CPM2004
Istanbul, Turkey

New results for the 2-interval pattern problem

Guillaume Blin[1], Guillaume Fertin[1] and Stéphane Vialette[2]

[1] LINA, FRE CNRS 2729
Université de Nantes,
2 rue de la houssinière
BP 92208 44322 Nantes Cedex 3
FRANCE
{blin,fertin}@lina.univ-nantes

[2] LRI, UMR CNRS 8623
Université Paris-Sud
Faculté des Sciences d'Orsay,
Bât 490, 91405 Orsay Cedex
FRANCE
vialette@lri.fr

Outline

- Preliminaries
  - Notations
  - The 2-interval pattern problem
- Two polynomial time algorithms
- Np-completeness proof

Outline

- Preliminaries
  - Notations
  - The 2-interval pattern problem
- Two polynomial time algorithms
- Np-completeness proof

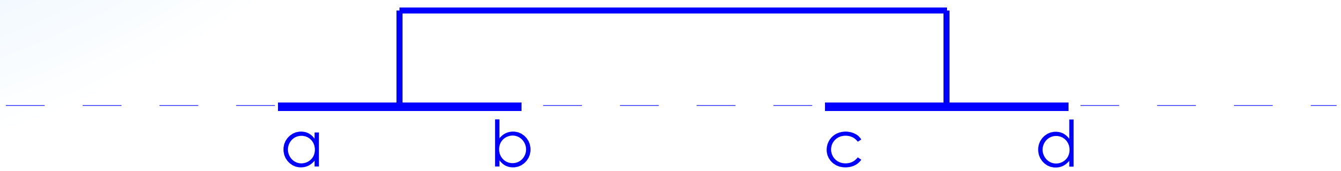
Preliminaries

- 1-interval
  - An interval (1-interval) is defined by its endpoints as follows:
    - M = [a : b], N = [c : d]
  - 1-intervals in precedence

    - M < N iff b < c
    - M and N are disjoint ($M \cap N = \emptyset$) iff M < N or N < M

Preliminaries

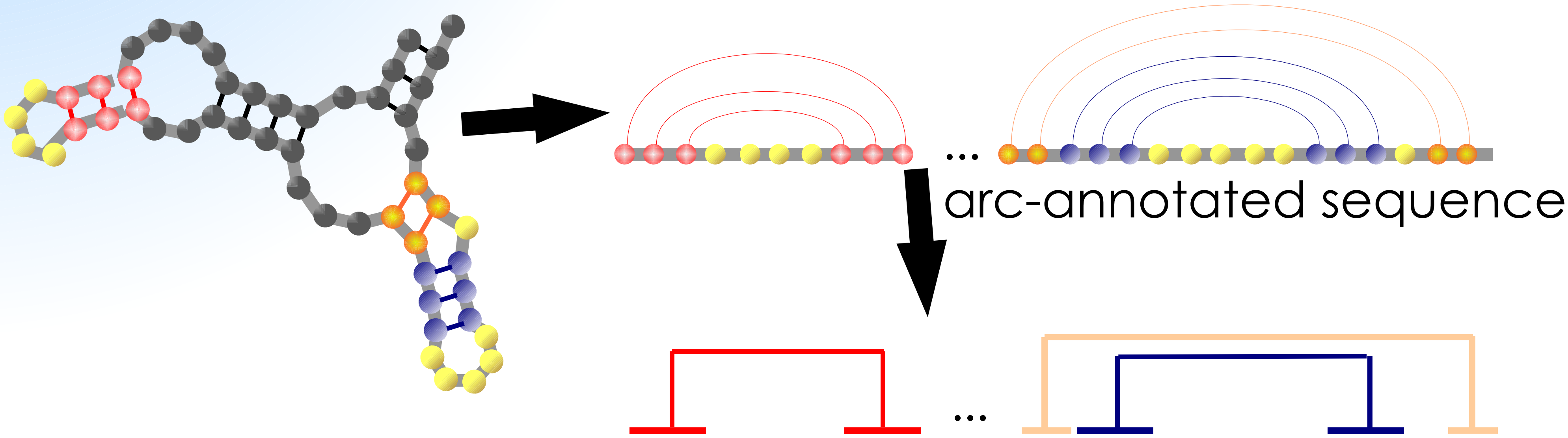
- 2-interval
  - the disjoint union of two 1-intervals on a line.
  - $D = (M,N)$ where $M=[\ a : b\ ] < N =[\ c : d\ ]$



