TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

The Complexity of Formal Language Decision Problems

Bachelor Thesis Talk

Philipp Johann Schepper

Department of Computer Science
Technische Universität Kaiserslautern

March 21, 2018

Context-Free Recognition (CFL recognition)

Let $G = (\Sigma, N, P, S)$ be a Context-Free Grammar (CFG) with:
$P \subseteq N \to (\Sigma \cup N)^*$
$S \in N$
Define $\mathcal{L}(G)$ as the language of G.

Definition (CFL recognition)

Decide for a input word w and a fixed grammar whether $w \in \mathcal{L}(G)$.

Theorem

CFL recognition can be decided in $\mathcal{O}\left(n^3\right)$ time, for a word $w = w_1 \cdots w_n$ and a fixed CFG G.

PDA Emptiness I

Let $A = (Q, \Sigma, \Gamma, \delta, q_0, Q_f)$ be a PDA with:
$\delta \subseteq (Q \times (\Sigma \cup \{\epsilon\}) \times (\Gamma \cup \{\epsilon\})) \to (Q \times \Gamma^*)$
$Q_f \subseteq Q$
Define $\mathcal{L}(A)$ as the language of the PDA A.

Definition (PDA Emptiness)

Decide for a PDA A whether $\mathcal{L}(A) = \emptyset$.

Acceptance by empty stack: $A := (Q, \Sigma, \Gamma, \delta, \S, q_0)$ with $\S \in \Gamma$ as the special initial stack symbol.

PDA Emptiness II

Definition (Size of a PDA)

- $|A|_1 := |Q|$
- $|A|_2 := |\delta|$, i.e. the number of transitions

We define further the standard size of a PDA as follows: $|A| := |A|_1 = |Q|$

Theorem

PDA Emptiness can be decided in $\mathcal{O}\left(|A|_2^3\right) = \mathcal{O}\left(|\delta|^3\right)$ *time, respectively* $\mathcal{O}\left(|A|^4\right)$ *time.*

Context-Free Reachability (CFR) I

Let $G = (\Sigma, N, P, S)$ be a fixed CFG. Let $H = (V, E)$ be a graph with:
$V = \{v_1, \ldots, v_n\}$
$E \subseteq V \times (\Sigma \cup \{\epsilon\}) \times V$

Definition (Definitions for CFR)

- A-path from s to t: Path from s to t labeled by a word w with $A \to^* w$.
- t is A-reachable from s. $\iff$ There is A-path from s to t.
- Write G-path for S-path and G-reachable for S-reachable.

Definition (CFR)

Decide for a graph H, two nodes $s, t \in V$ and a fixed CFG G whether there is a G-path from s to t.

Context-Free Reachability (CFR) II

Definition (Size of a CFR instance)

- Fix the grammar.
- Define the size of an CFR instance by the size of the graph.
- The size of a graph is its number of vertices.

Theorem

CFR can be decided in $\mathcal{O}\left(|H|^3\right)$ time. For a fixed CFG G.

Context-Free Reachability (CFR) III

Definition (Dyck-k Reachability (DkR))

For DkR we fix the CFG G_k producing the Dyck-k language Dk.
Define $G_k := (\Sigma_k \cup \overline{\Sigma_k}, \{S\}, P_k, S)$ with

$$\begin{aligned} \Sigma_k &:= \{(_1, (_2, \ldots, (_k\} \\ \overline{\Sigma_k} &:= \{\bar{a} \mid a \in \Sigma_k\} = \{)_1,)_2, \ldots,)_k\} \\ P_k &:= \{S \to \epsilon | SS | (_1 S)_1 | \ldots | (_k S)_k\}. \end{aligned}$$

G_k

k-CLIQUE

Let $H = (V, E)$ with $|V| = n$ be an undirected graph.

Definition

A k-clique is a set of k vertices, where each vertex is connected to each other vertex.

