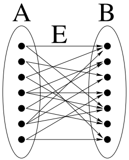
Descriptive complexity of translating an NFA into an ϵ-free NFA

Balázs Szörényi

Joint work with Judit Nagy-György, Szabolcs Iván and György Turán

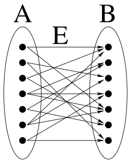
The problem of interest—basic notions

Given: a directed bipartite graph $G = (A, B, E)$, with all edges directed from A to B:

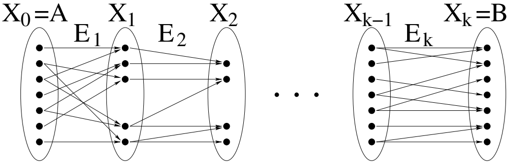


The problem of interest—basic notions

Given: a directed bipartite graph $G = (A, B, E)$, with all edges directed from A to B:



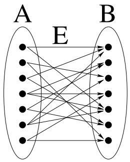
A *k-width representation* of G with is a G':



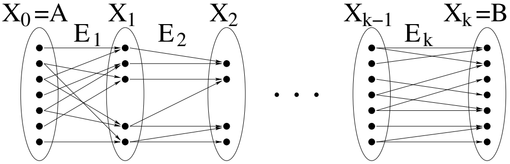
s.t. $\exists$ directed path $a \rightsquigarrow b$ in G' iff $(a, b) \in E$

The problem of interest—basic notions

Given: a directed bipartite graph $G = (A, B, E)$, with all edges directed from A to B:



A k-width representation of G with is a G':



s.t. $\exists$ directed path $a \rightsquigarrow b$ in G' iff $(a, b) \in E$

size $=$ number of edges
thus $\mathrm{size}(G) = |E|$ and $\mathrm{size}(G') = |E_1| + \cdots + |E_k|$

Assume $m = |A| = |B|$ throughout for simplicity

The problem of interest

Is there a large graph with small representation?
Yes: the complete bipartite graph $K_{m,m}$ has $\mathrm{size}(K_{m,m}) = m^2$, and has a trivial 2-width representation of size $2m$

The problem of interest

Is there a large graph with small representation?
Yes: the complete bipartite graph $K_{m,m}$ has $\mathrm{size}(K_{m,m}) = m^2$, and has a trivial 2-width representation of size $2m$

$k\text{-width-size}(G)$: size of the smallest k-width representation of G
$\text{unbounded-width-size}(G) = \min\{k\text{-width-size}(G) \ : \ k = 2, 3, \ldots\}$

The problem of interest

Is there a large graph with small representation?
Yes: the complete bipartite graph $K_{m,m}$ has $\mathrm{size}(K_{m,m}) = m^2$, and has a trivial 2-width representation of size $2m$

$k\text{-width-size}(G)$: size of the smallest k-width representation of G
$\text{unbounded-width-size}(G) = \min\{k\text{-width-size}(G) \ : \ k = 2, 3, \ldots\}$

Problem 1: determine the largest possible gap between the 2-width-size and the unbounded-width-size

Note: $\mathrm{size}(G) \leq k\text{-width-size}(G)^2$
(because $\mathrm{size}(G) \leq m^2$ and $\forall k \ m \leq k\text{-width-size}(G)$)

The problem of interest

Is there a large graph with small representation?
Yes: the complete bipartite graph $K_{m,m}$ has $\mathrm{size}(K_{m,m}) = m^2$, and has a trivial 2-width representation of size $2m$

$k\text{-width-size}(G)$: size of the smallest k-width representation of G
$\text{unbounded-width-size}(G) = \min\{k\text{-width-size}(G) \ : \ k = 2, 3, \ldots\}$

Problem 1: determine the largest possible gap between the 2-width-size and the unbounded-width-size

Note: $\mathrm{size}(G) \leq k\text{-width-size}(G)^2$
(because $\mathrm{size}(G) \leq m^2$ and $\forall k \ m \leq k\text{-width-size}(G)$)

