IOF, ACSys and WMSO+U

Stéphane Demri

CNRS – Marie Curie Fellow

Groupe de travail INFINI, January 2015
Overview

1. Marie Curie Fellowship IOF
2. ACSys Group
3. Temporal Logics on Strings
Marie Curie Fellowship IOF
International Outgoing Fellowship (IOF)

- Funding to carry out research abroad.
- IOFs are for researchers from EU member states.
- Minimal requirement: PhD.
- Outgoing phasis (1 or 2 years) + return phasis (1 year).
- Individual fellowships.
Non-flat (but flattable) system for Marie Curie fellowships

Marie Curie Actions Research Fellowship Program is a EU initiative to promote research and innovation.
Application form

- Research program ($\leq 8$ pages)
  This includes presentation of host institutions.

- Extended CV ($\leq 7$ pages).

- Training objectives ($\leq 2$ pages).

- Implementation ($\leq 6$ pages).

- Impact ($\leq 4$ pages).

- Deadline: so far early august (notification in december).
  Project can start up to 1 year after the final signature.

- Acceptance rate: $\sim 15\%$. 
ACSys Group
ACSys members

• Analysis of Computer Systems group (ACSys) is part of Courant Institute of Mathematical Sciences (CIMS), New York University.

• Faculty: Clark Barrett, Patrick Cousot, Ben Goldberg, Thomas Wies, Lenore Zuck.

• Research fellow / visiting positions: Morgan Deters, Dejan Jovanovic, Eric Koskinen, Daniel Schwartz-Narbonne.

CVC4 group

- CVC4: open-source automatic theorem prover for satisfiability modulo theories (SMT) problems. See Morgan’s slides or CVC4 web page.

- Members at NYU: Clark Barrett, Morgan Deters, Kshitij Bansal, Liana Hadarean, Tim King.

- Members at Iowa University and other places: Cesare Tinelli, Tianyi Liang, Andrew Reynolds, Dejan Jovanovic, François Bobot, etc.

- Leader among SMT solvers (performances, diversity of theories, participation to international standards such as SMT-LIB, etc.).
Other places in the area

- Courant Institute of Mathematical Sciences (CIMS).
- CUNY (S. Artemov, M. Fitting, R. Parikh).
- Yale University (R. Piskac)
- Columbia University
- Princeton (New Jersey), MIT (Boston, Main), UPenn (Philadelphia, Pennsylvania).
Overview of my research program there

- Temporal logics modulo theories. See the second part of the talk.

- Decision procedures for fragments of separation logic.
  1. Two-variable fragment. [Demri & Deters, CSL-LICS’14]
  2. One-variable fragment. [CSR’14]
  3. Survey paper. [Demri & Deters, AIML’14]

- Verification of integer programs with SMT solvers.
  1. Prototype: path schema enumeration. [Barrett & Demri & Deters, FROCOS’13]
  3. Survey paper.
Temporal Logics on Strings

Joint work with Morgan Deters (New York University)

See also recent LSV technical report online.
Reasoning about strings

- Need for string reasoning: program verification, analysis of web applications, etc.

- Theory solvers for strings.
  [Liang et al. – Abdulla et al., CAV’14; Hutagalung & Lange, CSR’14]

- Solving word equations.
  [Makanin, Math. 77; Plandowski, JACM 04]

- What about reasoning on sequences of strings?
LTL on strings: $\text{LTL}(\Sigma^*, \preceq_p)$

- **String variables** $\text{SVAR} = \{x_1, x_2, \ldots\}$.