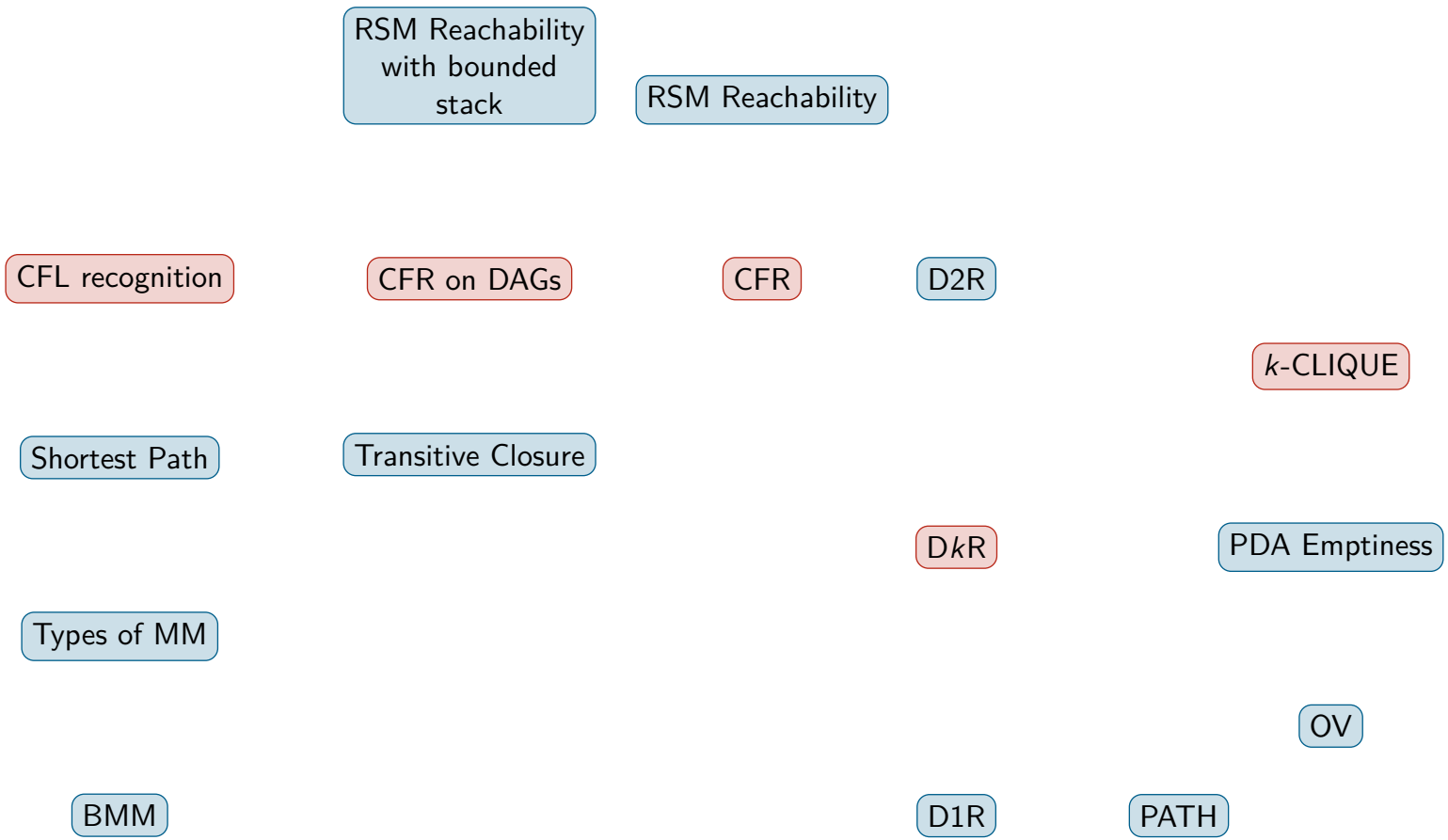
Definition (k-CLIQUE)

Decide for an undirected graph $G = (V, E)$ and a fixed $k \in \mathbb{N}$ whether there is a k-clique in G.

