Size-change termination in reverse mathematics

Emanuele Frittaion

(joint work with Silvia Steila, Keita Yokoyama, and Florian Pelupessy)

CiE 2016

Size-Change Termination (**SCT**) is a termination property that can be automatically verified.

SCT was introduced and studied by Lee, Jones, Ben-Amram 2001 in the context of **first-order functional programs**.

SCT first-order functional programs correspond to **multiply recursive** functions (Ben-Amram 2002).

Loosely speaking, a program is SCT if any infinite **call sequence** would cause an infinite descending sequence in some well-founded domain.

If a program does not terminate, then there exists an infinite call sequence. Thus SCT is a sufficient condition for **termination**.

Theorem (SCT criterion)

Let $\mathcal{G}$ be a **description** *of a program P.*

Then $\mathcal{G}$ is SCT iff every idempotent $G \in \mathrm{cl}(\mathcal{G})$ has an arc $x \xrightarrow{\downarrow} x$.

A description is a finite set of **size-change** graphs.

Theorem (Termination of SCT programs)

If P is SCT (i.e. P has a **safe** *SCT description) then P is terminating.*

For a reverse mathematics analysis...

Definition

- $\mathcal{G}$ is **MSCT** (Multipath-Size-Change Terminating) if $\mathcal{G}$ is SCT.
- $\mathcal{G}$ is **ISCT** (Idempotent-Size-Change Terminating) if every idempotent $G \in \mathrm{cl}(\mathcal{G})$ has an arc $x \xrightarrow{\downarrow} x$.

Definition

- P is **MSCT** if P has a **safe** MSCT decription.
- P is **ISCT** if P has a **safe** ISCT description.

Theorem (SCT criterion)

*Let $\mathcal{G}$ be a description of a program P.
Then $\mathcal{G}$ is MSCT iff $\mathcal{G}$ is ISCT.*

Theorem (SCT criterion for graphs)

*Let $\mathcal{G}$ be a finite set of size-change graphs.
Then $\mathcal{G}$ is MSCT iff $\mathcal{G}$ is ISCT.*

Theorem (Termination of SCT programs)

If P is M(I)SCT then P is terminating.

Summary of results

$\mathsf{I}\Sigma^0_2 =$ SCT criterion $=$ SCT criterion for graphs

$\mathrm{WO}(\omega^\omega) \leq$ Termination of MSCT programs $\leq$
Termination of ISCT programs $\leq$?

$\mathrm{WO}(\omega^{\omega^\omega}) \leq$ Termination of ISCT programs

Syntax

$$
\begin{aligned}
x &\in \mathrm{Var} \\
f &\in \mathrm{Fun} \\
op &\in \mathrm{Op} \\
a \in \mathrm{AExp} &::= x \mid x+1 \mid x-1 \mid op(a,\ldots,a) \mid f(a,\ldots,a) \\
b \in \mathrm{BExp} &::= x=0 \mid x=1 \mid x<y \mid x\leq y \mid b\wedge b \mid b\vee b \mid \neg b \\
e \in \mathrm{Exp} &::= a \mid \textbf{if } b \textbf{ then } e \textbf{ else } e \\
def \in \mathrm{Def} &::= f(x_0,\ldots,x_{n-1}) = e \\
p \in \mathrm{Prog} &::= def_0,\ldots,def_{m-1}
\end{aligned}
$$

A **program** P is a list of finitely many equations

$$f(x_0, \ldots, x_{n-1}) = e^f$$

where $f \in \mathrm{Fun}$ and e^f is an expression (the **body** of f).
We call $x_0, \ldots, x_{n-1}$ the **parameters** of f and denot it by $\mathrm{Var}(f)$.

Example (Peter-Ackermann)

$$\begin{aligned} A(x,y) = {} & \textbf{if } x = 0 \textbf{ then } y+1 \\ & \textbf{else if } y = 0 \textbf{ then } A(x-1,1) \\ & \textbf{else } A(x-1, A(x, y-1)) \end{aligned}$$

Definition (Size-Change graph)

Let $f, g \in \mathrm{Fun}$, and $\mathrm{Var}(f), \mathrm{Var}(g) \subseteq \mathrm{Var}$. A **size-change graph** $G : f \to g$ for P is a bipartite directed graph on $(\mathrm{Var}(f), \mathrm{Var}(g))$.