Theorem (main result): $\exists$ arbitrarily large G with
$2\text{-width-size}(G) = \Omega(m^{3/2})$ and $(\log m)\text{-width-size}(G) = O(m \log m)$

Motivation 1: rectifier networks

Boolean circuits containing only OR gates. Thus

- A correspond to the input gates
- B correspond to the output gates
- the inner layers correspond to the OR gates

Results are typically bounds (even) for the unbounded width size:

Motivation 1: rectifier networks

Boolean circuits containing only OR gates. Thus

- A correspond to the input gates
- B correspond to the output gates
- the inner layers correspond to the OR gates

Results are typically bounds (even) for the unbounded width size:

- Lupanov, 1956: $O(n^2/\log n)$ are enough to compute any $f : \{0,1\}^n \to \{0,1\}^n$
- Nechiporuk, 1969: example (projective plane) with $\Omega(n^{3/2})$ gates
- Pippenger, 1980 (based on Brown, 1966): example with $\Omega(n^{5/3})$ gates
- Melhorn, 1979 and Wegener, 1980: similar bounds
- Jukna (based on Kollár et al. 1996): $O(n^{2-\epsilon})$ for arbitrarily small ϵ

Motivation 1: rectifier networks

Boolean circuits containing only OR gates. Thus

- A correspond to the input gates
- B correspond to the output gates
- the inner layers correspond to the OR gates

Results are typically bounds (even) for the unbounded width size:

- Lupanov, 1956: $O(n^2/\log n)$ are enough to compute any $f : \{0,1\}^n \to \{0,1\}^n$
- Nechiporuk, 1969: example (projective plane) with $\Omega(n^{3/2})$ gates
- Pippenger, 1980 (based on Brown, 1966): example with $\Omega(n^{5/3})$ gates
- Melhorn, 1979 and Wegener, 1980: similar bounds
- Jukna (based on Kollár et al. 1996): $O(n^{2-\epsilon})$ for arbitrarily small ϵ

$\Rightarrow$ do not say anything about our problem

Motivation 2: formal languages and NFA

alphabet: some finite set Σ
word: sequence of the form $a_1 a_2 \dots a_k$ s.t. $a_i \in \Sigma$ and $k \in \mathbb{N}$
language: a set L of words

nondeterministic finite automaton (NFA) M for generating the words in some (restricted) language $L(M)$: $M = (V, R, e, \triangleright, F)$, where
V is the set of states
$R \subseteq V^2$ is the set of transitions
e: $R \to \Sigma \cup \{\epsilon\}$ is a labeling of the transitions (ϵ: "empty transition")
$\triangleright \in V$ is the initial state
$F \subseteq V$ is the set of finite states (notation: finite states will be boxed)

Motivation 2: formal languages and NFA

alphabet: some finite set Σ
word: sequence of the form $a_1 a_2 \ldots a_k$ s.t. $a_i \in \Sigma$ and $k \in \mathbb{N}$
language: a set L of words

nondeterministic finite automaton (NFA) M for generating the words in some (restricted) language $L(M)$: $M = (V, R, e, \triangleright, F)$, where
V is the set of states
$R \subseteq V^2$ is the set of transitions
e: $R \to \Sigma \cup \{\epsilon\}$ is a labeling of the transitions (ϵ: "empty transition")
$\triangleright \in V$ is the initial state
$F \subseteq V$ is the set of finite states (notation: finite states will be boxed)

example 1: $\Sigma = \{a, b\}$, $F = \{x_2\}$, $L(M) = \{ab\}$

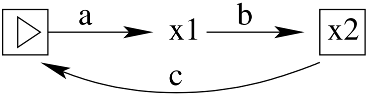
$$\triangleright \xrightarrow{\text{a}} \text{x1} \xrightarrow{\text{b}} \boxed{\text{x2}}$$

