

# Twenty years of CPM

Elena Yavorska Harris<sup>a</sup>, Thierry Lecroq<sup>b</sup>, Gregory Kucherov<sup>c</sup> and  
Stefano Lonardi<sup>a</sup>

<sup>a</sup>Dept. of Computer Science – University of California – Riverside, CA, USA

<sup>b</sup>University of Rouen, LITIS EA 4108, 76821 Mont-Saint-Aignan Cedex, France

<sup>c</sup>CNRS (LIFL, Lille and J.-V.Poncelet Lab, Moscow) and INRIA Lille – Nord Europe

Combinatorial Pattern Matching

22 – 24 June 2009 – Lille, France



# Locations, dates, program chairs, and number of PC members

Year	Location	Dates	PC size
1990	Paris, France	July	
1991	London, UK	April 17-19	
1992	Tucson, Arizona, USA	April 29 - May 1	10
1993	Padova, Italy	June 2-4	10
1994	Asilomar, California, USA	June 5-8	10
1995	Helsinki, Finland	July 5-7	10
1996	Laguna Beach, California, USA	June 10-12	12
1997	Aarhus, Denmark	June 30 - July 2	12
1998	Piscataway, New Jersey, USA	July 20-21	15
1999	Warwick, UK	July 22-24	15
2000	Montreal, Canada	June 21-23	14
2001	Jerusalem, Israel	July 1-4	20
2002	Fukuoka, Japan	July 3-5	17
2003	Morelia, Michocan, Mexico	June 25-27	19
2004	Istanbul, Turkey	July 5-7	16
2005	Jeju Island, Korea	June 19-22	20
2006	Barcelona, Spain	July 5-7	19
2007	London, Ontario, Canada	July 9-10	26
2008	Pisa, Italy	June 18-20	27
2009	Lille, France	June 22-24	31

# History

NATO-ASI Workshop in Maratea, Italy organized by A. Apostolico and Z. Galil in June (18th – 22nd) 1984



Alberto Apostolico

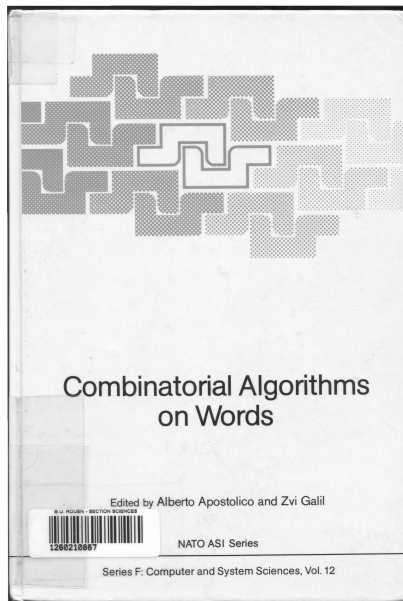


Zvi Galil

giving a talk at CPM 1990



# History



# First CPM??

In:  
Textual Substitution Techniques for Data Compression  
J.A. Storer

CPM: compressed pointer macro





Maxime Crochemore  
giving a talk at CPM 1990



# CPM 1990 (Paris)

## Call for participants

---

Advanced School  
**Combinatorial Pattern Matching**  
Paris, France, July 9-13, 1990

---

## Lecturers :

### Dr. A.V. Aho

A.T.& T., Bell Laboratories, U.S.A.

### Prof. Alberto Apostolico

Professor at University of l'Aquila, Italy, and  
Purdue University, USA

### Prof. Maxime Crochemore

Professor at University Paris 7, France

### Prof. Zvi Galil

Professor at University of Tel-Aviv, Israël,  
and Columbia University, USA.

### Prof. Esko Ukkonen

Professor at University of Helsinki, Finland

Laboratoire d'Informatique Théorique et Programmation  
INSTITUT BLAISE PASCAL

The school is funded by PRC "Mathématiques  
et Informatique"

## Aim of the school

The intention is to present to young researchers a wide variety of combinatorial methods used in the domain of Pattern Recognition. The school is also an occasion to gather specialists that are able to support interesting discussions on the subject.

