When Canadian Theoretical Computer Science was Born, a Personal Perspective

Andrew L. Szilard
Professor Emeritus
Department of Computer Science
Western University
London, Ontario, Canada

als@csd.uwo.ca

Invited Lecture to Honour Professor Janusz Brzozowski
on the occasion of his 80th birthday
University of Waterloo, Canada
June 24, 2015,

When Canadian Theoretical Computer Science was Born, a Personal Perspective

Abstract:

The sixties are remembered for many important revolutionary events in the world. From Mathematics and Electrical Engineering, Computer Science was born then.
Expensive mainframe computers started to be installed at some giant financial institutions and at large universities. Machine and programming languages were created to communicate algorithms, and a large number of analysts and programmers were needed to design and to code applications for the giant computers. Undergraduate and graduate programs were conceived to teach computer languages, file and data structures and advanced computer concepts. Good teachable undergraduate texts were nonexistent and the fundamentals of Computer Science were lacking. The first Canadian Theoretical Computer Science Summer School was held in Toronto, the main speakers were two young researchers Janusz Brzozowski and Arto Salomaa. This talk is an attempt to recreate some of Prof. Brzozowski's brilliant inspirational seminal lectures that became fundamental to the teaching of Computer Science Theory in Canada and beyond.

The 60's were defined & remembered by revolutionary events

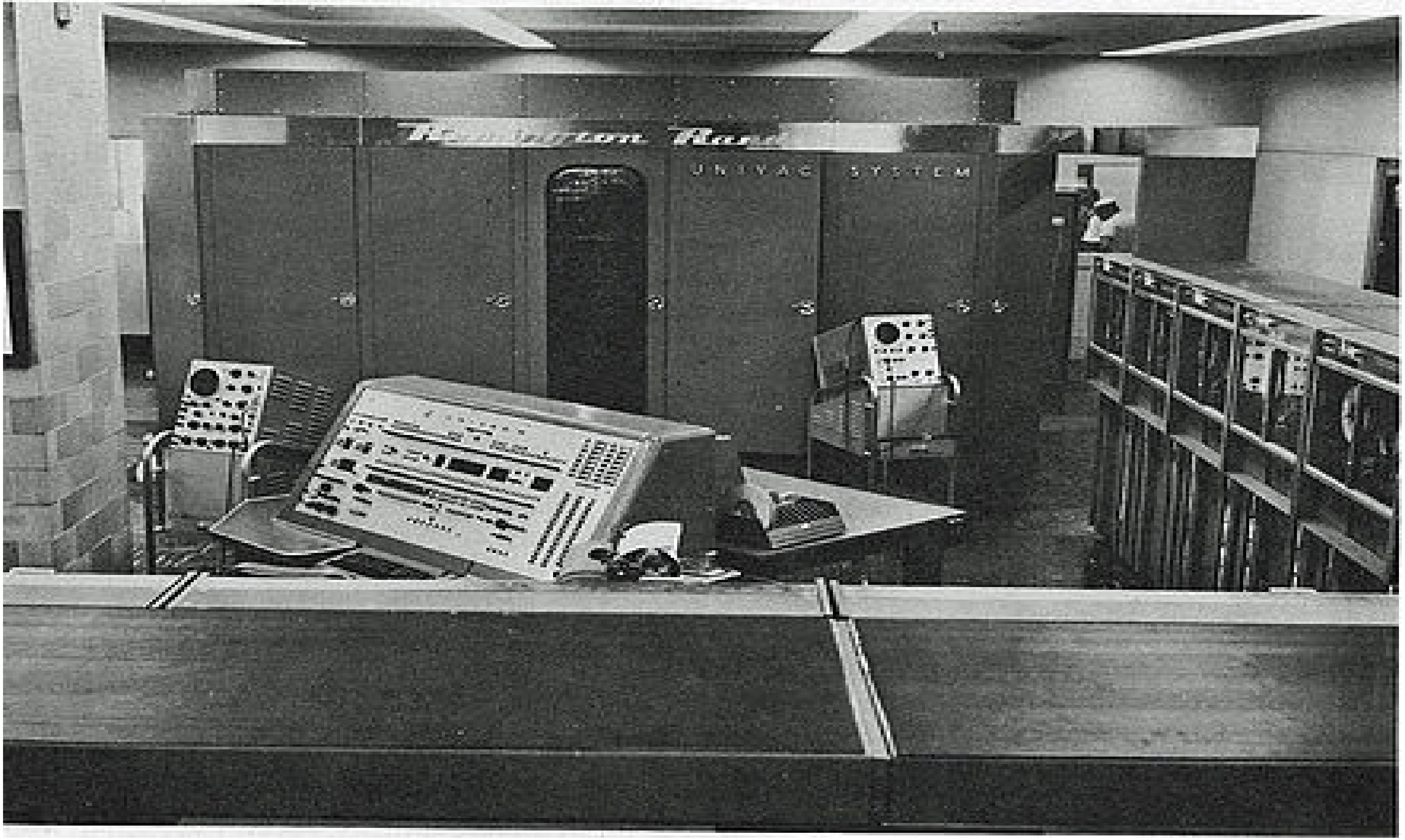
1. the introduction of commercial main-frame computing and the birth of Computer Science courses.
2. The civil rights movement in the U.S.
3. The arrival of the space age.
4. Cold war crises: Berlin Wall, Missiles in Cuba
5. Assassinations of political figures in the U.S.
6. Advances in office and computer technology
7. The Vietnam war and protests
8. Historical events in Canada
9. Music & festivals of the 60's

Some historical events that defined the sixties

1960, Remington Rand Corp installs commercial Univac II computers.
January 1960, At a meeting of the European language design group in Paris, Peter Naur presents the Algol 60 report, a formal language definition of Algol.
Febr. 1, 1960, 4 black University students sat down at the lunch counter in a Woolworth store in Greensboro, North Carolina, the first sit-in civil-right demonstration.
July 1960, IBM announces 7070/7074 computers with discrete transistors, no vacuum tubes.
Nov. 8, 1960, John F. Kennedy elected to be President of U.S.
1961, U.S. The FDA-approved oral contraceptive is made available.
April 12, 1961, Yuri Gagarin becomes the first human being in space.
April 17, 1961, CIA-trained operatives invade Cuba at Bay of Pigs, (were defeated in 3 days).
May 25, 1961, John F. Kennedy gives his we-will-put-a-man-on-the-Moon speech.
July 31, 1961, IBM introduces the IBM Selectric Typewriter.
Aug 13, 1961, The erection of the Berlin Wall
May 1962, Janusz Brzozowski graduates from Princeton University, Princeton, NJ
Oct. 22, 1962, President Kennedy's TV address, the Cuban missile crisis, risk of nuclear war, The Doomsday clock was to be set to at 1 minute till midnight.
April 1963, IBM ships its 7040 computers future customers incl. UWO, U. of Waterloo
Jun 12, 1963, Civil Rights advocate Medgar Evers assassinated in Jackson, Mississipi.
July 11, 1963, Nelson Mandela arrested, jailed, served 27 years.
Nov. 22, 1963, John F. Kennedy assassinated.
Nov. 24, 1963 Jack Ruby fatally shoots Lee Harvey Oswald.

February	1964,	The Beatles arrive to NY, JFK airport, Beatlemania comes to the U.S.
April 7,	1964,	IBM announces its family of System/360 computers
August 2,	1964,	Gulf of Tonkin incident that leads to the Authorization of the Vietnam War
September,	1964,	Western U. offers an undergraduate program in Computer Science
	1965,	4 undergraduate students, at Univ. of Waterloo, write a 100 statements/sec load-and-go, in-core, Fortran interpreter for the IBM 7040/7044 computer.
October 15,	1965,	First draft card burned, David Miller, New York, arrested, 2 years in prison
July	1966,	Arto Salomaa becomes Visiting Professor at Western
August	1966,	First Canadian Summer School on Theoretical Computer Science, featuring J. Brzozowski and A. Salomaa
October	1966,	the Black Panther Party is founded.
April 27,	1967,	EXPO 67 opens in Montreal
June 5 – 10,	1967,	The Six-Day War, Israel survives
June 16-18,	1967,	Monterey International Pop Festival
	1968,	Mass-mailed Chargex credit cards are introduced in Canada
April 4,	1968,	Dr. Martin Luther King, Jr. assassinated.
April 20,	1968,	Pierre Elliott Trudeau becomes Prime Minister of Canada
May 6,	1968,	Students-led revolution in Paris
Aug. 20,21	1968,	The Soviet Army supported by other Eastern Block countries invade Prague The Prague Spring is buried
	1969	Arto Salomaa's book Theory of Automata is published by Pergamon Press. 1st advanced undergraduate text on regular expressions & finite automata.
July 16,	1969,	Apollo 11 launch, lands men on the Moon, Sea of Tranquility July 20; returns July 24, 1969
Aug. 15-18,	1969	Woodstock Music Festival

Remington Rand Corp installs Univac II computers
magnetic core memory: 2,000-10,000 words
$1,500,000 – 3,000,000



1960 Remington Rand Corp installs Univac IIs.

Metropolitan Life Insurance Co., NY
United States Steel, Pittsburgh
Sun Life Insurance Co. Montreal, Que.
Pacific Mutual Life Insurance Co., LA
London Life Insurance Co. London, Ont.