The set of edges is a subset of $\mathrm{Var}(f) \times \mathrm{Var}(g) \times \{\downarrow, \Downarrow\}$ such that there is at most one edge to any $y \in \mathrm{Var}(g)$. We say that f is the **source** function of G and g is the **target** function of G.

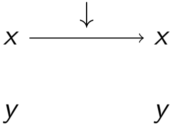
We write $x \xrightarrow{\downarrow} y$ for the **decreasing** edge $(x, y, \downarrow)$, and $x \xrightarrow{\Downarrow} y$ for the **nonincreasing** edge $(x, y, \Downarrow)$.

The idea is that a size-change graph describes the relations between a function f and a function g when f **calls** g in the execution of a program P.

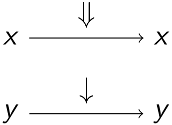
Example (Peter-Ackermann)

$$
\begin{aligned}
A(x,y) = \textbf{if } x = 0 \textbf{ then } & y+1 \\
& \textbf{else if } y = 0 \textbf{ then } \tau_0 : A(x-1,1) \\
& \qquad \textbf{else } \tau_1 : A(x-1, \tau_2 : A(x, y-1))
\end{aligned}
$$

Description of $\tau_0 : A \to A$



Description of $\tau_2 : A \to A$



Definition (composition)

Given two size-change graphs $G_0 : f \to g$ and $G_1 : g \to h$ we define their **composition** $G_0; G_1 : f \to h$. The composition of two edges $x \stackrel{\Downarrow}{\to} y$ and $y \stackrel{\Downarrow}{\to} z$ is one edge $x \stackrel{\Downarrow}{\to} z$. In all other cases the composition of two edges from x to y and from y to z is the edge $x \stackrel{\downarrow}{\to} z$.

We say that the size-change graph G is **idempotent** if $G; G = G$.

Given a finite set of size-change graphs $\mathcal{G}$, $\mathrm{cl}(\mathcal{G})$ is the smallest set which contains $\mathcal{G}$ and is closed by composition.

Definition (multipath)

A **multipath** $\mathcal{M}$ is a sequence $G_0, \ldots, G_n, \ldots$ of size-change graphs such that the target function of G_i is the source function of G_{i+1}. A **thread** is a connected path of edges in $\mathcal{M}$ that starts at some G_t, where $t \in \mathbb{N}$. A multipath $\mathcal{M}$ has an **infinite descent** if some thread in $\mathcal{M}$ contains infinitely many decreasing edges.

Definition (description)

A **description** of P is a set of size-change graphs

$$\mathcal{G} = \{G_\tau : \tau \text{ call of } P\}$$

Definition

A set of size-change graphs $\mathcal{G}$ is **MSCT** if every infinite multipath $G_0, \ldots, G_n, \ldots$ of graphs of $\mathcal{G}$ has an infinite descent.

A set of size-change graphs $\mathcal{G}$ is **ISCT** every idempotent $G \in \mathrm{cl}(\mathcal{G})$ has an arc $x \xrightarrow{\downarrow} x$.

Theorem (SCT criterion)

Let $\mathcal{G}$ be a description of program P. Then $\mathcal{G}$ is MSCT iff $\mathcal{G}$ is ISCT.

Theorem (SCT criterion for graphs)

Let $\mathcal{G}$ be a finite set of size-change graphs. Then $\mathcal{G}$ is MSCT iff $\mathcal{G}$ is ISCT.

SCT criterion is about graphs

Given a finite set $\mathcal{G}$ of size-change graphs, it is straighforward to define a program P such that $\mathcal{G}$ is a description of P.