Motivation 2: formal languages and NFA

alphabet: some finite set Σ
word: sequence of the form $a_1 a_2 \ldots a_k$ s.t. $a_i \in \Sigma$ and $k \in \mathbb{N}$
language: a set L of words

nondeterministic finite automaton (NFA) M for generating the words in some (restricted) language $L(M)$: $M = (V, R, e, \triangleright, F)$, where
V is the set of states
$R \subseteq V^2$ is the set of transitions
e: $R \to \Sigma \cup \{\epsilon\}$ is a labeling of the transitions (ϵ: "empty transition")
$\triangleright \in V$ is the initial state
$F \subseteq V$ is the set of finite states (notation: finite states will be boxed)

example 2: $\Sigma = \{a, b, c\}$, $F = \{x_2, \triangleright\}$, $L(M) = \{\lambda, ab, abc, abcab, \ldots\}$

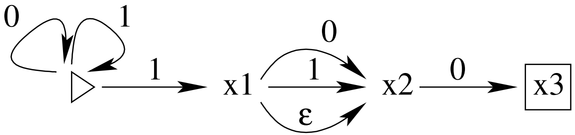


Motivation 2: formal languages and NFA

alphabet: some finite set Σ
word: sequence of the form $a_1 a_2 \dots a_k$ s.t. $a_i \in \Sigma$ and $k \in \mathbb{N}$
language: a set L of words

nondeterministic finite automaton (NFA) M for generating the words in some (restricted) language $L(M)$: $M = (V, R, e, \triangleright, F)$, where
V is the set of states
$R \subseteq V^2$ is the set of transitions
e: $R \to \Sigma \cup \{\epsilon\}$ is a labeling of the transitions (ϵ: "empty transition")
$\triangleright \in V$ is the initial state
$F \subseteq V$ is the set of finite states (notation: finite states will be boxed)

example 3: $\Sigma = \{0, 1\}$, $F = \{x_3\}$, $L =$ words ending with 10, 100 or 110

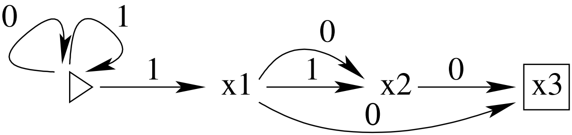


Motivation 2: formal languages and NFA

alphabet: some finite set Σ
word: sequence of the form $a_1 a_2 \dots a_k$ s.t. $a_i \in \Sigma$ and $k \in \mathbb{N}$
language: a set L of words

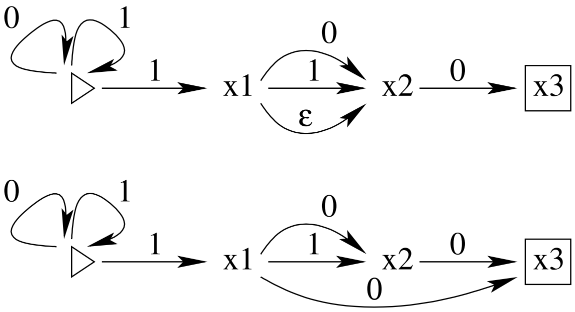
nondeterministic finite automaton (NFA) M for generating the words in some (restricted) language $L(M)$: $M = (V, R, e, \triangleright, F)$, where
V is the set of states
$R \subseteq V^2$ is the set of transitions
e: $R \to \Sigma \cup \{\epsilon\}$ is a labeling of the transitions (ϵ: "empty transition")
$\triangleright \in V$ is the initial state
$F \subseteq V$ is the set of finite states (notation: finite states will be boxed)

example 3': $\Sigma = \{0, 1\}$, $F = \{x_3\}$, $L =$ words ending with 10, 100 or 110



Motivation 2: formal languages and NFA

Esample 3 and example 3' generate the same language:



but the latter has no ϵ-transition: it is ϵ-free!

Problem 2: What is the largest blow-up when translating an NFA into an ϵ-free NFA?