PROGRAMMING AND NUMERICAL SYSTEM

Internal number system	Binary coded decimal	**Decimal digits/word**	12
Decimal digits/instruction	6	**Instructions per word**	2
Instructions used	54	**Arithmetic system**	Fixed point
Instruction type	One address	**Number range**	Between -1 and +1

Decimal point occurs at the right of the sign digit.

ARITHMETIC UNIT

Including Store Access in Microsec	**Add**	160	**Mult**	1,720	**Div**	3,030

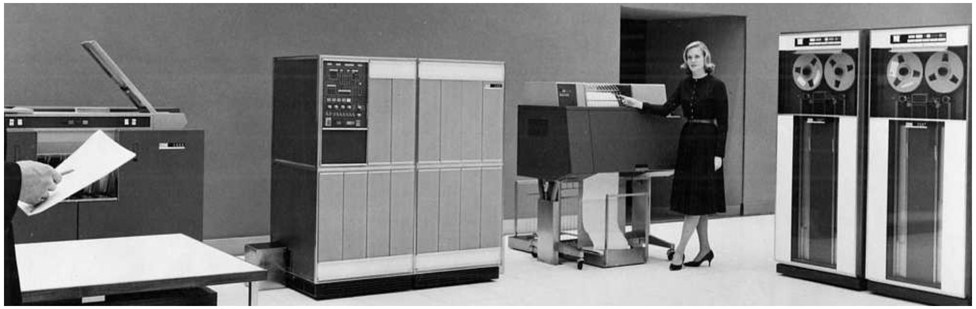
Construction	Vacuum tubes	**Arithmetic mode**	Serial
Timing	Synchronous	**Operation**	Sequential

COST, PRICE AND RENTAL RATES

	Monthly Rental 1 Shift	**Outright Sale Price**
Description	**5 Day Week**	**F.O.B. Factory**
Univac II Central Computer w/power supply & supervisory ctl desk	$18,540.00	$970,000
Uniservo II	450.00	20,000
Uniprinter	390.00	22,000
Extra Dolly for Uniprinter	122.50	7,000
Unityper II	90.00	4,500
High Speed Printer	3,300.00	185,000
Card-to-Tape Unit	2,520.00	142,100
Tape-to-Card Unit	2,300.00	130,000
Perforated Tape to Magnetic Tape Converter	1,800.00	108,000
Magnetic Tape to Perforated Tape Converter	1,500.00	90,000

The sixties   Business Computers

IBM 7070   50,000 'bytes'



IBM 1401   8,000 'bytes'

Programming Languages of the 60s

Fortran IV WATFOR

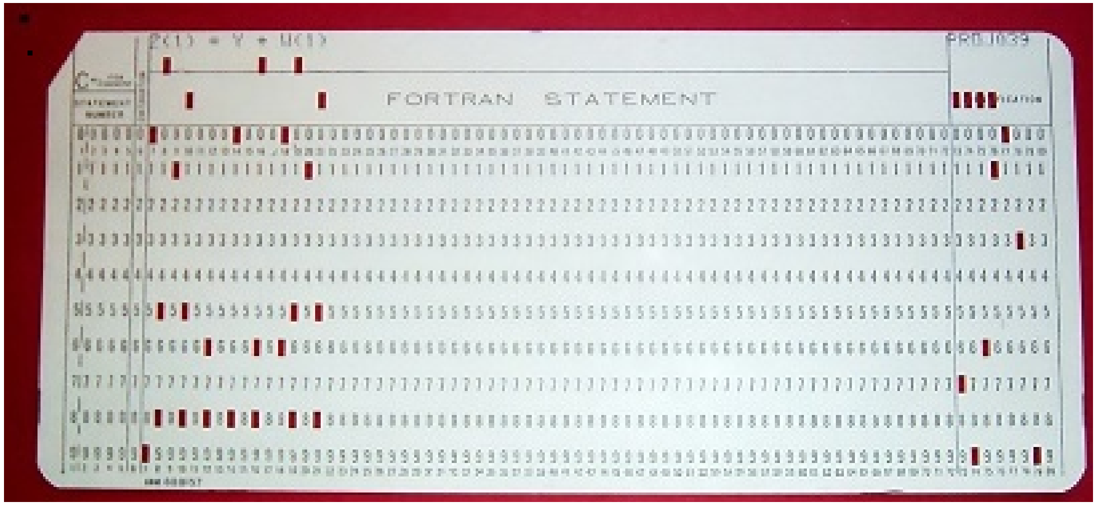
Algol 60 Simula

COBOL

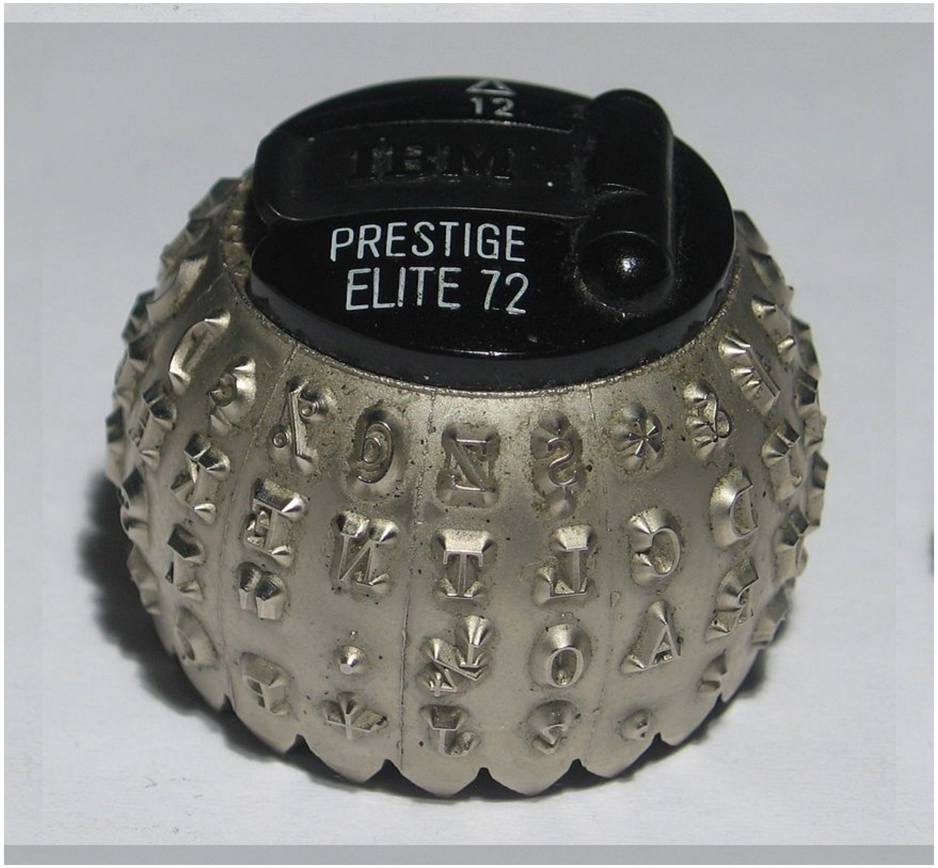
APL Lisp 1.5 LOGO

The IBM 7040 Scientific computer

16,384 36-bit words



IBM Selectric typewriter



Characters

A B C D E F G H I J K
L M N O P Q R S T U
V W X Y Z

a b c d e f g h i j k l m
n o p q r s t u v w x y z

0 1 2 3 4 5 6 7 8 9

@#! $ % & ( ): ; ' < > ?
+ - * / _ [ ] = " ° ¼ ½

A B C D E F G H I J K
L M N O P Q R S T U
V W X Y Z

a b c d e f g h i j k l m
n o p q r s t u v w x y z

0 1 2 3 4 5 6 7 8 9

@ # ! $ % & ( ): ; ' < > ?
+ - * / _ [ ] = " ° ¼ ½

Mathematical and composite characters

α β χ δ ε ϕ γ η ι φ κ
λ μ ν ο π θ ρ σ τ υ ϖ ζ

Α Β Χ Δ Ε Φ Γ Η Ι ϑ Κ
Λ Μ Ν Ο Π Θ Ρ Σ Τ ϒ
ς Ω Ξ Ψ Ζ

( ) [ ] { } ∂ ⊃ + − / | ∗
→

Over-strike, composition and underline tricks

$\underline{\leq}$ $\underline{\geq}$ $\underline{\equiv}$ $\underline{\pm}$

--> <> <-->

~~∀~~ ~~(~~ ~~θ~~ a^n A_n 0̸

$d(e^x)$
------ $= e^x$ $d(e^x)/dx = e^x$
dx

d(f(g(x))/dx = d(f(y)/dy)•d(g(x)/dx

d(f(x)•g(x)) = d(f(x))/dx•g(x) + f(x) •d(g(x))/dx

(f•g)' = f'•g + f•g'

The mimeograph machine

The print quality and inexpensive reproduction of print materials were limited.