Proposition (RCA_0)

The following are equivalent:

- *SCT criterion;*
- *SCT criterion for graphs.*

Proving SCT criterion

Proposition (RCA_0)

Let $\mathcal{G}$ be a finite set of size-change graphs. If every multipath $M = G_0, \ldots, G_n, \ldots$ has an infinite descent, then every idempotent $\mathcal{G} \in \mathrm{cl}(\mathcal{G})$ has an arc $x \xrightarrow{\downarrow} x$.

Proof.

Finite pigeonhole principle. □

For the other direction, we introduce:

STAR_k For every $c : [\mathbb{N}]^2 \to k$ there exist a number $t \in \mathbb{N}$ and a color $i \in k$ such that

$$(\exists^\infty n, m)(t < n < m \wedge c(t, n) = c(t, m) = c(n, m) = i)$$

Let $\mathrm{STAR} = \forall k \mathrm{STAR}_k$. STAR_k is a consequence of:

RT^2_k For every $c : [\mathbb{N}]^2 \to k$ there exist an infinite set $H \subseteq \mathbb{N}$ and a color $i \in k$ such that

$$(\forall t, n \in X)(t < n \to c(t, n) = i)$$

Remark

RCA_0 proves STAR_k for any standard k.

Theorem

Ove RCA_0 *the following are equivalent:*

- $\mathsf{I}\Sigma^0_2$*;*
- STAR*;*
- *SCT criterion.*

By Yokoyama and Slaman (unpublished) RT^2 is Π^1_1-conservative over $\mathsf{B}\Sigma^0_3$. It is known that $\mathsf{B}\Sigma^0_3$ is $\widetilde{\Pi}^0_4$-conservative over $\mathsf{I}\Sigma^0_2$ (Parsons 1970). The principle STAR is $\widetilde{\Pi}^0_4$.

Termination of SCT programs

Definition

- P is MSCT if P has a **safe** MSCT decription $\mathcal{G}$.
- P is ISCT if P has a **safe** ISCT description $\mathcal{G}$.

Over RCA_0, MSCT implies ISCT. Over $\mathsf{RCA}_0 + \mathsf{I}\Sigma^0_2$, they are equivalent.

Theorem (Termination of SCT programs)

If P is M(I)SCT then P is terminating.

Over RCA_0, "termination of ISCT programs implies termination of MSCT programs". Over $RCA_0 + I\Sigma^0_2$, they are equivalent.

Semantics

To define **safety** and **termination** we need a **semantics**.

- **Denotational**: can't do in RCA_0
- **Operational**: can do in RCA_0 and straighforward
- **TRS** (Term Rewrite System): can do in RCA_0, straighforward, but not "natural"

An operational semantics consists of **rules** to derive statements of the form $f(u_0, \dots, u_{n-1}) \downarrow v$, with $u_i, v \in \mathbb{N}$.

The rules depends on the program P and the **interpretation** of the functions in $\mathrm{Op}(P)$.

Termination is a Π^0_2 property.

The following rules give a **call-by-value big-step operational semantics**.

$$\frac{}{v \downarrow v} \ (value)$$

$$\frac{t_0 \downarrow v_0 \quad \dots \quad t_{n-1} \downarrow v_{n-1}}{op(t_0, \dots, t_{n-1}) \downarrow \mathbf{op}(v_0, \dots, v_{n-1})} \ (op)$$

$$\frac{t_i \downarrow v_i \quad f(v_0, \dots, v_i, \dots, t_{n-1}) \downarrow v}{f(v_0, \dots v_{i-1}, t_i, \dots, t_{n-1}) \downarrow v} \ (fun)$$

$$\frac{e^f[v_0, \dots, v_{n-1}] \downarrow v}{f(v_0, \dots, v_{n-1}) \downarrow v} \ (call)$$

All v's are natural numbers. In (fun) we have $t_i \notin \mathbb{N}$.

The **terms** t on the left-hand side of a statement $t \downarrow v$ are built up from natural numbers, function and operation variables. We can dispense with boolean terms because they are decidable.

For instance the term $(\textbf{if } 2 < 0 \textbf{ then } t_0 \textbf{ else } t_1)$ is just t_1.

P **terminates** on **input u** if there exists a derivation of $f(\mathbf{u}) \downarrow v$ for some $v \in \mathbb{N}$.