- **Terms**: $t ::= w \mid x \mid Xx$ \hspace{1cm} ($x \in \text{SVAR}, w \in \Sigma^*$)

- **Formulae**:

  $\phi ::= t \preceq_p t' \mid \neg \phi \mid \phi \land \phi \mid X\phi \mid \phi \cup \phi$

- **Example**:

  $$\text{GF}((001 \preceq_p x) \lor (x \preceq_p 1001)) \land \text{G}(\neg (x \preceq_p Xx))$$
A model with $\Sigma = \{0, 1\}$

\[
\begin{align*}
  x_1 & \quad 000 & 011110 & \varepsilon & 1111 & \ldots \\
  x_2 & \quad 101 & 010001 & 010001 & 00 & \ldots & \models F(x_2 \preceq_p X x_3) \\
  x_3 & \quad 00 & 111 & 010001101 & \varepsilon & \ldots
\end{align*}
\]
The case $\Sigma = \{0\}$

- $LTL(\mathbb{N}, \leq) \overset{\text{def}}{=} LTL(\Sigma^*, \preceq_p)$ with $\Sigma = \{0\}$.

- Satisfiability problem for $LTL(\mathbb{N}, \leq)$ is $\text{PSPACE}$-complete.
  
  [Demri & D’Souza, IC 07; Demri & Gascon, TCS 08]
  
  See also [Segoufin & Torunczyk, STACS’11]

- The $\text{PSPACE}$ upper bound is preserved with several LTL extensions or with richer numerical constraints (but no successor relation).
**Logic** \(\text{LTL}(\Sigma^*, \text{clen})\)

- \(\text{clen}(w, w')\): length of the longest common prefix between \(w\) and \(w'\) in \(\Sigma^*\).

\[
\sigma, i \models \text{clen}(t_0, t'_0) \leq \text{clen}(t_1, t'_1) \\
\text{def} \\
\iff \\
\text{clen}([t_0]_i, [t'_0]_i) \leq \text{clen}([t_1]_i, [t'_1]_i)
\]

- Reduction from \(\text{LTL}(\Sigma^*, \preceq_p)\) to \(\text{LTL}(\Sigma^*, \text{clen})\).

\(t \preceq_p t' \implies \text{clen}(t, t) \leq \text{clen}(t, t')\).

- In the sequel either \(\Sigma = [0, k - 1]\) for some \(k \geq 1\) or \(\Sigma = \mathbb{N}\).
Symbolic models for $\text{LTL}(\mathbb{N}, \leq)$

+ Local consistency between two consecutive positions.
Rephrasing the satisfiability property

\[ \phi \text{ is } \text{LTL}(\mathbb{N}, \leq) \text{ satisfiable} \]

iff

there is a symbolic model \( \sigma \) such that

\[ \sigma \models_{\text{symb}} \phi \text{ and } \sigma \text{ has a concrete interpretation in } \mathbb{N} \]
Characterisation for \( \text{LTL}(\mathbb{N}, \leq) \)

- Usual notion of path \( \pi \) between two nodes.

- Strict length of the path \( \pi \): \( \text{slen}(\pi) = \) number of edges labelled by \(<\).

- Strict length between \( \langle x, i \rangle \) and \( \langle x', i' \rangle \):

\[
\text{slen}(\langle x, i \rangle, \langle x', i' \rangle) \overset{\text{def}}{=} \sup \{ \text{slen}(\pi) : \text{path } \pi \text{ from } \langle x, i \rangle \text{ to } \langle x', i' \rangle \}
\]

- Symbolic model \( \sigma \) has a concrete interpretation iff any pair of nodes has a finite strict length.

\[ [\text{Cerans, ICALP’94; Demri & D’Souza, IC 07}] \]
\[ [\text{Gascon, PhD thesis 07; Carapelle & Kartzow & Lohrey, CONCUR’13}] \]
When WMSO+U enters into the play

- There are formulae $\phi$ in $\text{LTL}(\mathbb{N}, \leq)$ for which the set of symbolic models satisfying $\phi$ symbolically and having a concrete interpretation is not $\omega$-regular.  
  [Demri & D’Souza, IC 07]

- $\sigma \models U X \phi \iff$ for every $b \in \mathbb{N}$, there is a finite $Y$ with $\text{card}(Y) \geq b$ such that $\sigma \models \phi(Y)$.
  $B X \phi \overset{\text{def}}{=} \neg U X \phi$.
  [Bojańczyk, CSL'04; Bojańczyk & Colcombet, LICS’06]

- Symbolic models for $\text{LTL}(\mathbb{N}, \leq)$ having a concrete interpretation can be characterized by a formula in $\text{Bool}(\text{MSO}, \text{WMSO}+\text{U})$.