For each page a “master” was first produced on a special wax-covered stencil by a ribonless typewriter.

The typewriter thus made impressions in the stencil, which were filled with ink and squeezed onto paper by the mimeograph’s roller. The stencils could also be marked with drawings made by hand.

n copies were made by rolling each master *n* times on pages fed to the machine whose purple, hallucinogenic, indelible ink was hated by our secretaries. The copies had to be collated by hand in the correct order and stapled to produce the *n* complete sets.

Lecture notes were produced through this painful way,
corrections were especially difficult.

Hand-outs were restricted mostly to exam papers and home-work assignments.

The Xerox revolution

Xerox corporation introduces affordable photocopying machines:

5-10 cents/page

Minimum wage in the U.S $1.00-1.60/hour

throughout the 60s

Journals for automata theory in the 60s

Information and Control

Communications of the ACM

Journal of the ACM

IBM Journal of Research & Development

Pacific Journal of Mathematics

Michigan Mathematical Journal

IEEE Transactions Electronic Computers

Authors on Automata Theory before 1965 publishing in English

J. Brzozowski
J. Büchi
N. Chomsky
L.C. Eggan
S. Ginsburg
J. Hartmanis
R. McNaughton
S.C. Kleene
M.O. Rabin
A. Salomaa
M.P. Schützenberger
D.S. Scott

Graduate Textbooks

on Automata Theory before 1965

Automata Studies C. Shannon & J. McCarthy

Theory of Self-Reproducing Automata
John Von Neumann

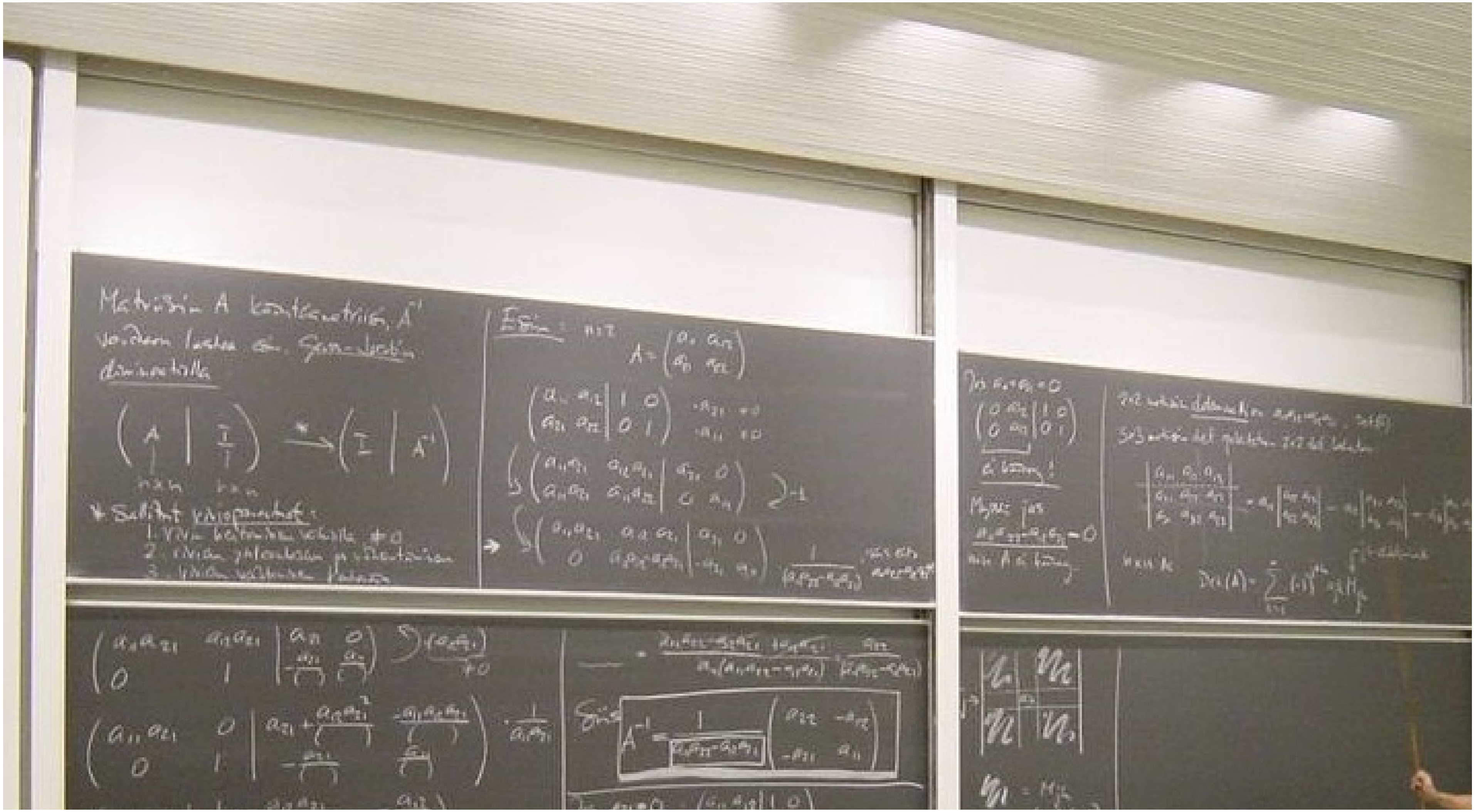
Undergraduate Textbooks on Automata Theory before 1965

$\emptyset$

A small sample of the 100s of books on automata theory available today

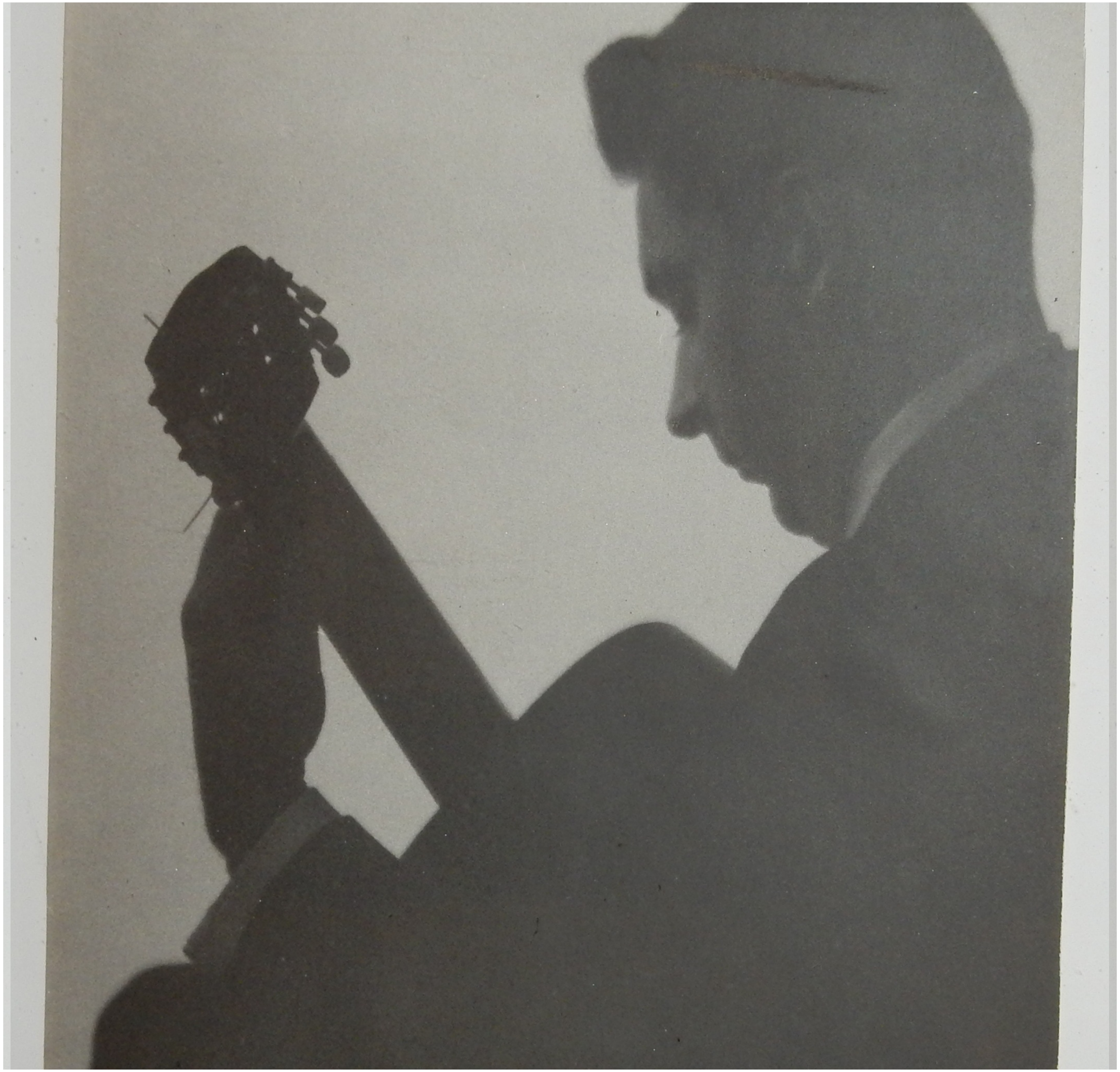
Introduction to
Automata Theory & Formal Languages
Adesh K. Pandey
A Textbook on Automata Theory
Nasir S.F.B. • P.K. Srimani
An Introduction to AUTOMATA THEORY & FORMAL LANGUAGES
ELEMENTS OF AUTOMATA THEORY
JACQUES SAKAROVITCH
CAMBRIDGE
Theory of Automata & Formal Languages
(As per JNTU Syllabus)
A.M. Natarajan
A. Tamilarasi
P. Balasubramani
NEW AGE INTERNATIONAL PUBLISHERS
K V N SUNITHA | N KALYANI
FORMAL LANGUAGES AND AUTOMATA THEORY
JEFFREY SHALLIT
A Second Course in Formal Languages and Automata Theory