Proposition (RCA_0)

- *If d is a derivation of $t \downarrow v$ and d' is a derivation of $t \downarrow v'$, then $d = d'$.*
- *If $t \downarrow$ and s is a subterm of t, then $s \downarrow$.*

The semantics is pretty standard. See for instance Winskel: *The formal semantics of programming languages: An Introduction*, 1993).

Our rules are slightly different from Winskel cause we want to make the **call relation** more "explicit".

The **call relation** $t \to s$ is defined between **terms**.

For instance, we want $A(2,3) \to A(1, A(2,2))$.

The call relation is a Σ^0_1 property.

A **state transition** is $(f, \mathbf{u}) \to (g, \mathbf{v})$, where $f, g \in \mathrm{Fun}(P)$ and $\mathbf{u}, \mathbf{v}$ are (tuples of) natural numbers, and $f(\mathbf{u}) \to g(\mathbf{v})$.

Let P be a program. Given $\mathbf{u}$, the **activation tree** $T = T_P^{\mathbf{u}}$ consists of all finite sequences of state transitions

$$(f, \mathbf{u}) \to (g, \mathbf{v}) \to \ldots \to (h, \mathbf{w}) \in T,$$

where f is the initial function of P.

The activation tree is Σ_1^0.

Proposition (RCA_0)

The activation tree $T_P^{\mathbf{u}}$ is finite iff P terminates on $\mathbf{u}$.

Definition

Let P be a program, $f, g \in \mathrm{Fun}(P)$

$$f(x_0, \dots, x_{n-1}) = e^f$$
$$g(y_0, \dots, y_{k-1}) = e^g$$

Let $G\colon f \to g$ be a size-change graph on $(\mathrm{Var}(f), \mathrm{Var}(g))$. Then G is **safe** if whenever $(f, \mathbf{u}) \to (g, \mathbf{v})$,

- $x_i \xrightarrow{\downarrow} y_j \in G$ implies $u_i > v_j$
- $x_i \xrightarrow{\Downarrow} y_j \in G$ implies $u_i \geq v_j$.

A description $\mathcal{G}$ of a program P is **safe** if every $G \in \mathcal{G}$ is safe.

If a program is SCT then there are no infinite sequences of state transitions $(f, \mathbf{u}) \to (g, \mathbf{v}) \to \ldots \to (h, \mathbf{w}) \to \ldots$.

Proposition (RCA_0)

If P is MSCT, then $T_P^{\mathbf{u}}$ has no infinite branches for all $\mathbf{u}$.

Proposition (RCA_0)

TFAE:

- ACA_0
- *If P does not terminate on $\mathbf{u}$ then $T_P^{\mathbf{u}}$ has an infinite branch.*

Lower bounds for termination

Kreuzer and Yokoyama 2016 showed that, over RCA_0,

$$\mathrm{WO}(\omega^\omega) \text{ is equivalent to } \forall f \, \mathrm{Tot}(A_f),$$

where $A_f : \mathbb{N}^2 \to \mathbb{N}$ is the Péter-Ackermann function relativized to $f : \mathbb{N} \to \mathbb{N}$:

$$\begin{aligned} A_f(x, y) = \ &\textbf{if } x = 0 \textbf{ then } f(y) \\ &\textbf{else if } y = 0 \textbf{ then } A_f(x-1, 1) \\ &\textbf{else } A_f(x-1, A_f(x, y-1)) \end{aligned}$$

For all $n \geq 2$ and for all $f : \mathbb{N} \to \mathbb{N}$ we define $A_f^n : \mathbb{N}^n \to \mathbb{N}$ such that

Proposition (RCA_0)

- $(\forall f)(A_f^2$ *is MSCT*$)$
- $(\forall n \geq 2)(\forall f)(A_f^n$ *is ISCT*$)$
- $(\forall n \geq 1)$, $\mathrm{WO}(\omega^{\omega^n})$ *is equivalent to* $\forall f \, \mathrm{Tot}(A_f^{n+1})$

Theorem (RCA_0)

- *"Every ISCT program terminates" implies* $\mathrm{WO}(\omega^{\omega^\omega})$
- *"Every MSCT program terminates" implies* $\mathrm{WO}(\omega^\omega)$

Upper bounds for termination

Question

What do we need to prove that M(I)SCT programs terminate?

