorem (Muller-Schupp 1987)

Every alternating finite tree automaton can be simulated by some non-deterministic finite tree-automaton.

Key-idea : use Muller deterministic automata over branches.
Extension to trees with infinite arity : [Walukiewicz 1996].

Tools :games

Games :

- Parity games :

Arena : a bipartite graph $(V_0 \cup V_1, E, \Omega)$ where $\Omega : V_0 \cup V_1 \to [0, n]$ is the *priority* map. Play :

$$v_0, v_1, \ldots, v_m, \ldots$$

which is a path in the arena ; either it is infinite or its last vertex is a dead-end.

The winner is J0 iff

$$\max\{r \mid v_i = r \text{ i.o. }\} \equiv 0 \pmod 2$$

or the last vertex is a position of V_1 which is a dead-end. Otherwise $J1$ is the winner.

Tools :games

Strategy for Jj : a map $S_j : V^* V_j \to V_{1-j}$ such that

$$S_j \subseteq E \quad \text{and} \quad \mathrm{dom}(S_j) = \mathrm{dom}(E) \cap V_j$$

It is said positional if $S_j(u \cdot v)$ depends on v only.

Theorem (Emerson-Jutla 91)

Let G be a parity game.
1- Either player 0 or player 1 has a winning strategy.
2- The winner has a positional winning strategy.

Other games :
- Muller games
- Ehrenfeucht-Fraïssé games ([1961],[1954])

Tools :interpretations

We call MSO-*interpretation* of the structure $\mathcal{M}$ into the structure $\mathcal{M}'$ every injective map $\varphi : D_{\mathcal{M}} \to D_{\mathcal{M}'}$ such that,
1- There exists a formula $\Phi'(X) \in \mathcal{L}'$, with one free-variable X, which is second-order, fulfilling that, for every subset $X_{\mathcal{M}'} \subseteq D_{\mathcal{M}'}$

$$X_{\mathcal{M}'} = \varphi(D_{\mathcal{M}}) \Leftrightarrow \mathcal{M}' \models \Phi'(X_{\mathcal{M}'})$$

2- For every $i \in [1, n]$, there exists a formula $\Phi'_i(x_1, \ldots, x_{\rho_i})$, fulfilling that, for every valuation ν

$$(\mathcal{M}, \nu) \models r_i(x_1, \ldots, x_{\rho_i}) \Leftrightarrow (\mathcal{M}', \varphi \circ \nu) \models \Phi'_i(x_1, \ldots, x_{\rho_i}).$$

Tools :interpretations

Theorem

Suppose that there exists a MSO-interpretation *of the structure* $\mathcal{M}$ *into the structure* $\mathcal{M}'$*. Then, there exists a computable map from* $\mathcal{L}$ *to* $\mathcal{L}'$ *:* $\Phi \mapsto \Phi'$ *such that*

$$\mathcal{M} \models \Phi \textit{ iff } \mathcal{M}' \models \Phi'.$$

In particular, if $\mathcal{M}'$ *has a decidable MSO-theory, then* $\mathcal{M}$ *has a decidable MSO-theory too.*

Infinite binary tree

Theorem (Rabin 1969)

Let $\mathcal{S}$ be the signature $\langle S_a, S_b \rangle$. The MSO-theory of the structure $\langle \{a, b\}^*, S_a, S_b \rangle$ is decidable.

Theorem (Rabin 1969)

A subset $R \subset \{a, b\}^*$ is MSO-definable iff it is recognizable by a deterministic f.automaton.

Theorem (Rabin 1969)

The MSO-theory of the structure $\langle \{a, b\}^*, S_a, S_b \rangle$ has the selection property.