- This leads to decidability of $\text{CTL}^*(\mathbb{N}, \leq)$.
  [Carapelle & Kartzow & Lohrey, CONCUR’13]
  (based on [Bojańczyk & Toruńczyk, STACS’12])
Back to strings
Simple but essential properties for $\text{clen}(\cdot)$

\[
\begin{align*}
\omega_1 & \quad 0 \ 0 \ 0 \ 1 \ 0 \ 2 \\
\omega_2 & \quad 0 \ 0 \ 0 \ 0 \\
\implies & \quad \text{clen}(\omega_1, \omega_2) \leq \text{len}(\omega_1)
\end{align*}
\]

\[
\begin{align*}
\omega_0 & \quad 0 \ 0 \ 0 \ 1 \ 0 \ 2 \\
\omega_1 & \quad 0 \ 0 \ 0 \ 0 \ 1 \ 3 \ 5 \ 6 \\
\omega_2 & \quad 0 \ 0 \ 0 \ 2 \ 1 \ 4 \\
\ldots \\
\omega_k & \quad 0 \ 0 \ 0 \ 3 \ 1 \ 3 \\
\implies & \quad \exists i, j \in [1, k] \text{ such that } \text{clen}(\omega_0, \omega_1) < \text{clen}(\omega_i, \omega_j)
\end{align*}
\]
(Pigeonhole Principle – $\text{card}(\Sigma) = k \geq 2$)

\[
\begin{align*}
\omega_0 & \quad 0 \ 0 \ 0 \ 1 \ 0 \ 2 \quad \text{and} \quad \omega_1 & \quad 0 \ 0 \ 0 \ 0 \ 1 \ 3 \ 5 \\
\omega_1 & \quad 0 \ 0 \ 0 \ 0 \ 1 \ 3 \ 5 \quad \text{and} \quad \omega_2 & \quad 0 \ 0 \ 0 \ 0 \ 1 \ 4 \\
\implies & \quad \text{clen}(\omega_0, \omega_1) = \text{clen}(\omega_0, \omega_2)
\end{align*}
\]
String compatible counter valuations

- Counter valuation $c : \{\text{clen}(t, t') : t, t' \in T\} \rightarrow \mathbb{N}$.

- String-compatibility:

$$\bigwedge_{t, t' \in T} (\text{clen}(t, t) \geq \text{clen}(t, t'))$$

$$\bigwedge_{t_0, \ldots, t_k \in T} \big( \bigwedge_{i \in [0, k]} (\text{clen}(t_0, t_1) < \text{clen}(t_i, t_i)) \land \text{clen}(t_0, t_1) = \cdots = \text{clen}(t_0, t_k) \big)$$

$$\Rightarrow \big( \bigvee_{i \neq j \in [1, k]} (\text{clen}(t_0, t_1) < \text{clen}(t_i, t_j)) \big)$$

$$\bigwedge_{t, t', t'' \in T} (\text{clen}(t, t') < \text{clen}(t', t'')) \Rightarrow (\text{clen}(t, t') = \text{clen}(t, t''))$$

- Size in $\mathcal{O}((q + r)^{k+2})$ with $\text{card}(T) = q + r$. 

Temporal Logics on Strings

23
Characterisation

- String compatibility is equivalent to the existence of a string valuation witnessing the values of the counters $\text{clen}(t, t')$.

- The exact statement is a bit more complex to be used after in the translation from $\text{LTL}(\Sigma^*, \text{clen})$ to $\text{LTL}(\mathbb{N}, \leq)$.