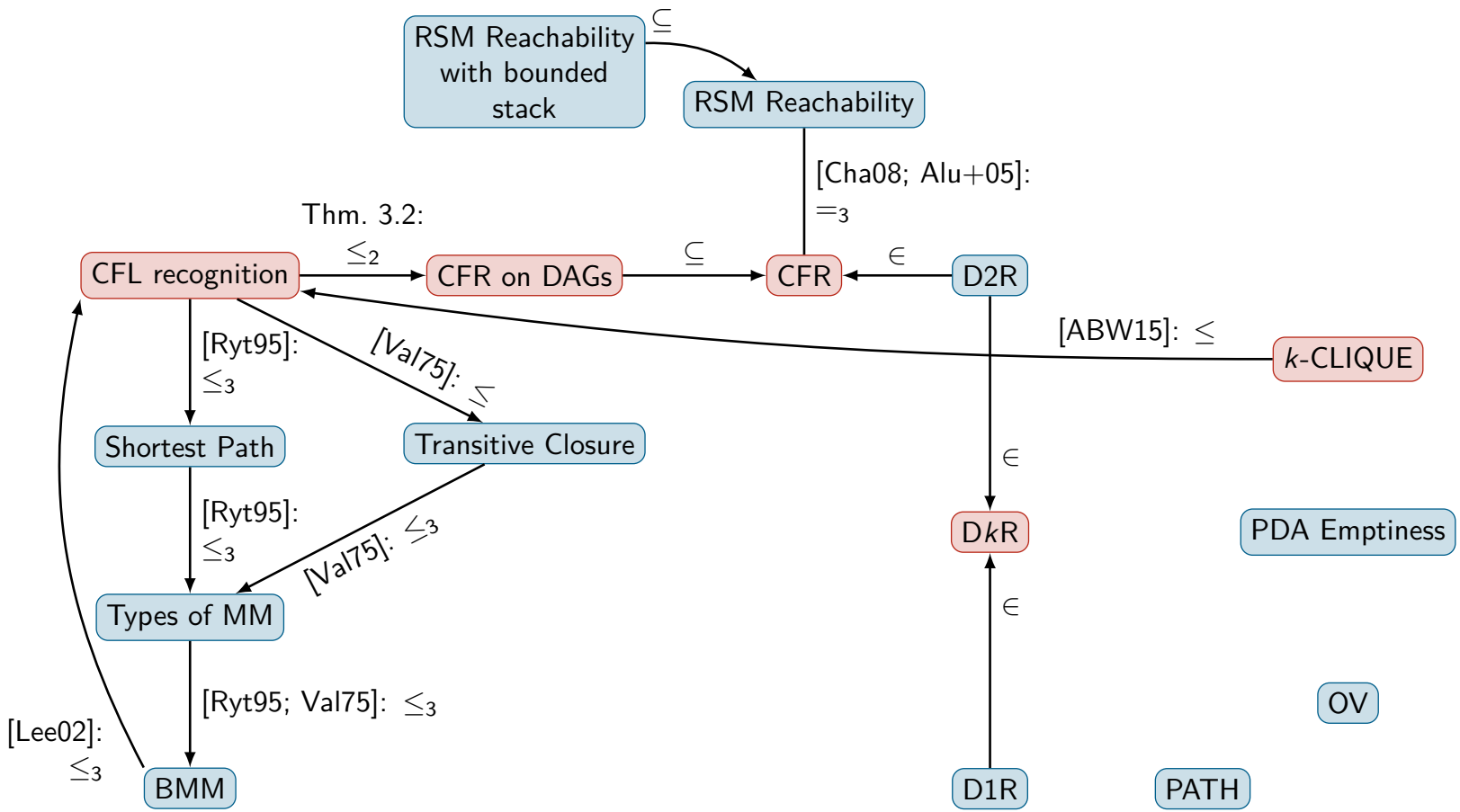
Theorem

k-CLIQUE can be decided in $\mathcal{O}\left(n^k\right)$ time.