Decidability versus selection

Example 1 : [Rabinovitch 05], [Lifsches-Shellah 98].
- UNdecidable MSO
- Selection property

Example 2 : A structure definable inside an algebraic tree (computation-tree of some pushdown automaton) :
- Decidable MSO (MSO-interpretation into the unravelling of a c.f. graph)
- NO selection property (use Ehrenfeucht-Fraïssé games)

Context-free graphs

spherical ends

Let Γ be a graph, labelled over an alphabet X. Given some vertex $v \in V_\Gamma$, and some radius $n \in \mathbb{N}$, we call *(v, n)-end* of Γ (relative to the ball $\mathrm{B}(v, n)$) any connected component of $\Gamma - \mathrm{B}(v, n)$.

Definition

A graph Γ is said context-free iff it is connected and has only finitely many isomorphism classes of ends.

C.f. graphs : definability

The *canonical automaton* (Muller-Schupp) :
- pushdown symbols $Z_{i,j}$: j is the number of one $(n+1)$-end inside the $n - end$ with number i
- states q_ℓ : point number ℓ on the frontier of the end
- transitions : edges from the frontier of a (v, n)-end to the frontier of a $(v, n+1)$-end.

Theorem (Muller-Schupp 85)

If Γ is c.f. then it is isomorphic with the computation-graph of its canonical automaton.

Idea : the accessible configurations of the canonical automaton correspond bijectively to the vertices of Γ.

C.f. graphs : decidability

Theorem (Muller-Schupp 1985)

For every c.f. graph Γ, the MSO-theory of Γ is decidable

Idea : The structure Γ is MSO-interpretable in $Q \times \{z_{i,j}\}^*$.

C.f. graphs : definability

A subset R of V_Γ is said recognizable iff its set of "coordinates" in the canonical automaton is recognized by some f.automaton

Theorem

A subset R of V_Γ is definable iff it is recognizable

Idea : the accessible configurations of the canonical automaton correspond bijectively to the vertices of Γ.

C.f. graphs : selection

MSO Formula defining the automorphisms.
Let us choose a vertex v_0 : it defines ends relative to (v_0, n) and a canonical automaton.

$$\theta(v, V_{0,0,0}, \ldots, V_{i,j,\ell}, \ldots, D_0, \ldots, D_m, \ldots, D_{c-1})$$

Property :
- one model of the formula is :

$$\nu(v) = v_0, \nu(V_{i,j,\ell}) = \{q_\ell z_{0,j_0} \cdots z_{i,j}\},$$
$$\nu(D_m) = \{v \mid \mathrm{d}(v_0, v) \equiv m \pmod{c}\}$$

- every model has the form

$$\nu(v) = h(v_0), \quad \nu(V_{i,j,\ell}) = h(\{q_\ell z_{0,j_0} \cdots z_{i,j}\}),$$
$$\nu(D_m) = \{v \mid \mathrm{d}(v_0, v) \equiv m \pmod{c}\}$$

for some automorphism h of Γ.

C.f. graphs : selection

- The map $Aut(\Gamma) \rightarrow$ models of θ

$$h \mapsto (h(v_0), \ldots, h(\{q_\ell z_{0,j_0} \cdots z_{i,j}\}), \ldots, D_m)$$

is a bijection.

C.f. graphs : selection

Selection property for c.f. graphs.

Theorem

For every context-free graph $\mathcal{M}$, whith trivial automorphism-group, the MSO-theory of $\mathcal{M}$ has the selection-property.

Idea : given a formula $\Phi(X)$
- translate it as a formula $\Psi(Z)$ over $Q \times \{z_{i,j}\}^*$
- selector $\hat{\Psi}(Z)$ over $Q \times \{z_{i,j}\}^*$
- formula $\hat{\Psi}'(Z)$ obtained from $\hat{\Psi}(Z)$ by :

replacing $S_{z_{i,j}}(x,y)$ by

$$\bigvee_m x \in D_m \wedge y \in D_{m+1} \wedge (\bigvee_\ell y \in V_{i,j,\ell})$$

replacing $q_\ell(x)$ by $\bigvee_{i,j} x \in V_{i,j,\ell}$

C.f. graphs : selection

- we define the selector $\hat{\Phi}(X)$ by :

$$\exists v, \ldots, V_{i,j,\ell}, \ldots, D_0, \ldots, D_{c-1}$$
$$\theta(v, \ldots, V_{i,j,\ell}, \ldots, D_0, \ldots, D_{c-1}) \wedge \hat{\Psi}'(Z).$$

Extension : when the automorphism group is arbitrary, the set of models of $\hat{\Phi}(X)$ is one orbit of $Aut(\Gamma)$ (acting on $\mathcal{P}(\Gamma)$).