- Two 2-intervals $D_1=(M_1,N_1)$ and $D_2=(M_2,N_2)$ are comparable iff:
  - $(M_1 \cap M_2) \cup (M_1 \cap N_2) \cup (N_1 \cap M_2) \cup (N_1 \cap N_2) = \varnothing$

Preliminaries

- 2-intervals are macroscopic describers of RNA secondary structure

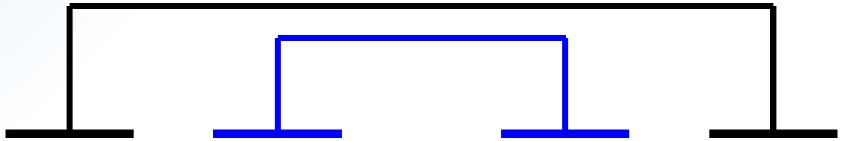


  - 1-interval ≡ sequence of bases
  - 2-interval ≡ a sequence of consecutive pair of bases, i.e. a stem

Preliminaries

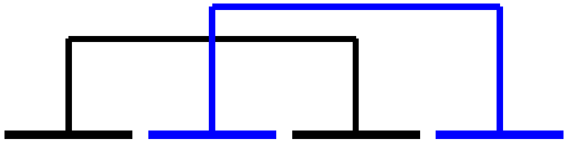
- Relationship between 2-intervals:
  - $D_1 < D_2$

  - $D_2 \sqsubset D_1$



transitive

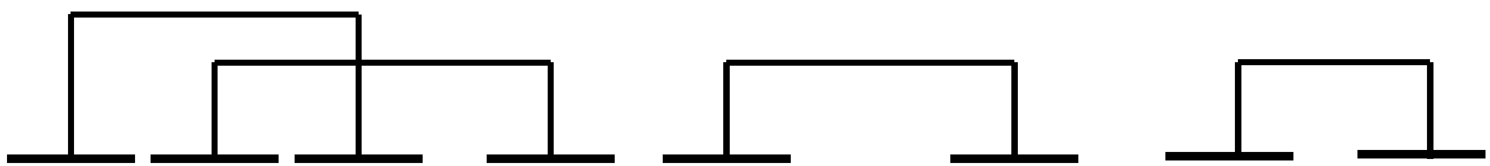
  - $D_1 \between D_2$



  - If D_1 and D_2 are comparable by $<$, $\sqsubset$ or $\between$ then they are disjoint !