## Outline of the school

The advent of large mass memory devices has made possible to store important amount of data. Information retrieval systems applied to these data require that powerful methods must be applied. Solutions to Pattern Matching problems provide efficient algorithms for a large variety of applications such as textprocessing, image analysis or genome investigations. This seminar will present a panorama of combinatorial pattern matching methods including up-to-date discoveries on the subject.

## Topics

*Exact string-searching : on-line and off-line algorithms, space considerations, index.*  
*Approximate string-searching : Hamming distance, edit distance, distances based on subsequences.*  
*Dynamic programming on strings.*  
*Parallel computations on strings*  
*Applications to linguistics.*

## Dates

The school will be held the second week of July 90, 9-13 at Ministère de la Recherche et de la Technologie in Paris. Next week is ICALP' 90 in Warwick. July 14 is the national day in France. The end of the school coincides with the open-air dances organised in several places in Paris for this celebration.

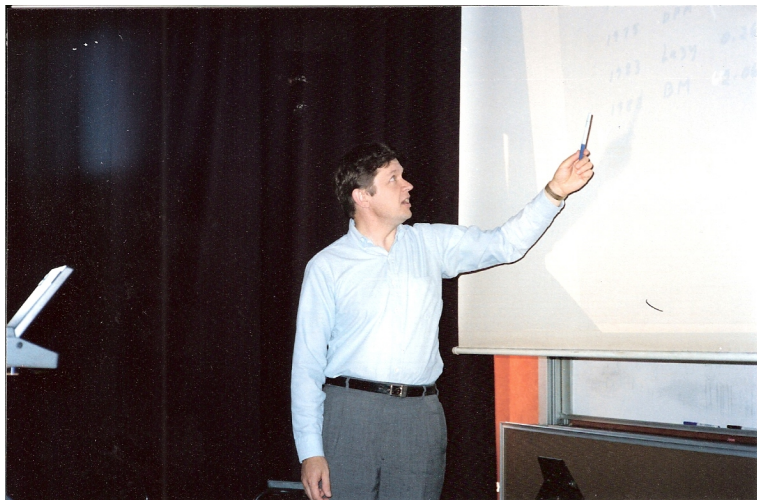
## Application

Enrollment to the school is limited to about twenty qualified participants. There are no fees and no strict deadlines. Participants will be encouraged to present and discuss their own research. A few guest lectures by leading specialists will also be scheduled.

To apply, send a short résumé to :

Prof. M. Crochemore, CPM School, LITP,  
Université Paris 7, 2 Place Jussieu, 75251  
Paris Cedex 05, France.  
Tel. : (33) 1 43 29 52 96 or (33) 1 43 25 98 74.  
E-mail : mac@litp.ibp.fr

# CPM 1990 (Paris)



Alfred Aho  
giving the 1st talk of the 1st CPM in 1990





## CPM 1991 (London)



Costas Iliopoulos  
giving a talk at CPM 1990



# CPM 1991 (London)

## CALL FOR PARTICIPATION

### COMBINATORIAL PATTERN MATCHING

London, 17-19 April 1991

The Combinatorial Pattern Matching School was first held at University of Paris 7 in June 1989 and 1990. The 1991 Conference will be held at Royal Holloway and Bedford New College (University of London), located West of London near Windsor Great Park, it is the only rural campus college of the University of London.

Papers presenting original, fundamental advances in algorithmic and combinatorial problems arising in applications which involve extensive manipulation of strings will be presented. Typical application areas include pattern recognition and computer vision, molecular sequence analysis, data compression and encoding, speech synthesis, text retrieval and editing etc. Also results that expose and study conceptual advances in parallel architectures or describe algorithmic work that enhances or exploits features of these architectures within the above framework will be presented.