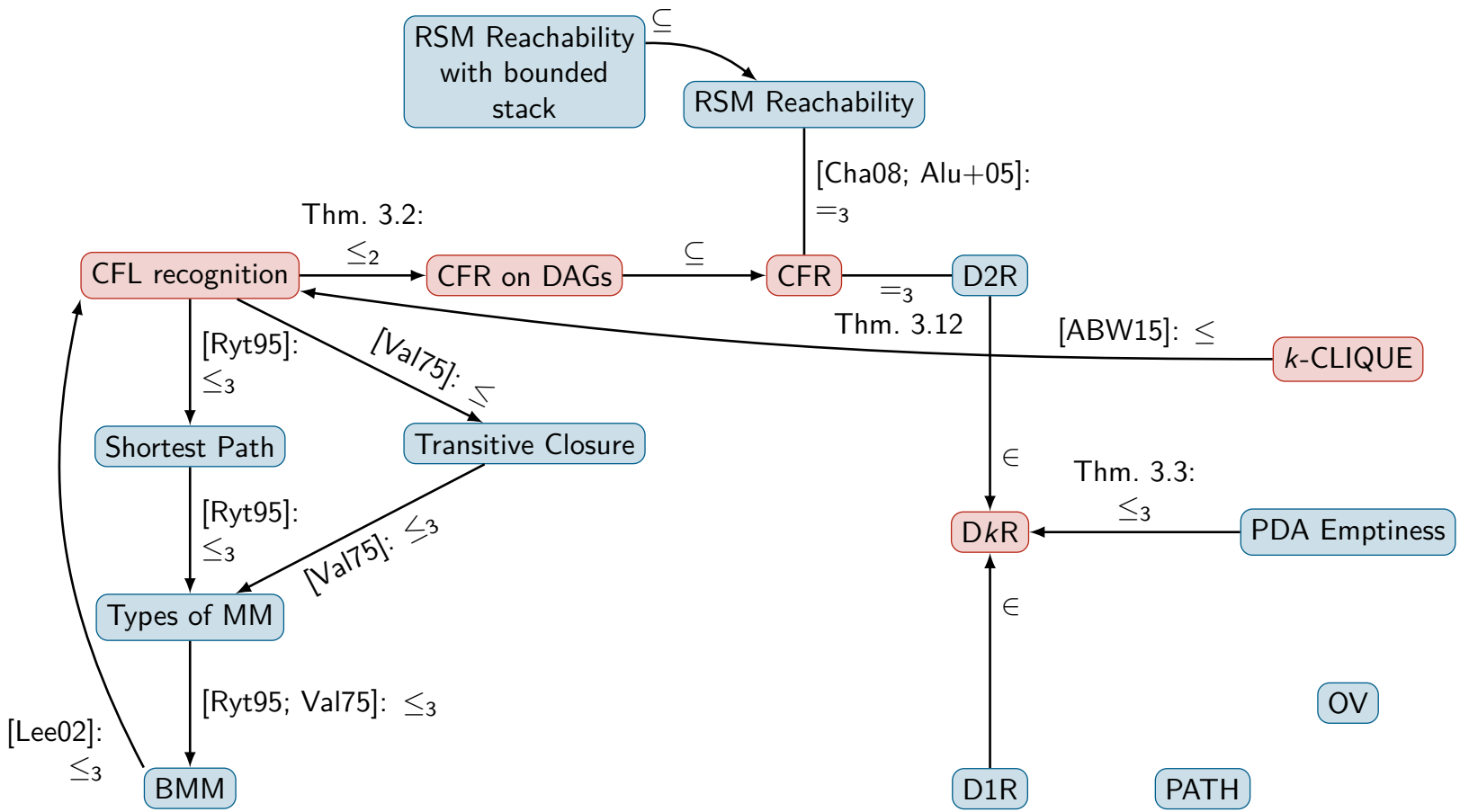
Overview


RSM Reachability with bounded stack
RSM Reachability
CFL recognition
CFR on DAGs
CFR
D2R
k-CLIQUE
Shortest Path
Transitive Closure
DkR
PDA Emptiness
Types of MM
OV
BMM
D1R
PATH