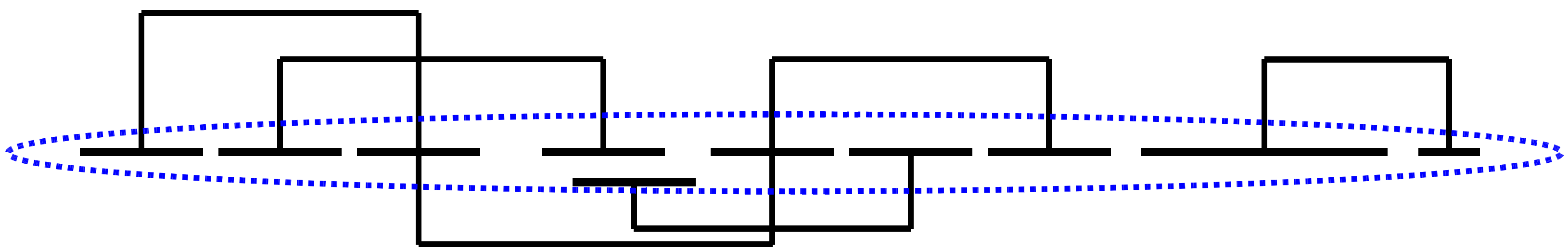
Preliminaries

- R-comparability:
  - Two 2-intervals D_1 and D_2 are α-comparable for some $\alpha \in \{<, \sqsubset, \between\}$ if $D_1 \ \alpha \ D_2$ or $D_2 \ \alpha \ D_1$.
    - Example:
      D_1 *and* D_2 are <-comparable if $D_1 < D_2$ or $D_2 < D_1$
  - Extension to a set of 2-intervals
    - Example:
      A set of $\{<, \between\}$-comparable 2-intervals

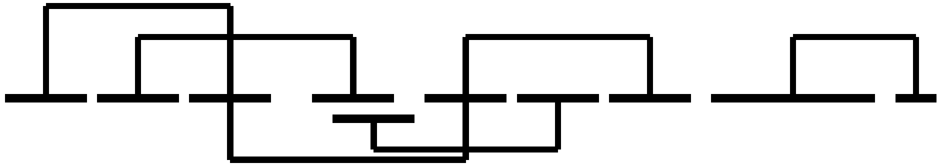


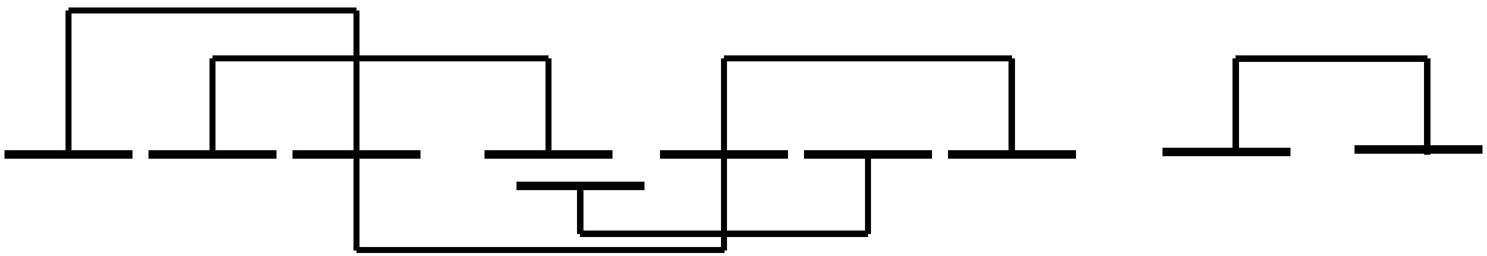
Preliminaries

- Support:
  - Given a set *D* of 2-intervals, the support of *D* is the set of all 1-intervals involved in *D*.
  - **Unlimited support** : no restrictions

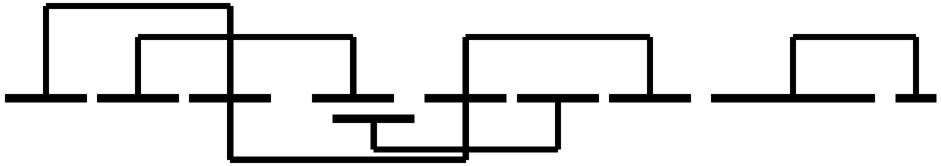
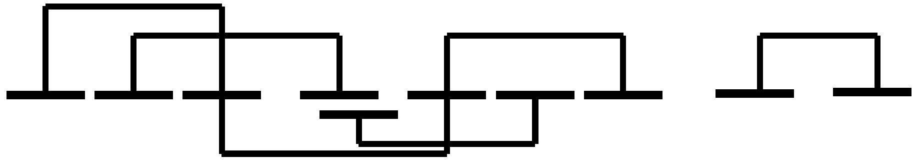