Chalk board



Some of us are still using chalk boards

R ∈ URat M α inj on R ⇒ Rα ∈ URat N
α surj. U ∈ URat N ∃ R ∈ URat M Rα = U α inj. on R
Rec M ∩ Rat M ⊆ Rat M
Thm (Eilenberg & Schützenberger 70)
Rat N^k = URat N^k
Thm A congruence of N^k
N^k
d A
66)
algebra
Presburger def
N^k



In spite of his enormous musical talent,
Janusz decided to pursue studies in Electrical Engineering

1959 Master of E.E Toronto

.

PhD at Princeton '62

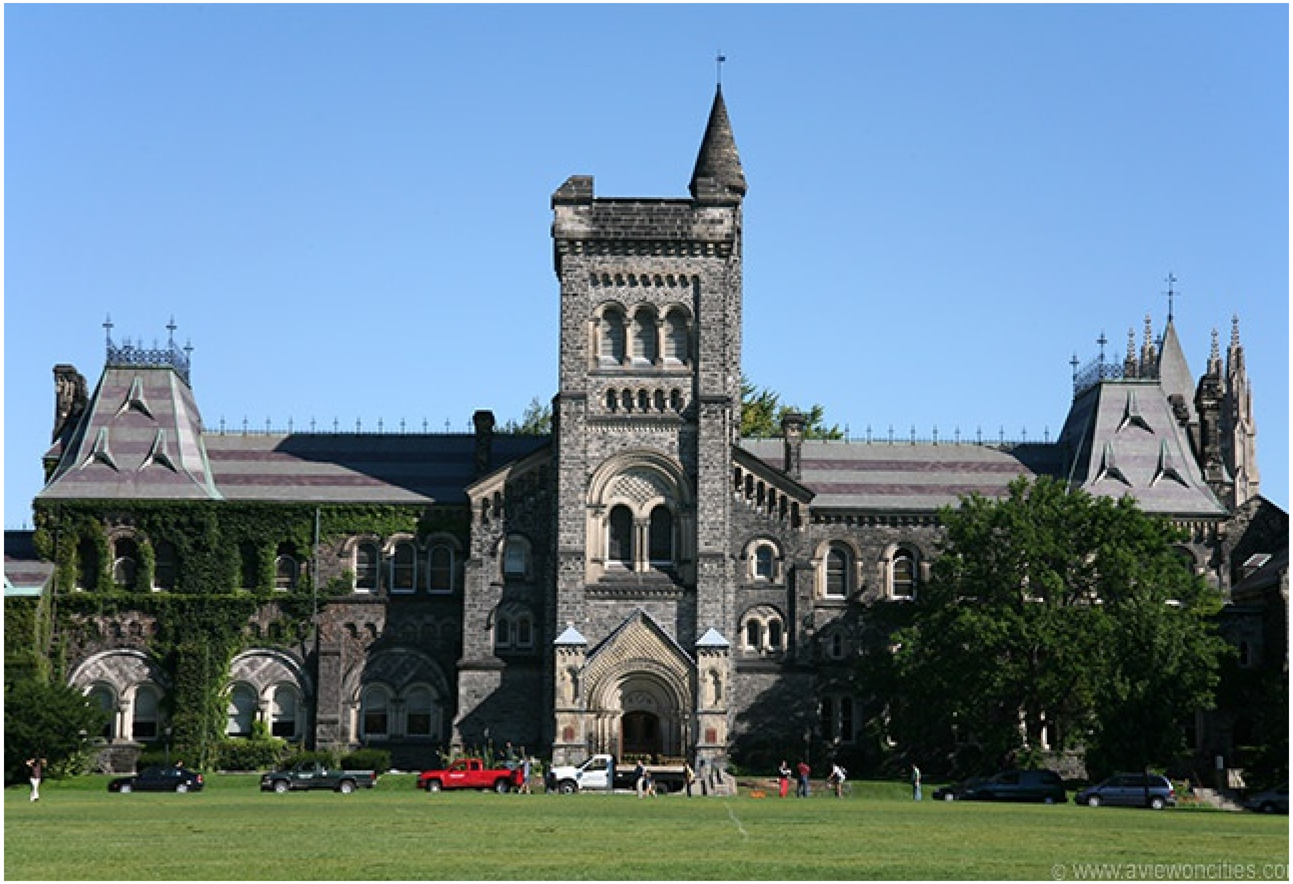
1962 Appointed to the Faculty of Electrical Engineering and Computer Science, at the University of Ottawa

School of Electrical Engineering & Computer Science, U. of Ottawa

University of Turku

Arto Salomaa & Janusz Brzozowski

The main speakers
at the first Canadian Theoretical Computer Science
summer school / workshop conference
at the University of Toronto



This is how Janusz Brzozowski introduced me to regular expressions & languages

Ø denotes the ***empty set***

Σ denotes an ***alphabet,*** a finite set of letters, for example if Σ = {0, 1}
then Σ is the binary alphabet of characters 0 and 1.

A finite letter sequence, where the letters are from Σ, is called a ***word*** *over* Σ.
The length of the sequence is called the ***length*** of the word.

Σ* denotes the ***set of all words*** over Σ.

λ denotes the ***empty word****,* namely the empty sequence of letters.

Subsets of Σ* are called ***languages over*** Σ

Λ or Ø* denote the ***empty-word language****,*
the singleton language that contains **only** the empty word, Λ = {λ} = Ø*

Operation on words over Σ,
*word **concatenation product**,* or just ***product*** of two words x, y: $x \cdot y$

let $x = x_1 \ldots x_n$, $y = y_1 \ldots y_m$ then $x \cdot y = x_1 \ldots x_n \cdot y_1 \ldots y_m$, where all x_i, $y_j \in \Sigma$

Note: for any word x, $\lambda \cdot x = x \cdot \lambda = x$ and for any three words x,y,z $(x \cdot y) \cdot z = x \cdot (y \cdot z) = x \cdot y \cdot z = xyz$.

Operation on languages
Boolean set operation on languages A and B: union $\cup$, intersection $\cap$, set difference $-$, etc.

$A \cup B = \{x \mid x \in A \text{ or } x \in B\}$, $A \cap B = \{x \mid x \in A \text{ and } x \in B\}$, $A - B = \{x \mid x \in A \text{ and } x \notin B\}$.

All ***identities from Boolean algebra*** of sets apply to all languages A, B, C over Σ.

$A \cup A = A \cap A = A \cup \varnothing = A - \varnothing = A \cap \Sigma^* = A \cup (A \cap B) = A \cap (A \cup B) = \Sigma^* - (\Sigma^* - A) = A$,

$A \cup B = B \cup A$, $A \cap B = B \cap A$, $\Sigma^* - (A \cup B) = (\Sigma^* - A) \cap (\Sigma^* - B)$,

$A \cup \Sigma^* = \Sigma^*$, $A - B = A \cap (\Sigma^* - B)$, $A - \Sigma^* = A \cap \varnothing = A - A = \varnothing - A = \varnothing$,

$(A \cup B) \cup C = A \cup (B \cup C)$, $(A \cap B) \cap C = A \cap (B \cap C)$, $(A - B) - C = A - (B \cup C)$,

$(A \cup B) - C = (A - C) \cup (B - C)$, $(A \cap B) - C = (A - C) \cap (B - C)$,

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

The ***concatenation product***, or just ***product*** of languages A, B: $A \bullet B = \{x \cdot y \mid x \in A, y \in B\}$

Basic identities for products of languages A, B, C over Σ :

$$A \bullet (B \bullet C) = (A \bullet B) \bullet C, \qquad A \bullet (B \mathbf{\,U\,} C) = A \bullet B \mathbf{\,U\,} A \bullet C, \qquad (B \mathbf{\,U\,} C) \bullet A = B \bullet A \mathbf{\,U\,} C \bullet A$$

$$A \bullet \varnothing^* = \varnothing^* \bullet A = A, \qquad A \bullet \varnothing = \varnothing \bullet A = \varnothing$$

Iterated concatenation product, or just *the* ***n***

$$A^0 = \varnothing^* = \{\lambda\}, \quad A^n = A \bullet A^{n-1} = A^{n-1} \bullet A \text{ for all } n > 0$$

The Kleene-star closure

Kleene **star** of a language A over Σ, A*: set of all finite-sequence products of words from A

$$A^* = \varnothing^* \mathbf{\,U\,} A \mathbf{\,U\,} A^2 \mathbf{\,U\,} A^3 \mathbf{\,U\,} A^4 \mathbf{\,U\,} \ldots = \{w_1 w_2 \ldots w_n \mid w_i \in A, n \in \mathbb{N}, i \in [1 \ldots n]\} \mathbf{\,U\,} \{\lambda\},$$

including λ, the empty sequence, when *n* = 0

Important identities for star closure

$$A^* = \varnothing^* \mathbf{\,U\,} A \bullet A^* = A^* \bullet A^* = (A^*)^* = (A \mathbf{\,U\,} \varnothing^*)^* = (A - \varnothing^*)^*, \qquad A \bullet A^* = A^* \bullet A$$

Regular expressions, regular languages, Brzozowski derivatives.