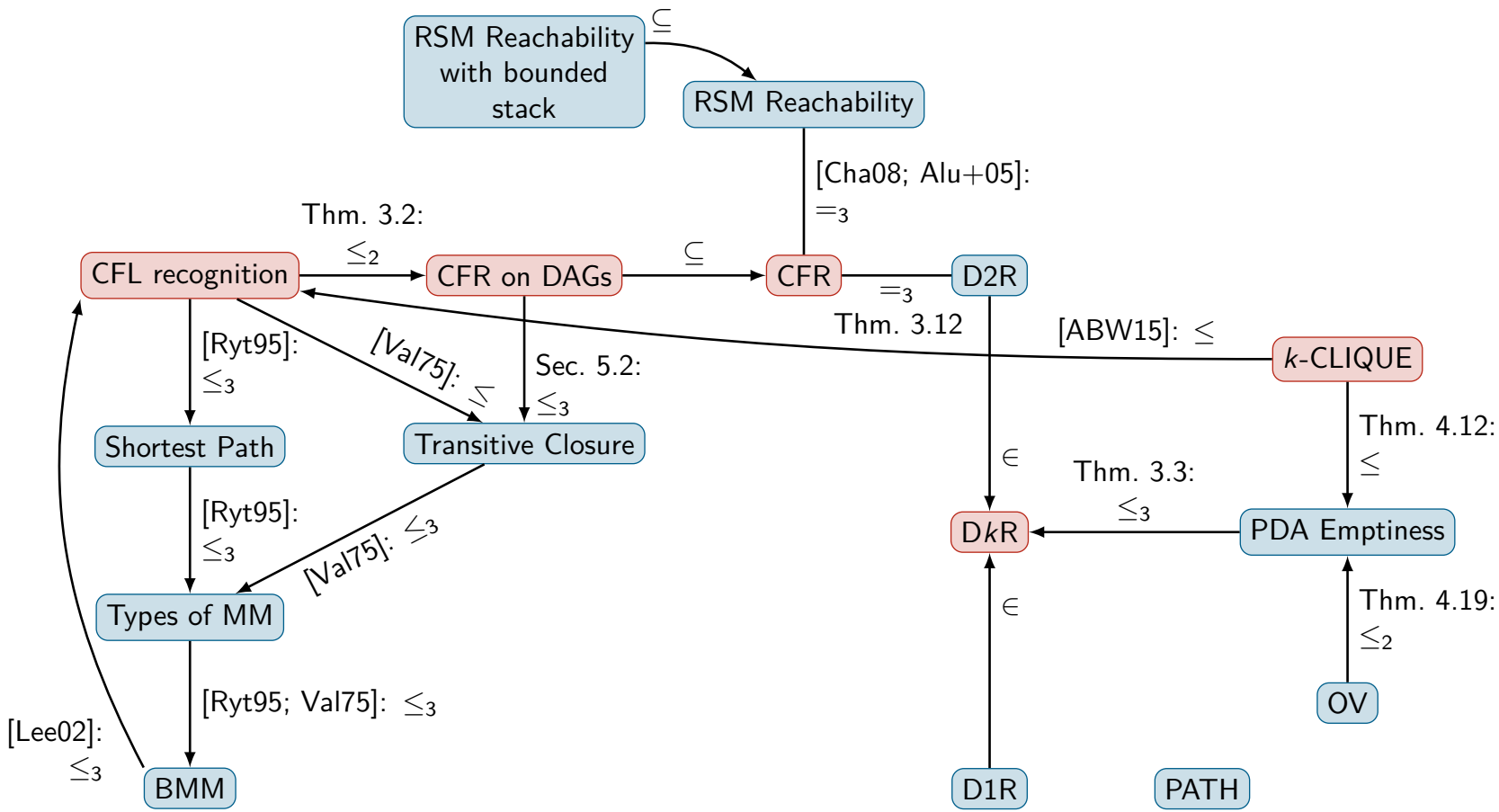
Overview


RSM Reachability with bounded stack
⊆
RSM Reachability
[Cha08; Alu+05]: =3
Thm. 3.2: ≤2
CFL recognition
CFR on DAGs
⊆
CFR
∈
D2R
[ABW15]: ≤
k-CLIQUE
[Ryt95]: ≤3
[Val75]: ≤
Shortest Path
Transitive Closure
∈
[Ryt95]: ≤3
[Val75]: ≤3
DkR
PDA Emptiness
Types of MM
∈
OV
[Ryt95; Val75]: ≤3
[Lee02]: ≤3
BMM
D1R
PATH