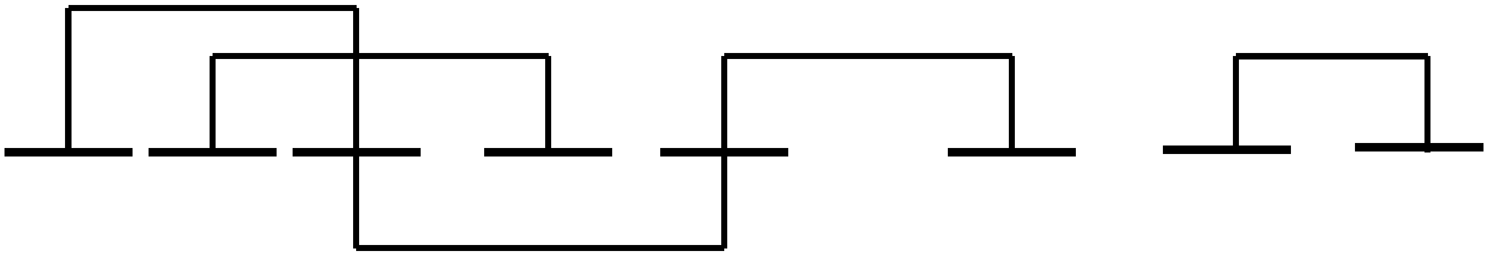
Preliminaries

- Support:
  - Given a set *D* of 2-intervals, the support of *D* is the set of all 1-intervals involved in *D*.
  - Unlimited support 
  - **Unitary support** : all 1-intervals in *D* are of the same size



Preliminaries

- Support:
  - Given a set *D* of 2-intervals, the support of *D* is the set of all 1-intervals involved in *D*.
  - Unlimited support 
  - Unitary support 
  - **Disjoint support** : all 1-intervals in *D* are of the same size and disjoint



Preliminaries

- 2-interval Pattern Problem:
  - Instance :
    - A set of 2-intervals *D* over a support, a model R $\subseteq$ {<, ⊏, ≬} and a positive integer *k*
  - Question :
    - Is it possible to find a subset $D' \subseteq D$ such that D' is R-comparable and $|D'| \geq k$?

Results

- Seven models for R and three restrictions on the support

2-INTERVAL PATTERN PROBLEM			
	SUPPORT		
MODEL	UNLIMITED	UNITARY	DISJOINT
$\{<, \sqsubset, \between\}$	**NP**-complete		$O(n\sqrt{n})$[MV80]
$\{\sqsubset, \between\}$	**NP**-complete		$O(n^2\sqrt{n})$ $\star$
$\{<, \sqsubset\}$	$O(n^2)$		
$\{<, \between\}$	**NP**-complete $\star$		?
$\{<\}$	$O(n \log n)$		
$\{\sqsubset\}$	$O(n \log n)$ $\star$ $\bullet$		
$\{\between\}$	$O(n^2 \log n)$		

with n=|D|

Outline

- Preliminaries
- Two polynomial time algorithms
  - 2-IP-UNL-$\{\sqsubset\}$
  - 2-IP-UNI-$\{\sqsubset, \between\}$
- Np-completeness proof

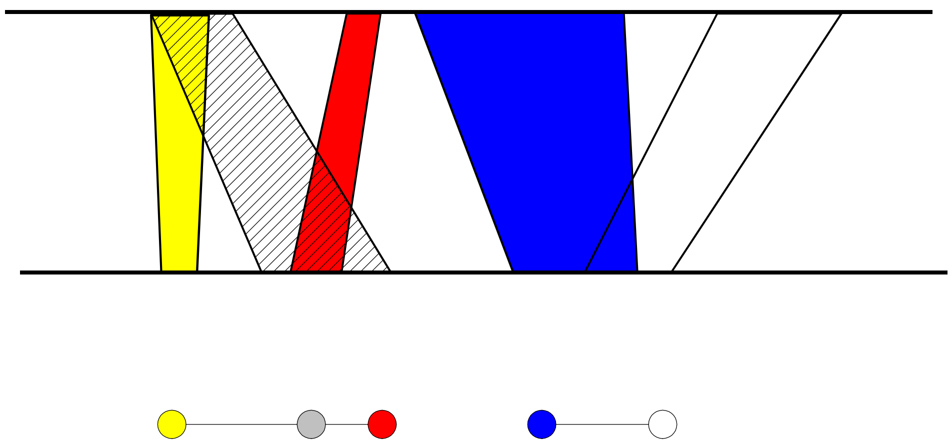
Improving the complexity of 2-IP-UNL-{□}