- Checking satisfiability of Boolean combinations of prefix constraints is NP-complete. (upper bound by reduction into QF Presburger arithmetic)

- PSPACE can be obtained using word equations and Plandowski’s PSPACE upper bound. (suffix constraints can be added at no cost)
Translation

- Formula $\phi$ with constant strings $w_1, \ldots, w_q$ and, string variables $x_1, \ldots, x_r$.

- For all $i, j \in [1, q]$, $c_{i,j} \overset{\text{def}}{=} \text{clen}(w_i, w_j)$.

- $T \overset{\text{def}}{=} \{y_1, \ldots, y_q\} \cup \{x_1, \ldots, x_r\} \cup \{Xx_1, \ldots, Xx_r\}$.

- $\phi^{\text{subst}}_1$: replace each $w_i$ by $y_i$.

- $\phi^{\text{rig}}_2 \overset{\text{def}}{=} G(\bigwedge_{i,j \in [1, q]}(\text{clen}(y_i, y_j) = c_{i,j}))$. 

Temporal Logics on Strings
Translation (II)

- Formula $\phi_{next}^3$:

$$G \left( \bigwedge_{t, t' \in \{y_1, \ldots, y_q\} \cup \{x_{x_1}, \ldots, x_{x_r}\}} \text{clen}(t, t') = X \text{clen}(t \setminus X, t' \setminus X) \right)$$

- Formulae $\psi_I, \psi_{II}$ and $\psi_{III}$ related to string-compatible counter valuations over $\mathbb{T}$.

- $\phi$ is satisfiable in $\text{LTL}(\Sigma^*, \text{clen})$ iff

$$\phi_{subst} \land \phi_{rig} \land \phi_{next}^3 \land \psi_I \land \psi_{II} \land \psi_{III}$$

is satisfiable in $\text{LTL}(\mathbb{N}, \leq)$. 
Complexity and decidability

- Satisfiability problems for $\text{LTL}(\Sigma^*, \preceq_p)$ and $\text{LTL}(\Sigma^*, \text{clen})$ are $\text{PSPACE}$-complete.

- This also holds for any LTL extension that behaves as LTL as far as the translation into Büchi automata is concerned (Past LTL, linear $\mu$-calculus, ETL, etc.).

- For any satisfiable $\phi$ in $\text{LTL}(\mathbb{N}^*, \text{clen})$, models with letters in $[0, N + 2 \times \text{size}(\phi)]$ are sufficient ($N$ max. letter in $\phi$).
Lifting to branching-time temporal logics

- $\text{CTL}^*(\Sigma^*, \text{clen})$: branching-time extension of $\text{LTL}(\Sigma^*, \text{clen})$.

- Translation can be extended for $\text{CTL}^*(\Sigma^*, \text{clen})$.

- Proof is a bit more complex but the string characterisation is used similarly.

- The satisfiability problem for $\text{CTL}^*(\Sigma^*, \text{clen})$ is decidable. By reduction into $\text{CTL}^*([\mathbb{N}, \leq])$ shown decidable in

[Carapelle & Kartzow & Lohrey, CONCUR’13]
A selection of open problems

- Complexity characterisation for uniform sat. problem.
  
  **input:** alphabet $\Sigma = [0, k - 1]$ ($k$ in unary) or $\Sigma = \mathbb{N}$, and a formula $\phi$ in $\text{LTL}(\Sigma^*, \text{clen})$

  **question:** is $\phi$ satisfiable in $\text{LTL}(\Sigma^*, \text{clen})$?

- Dec. status of $\text{LTL}([0, 1]^*, \preceq_p, \preceq_s)$.

- Dec. status of $\text{LTL}([0, 1]^*, \preceq_p, \text{REG})$ with regularity tests.

- Decidability status of $\text{LTL}([0, 1]^*, \sqsubseteq)$.