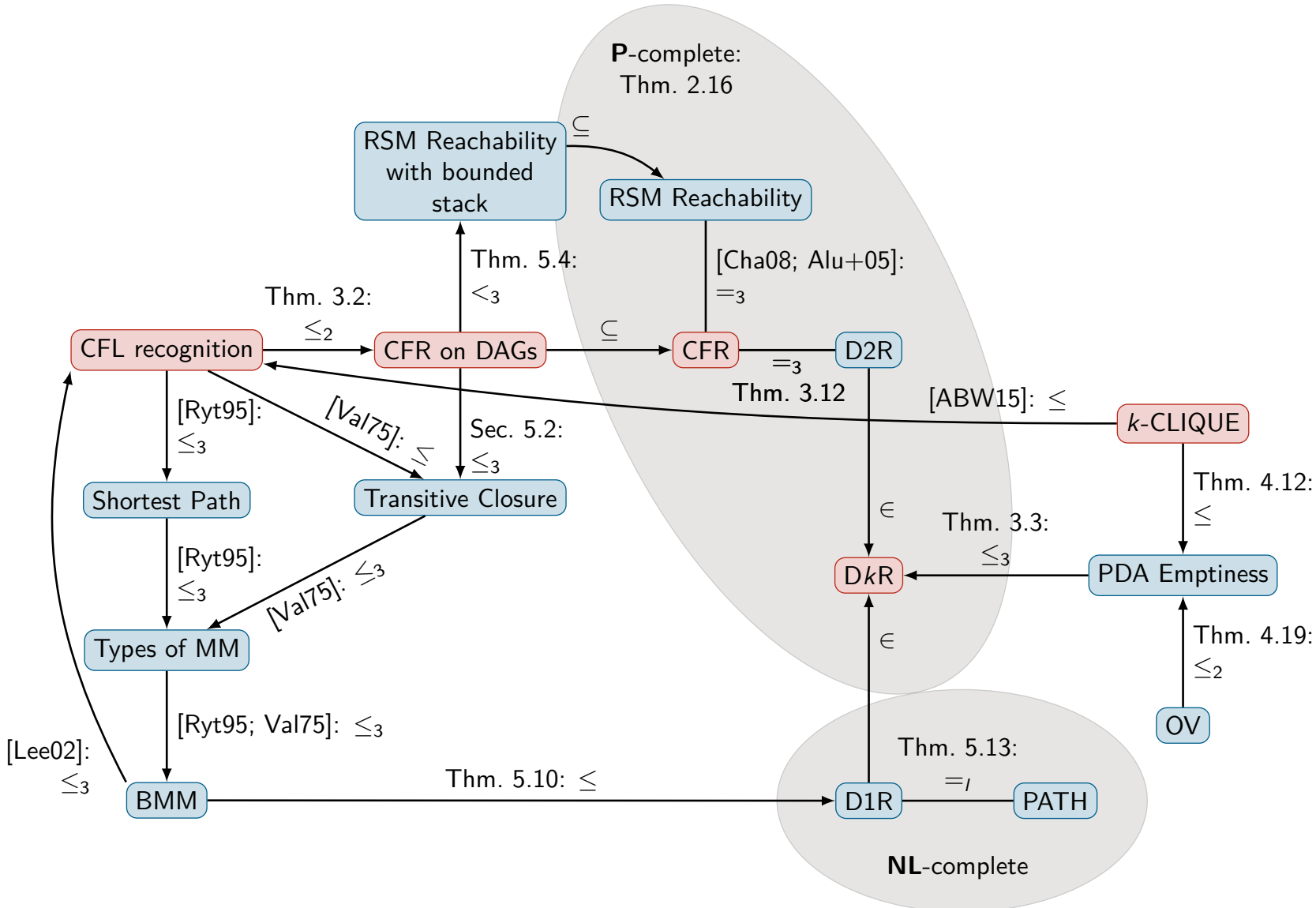
Overview


RSM Reachability with bounded stack
⊆
RSM Reachability
[Cha08; Alu+05]:
=3
Thm. 3.2:
≤2
CFL recognition
CFR on DAGs
⊆
CFR
=3
D2R
Thm. 3.12
[ABW15]: ≤
k-CLIQUE
[Ryt95]:
≤3
[Val75]:
Shortest Path
Transitive Closure
∈
Thm. 3.3:
≤3
DkR
PDA Emptiness
[Ryt95]:
≤3
[Val75]:
Types of MM
∈
OV
[Ryt95; Val75]: ≤3
[Lee02]:
≤3
BMM
D1R
PATH

Overview


RSM Reachability with bounded stack
⊆
RSM Reachability
[Cha08; Alu+05]:
=3
Thm. 3.2:
≤2
CFL recognition
CFR on DAGs
⊆
CFR
=3
D2R
Thm. 3.12
[ABW15]: ≤
k-CLIQUE
[Ryt95]:
≤3
[Val75]:
Sec. 5.2:
≤3
Shortest Path
Transitive Closure
Thm. 4.12:
≤
∈
Thm. 3.3:
≤3
[Ryt95]:
≤3
[Val75]:
DkR
PDA Emptiness
Types of MM
Thm. 4.19:
≤2
∈
[Ryt95; Val75]: ≤3
[Lee02]:
≤3
OV
BMM
D1R
PATH

Overview


P-complete:
Thm. 2.16
RSM Reachability with bounded stack
⊆
RSM Reachability
Thm. 5.4:
$<_3$
[Cha08; Alu+05]:
$=_3$
Thm. 3.2:
$\leq_2$
CFL recognition
CFR on DAGs
⊆
CFR
$=_3$
D2R
Thm. 3.12
[ABW15]: ≤
k-CLIQUE
[Ryt95]:
$\leq_3$
[Val75]:
$\leq_3$
Sec. 5.2:
$\leq_3$
Shortest Path
Transitive Closure
∈
Thm. 3.3:
$\leq_3$
Thm. 4.12:
≤
DkR
PDA Emptiness
[Ryt95]:
$\leq_3$
[Val75]:
$\leq_3$
Thm. 4.19:
$\leq_2$
Types of MM
∈
OV
[Ryt95; Val75]: $\leq_3$
Thm. 5.13:
$=_l$
[Lee02]:
$\leq_3$
BMM
Thm. 5.10: ≤
D1R
PATH
NL-complete