Conjecture

$\mathsf{RCA}_0 + \mathrm{WO}(\omega^{\omega^{\omega}})$.

SCT programs compute exactly the multiply recursive functions.

Every multiply recursive function is bounded by F_α for some $\alpha < \omega^\omega$.

The **fast-growing** hierarchy is

$$F_0(x) = x + 1$$
$$F_{\alpha+1}(x) = F_\alpha^{(x+1)}(x)$$
$$F_\lambda(x) = F_{\lambda[x]}(x)$$

Theorem (RCA_0)

For every $\alpha < \varepsilon_0$, $\mathrm{WO}(\omega^\alpha)$ is equivalent to $\forall f \, \mathrm{Tot}(F_{\alpha,f})$, where $\{F_{\alpha,f}\}$ is the fast-growing hierachy relative to $f : \mathbb{N} \to \mathbb{N}$.

Corollary (RCA_0)

$\mathrm{WO}(\omega^{\omega^\omega})$ is equivalent to $\forall f \, \mathrm{Tot}(F_{\omega^\omega,f})$.

Conjecture (RCA_0)

$\mathrm{WO}(\omega^{\omega^\omega})$ implies "every ISCT program terminates"

The original proof goes as follows: if P does not terminate then there exists an infinite call sequence. Such a proof requires ACA_0.

Tait 1961 proved that nested ordinal recursion of type α can be reduced to unnested ordinal recursion of type ω^α.

This is related to the fact that $F_\alpha = H_{\omega^\alpha}$, where the **Hardy** hierarchy is defined by

$$
\begin{aligned}
H_0(x) &= x \\
H_{\alpha+1}(x) &= H_\alpha(x) \\
H_\lambda(x) &= H_{\lambda[x]}(x)
\end{aligned}
$$

Clearly, $\mathrm{WO}(\alpha)$ implies $\mathrm{Tot}(H_\alpha)$.

Skipping the details, the proof of Tait shows how to assign descending sequences of ordinals to reduction sequences.

For instance, to the sequence

$$A(2,3) \to A(1, A(2,2)) \to A(1, A(1, A(2,1))) \to A(1, A(1, A(1, A(1,0)))) \to A(1, A(1, A(1, A(0,1)))) \to A(1, A(1, A(1, 2)))$$

we assign the ordinals

$$\omega^{\omega 2+3} \cdot 2 > \omega^{\omega 2+3} + \omega^{\omega 2+2} \cdot 2 > \omega^{\omega 2+3} + \omega^{\omega 2+2} + \omega^{\omega 2+1} \cdot 2 > \dots$$

Let $f \in \mathrm{Fun}(P)$, $\mathrm{Var}(f) = \{x_0, \ldots, x_{n-1}\}$.

We say that f has **semantic lexicographic descent** if we can order the parameters $x_{k_0}, \ldots, x_{k_{n-1}}$ so that whenever $(f, \mathbf{u}) \to (f, \mathbf{v})$ we have $(u_{k_0}, \ldots, u_{k_{n-1}}) > (v_{k_0}, \ldots, v_{k_{n-1}})$ in the standard lexicographic ordering of $\mathbb{N}^n$.

We say that P has **semantic lexicographic descent** if every function in P has semantic lexicographic descent.

semantic lexicographic descent $\approx$ multiply recursive (programming oriented definition)

Ben-Amram 2002 showed that every SCT program P can be effectively transformed into an **equivalent** program P' of the form

(!) $f(x_0, \ldots, x_{n-1}) = e^f$

with semantic lexicographic descent.

Proposition ($\mathrm{RCA}_0 + \mathrm{WO}(\omega^{\omega^\omega})$)

If P is (!) and has semantic lexicographic descent, then P terminates.

Proof.

Tait. □

Thank you very much for listening