10 Years DCFS

Jürgen Dassow

Otto-von-Guericke-Universität Magdeburg
Fakultät für Informatik

Workshop on Descriptional Complexity of Formal Systems,
Magdeburg, July 6 – 9, 2009

What is DCFS?

DCFS = Descriptional Complexity of Formal Systems

What is DCFS?

DCFS = Descriptional Complexity of Formal Systems

DCFS = Daimler Chrysler Finance Service

What is DCFS?

DCFS = Descriptional Complexity of Formal Systems

DCFS = Daimler Chrysler Finance Service

DCFS = Department for Children and Family Service

Some History I

Formal Descriptions and Software Reliability

Paderborn	(Germany)	November 3, 1998
Boca Raton	(USA)	October 31, 1999
San Jose	(USA)	October 7, 2000

Descriptional Complexity of Automata, Grammars and Related Structures

Magdeburg	(Germany)	July 20 – 23, 1999
London	(Canada)	July 27 – 29, 2000
Vienna	(Austria)	July 20 – 22, 2001

Some History II

Descriptional Complexity of Formal Systems

London	(Canada)	August 21 – 24, 2002
Budapest	(Hungary)	July 12 – 14, 2003
London	(Canada)	July 26 – 28, 2004
Como	(Italy)	June 30 – July 2, 2005
Las Cruces	(USA)	June 21 – 23, 2006
Novy Smokovec	(Slovakia)	July 20 – 22, 2007
Charlottetown	(Canada)	July 16 – 18, 2008
Magdeburg	(Germany)	July 6 – 9, 2009

Some Statistics: Paper Distribution to Workshops

year	# paper
1999	20
2000	21
2001	19
2002	23
2003	26
2004	23
2005	30
2006	26
2007	16
2008	21
2009	20
	245

192 authors from 27 countries

Some Statistics: Paper Distribution to Authors
Share of Papers

1.	Martin Kutrib (Germany)	6.16
2.	Andreas Malcher (Germany)	5.83
3.	Alexander Okhotin (Russia)	5.09
4.	Markus Holzer (Germany)	4.91
5.	Kai Salomaa (Canada)	4.16
6.	Henning Bordihn (Germany)	3,91
7.	Michael Domaratzki (Canada)	3.75
	Jeffrey Shallit (Canada)	3.75
	György Vaszil (Hungary)	3.75
10.	Bettina Sunckel (Germany)	3.50

Some Statistics: Paper Distribution to Authors
Number of Papers

1.	Martin Kutrib (Germany)	11
	Kai Salomaa (Canada)	11
3.	Markus Holzer (Germany)	10
4.	Henning Bordihn (Germany)	8
	Michael Domaratzki (Canada)	8
6.	Rudolf Freund (Austria)	7
	Andreas Malcher (Germany)	7
	Alexander Okhotin (Russia)	7
	Jeffrey Shallit (Canada)	7
	Sheng Yu (Canada)	7

Some Statistics: Paper Distribution to Countries

1.	Germany	52,00
2.	Canada	26.75
3.	Czech Republic	14.50
4.	Romania	14.00
5.	India	13.59
6.	France	12.33
7.	Italy	11.16
8.	Slovakia	10.00
9.	USA	9.50
10.	Hungary	9.83

Descriptional Complexity and Transformations
DFA to NFA I

Theorem: Let $\mathcal{A}$ be a minimal nondeterministic finite automaton with n states and $\mathcal{B}$ a deterministic finite automaton $\mathcal{B}$ with $T(\mathcal{A}) = T(\mathcal{B})$. If $\mathcal{B}$ has d states, then $n \leq d \leq 2^n$.

Theorem: For any n, there are minimal nondeterministic finite automata $\mathcal{A}$ and $\mathcal{A}'$ with n states such that the minimal deterministic automata $\mathcal{B}$ and $\mathcal{B}'$ with $T(\mathcal{A}) = T(\mathcal{B})$ and $T(\mathcal{A}') = T(\mathcal{B}')$ have n and 2^n states, respectively.

$$B(n,k) = \{d \mid \text{there is a minimal NFA } \mathcal{A} \text{ with } n \text{ states over a } k\text{-letter alphabet such that the minimal DFA } \mathcal{B} \text{ with } T(\mathcal{A}) = T(\mathcal{B}) \text{ has } d \text{ states}\}$$

Descriptional Complexity and Transformations DFA to NFA II

Theorem: (Iwama/Kombayashi/Takaki, 2000)
$\{2^n - 2^k \mid 0 \leq k \leq \frac{n}{2} - 2\} \cup \{2^n - 2^k - 1 \mid 0 \leq k \leq \frac{n}{2} - 2\} \subseteq B(n, 2)$.