Let Σ be a finite set of symbols with no elements from {**(**, **)**, **+**, •, **Ø**,*, ∩, **—**, =, Σ, δ, ∂}

Syntax

The following syntax rules define the form of *regular expressions over* Σ, $R_{eg}\Sigma$:

S1. **Ø** ϵ $R_{eg}\Sigma$

S2. if x ϵ Σ then x ϵ $R_{eg}\Sigma$

S3. if A, B ϵ $R_{eg}(\Sigma)$ then so are **(**A **+** B**)**, **(**A•B**)**, **(**A***)**

S4. nothing else is in $R_{eg}\Sigma$
unless its being is the result of a finite no. of applications of steps S1., S2., and S3.

Denotations. The **meaning** of a regular expression A, **|**A**|**

M1. Ø denotes the empty set { } , **|**Ø**|** = { }

M2. for all x ϵ Σ, x denotes the **singleton** {x}, **|**x**|** = {x}
the language of a one-letter word, namely x

M3. for all A, B ϵ $R_{eg}\Sigma$,

(A **+** B**)** denotes the **union** of the two sets denoted by A and B, |**(**A **+** B**)|** = **|**A**| U |**B**|**

(A•B**)** denotes the **product** of the two sets denoted by A and B, |**(**A•B**)|** = **|**A**|** • **|**B**|**

(A***)** denotes the Kleene **star** closure of the set denoted by A |**(**A***)|** = **|**A**|***

Simplifications and abbreviations.

We may omit the • and the parentheses where possible:

we write $A+B+C$ for $((A+B)+C)$ and for $(A+(B+C))$
we write ABC for $((AB)\bullet C)$ and for $(A\bullet(B\bullet C))$

we assume the • has higher precedence than the +, and the * has the highest precedence
for example, we write $A+BC^*$ for $(A+(B\bullet(C)^*))$

Let $\Sigma = \{x_1, x_2, ..., x_n\}$, then we write $\mathbf{\Sigma}$ for $(x_1 + x_2 + ... + x_n)$
For $|A| \cap \{\lambda\}$ we write $\delta|A|$ and we write $\boldsymbol{\delta}\mathbf{A}$ to denote this set.

Brzozowski's X-regular expressions, Boolean operations ∩ and ─ are included:

The syntax of X-regular expressions over Σ, $XR_{eg}\Sigma$, is defined as follows:

XS1. $\mathbf{Ø} \in XR_{eg}\Sigma$
XS2. if $x \in \Sigma$ then $x \in XR_{eg}\Sigma$
XS3. if $A, B \in XR_{eg}(\Sigma)$ then so are $(A + B)$, $(A \cap B)$, $(A - B)$, $(A\bullet B)$, $(A)^*$
XS4. nothing else is in $XR_{eg}\Sigma$
unless its being is the result of a finite no. of applications of steps XS1., XS2., and XS3.

We abbreviate $A \cap \mathbf{Ø}^*$ as $\boldsymbol{\delta}A$

Meaning. The **meaning** of an X-regular expression

XM1. **Ø** denotes the empty set { }

XM2. for all $x \in \Sigma$, x denotes the **singleton** $\{x\}$, a language of a one-letter word, namely x

XM3. for all $A, B \in XR_{eg}\Sigma$,

(A + B) denotes the **union** of the two sets denoted by **A** and **B** $\quad |(A + B)| = |A| \cup |B|$

(A ∩ B) the **intersection** of the sets denoted by **A** and **B** $\quad |(A \cap B)| = |A| \cap |B|$

(A ─ B) the **difference** of the sets denoted by **A** and **B** $\quad |(A - B)| = |A| - |B|$

(A•B) the **product** of the two sets denoted by **A** and **B** $\quad |(A \bullet B)| = |A| \bullet |B|$

(A*) the Kleene **star** closure of the set denoted by **A** $\quad |(A^*)| = |A|^*$

For all $A \in XR_{eg}\Sigma$, **δA** denotes the X-regular expression **(A ∩ Ø*)**, $\delta A = (A \cap Ø^*)$

i.e., $\delta A = Ø^*$ iff $\lambda \in |A|$, and $\delta A = Ø$ iff $\lambda \notin |A|$.

Therefore for any $B \in XR_{eg}\Sigma$ $(\delta A) \bullet B = B$ iff $\lambda \in |A|$ **and** $(\delta A) \bullet B = Ø$ iff $\lambda \notin |A|$

The Brzozowski derivatives of X-Regular Expressions.

For each word $w \in \Sigma^*$, we define a mapping $\partial_w : XR_{eg}\Sigma \rightarrow XR_{eg}\Sigma$ recursively as follows:

For all $A \in XR_{eg}\Sigma$, $\partial_\lambda(A) = A$, for all $x \in \Sigma$, $\partial_x(Ø) = Ø$.

For all $x \in \Sigma$, $\partial_x(x) = Ø^*$, and for all $x, y \in \Sigma$, where $x \neq y$ $\partial_x(y) = Ø$.

For all $A, B \in XR_{eg}\Sigma$ and for all $x \in \Sigma$,

$$\partial_x((A + B)) = (\partial_x(A) + \partial_x(B))$$

$$\partial_x((A \cap B)) = (\partial_x(A) \cap \partial_x(B))$$

$$\partial_x((A - B)) = (\partial_x(A) - \partial_x(B))$$

$$\partial_x((A \bullet B)) = (\partial_x(A) \bullet B + \delta A \bullet \partial_x(B))$$

$$\partial_x(A^*) = (\partial_x(A) \bullet A^*)$$

For all $A \in XR_{eg}\Sigma$ and for all $x \in \Sigma$ and all words $w \in \Sigma^*$, we define $\partial_{xw}(A) = \partial_w(\partial_x(A))$

Exercise 1, Show that for all **A**, **B** ϵ $XR_{eg}\Sigma$ and for all w ϵ Σ^*, $\partial_w((\mathbf{A} + \mathbf{B})) = (\partial_w(\mathbf{A}) + \partial_w(\mathbf{B}))$,
$\partial_w((\mathbf{A} \cap \mathbf{B})) = (\partial_w(\mathbf{A}) \cap \partial_w(\mathbf{B}))$ and $\partial_w((\mathbf{A} - \mathbf{B})) = (\partial_w(\mathbf{A}) - \partial_w(\mathbf{B}))$.

Exercise 2, the meaning of the Brzozowski derivative

Show that the language denoted by $\partial_w(\mathbf{A})$, where w ϵ Σ^* and **A** ϵ $XR_{eg}\Sigma$, is the following:

$$|\partial_w(\mathbf{A})| = \{z \epsilon\ \Sigma^* |\ wz\ \epsilon\ |\mathbf{A}|\}$$

Similarity and equivalence of Extended Regular expressions.

For **A**, **B** ϵ $XR_{eg}\Sigma$, we say they are ***similar***, **A** ≈ **B**,
if starting from **A** and applying a finite sequence of the Boolean-, product- and star identities given below, one can obtain **B**.