Stupp's expansion

S-expansion :definition

Definition (S-expansion, Stupp 1975)

Let $Sig = (r_1, \ldots)$ be a signature containing only relational symbols and $\mathcal{M} = \langle A, r_1, \ldots \rangle$ a structure over Sig.
The S-extended signature is $Sig^* = Sig \cup \{son\}$
The S-expansion is

$$\mathcal{M}^* = \langle A^*, son, {r_1}^*, \ldots \rangle$$

where A^* is the set of all finite sequences of elements of A and the relations are defined by :

$$\begin{array}{rcl} son & = & \{(u, ud) : \quad u \in A^*, \quad d \in A\} \\ r^* & = & \{(ud_1, \ldots, ud_k) : \quad u \in A^*, \quad (d_1, \ldots, d_k) \in r_{\mathcal{M}}\}, \end{array}$$

(for all $r \in Sig$, of arity k).

S-expansion :decidability

Theorem (Stupp 1975)

If the MSO theory of $\mathcal{M}$ is decidable, then the MSO theory of $\mathcal{M}^*$ is decidable.

Idea of proof :
- formula -> non-deterministic tree-automaton
(extends [Rabin 71]) :
transitions are defined by logical formulas :

$$\psi_{p,\sigma}(\ldots, X_p, \ldots, X_q, \ldots)$$

expressing that a run, with state p at vertex u, with label $t(u) = \sigma$, has a correct choice of states $\ldots, p, \ldots, q, \ldots$ for its sons
- boolean operations on languages translate into effective operations on automata
- emptiness of the recognized language is decidable.

S-expansion :definability

$$\text{Let} \quad \mathcal{M} = \langle A, r_1, \ldots \rangle.$$

Definition

A logical finite automaton over A^* is a tuple

$$\mathbb{A} = \langle Q, q_0, Q_f, \delta \rangle$$

where Q is a finite set, $q_0 \in Q$, $Q_f \subseteq Q$ and

$$\delta = (\Phi_{p,q})_{(p,q) \in Q \times Q}$$

is a family of MSO-formulas with one free individual variable.

S-expansion :definability

A computation of $\mathbb{A}$ is a sequence

$$q_0, a_1, q_1, \ldots, a_n, q_n, \ldots, a_\ell, q_\ell$$

such that, for every $n \in [0, \ell - 1]$,

$$\mathcal{M} \models \Phi_{q_n, q_{n+1}}(a_{n+1}).$$

The language recognized by $\mathbb{A}$ is the set of finite sequences $a_1 \cdots a_n \cdots a_\ell$ that admit a computation ending in $q_\ell \in Q_f$.

S-expansion :definability

Theorem (J. Pablo, Master thesis 2016)

Let $R \subseteq A$. It is S-definable iff it is recognized by some finite logical automaton over $\mathcal{M}$.

Idea of proof :
1- Suppose $R \subseteq A$ is S-definable.
- Stupp's tree automaton -> game over $\mathcal{M}^*$
general position of J0 : $(q, u) \in Q \times A^*$
general position of J1 : partition $(S_q)_{q \in Q}$ of $u \cdot A$.
- game over $\mathcal{M}^*$ -> game over $\mathcal{M}$
general position of J0 : $q \in Q$
general position of J1 : partition $(S_q)_{q \in Q}$ of A.
- winning positions of J0 are expressible within MSO.

S-expansion :definability

- logical automaton : transition $\Phi_{p,q}(x)$ expresses the fact that there exists some partition $(S_q)_{q \in Q}$ of A that can be played from p by J0, and is winning for J0 and $x \in S_q$.

2- Converse : express the language as a least fixpoint.