Theorem: (Iwama/Matsuura/Paterson, 2003)
$\{2^n - k \mid 5 \leq k \leq 2n - 2 \wedge$ some coprimality condition for $k\} \subseteq B(n, 2)$.

Theorem: (Jirásková, 2001, 2006,
Jirásek/Jirásková/Szabari, 2007)
For $k \geq 4$, $B(n, k) = \{d \mid n \leq d \leq 2^n\}$.

Descriptional Complexity and Transformations DFA to NFA III

Theorem: (Geffert, DCFS 2005)
For sufficiently large n and any arbitrarily small real $\varepsilon > 0$,

$$\{d \mid n < d \leq e^t,\ t = \frac{1-\varepsilon}{\sqrt[3]{9}}\sqrt[3]{n^4}\sqrt[3]{(\ln(n))^2}\} \subseteq B(n,2).$$

Theorem: (Matsuura/Saito, DCFS 2008)

i) For any $n \geq 11$ and any α with $28 \leq \alpha \leq 3n-3$, $\lfloor \frac{\alpha-1}{3} \rfloor$ is odd and relatively prime with n, $2^n - \alpha \in B(n,2)$.
ii) For any $n \geq 11$ and any α with $37 \leq \alpha \leq 4n-5$, α is not divisible by 4 and $\lfloor \frac{\alpha-1}{4} \rfloor$ is relatively prime with n, $2^n - \alpha \in B(n,2)$.

Descriptional Complexity and Transformations DFA to NFA III

Theorem: (Geffert, DCFS 2005)
For sufficiently large n and any arbitrarily small real $\varepsilon > 0$,

$$\{d \mid n < d \leq e^t,\ t = \frac{1-\varepsilon}{\sqrt[3]{9}} \sqrt[3]{n^4} \sqrt[3]{(\ln(n))^2}\} \subseteq B(n,2).$$

Theorem: (Matsuura/Saito, DCFS 2008)

i) For any $n \geq 11$ and any α with $28 \leq \alpha \leq 3n-3$, $\lfloor \frac{\alpha-1}{3} \rfloor$ is odd and relatively prime with n, $2^n - \alpha \in B(n,2)$.

ii) For any $n \geq 11$ and any α with $37 \leq \alpha \leq 4n-5$, α is not divisible by 4 and $\lfloor \frac{\alpha-1}{4} \rfloor$ is relatively prime with n, $2^n - \alpha \in B(n,2)$.

Open Problem: Determination of $B(n,2)$ and $B(n,3)$ and $B(n,1)$.

Descriptional Complexity and Transformations

topic	author(s)	year
from reg. canonical systems to DFA	Petersen	1999
from k-entry DFA to DFA	Kappes	1999
	Holzer/Salomaa/Yu	2000
	Polák	2005
DFA vs. NFA vs.	Mereghetti/Palano/Pighizzini	1999
probabilisitic FA vs. quantum FA	Milano/Pighizzini	2000
NFA/DFA vs. regular expression	Ellul/Shallit/Wang	2002
from k-head DFA to DFA	Kutrib	2003
from $k+1$-head DFA to k-head DFA	Kapoutsis	2004
from NFA to cover DFA	Câmpeanu/A. Păun	2004
subregular languages	Bordihn/Holzer/Kutrib	2008
from pebble 2-way to 2-way	Geffert/Ištoňová	2009

Descriptional Complexity and Operations

Let L be a regular language. Then we define $s(L)$ as the number of states of an minimal deterministic automaton accepting L.

Let τ be an m-ary operation on languages which preserves regularity.

We define $s_\tau(n_1, n_2, \ldots, n_m)$ as the maximal number d such that there are languages $L_1, L_2, \ldots, L_m$ with $s(L_i) = n_i$ for $1 \leq i \leq m$ and $s(\tau(L_1, L_2, \ldots, L_m)) = d$.