$A + A \equiv A \cap A \equiv A + \emptyset \equiv A - \emptyset \equiv A \cap \Sigma^* \equiv A + (A \cap B) \equiv A \cap (A + B) \equiv \Sigma^* - (\Sigma^* - A) \equiv A,$

$A + B \equiv B + A, \quad A \cap B \equiv B \cap A, \quad A + \Sigma^* \equiv \Sigma^*, \quad \Sigma^* - (A \cup B) \equiv (\Sigma^* - A) \cap (\Sigma^* - B),$

$A - B \equiv A \cap (\Sigma^* - B) \quad A - \Sigma^* \equiv A \cap \emptyset \equiv A - A \equiv \emptyset - A \equiv \emptyset,$

$(A + B) + C \equiv A + (B + C), \quad (A \cap B) \cap C \equiv A \cap (B \cap C), \quad (A - B) - C \equiv A - (B + C),$

$(A + B) - C \equiv (A - C) + (B - C), \quad (A \cap B) - C \equiv (A - C) \cap (B - C),$

$A \cap (B + C) \equiv (A \cap B) + (A \cap C), \quad A + (B \cap C) \equiv (A + B) \cap (A + C),$

$A \bullet (B \bullet C) \equiv (A \bullet B) \bullet C, \quad A \bullet \emptyset^* \equiv \emptyset^* \bullet A \equiv A, \quad A \bullet (B + C) \equiv A \bullet B + A \bullet C, \quad (B + C) \bullet A \equiv B \bullet A + C \bullet A$

$A \bullet \emptyset \equiv \emptyset \bullet A \equiv \emptyset \quad A \bullet A^* \equiv A^* \bullet A \quad \emptyset^* + A \bullet A^* \equiv A^* \bullet A^* \equiv (A^*)^* \equiv (A + \emptyset^*)^* \equiv (A - \emptyset^*)^* \equiv A^*$

Exercise 3, Give three languages A, B, C over {0,1} such that $A \bullet (B - C) \neq A \bullet B - A \bullet C$

A solution: $A = \{0, 01\}$, $B = \{0, 10\}$, $C = \{\lambda, 0, 1\}$

$$A \bullet (B - C) = \{0, 01\} \bullet (\{0, 10\} - \{\lambda, 0, 1\}) = \{0, 01\} \bullet \{10\} = \mathbf{\{010, 0110\}}$$

$$A \bullet B - A \bullet C = \{0, 01\} \bullet \{0, 10\} - \{0, 01\} \bullet \{\lambda, 0, 1\} = \{00, 010, 0110\} - \{0, 00, 01, 010, 011\} =$$
$$= \mathbf{\{0110\}}$$

Equivalence

For **A**, **B** ϵ $XR_{eg}\Sigma$, we say **A** and **B** are ***equivalent***, **A ≡ B**, if **|A|** = **|B|.**

Note: **A ≈ B** implies **A ≡ B.**

Dissimilarity

Two (extended) regular expressions **A** and **B** are termed ***dissimilar*** if they are NOT similar**.**

Similarity simplifications

In what follows, we assume that regular expressions are expressed in a form that is the result of a scan from left to right and any simplifying identities that are applicable are applied.

We assume therefore that all singletons {w} where w ϵ Σ^* are represented
by the regular expression simply as **w**.

Exercise 4: Show that for any $w \epsilon \Sigma^*$, and any $\mathbf{A} \epsilon XR_{eg}(\Sigma)$, $\partial_w(\mathbf{w}\bullet\mathbf{A}) = \mathbf{A}$

Solution: If $w = \lambda$ then $\mathbf{w} = \mathbf{Ø^*}$ and then $\partial_\lambda(\mathbf{w}\bullet\mathbf{A}) = \partial_\lambda(\mathbf{Ø^*}\bullet\mathbf{A}) = \partial_\lambda(\mathbf{A}) = \mathbf{A}$

We proceed by induction on the length of the letter sequence making up the word w.

Because of the first line, the statement of Exercise 4 holds for length 0.

Assume that the statement $\partial_w(\mathbf{w}\bullet\mathbf{A}) = \mathbf{A}$ holds for all words $w \epsilon \Sigma^*$ of length $\leq n$,

then for any word w of length n+1, we have $w = xv$, where $x \epsilon \Sigma$ and $v \epsilon \Sigma^n$,

then $\partial_w(\mathbf{w}\bullet\mathbf{A}) = \partial_{xv}(\mathbf{x}\bullet\mathbf{v}\bullet\mathbf{A}) = \partial_v(\partial_x(\mathbf{x}\bullet\mathbf{v}\bullet\mathbf{A})) = \partial_v(\partial_x(\mathbf{x})\bullet\mathbf{v}\bullet\mathbf{A} + \boldsymbol{\delta}\mathbf{x}\bullet\partial_x(\mathbf{v}\bullet\mathbf{A})) =$

$= \partial_v(\mathbf{Ø^*}\bullet\mathbf{v}\bullet\mathbf{A} + \mathbf{Ø}\bullet\partial_x(\mathbf{v}\bullet\mathbf{A})) = \partial_v(\mathbf{v}\bullet\mathbf{A} + \mathbf{Ø}) = \partial_v(\mathbf{v}\bullet\mathbf{A}) = \mathbf{A}$, since $v \epsilon \Sigma^n$

This completes the inductive proof for words w of any length.

Exercise 5:

Show that for any $u,v \epsilon \Sigma$, $u \neq v$ and any $\mathbf{A} \epsilon XR_{eg}\Sigma$, $\partial_u(\mathbf{v}\bullet\mathbf{A}) = \mathbf{Ø}$,

Solution: $\partial_u(\mathbf{v}\bullet\mathbf{A}) = (\partial_u(\mathbf{v})\bullet\mathbf{A} + \boldsymbol{\delta}\mathbf{v}\bullet\partial_u(\mathbf{v}\bullet\mathbf{A})) = (\mathbf{Ø}\bullet\mathbf{A} + \mathbf{Ø}\bullet\partial_u(\mathbf{v}\bullet\mathbf{A})) = (\mathbf{Ø} + \mathbf{Ø}) = \mathbf{Ø}$

Exercise 6: Show that for all words $w \in \Sigma^*$, a) $\partial_w(\Sigma^*) = \Sigma^*$ and b) $\partial_w(\boldsymbol{\emptyset}) = \boldsymbol{\emptyset}$.

Solution: The b) part is obvious.
a) We note $\partial_\lambda(\Sigma^*) = \Sigma^*$
Because of this, the statement of Exercise 6 holds for words of length 0.
We proceed with the following induction hypothesis, **IH** on the length w.
Assume that the statement $\partial_w(\Sigma^*) = \Sigma^*$ holds for all words w of length up to *n*,
then for any word w of length *n*+1, we have
w = xv, where $x \in \Sigma$ and $v \in \Sigma^n$, then let $|\mathbf{x}| = \{x\}$ and $|\mathbf{v}| = \{v\}$

$$\partial_{xv}(\Sigma^*) = \partial_v(\partial_x(\Sigma^*)) = \partial_v(\partial_x(\Sigma)\Sigma^*) = \partial_v(\partial_x(\mathbf{(}\Sigma - x\mathbf{)} \mathbf{+} x)\Sigma^*) = \partial_v(\mathbf{(}\partial_x(\Sigma - x) \mathbf{+} \partial_x(x)\mathbf{)}\Sigma^*) =$$

$= \partial_v(\mathbf{(}\boldsymbol{\emptyset} + \boldsymbol{\emptyset}^*\mathbf{)}\Sigma^*) = \partial_v(\boldsymbol{\emptyset}^*\Sigma^*) = \partial_v(\Sigma^*) = \Sigma^*$. The last step follows from **IH**.

Exercise 7: Show that for any $\mathbf{A} \in XR_{eg}\Sigma$, the cardinality of the set of dissimilar (distinct) derivatives of **A**, is the same as that of $\Sigma^* - \mathbf{A}$.

Hint: 1-1 correspondence exists between the set of dissimilar (distinct) derivatives of A and those of $\Sigma^* - \mathbf{A}$ **,**
for each $w \in \Sigma^*$, $\partial_w(\mathbf{A})$ goes to $\partial_w(\Sigma^* - \mathbf{A}) = \partial_w(\Sigma^*) - \partial_w(\mathbf{A}) = \Sigma^* - \partial_w(\mathbf{A})$.

Exercise 8:
Show that there are three dissimilar derivatives of **L**, the language of all binary (**0**,**1**) strings without consecutive **1**s, *i.e.*, $\mathbf{L} = \Sigma^* - \Sigma^*\mathbf{11}\Sigma^*$ and $\Sigma = \{\mathbf{0},\mathbf{1}\}$.

Solution: We try to obtain dissimilar derivatives by taking derivatives with respect to words of increasing lengths.