Questions?

References I

[ABW15] Amir Abboud, Arturs Backurs, and Virginia Vassilevska Williams. "If the Current Clique Algorithms Are Optimal, So is Valiant's Parser". In: *Proceedings of the 2015 IEEE 56th Annual Symposium on Foundations of Computer Science (FOCS)*. FOCS '15. Washington, DC, USA: IEEE Computer Society, 2015, pp. 98–117.

[Alu+05] Rajeev Alur, Michael Benedikt, Kousha Etessami, Patrice Godefroid, Thomas Reps, and Mihalis Yannakakis. "Analysis of Recursive State Machines". In: *ACM Trans. Program. Lang. Syst.* 27.4 (July 2005), pp. 786–818.

[Cha08] Swarat Chaudhuri. "Subcubic Algorithms for Recursive State Machines". In: *Proceedings of the 35th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages*. POPL '08. San Francisco, California, USA: ACM, 2008, pp. 159–169.

References II

[Lee02] Lillian Lee. "Fast Context-free Grammar Parsing Requires Fast Boolean Matrix Multiplication". In: *J. ACM* 49.1 (Jan. 2002), pp. 1–15.

[Ryt95] Wojciech Rytter. "Context-free recognition via shortest paths computation: a version of Valiant's algorithm". In: *Theoretical Computer Science* 143.2 (1995), pp. 343–352.

[Val75] Leslie G. Valiant. "General context-free recognition in less than cubic time". In: *Journal of Computer and System Sciences* 10.2 (1975), pp. 308–315.