Descriptional Complexity and Operations
Determination/Estimation of s_τ

operation	author(s)	year
union, intersection, complement,	Yu	1999
reversal, catenation, star	Jirásková	2003
	Mera/Pighizzini	2003
proportional removals	Domaratzki	2001
decimations	Shallit	2006
shuffle	Domaratzki/Salomaa	2002
	Biegler/Daley/McQuillan	2009
cyclic shift	Jirásková/Okhotin	2005
combined operations	Gao/Salomaa/Yu	2006
	Esik/Gao/Liu/Yu	2008
orthogonal catenation	Daley/Domaratzki/Salomaa	2008

Descriptional Complexity and Operations
Problem 1: Range instead of Maximum

$$B_\tau(n_1, n_2, \ldots, n_m) = \{d \mid \text{there are languages } L_1, L_2, \ldots, L_m \text{ with } s(L_i) = n_i \text{ for } 1 \le i \le m \text{ and } s(\tau(L_1, L_2, \ldots, L_m)) = d\}$$

Trivial fact: $B_{complement}(n) = \{n\}$.

Descriptional Complexity and Operations
Problem 1: Range instead of Maximum

$$B_\tau(n_1, n_2, \ldots, n_m) = \{d \mid \text{there are languages } L_1, L_2, \ldots, L_m \text{ with } s(L_i) = n_i \text{ for } 1 \leq i \leq m \text{ and } s(\tau(L_1, L_2, \ldots, L_m)) = d\}$$

Trivial fact: $B_{complement}(n) = \{n\}$.

Theorem: (Hricko/Jirásková/Szabari, DCFS 2005)
i) $B_\cup(n, m) = \{d \mid 1 \leq d \leq s_\cup(n, m)\}$.
ii) $B_\cap(n, m) = \{d \mid 1 \leq d \leq s_\cap(n, m)\}$.

Descriptional Complexity and Operations
Problem 2: Average instead of Worst Case

s_τ gives the worst case complexity of languages obtained by τ

Problem: Investigate the average complexity

$$\overline{s}(n_1, n_2, \ldots, n_m) = \frac{1}{\#(U(n_1, \ldots, n_m))} \sum_{L \in U(n_1, \ldots, n_m)} s(L)$$

where $U(n_1, n_2, \ldots, n_m)$ is the set of languages $\tau(L_1, \ldots, L_m)$ with $n_i = s(L_i)$ for $1 \leq i \leq m$ (and L_i over a fixed alphabet).

Note: There is only a very small knowledge on $U(n_1, \ldots, n_m)$ or $\#(U(n_1, \ldots, n_m))$ for alphabets with at least two letters.

Descriptional Complexity and Operations
Problem 2: Average instead of Worst Case

For the case of a unary alphabet we have the following results:

Theorem: (Nicaud 1999)

i) $\overline{s}_*(n) \leq C_1$ for some constant C_1.

ii) If $n < m < P(n)$ for some polynomial P, then $\overline{s}_\cdot(n, m) \leq C_2$ for some constant C_2.

iii) $\overline{s}_\cap(n, m) = \frac{3\zeta(3)}{2\pi^2} nm$.

Let $\frac{1}{2}L = \{x \mid \text{there is a } y \text{ with } |x| = |y| \text{ and } xy \in L\}$.

Theorem: (Domaratzki, DCFS 2003)

$\overline{s}_{\frac{1}{2}L}(n) = (\frac{5}{8}n + c)(1 + \lambda(n))$ for some constant c and some function λ which tends to 0 if n tends to infinity.

Results on $\overline{s}_\tau$ for some τ and $nondeterministic$ state complexity by Gruber/Holzer.

Descriptional Complexity and Operations
Problem 3: Special Cases

topic	author(s)	year
finite languages	Yu	1999
prefix-free languages	Han/Salomaa/Wood	2006

operation	general case	finite	prefix-free
star	$3 \cdot 2^{n-2}$	$3 \cdot 2^{n-4}$	n
catenation	$(2n-1)2^{m-1}$	$O(mn^t + n^t)$	$m+n+2$
intersection	nm	$O(nm)$	$mn - 2(n+m) + 6$

Descriptional Complexity and Operations
Problem 4: Other Complexity Measures

complexity measure	author(s)	year
transition complexity of NFA	Domaratzki/Salomaa	2006
state compl. of alternating FA (very restricted operations)	Kavitha/Jeganathan/Sethuraman	2006
size of regular expressions		
unary alphabet	Ellul/Shallit/Wang	2002
quotient, shift	Gruber/Holzer	2008
quotient complexity	Brzozowski	2009

Descriptional Complexity of Devices and Languages

Let $\mathcal{G}$ be a class of devices (automata, grammars, etc.) and $\mathcal{L}$ the set of languages generated/accepted/etc. by element of $\mathcal{G}$. Let $\kappa : \mathcal{G} \to \mathbb{N}$ be a parameter measuring the "size" of elements of $\mathcal{G}$.