$\partial_\lambda(\mathbf{L}) = \mathbf{L} = \Sigma^* - \Sigma^*\mathbf{11}\Sigma^*$ **(0)**

$\partial_0(\mathbf{L}) = \partial_0(\Sigma^* - \Sigma^*\mathbf{11}\Sigma^*) = \partial_0(\Sigma^*) - \partial_0(\Sigma^*\mathbf{11}\Sigma^*) = \partial_0(A) - \partial_0(B)$ where $A = \Sigma^*$ and $B = \Sigma^*\mathbf{11}\Sigma^*$ (1)

$\partial_0(A) = \partial_0(\Sigma^*) = \Sigma^*$ (2)

$\partial_0(B) = \partial_0(\Sigma^*\mathbf{11}\Sigma^*) = \mathbf{(}\partial_0(\Sigma^*)\mathbf{11}\Sigma^* \mathbf{+} \boldsymbol{\delta}A\bullet\partial_0(\mathbf{11}\Sigma^*)\mathbf{)} = \mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+} \boldsymbol{\varnothing^*}\bullet\partial_0(\mathbf{11}\Sigma^*)\mathbf{)} =$
$= \mathbf{(}B \mathbf{+} \partial_0(\mathbf{11}\Sigma^*)\mathbf{)} = \mathbf{(}B \mathbf{+} \mathbf{(}\partial_0(\mathbf{1})\mathbf{1}\Sigma^* \mathbf{+} \boldsymbol{\delta}\mathbf{1}\bullet\partial_0(\mathbf{1}\Sigma^*)\mathbf{))} = \mathbf{(}B \mathbf{+} \mathbf{(}\boldsymbol{\varnothing}\bullet\mathbf{1}\Sigma^* \mathbf{+} \boldsymbol{\varnothing}\bullet\partial_0(\mathbf{1}\Sigma^*)\mathbf{))}$
$= \mathbf{(}B \mathbf{+} \mathbf{(}\boldsymbol{\varnothing} \mathbf{+} \boldsymbol{\varnothing}\mathbf{))} = B$ (3)

From (1), (2) and (3) we have $\partial_0(\mathbf{L}) = \partial_0(A) - \partial_0(B) = \Sigma^* - B = \Sigma^* - \Sigma^*\mathbf{11}\Sigma^* = \mathbf{L}$ (4)

$\partial_1(\mathbf{L}) = \partial_1(\Sigma^* - \Sigma^*\mathbf{11}\Sigma^*) = \partial_1(\Sigma^*) - \partial_1(\Sigma^*\mathbf{11}\Sigma^*) = \partial_1(A) - \partial_1(B)$, where $A = \Sigma^*$ and $B = \Sigma^*\mathbf{11}\Sigma^*$ (5)

$\partial_1(A) = \partial_1(\Sigma^*) = \Sigma^*$ (6)

$\partial_1(B) = \partial_1(\Sigma^*\mathbf{11}\Sigma^*) = \partial_1(A\bullet\mathbf{11}\Sigma^*) = \mathbf{(}\partial_1(A)\mathbf{11}\Sigma^* \mathbf{+ \delta}A\bullet\partial_1(\mathbf{11}\Sigma^*)\mathbf{)} = \mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ \varnothing^*\bullet}\partial_1(\mathbf{11}\Sigma^*)\mathbf{)} =$

$= \mathbf{(}B \mathbf{+ \varnothing^*\bullet}\partial_1(\mathbf{11}\Sigma^*)\mathbf{)} = \mathbf{(}B \mathbf{+ 1}\Sigma^*\mathbf{)} = \mathbf{(}\Sigma^*\mathbf{11}\Sigma^*\mathbf{+ 1}\Sigma^*\mathbf{)}$ (7)

From (5), (6) and (7) we have $\partial_1(\mathbf{L}) = \partial_1(A) - \partial_1(B) = \Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^*\mathbf{+ 1}\Sigma^*\mathbf{)}$ **(8)**

From (4) we see that for any word w ϵ Σ^*, $\partial_{0w}(\mathbf{L}) = \partial_w(\partial_0(\mathbf{L})) = \partial_w(\mathbf{L})$ (9)

i.e., No new dissimilar Brzozowski derivative is obtained
by taking derivatives with respect to a word w lengthened by a prefix 0.

In particular $\partial_{00}(\mathbf{L}) = \partial_0(\mathbf{L}) = \mathbf{L}$ and $\partial_{01}(\mathbf{L}) = \partial_1(\mathbf{L}) = \Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^*\mathbf{+ 1}\Sigma^*\mathbf{)}$ (10)

Dissimilar derivatives might still be obtained, however, from $\partial_1(\mathbf{L})$.

$\partial_{10}(\mathbf{L}) = \partial_0(\partial_1(\mathbf{L})) = \partial_0(\Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ 1}\Sigma^*\mathbf{)}) = \partial_0(\Sigma^*) - \partial_0(\mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ 1}\Sigma^*\mathbf{)}) =$

$= \Sigma^* - \partial_0(\mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ 1}\Sigma^*\mathbf{)}) = \Sigma^* - \mathbf{(}\partial_0(\Sigma^*\mathbf{11}\Sigma^*) \mathbf{+} \partial_0(\mathbf{1}\Sigma^*)\mathbf{)} = \Sigma^* - \mathbf{(}\partial_0(\Sigma^*\mathbf{11}\Sigma^*) \mathbf{+ \varnothing)} =$

$= \Sigma^* - \partial_0(\Sigma^*\mathbf{11}\Sigma^*) = \Sigma^* - \mathbf{(}\partial_0(\Sigma^*)\mathbf{11}\Sigma^* \mathbf{+ \delta}\Sigma^*\mathbf{\bullet}\partial_0(\mathbf{11}\Sigma^*)\mathbf{)} =$

$= \Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ \varnothing}^*\mathbf{\bullet}\partial_0(\mathbf{11}\Sigma^*)\mathbf{)} = \Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ \varnothing}^*\mathbf{\bullet\varnothing)} = \Sigma^* - \Sigma^*\mathbf{11}\Sigma^* = \mathbf{L}$ (11)

$$\partial_{11}(\mathbf{L}) = \partial_1(\partial_1(\mathbf{L})) = \partial_1(\Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ 1}\Sigma^*\mathbf{)}) = \Sigma^* - \partial_1\mathbf{(}\Sigma^*\mathbf{11}\Sigma^* \mathbf{+ 1}\Sigma^*\mathbf{)} =$$

$$= \Sigma^* - \mathbf{(}\partial_1(\Sigma^*\mathbf{11}\Sigma^*) \mathbf{+} \partial_1(\mathbf{1}\Sigma^*)\mathbf{)} = \Sigma^* - \mathbf{(}\partial_1(\Sigma^*\mathbf{11}\Sigma^*) \mathbf{+} \Sigma^*\mathbf{)} = \Sigma^* - \Sigma^* = \mathbf{Ø} \qquad \mathbf{(12)}$$

$$\partial_{110}(\mathbf{L}) = \partial_0(\partial_{11}(\mathbf{L})) = \partial_0(\mathbf{Ø}) = \mathbf{Ø} \quad \text{and} \quad \partial_{111}(\mathbf{L}) = \partial_1(\partial_{11}(\mathbf{L})) = \partial_1(\mathbf{Ø}) = \mathbf{Ø} \qquad (13)$$

i.e., No new dissimilar Brzozowski derivative can be obtained by further taking derivatives.

$\partial_\lambda(\mathbf{L}) = \mathbf{L} = \Sigma^* - \Sigma^*\mathbf{11}\Sigma^*$, $\partial_1(\mathbf{L}) = \Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^*\mathbf{+ 1}\Sigma^*\mathbf{)}$ and $\partial_{11}(\mathbf{L}) = \mathbf{Ø}$
this shows that there are at most three distinct dissimilar Brzozowski derivatives of **L**.

It is easy to show that these derivatives represent three distinct languages.

The deterministic finite automaton **M** that accepts **L** can be given by a **state table** as follows:

States of M	**0**	**1**	**δ**
$\partial_\lambda(\mathbf{L}) = \mathbf{L} = \Sigma^* - \Sigma^*\mathbf{11}\Sigma^*$	**L**	$\partial_1(\mathbf{L})$	**Ø***
$\partial_1(\mathbf{L}) = \Sigma^* - \mathbf{(}\Sigma^*\mathbf{11}\Sigma^*\mathbf{+ 1}\Sigma^*\mathbf{)}$	**L**	$\partial_{11}(\mathbf{L})$	**Ø***
$\partial_{11}(\mathbf{L}) = \mathbf{Ø}$	$\partial_{11}(\mathbf{L})$	$\partial_{11}(\mathbf{L})$	**Ø**

The states of M are languages over Σ^*,

the ***initial state*** is **L**;

a state **S** is a ***final state*** iff **δS** = **$\emptyset^*$**.

In a state **S** reading a letter x ϵ Σ , the automaton **M** goes into the state ∂_X(**S**).

A word w ϵ Σ^* is ***accepted***

if **M** ends up a final state on reading the sequence of letters of w from start to end.

Exercise 9: Show that there are only a finite number of dissimilar derivatives for any regular expression **E** ϵ X$R_{eg}\Sigma$.

Solution: We proceed by induction on the *depth of parenthetical nestedness* of in **E,** Δ(E).

The basis of the induction: when Δ(E) = 0.

If **E** = **$\emptyset$**, then the cardinality of the set of dissimilar derivatives of **E** is one,

since for any w ϵ Σ^*, ∂_w(**E**) = **$\emptyset$**.

If **E** = **$\emptyset^*$**, then the cardinality of the set of dissimilar derivatives of **E** is two,

1. ∂_λ(**$\emptyset^*$**) = **$\emptyset^*$** 2. for any w ϵ Σ^* — {λ} , ∂_w(**$\emptyset^*$**) = **$\emptyset$**

If **E** = x, where x ϵ Σ, then there are at most three distinct/dissimilar derivatives of **E,** as follows:

1. ∂_λ(**x**) = **x**, 2. ∂_x(**x**) = **$\emptyset^*$**, 3. ∂_w(**x**) = **$\emptyset$**, for all w ϵ Σ^* — {λ, x}.

If Card(Σ) > 1, then these are, in fact, exactly three derivatives.