S-expansion :selection

Theorem (J. Pablo and G.S today)

If the MSO theory of $\mathcal{M}$ has the selection property, then the MSO theory of $\mathcal{M}^*$ has the selection property,

Ideas :
Game over $\mathcal{M}$.
Use the existence of a winning positional strategy for every parity game.

Muchnik's expansion

M-expansion : definition

Definition (M-expansion, Muchnik-Semenov 1984)

Let $Sig = (r_1, \ldots)$ be a signature containing only relational symbols and $\mathcal{M} = \langle A, r_1, \ldots \rangle$ a structure over Sig.
The M-extended signature is $Sig^* = Sig \cup \{son, clone\}$
The M-expansion is

$$\mathcal{M}^* = \langle A^*, son, clone, r_1^*, \ldots \rangle$$

where the relations are defined by :

$$\begin{array}{rcl} son & = & \{(u, dd) : \quad u \in A^*, \quad d \in A\} \\ clone & = & \{udd : \quad u \in A^*, \quad d \in A\} \\ r^* & = & \{(ud_1, \ldots, ud_k) : \quad u \in A^*, \quad (d_1, \ldots, d_k) \in r_{\mathcal{M}}\}, \end{array}$$

(for all $r \in Sig$, of arity k).

M-expansion :decidability

Theorem (Muchnik-Semenov 84, Walukiewicz 96)

For every MSO formula Φ over the signature Sig^* one can effectively find a MSO formula $\widehat{\Phi}$ over the signature Sig such that, for every structure M :

$$M \models \widehat{\Phi} \quad iff \quad M^* \models \Phi$$

Proof :

- alternating automata -> non-deterministic automata
- transitions are defined by logical formulas :

$$\psi_{p,\sigma}(x, \ldots, X_p, \ldots, X_q, \ldots)$$

expressing that a run, with state p at vertex ux, with label $t(ux) = \sigma$, has a correct choice of states $\ldots, p, \ldots, q, \ldots$ for its sons

M-expansion :decidability

- boolean operations on languages translate into effective operations on automata
- emptiness of the recognized language is decidable.

M-expansion :definability

Definition

A logical finite di-automaton over A^* is a tuple

$$\mathbb{A} = \langle Q, a_0, q_0, Q_f, \delta\rangle$$

where Q is a finie set, $q_0 \in Q$, $Q_f \subseteq Q$ and

$$\delta = (\Phi_{p,q})_{(p,q)\in Q\times Q}$$

is a family of MSO-formulas with two free individual variables.

A computation of $\mathbb{A}$ is a sequence

$$a_0, q_0, a_1, q_1, \dots, a_n, q_n, \dots, a_\ell, q_\ell$$

such that, for every $n \in [0, \ell - 1]$,

$$\mathcal{M} \models \Phi_{q_n,q_{n+1}}(a_n, a_{n+1}).$$

M-expansion :definability

Theorem (G.S. today)

Let $R \subseteq A$ be M-definable. It is M-definable iff it is recognized by some finite logical di-automaton over $\mathcal{M}$.

Adapt the case of S-definable subsets to the transitions of the M-tree-automata.

Decidability of definability

The question

Input : A M-definable subset R of $\mathcal{M}^*$?
Question : Is R also S-definable ?

Algebraic Characterisation

A “logical di-automaton” can be turned into a deterministic automaton

$$AQ \times A \to Q.$$

Proposition

Let $R \subseteq A$ be M-definable. It is S-definable iff the minimal automaton of R is finite.

Idea : If the minimal automaton is finite then its transitions can be expressed within MSO over $\mathcal{M}$; hence it is a logical finite automaton.

Decision procedure

Theorem (G.S. today)

Let us assume that $\mathcal{M}$ has a decidable MSO and that the finiteness property for a subset of $\mathcal{M}$ is expressible in MSO.
Given a M-definable subset of $\mathcal{M}$ one can decide whether this subset is S-definable.

Ideas :
- use the NSC from proposition ??
- express existence of a finite transversal for $\equiv_R$ by a MSO-formula.