For $L \in \mathcal{L}$, let $\kappa_{\mathcal{G}}(L) = \min\{\kappa(G) \mid G \in \mathcal{G},\ L(G) = L\}$.

$\mathcal{L}_\kappa(n) = \{L \mid \kappa(L) \leq n,\ L \in \mathcal{L}\}$

Question:
Does κ induce an infinite hierarchy (i.e., for any $n \geq 1$, there are n_1 and n_2 such that $n \leq n_1 < n_2$ and $\mathcal{L}_\kappa(n_1) \subset \mathcal{L}_\kappa(n_2)$)
or
does there exist a number m such that $\kappa(L) \leq m$ for all $L \in \mathcal{L}$ (especially, what is the minimal m)?

Descriptional Complexity of Devices and Languages
Upper Bounds I

parameter	author(s)	year
diameter of splicing systems	A. Păun	1999
number of components	Csuhaj-Varjú	1999
in grammar systems	Vaszil	2005
	Csuhaj-Varjú/Vaszil	2009
number of active symbols in	Bordihn/Holzer	2000
CD grammar systems/L systems	Dassow	2004
	Bordihn/Sunckel	2006
number of non-context-free rules/	Meduna	2001
non-context-free nonterminals	Vaszil	2003
in some controlled grammars	Masopust/Meduna	2007
	Masopust/Meduna	2009

Descriptional Complexity of Devices and Languages
Upper Bounds II

parameter	author(s)	year
number of nodes	Mitrana	2003
in evolutionary systems	Alhazov et al.	2008
	Loos/Manea/Mitrana	2009
number of membranes/catalysts	Gh. Păun	2003
in membrane systems	Freund/Oswald/Sosík	2003
	Freund/Oswald/A. Păun	2004
weight in insertion systems	Kari/Sosík	2006

Descriptional Complexity of Devices and Languages
Nonterminal Complexity of Controlled Grammars

Theorem: (Fernau/Freund/Oswald/Reinhardt, DCFS 2005)

$$nt_{GC}(2) = nt_{Mat}(3) = \mathcal{L}(RE).$$

Theorem: (Fernau/Freund/Oswald/Reinhardt, DCFS 2005)
Any graph controlled grammar with one nonterminal generates a recursive language.
Corollary: $nt_{GC}(1) \subset nt_{GC}(2)$.

Theorem: (Dassow/Gh. Păun, 1985)
There is a language L which cannot be generated by a matrix grammar with two nonterminals.
Corollary: $nt_{Mat}(2) \subset nt_{Mat}(3)$.

Descriptional Complexity of Devices and Languages
Insertion/Deletion Systems I

Insertion/deletion system $G = (V, T, A, D, I)$ of degree (n, m)
with alphabet V, $T \subseteq V$, A, D, I finite subsets of V^*
and $|x| \leq n$ for all $x \in D$ and $|y| \leq m$ for all $y \in I$

$u \Longrightarrow v$ if and only if there are $u_1, u_2 \in V^*$, $x \in D$ or $y \in I$ such that

$$u = u_1 x u_2, v = u_1 u_2 \quad \text{or} \quad u = u_1 u_2, v = u_1 y u_2$$

$$L(G) = \{w \mid z \Longrightarrow^* w \text{ for some } z \in A\}$$

Descriptional Complexity of Devices and Languages
Insertion/Deletion Systems II

Theorem: (Margenstern/Gh. Păun/Rogozhin/Verlan, 2005)
Any recursively enumerable language can be generated by an insertion/deletion system of degree $(3, 3)$.

Theorem: (Verlan, DCFS 2005)

i) A insertion/deletion system of degree $(2, 2)$ generates a context-free language.
ii) For $n \geq 1$, an insertion/deletion system of degree $(n, 1)$ generates a context-free language.
iii) For $m \geq 1$, an insertion/deletion system of degree $(1, m)$ generates a regular language.

Topics which I Miss(ed)

1. Kolmogorov complexity
   5 invited lecture, at most 2 contributions

2. complexity of Boolean functions described by circuits, branching programs, etc.
   only 2 papers on descriptions of Boolean functions

3. descriptional complexity of algorithms

4. "good" (not "small") descriptions for some purpose
   1 invited lecture