We are inviting you to participate at CPM '91. People who wish to attend, please return the registration form by 20th March to the address below:

CPM '91,  
RHBNC, University of London,  
Dept of Computer Science,  
Egham, Surrey TW20 0EX, England

*Invited Speakers:*

A. Apostolico    M. Crochemore    A. Ehrenfeucht    Z. Galil

*Organising Committee:*

A. Davies, P. Hoare, C. Iliopoulos, A. Johnstone  
e-mail: csi@oak.cc.kcl.ac.uk Fax: (+44 0784) 437520

# History

1990 and 1991: Schools

From 1992: submission, referring and LNCS proceedings



## Number of papers

Year	Accepted	Submitted	Ratio	Authors	New	authors/paper
1992	22	39	56.4	43	38	2.04
1993	19	34	55.9	33	16	1.89
1994	26	41	63.4	52	38	2.15
1995	29	44	65.9	52	35	2.03
1996	28	48	58.3	61	34	2.28
1997	20	32	62.5	51	33	2.6
1998	17	49	34.7	42	22	2.52
1999	21	26	80.8	44	27	2.43
2000	29	44	65.9	64	38	2.03
2001	22	35	62.9	46	23	2.5
2002	23	37	62.2	58	33	2.78
2003	28	57	49.1	63	36	2.36
2004	36	79	45.6	95	62	2.94
2005	37	129	28.7	98	63	2.73
2006	33	88	37.5	82	38	2.67
2007	32	64	50	84	36	2.91
2008	25	78	32.1	67	27	2.84
2009	27	63	42.9	78	35	3.11

# CPM 2009

Submitted: 63      accepted: 27      ratio: 42.9  
Authors: 84      different: 74      new authors: 35  
Authors/paper in average: 3.11

## Contributing countries

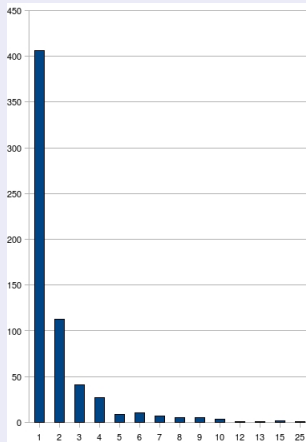
Israel	16	UK	3
France	16	Denmark	2
Germany	9	Canada	2
USA	7	Algeria	1
Switzerland	6	China	1
Taiwan	5	Hong Kong	1
Italy	4	Finland	1
Australia	3	Poland	1



# Number of authors

487 articles: 632 authors

# articles	1	2	3	4	5	6	7	8	9	10	12	13	15	25
# authors	406	113	41	27	9	10	7	5	5	3	1	1	2	1

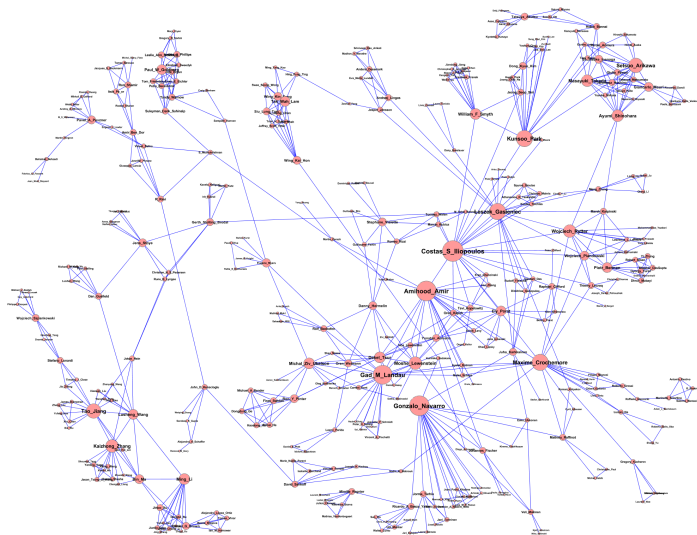


## Main contributors

Gonzalo Navarro	25
Gad M. Landau	15
Kaizhong Zhang	15
Amihood Amir	13
Kunsoo Park	12
Leszek Gasieniec	10
Wojciech Rytter	10
Masayuki Takeda	10
Bin Ma	9
Costas S. Iliopoulos	9
Maxime Crochemore	9
Veli Makinen	9
Michal Ziv-Ukelson	9



