Weak Synchronization and Synchronizability of Multi-tape Automata and Machines

Oscar H. Ibarra\textsuperscript{1} and Nicholas Tran\textsuperscript{2}

\textsuperscript{1}Department of Computer Science
University of California at Santa Barbara
ibarra@cs.ucsb.edu

\textsuperscript{2}Department of Mathematics & Computer Science
Santa Clara University
ntran@math.scu.edu

Language and Automata Theory and Applications, 2012
Outline

1. Problems and Motivation
2. Main Results
3. Proof Techniques
4. Summary
Basic Model: Multi-tape PDA + Reversal-bounded Counters
Problem I: Is machine \textit{weakly} $k$-synchronized?

\textit{weakly} $k$-synchronized: If $(x_1, x_2, \ldots, x_l)$ is accepted, then there is an \textit{accepting} computation in which no two input heads \textit{not reading} $\$$ are ever more than $k$ cells apart.
Problem II: Is machine *weakly synchronizable*?

**weakly synchronizable:** Given a machine $M$, is there an $M'$ such that $L(M) = L(M')$ and $M'$ is weakly k-synchronized?
Motivation: SQL Injection Attack

```php
<?php
$query = 'SELECT * FROM users WHERE
email = '' . $_POST['email'] . ''
AND pwdhash = ''
hash('sha256', $_POST['password']) . '';
?>
```

**Input:**
- email: "OR 1=1;--
- password: anything

**Output:**
```
SELECT * FROM users WHERE email='' OR 1=1;-- 
AND pwdhash=''...
```
Previous Works

- \textit{single-track} DFA to represent individual string variables [Xu et al.], [Shannon et al.]
- \textit{multi-track} DFA to represent groups of string variables [Yu et al.]
- \textit{strongly synchronized multi-tape} and multi-head FA [Ibarra et al.]
- \textit{weakly synchronized multi-tape} FA [Egecioglu et al.]
Variants of Basic Model

- **$k$-ambiguous**: at most $k$ accepting computations per input
- **unambiguous**: 1-ambiguous
- **bounded input**: input $x \in a_{i1}^* a_{i2}^* \ldots a_{ik}^*$ ($a_{ij}$’s distinct letters)
- **$NCM$**: NPDA with unary stack alphabet
- **reversal-bounded counters**: the stack height function has a bounded number of local maxima/minima
Main Results: multi-tape PDA

2-ambiguous 2-tape 1-reversal NCM (I, II)
2-ambiguous 2-tape 3-reversal NPDA over \( \{a, b\}^* \times c^* \) (I, II)
2-tape 1-reversal NCM over \( \{a, b\}^* \times c^* \) (I, II)
3-ambiguous 2-tape NCM over \( \{a, b\}^* \times c^* \) (I, II)

Unambiguous \( n \)-tape NPDA + reversal-bounded counters (I)
\( n \)-tape NPDA + reversal-bounded counters over bounded inputs (I)
Technique 1: Infiniteness of one-tape NPDA

Theorem

*It is decidable whether an unambiguous n-tape NPDA M with reversal-bounded counters is weakly k-synchronized for some k.*

Proof.

*M is not weakly synchronized iff two of its heads get separated by more than k cells for infinitely many k’s on accepting computations.*

Construct a one-tape NPDA $M'$ with 2 counters to catch these offending *accepting* computations.

It is decidable whether $L(M')$ is infinite!
Problems and Motivation
Main Results
Proof Techniques
Summary

Guessing offending accepting computations

Input 1

Input 2

Stack...

M

0 0 0

M'

Stack...

1 1 1 1 1

1 1
**Theorem**

Given a (2-ambiguous) 2-tape 1-reversal NCM $M$, it is undecidable whether there is a 2-tape 0-synchronized NPDA $M'$ such that $L(M) = L(M')$.