- In [Via04], this problem is solvable in a **$O(n^2)$** time algorithm

- We propose a **O(n log n)** time algorithm using *trapezoid graph*.

Improving the complexity of 2-IP-UNL-{□}

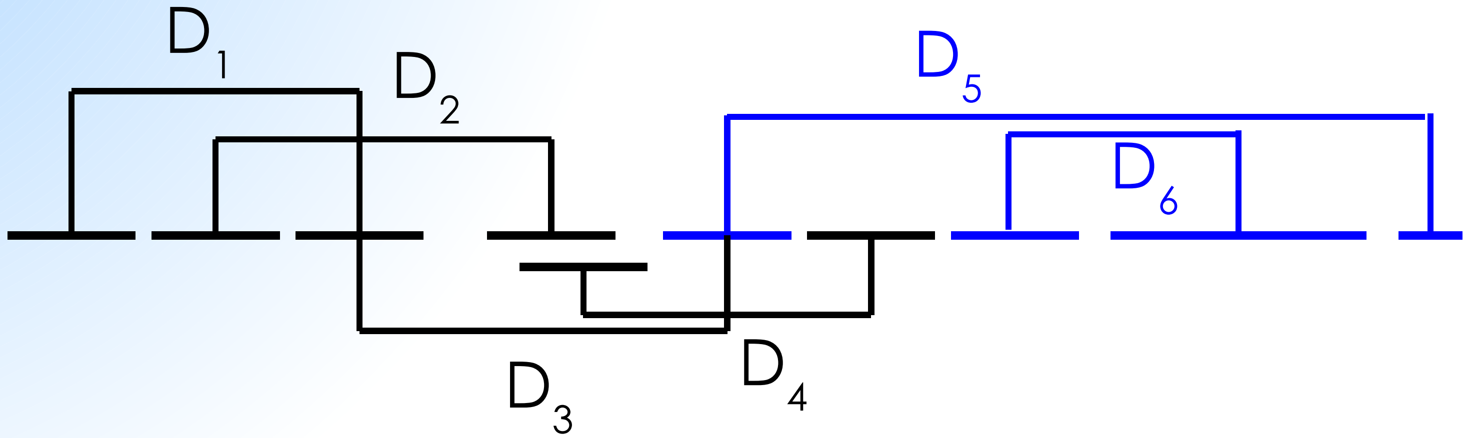
- A trapezoid graph is the **intersection graph** of a set of trapezoids between two lines.