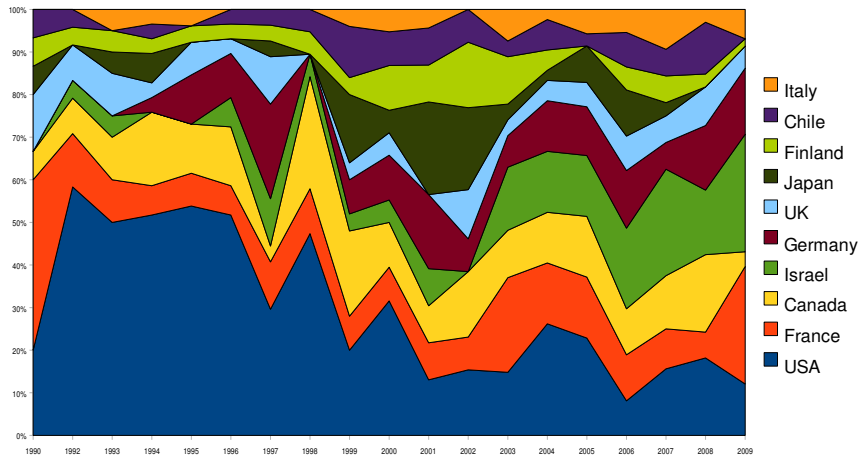
# Author network



<http://www.cs.ucr.edu/~stelo/cpm>



# Contributing countries





# Mining the titles

## 1 word

matching	115		
algorithm	90		
string	82	BUT	combinatorial 3
tree	57		
pattern	50		
approximate	47		

## 2 words

string – matching	48
pattern – matching	37
approximate – matching	37
matching – algorithm	26
approximate – string	24

# Mining the titles

## 3 words

approximate – string – matching	23
dimensional – pattern – matching	11
two – dimensional – matching	10
longest – common – subsequence	10
string – matching – algorithm	10

## 4 words

two – dimensional – matching – pattern	7
linear – time – suffix – construction	6
linear – suffix – array – time	5
approximate – compressed – string – matching	5



# Citations

Authors	Title	# Cit.	Year
E. Ukkonen	Approximate string-matching with $q$ -grams and maximal matches	249	1990
E. Myers	A Fast Bit-Vector Algorithm for Approximate String Matching Based on Dynamic Programming	180	1998
P. Ko, S. Aluru	Space Efficient Linear Time Construction of Suffix Arrays	164	2003
T. Jiang, L. Wang, K. Zhang	Alignment of Trees – An Alternative to Tree Edit	153	1994
G. M. Landau, J. P. Schmidt	An Algorithm for Approximate Tandem Repeats	140	1993
T. Kasai, G. Lee, H. Arimura, S. Arikawa, K. Park	Linear-Time Longest-Common-Prefix Computation in Suffix Arrays and Its Applications	135	2001
E. Ukkonen	Approximate String-Matching over Suffix Trees	129	1993
D. K. Kim, J. S. Sim, H. Park, K. Park	Linear-Time Construction of Suffix Arrays	124	2003
R. A. Baeza-Yates, W. Cunto, U. Manber, S. Wu	Proximity Matching Using Fixed-Queries Trees	121	1994
A. Bergeron	A Very Elementary Presentation of the Hannenhalli-Pevzner Theory	111	2001

Google Scholar, June 19th, 2009

# Acknowledgements

Many thanks to Alberto, Maxime & Costas for providing information about the history and the first CPMs



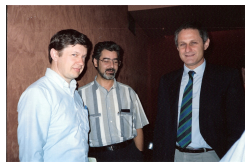
# CPM 1990



Maxime & Alberto



Gadi



Alfred, Alberto & Zvi



Esko



Alberto, Maxime & Mike



Maxime



Alberto



Alfred Aho



Esko & Maxime