**Proof.**

Reduction from Post Correspondence Problem:

**INPUT:** $I = \{x_1, \ldots, x_m\}, \{y_1, \ldots, y_n\}$

**OUTPUT:** yes iff $x_{i_1} \ldots x_{i_k} = y_{i_1} \ldots y_{i_k}$ for some indices $i_1, \ldots, i_k$.

Reduction function: $f(I) = M$, which accepts $L_1 \cup L_2$:

- $L_1 = \{(xa^ib^j, yc^k) : i, j, k > 0, x \neq y\}$
- $L_2 = \{(xa^{3i}b^{2i}, xc^i) : i > 0, x \text{ encodes a solution of } I\}$
Construction of 2-tape 1-reversal NCM $M$

$L = \{ (xa^ib^j, yc^k) : i, j, k > 0, x \neq y \} \cup \{ (xa^{3i}b^{2i}, xc^i : i > 0, x \text{ encodes a solution of } I \}$

On input $xa^rb^s, yc^t$, $M$ nondeterministically checks:
- $x \neq y$ (0-synchronized DFA);
- $x = x_{i_1} \ldots x_{i_k}$, $y = y_{i_1} \ldots y_{i_k}$ for some $i_1, \ldots i_k$, and $r = 3t$, $s = 2t$
I has a solution iff

$L$ is accepted by a 0-synchronized 2-tape NCM

- $I$ has no solutions:
  
  $L = L_1$ and hence $L$ is accepted by a 0-synchronized 2-tape DFA

- $I$ has a solution $x = i_1, i_2, \ldots, i_k$:
  
  if $L$ is accepted by a 0-synchronized 2-tape NCM $M$, convert $M$ into a one-tape two-track NCM $M'$

  apply Ogden’s lemma to $\binom{xa^3i b^{2i}}{xc^i}$ to obtain a string not in $L$, contradiction.
Technique 3: Halting Problem

**Theorem**

Given a 2-tape 1-reversal NCM $M$ over $\Sigma^* \times c^*$, where $|\Sigma| \geq 2$, it is undecidable whether $M$ is weakly $k$-synchronized for some $k$.

**Proof.**

Reduction from the halting problem for one-tape Turing machines $T$ on blank input.

Define $H(T) = I_1 \# I_2 \# \ldots \# I_m \#$ to be an encoding of the halting computation of $T$ on blank input, if exists.

Reduction function: $f(T) = M$, a 2-tape NCM that accepts:

$$L = \{(xa^ib^j, c^{\lfloor x \rfloor + k}) : x \in \{0, 1, \#\}^* \land (i, j, k > 0) \land [x \neq H(T) \lor (i = 3k \land j = 2k)]\}$$
Construction of 2-tape 1-reversal NCM $M$

$$L = \{(xa^ib^j, c^{\mid x \mid + k}) : x \in \{0, 1, \#\}^* \land (i, j, k > 0) \land [x \neq H(T) \lor (i = 3k \land j = 2k)]\}$$

On input $xa^rb^s$, $c^t$, $M$ nondeterministically checks:

- $r = 3k$ and $s = 2k$, where $k = t - |x|$ (one reversal)
- $x \neq H(T)$: 0-synchronized DFA
$T$ halts on $\lambda$ iff
$M$ is a 0-synchronized 2-tape NCM

- $T$ does not halt on $\lambda$:
  \[ L = \{(xa^ib^i, c^k) : x \in \{0, 1, \#\}^*, i, j, k > 0\} \] so there is a 0-synchronized accepting computation for all $x \in L$

- $T$ halts on $\lambda$:
  if $M$ is 0-synchronized, convert $M$ into a one-tape two-track NCM $M'$
  apply Ogden's lemma to $\left( \frac{H(T)a^{3i}b^{2i}}{c^{|H(T)|+i}} \right)$ to obtain a string not in $L$, contradiction.
studied synchronization and synchronizability (motivated by web security)

main result: synchronization is undecidable for 2-ambiguous, 2-tape, 1-reversal NCM

becomes decidable for unambiguous or bounded input (even with additional tapes, reversal bounded counters and one pushdown stack)

when one tape is unary, synchronization is undecidable for ambiguous NCM with one reversal, 3-ambiguous NCM, 2-ambiguous NPDA with three reversals