The induction hypothesis.
Assume that the statement of **Exercise 9** is true for all **A**, **B** ϵ $XR_{eg}\Sigma$ if Δ(**A**), Δ(**B**) < k ,
then we show that the statement is true also for **E** ϵ $XR_{eg}\Sigma$, where Δ(**E**) = k.

Let $\mathbb{D}$(**A**), $\mathbb{D}$(**B**) and $\mathbb{D}$(**E**) be the set of dissimilar derivatives of **A**, **B** and **E** resp.
Let $\mathbb{D}$(**A**) = {$\mathbf{A_1}$, $\mathbf{A_2}$, ... $\mathbf{A_n}$} and let $\mathbb{D}$(**B**) = {$\mathbf{B_1}$, $\mathbf{B_2}$, ... $\mathbf{B_m}$}, for some positive integers n and m.

Case A.
Assume **E** = **(A + B)** then we have ∂_w(**(A + B)**) = **(**∂_w(**A**) **+** ∂_w(**B**)**)**, for all w ϵ Σ*
let *n* and *m* be the number of dissimilar derivatives of **A** and **B** respectively,
then that of **E** will not exceed *n·m*.
A similar statement can be said when **E** = **(A ∩ B)** and also when **E** = **(A ─ B)**.

Case B
Assume **E** = **(A • B)** then ∂_λ(**E**) = **(A • B)** and for all x ϵ Σ, then
we have ∂_x**(A • B)** = **(**∂_x(**A**) **• B)**, or ∂_x**(A • B)** = **((**∂_x(**A**) **• B) +** ∂_x(**B**)**)**

Then all further derivatives of **E** are also in the form
($\mathbf{A_i}$ **•, B)** or **((**$\mathbf{A_i}$ **• B) +** $\mathbf{B_j}$**)** where i ϵ [1.. *n*] and j ϵ [1.. *m*] .

Thus $\mathbb{D}$(**E**) is a subset of { **(**$\mathbf{A_i}$ **• B) |** i ϵ [1.. *n*] } **U** { **((**$\mathbf{A_i}$ **• B) +** $\mathbf{B_j}$**) |** i ϵ [1.. *n*], j ϵ [1.. *m*] }
which is a finite set whose cardinality is not more than *n* + *n·m*.

Case C

Assume $\mathbf{E} = (\mathbf{A}^*)$, since $\mathbf{A}^* = (\mathbf{A} + \emptyset^*)^* = (\mathbf{A} - \emptyset^*)^*$,

without loss of generality we further assume that $\delta\mathbf{A} = \emptyset$,

then $\partial_\lambda(\mathbf{E}) = (\mathbf{A}^*)$ and for all $x \in \Sigma$, then we have $\partial_x(\mathbf{A}^*) = (\partial_x(\mathbf{A}) \bullet (\mathbf{A}^*))$.

Then all further derivatives of **E** are also in the same form

$$((\mathbf{A}_i \bullet (\mathbf{A}^*)) + (\delta\mathbf{A} \bullet (\mathbf{A}^*)) = ((\mathbf{A}_i \bullet (\mathbf{A}^*)) + (\emptyset \bullet (\mathbf{A}^*)) = (\mathbf{A}_i \bullet (\mathbf{A}^*)).$$

Thus $\mathbb{D}(\mathbf{E})$ is a subset of $\{ (\mathbf{A}_i \bullet (\mathbf{A}^*)) \mid i \in [1..\, n] \}$

which is a finite set whose cardinality is not more than *n.*

This completes the inductive proof.

Formal power series associated with languages.

Let **L** be a language over Σ, the formal power series, $\mathbb{L}(x)$, associated with **L** is defined as follows:

$$\mathbb{L}(x) = \boldsymbol{\delta}\mathbf{L} + (\mathbf{L} \cap \Sigma)x + (\mathbf{L} \cap \Sigma^2)x^2 + (\mathbf{L} \cap \Sigma^3)x^3 + (\mathbf{L} \cap \Sigma^4)x^4 + \ldots$$

Example 1: Let $\Sigma = \{\mathbf{0}, \mathbf{1}\}$ and $\mathbf{L} = \Sigma^* - \Sigma^*\mathbf{11}\Sigma^*$, then

$\boldsymbol{\delta}\mathbf{L} = \boldsymbol{\emptyset^*}$, $\mathbf{L} \cap \Sigma = \{0,1\}$, $\mathbf{L} \cap \Sigma^2 = \{00,01,10\}$, $\mathbf{L} \cap \Sigma^3 = \{000,001,010,100,101\}$

$\mathbf{L} \cap \Sigma^4 = \{0000,0001,0010,0100,0101,1000,1001,1010\}$

$$\mathbb{L}(x) = \boldsymbol{\emptyset^*} + \{0,1\}x + \{00,01,10\}x^2 + \{000,001,010,100,101\}x^3 + \{0000,0001,0010,0100,0101,1000,1001,1010\}x^4 + \ldots$$

Let us look at the following two formal power series $\{\mathbf{0}\} \circ \mathbb{L}(x)x$ and $\{\mathbf{10}\} \circ \mathbb{L}(x)x^2$,
where the scalar multiplication ○ and the multiplication by powers of x are defined as component-wise multiplication on each coefficient and powers of x.

$$\{0\}\circ\mathbb{L}(x)x = \{0\}\varnothing^{*}x + \{0\}\bullet\{\{0,1\}x^{2} + \{0\}\bullet\{00,01,10\}x^{3} + \{0\}\bullet\{\{000,001,010,100,101\}x^{4} + \\ + \{0\}\bullet\{\{0000,0001,0010,0100,0101,1000,1001,1010\}x^{5} + \ldots$$

$$\{10\}\circ\mathbb{L}(x)x^{2} = \{10\}\bullet\varnothing^{*}x^{2} + \{10\}\bullet\{0,1\}x^{3} + \{10\}\bullet\{00,01,10\}x^{4} + \{10\}\bullet\{000,001,010,100,101\}x^{5} + \\ + \{10\}\bullet\{0000,0001,0010,0100,0101,1000,1001,1010\}x^{6} + \ldots$$

$$\{0\}\circ\mathbb{L}(x)x = \{0\}x + \{00,01\}x^{2} + \{000,001,010\}x^{3} + \{0000,0001,0010,0100,0101\}x^{4} + \\ + \{00000,00001,00010,00100,00101,01000,01001,01010\}x^{5} + \ldots$$

$$\{10\}\circ\mathbb{L}(x)x^{2} = \{10\}x^{2} + \{100,101\}x^{3} + \{1000,1001,1010\}x^{4} + \{10000,10001,10010,10100,10101\}x^{5} \\ + \{100000,100001,100010,100100,100101,101000,101001,101010\}x^{6} + \ldots$$

Let us form the formal power series $\{0\}\circ\mathbb{L}(x)x$ ✚ $\{10\}\circ\mathbb{L}(x)x^{2}$,
where the operation ✚ is defined by taking the union of the coefficients of like powers of x.

$$\{0\}\circ\mathbb{L}(x)x \text{ ✚ } \{10\}\circ\mathbb{L}(x)x^{2} = \{0\}x + \{00,01,10\}x^{2} + \{000,001,010,100,101\}x^{3} + \\ + \{0000,0001,0010,0100,0101,1000,1001,1010\}x^{4} + \\ + \{00000,00001,00010,00100,00101,01000,01001,01010,10000,10001,10010,10100,10101\}x^{5} + \ldots$$

Let us form the formal power series as the difference $\mathbb{L}(x) - (\{0\} \circ \mathbb{L}(x)x + \{10\} \circ \mathbb{L}(x)x^2)$, where the operation $-$ is defined by taking the set difference of the coefficients of like powers of x.

$$\mathbb{L}(x) - (\{0\} \circ \mathbb{L}(x)x + \{10\} \circ \mathbb{L}(x)x^2) = \textbf{Ø*} \mathbf{+} \{1\}x \mathbf{+} \textbf{Ø}x^2 \mathbf{+} \textbf{Ø}x^3 \mathbf{+} \textbf{Ø}x^4 \mathbf{+} \mathbf{\ldots} = \textbf{Ø*} \mathbf{+} \{1\}x$$

Let $\mathbb{1}$ denote the formal power series where the constant is **Ø***
and the coefficients of positive powers of x are all **Ø**.

$\mathbb{L}(x)$ is expressed as a formal rational polynomial.

$$(\mathbb{1} - (\{0\}x + \{10\}x^2)) \circ \mathbb{L}(x) = \mathbb{1} \mathbf{+} \{1\}x$$

$$\mathbb{L}(x) = (\mathbb{1} \mathbf{+} \{1\}x) / (\mathbb{1} - (\{0\}x + \{10\}x^2)) \quad ??$$

The meaning of these operations needs more explanations.
We may note that if we replace the coefficient sets by the cardinality of these sets
we obtain the rational polynomial L(x):

$$L(x) = (1 + 1x)/(1 – (1x +1x^2)) = (1 + x)/(1 – x – x^2)) = 1 + 2x + 3x^2 + 5x^3 + 8x^4 + 13x^5 + 21x^6 \ldots$$

$$L(0.001) = 1.001/0.998999 = 1.002003005008013021034055089144233377610988$$