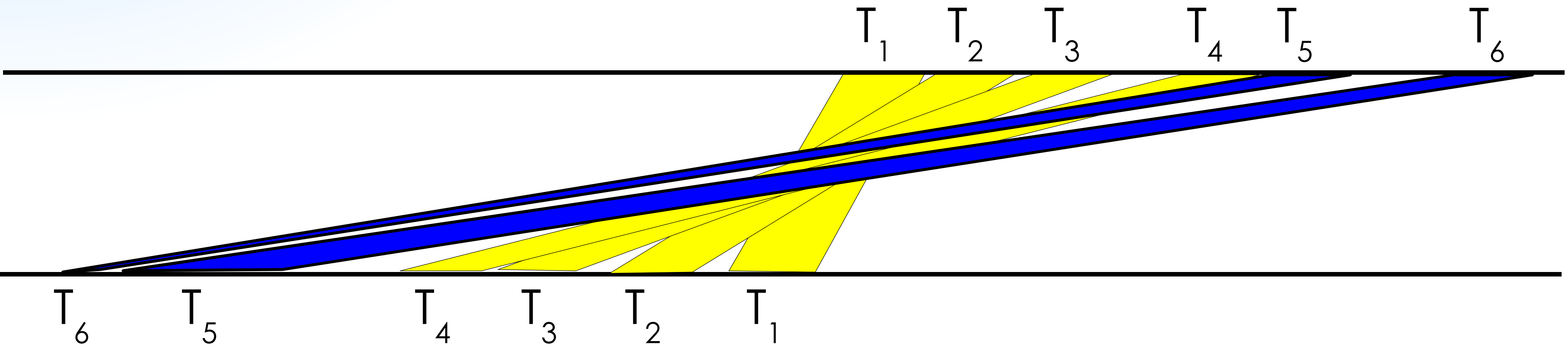


Improving the complexity of 2-IP-UNL-{□}

- Algorithm:
  - Given $D=\{D_1, D_2, \ldots, D_n\}$
  - For each $D_i = ([x : y], [x' : y']) \in D$
    create $T_i = ([x : y], [-y' : -x'])$
  - Let $T=\{T_1, T_2, \ldots, T_n\}$
  - T is obtained in O(n) time

Improving the complexity of 2-IP-UNL-{⊏}





D_k and D_l are {⊏}-comparable iff T_k and T_l are non-intersecting

Improving the complexity of 2-IP-UNL-{□}

- [Felsner et al 97] have designed a ***O(n log n)*** algorithm.

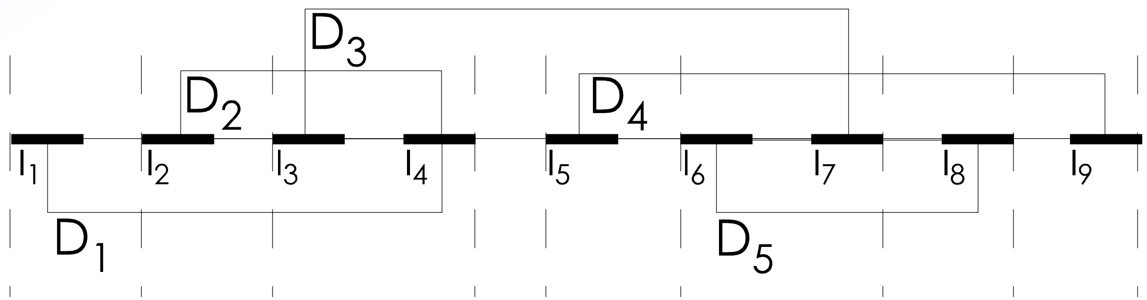
- Complexity analysis:
  - Construction in O(n) time
  - Find a solution in O(n log n) time
  - Global complexity: O(n) + O(n log n)
    -> **O(n log n)**

Outline

- Preliminaries
- Two polynomial time algorithms
  - 2-IP-UNL-$\{\sqsubset\}$
  - 2-IP-UNI-$\{\sqsubset, \between\}$
- Np-completeness proof