Motivation 2: formal languages and NFA

Similar questions considered in the literature:

translating from	to	the blow-up is	result by
regular expression	ϵ-free NFA	$O(n \log^2(n))$	Hromkovic et al, 1997
regular expression	ϵ-free NFA	$\Omega\left(\frac{n \log^2(n)}{\log \log n}\right)$	Lifshits, 2003
regular expression	ϵ-free NFA	$\Omega(nlog^2 n)$	Schnitger, 2006
CFG	chain-rule-free CFG	$\Omega(n \log \log n)$	Blum, 1982
CFG	chain-rule-free CFG	$\Omega(n^{3/2-\epsilon})$	Blum, 1983
CFG	chain-rule-free CFG	$O(n^2)$	folclore
NFA	ϵ-free NFA	$O(n^2)$	folclore
NFA	ϵ-free NFA	$\Omega(nlog^2 n)$	Schnitger, 2006
chain-rule-free CFG	Chomsky-for CFG	$\Theta(n)$	Folclore
chain-rule-free CFG	Greibach-form CFG	$O(n3)$	Blum et al, 1997
CFG	Greibach-form CFG	$O(n^4)$	Blum et al, 1997
CFG	Greibach-form CFG	$\Omega(n^2)$	Kelemenova, 1984

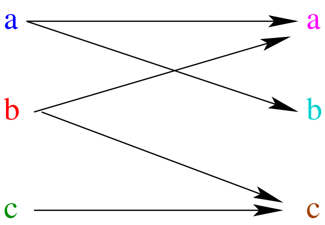
Importance: programming and script languages are typically context free or regular languages

Motivation 2: language containing only words of length 2

Σ: some alphabet
L: some language over Σ containing only words of length 2

Graph representation:
the bipartite graph $(\Sigma, \Sigma, \vec{E}_L)$ s.t. $(a, b) \in \vec{E}_L$ iff $ab \in L$

Example: $L = \{aa, ab, ba, bc, cc\}$

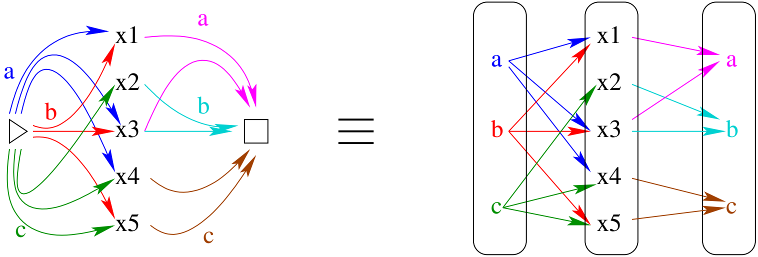


Motivation 2: language containing only words of length 2

Σ: some alphabet
L: some language over Σ containing only words of length 2

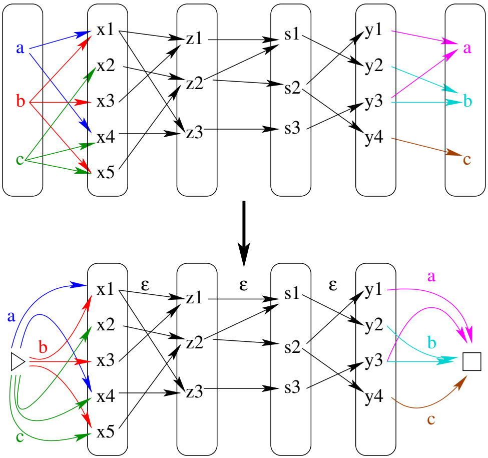
Graph representation:
the bipartite graph $(\Sigma, \Sigma, \vec{E}_L)$ s.t. $(a, b) \in \vec{E}_L$ iff $ab \in L$

Conversion between ϵ-free NFA and width-2 representation:



Motivation 2: language containing only words of length 2

Converting a width-k representation to an NFA:



Motivation 2: formal languages and NFA

Thus a lower bound for Problem 1 translates to a lower bound for Problem 2

Proof outline for the main result

Let q be some prime power, and d a positive integer.

The graph $G = (A, B, E)$ is constructed as follows.
A and B are two distinct copies of $(\mathbb{Z}_q)^d$ and
$E = \{(a, b) : \langle a, b \rangle = 0\}$, where $\langle a, b \rangle = \sum_{i=1}^{d} a_i b_i$

Lemma: $d\text{-width-size}(G) = O(d \cdot q^{d+2})$

Lemma: $2\text{-width-size}(G) = \Omega(q^{3d/2})$