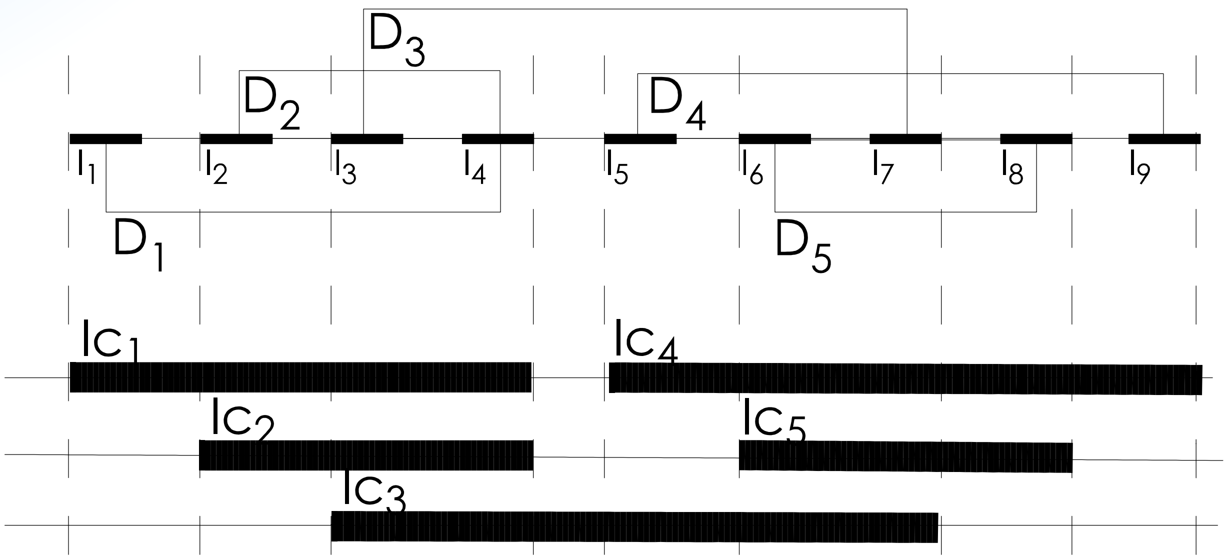
A polynomial time algorithm for 2-IP-DIS-$\{\sqsubset, \between\}$

Given a set of 2-intervals **D** on a disjoint support



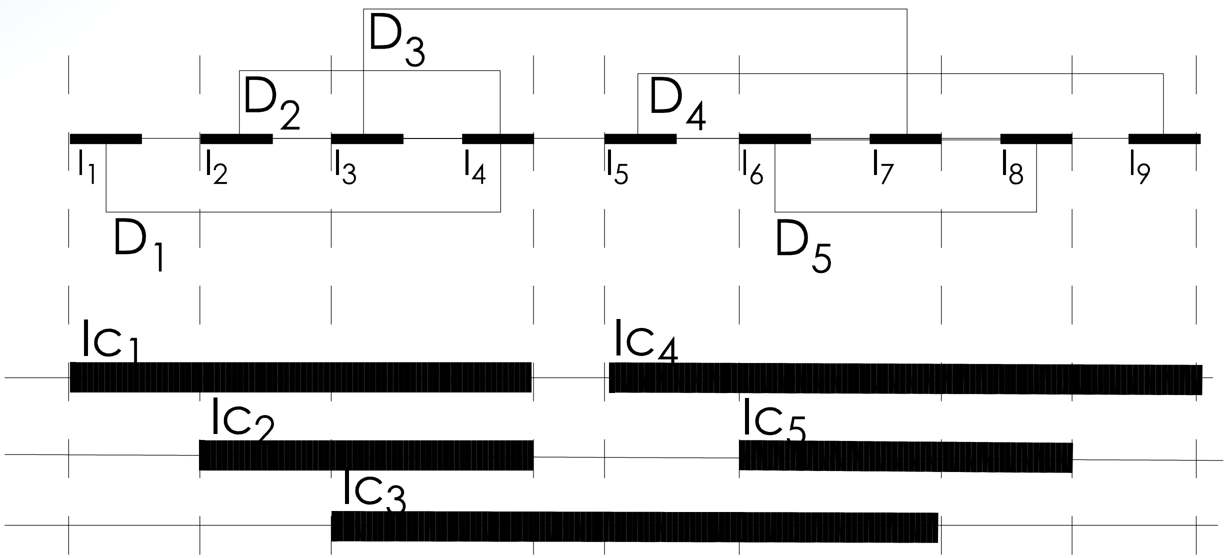
A polynomial time algorithm for 2-IP-DIS-{$\sqsubset$, $\between$}

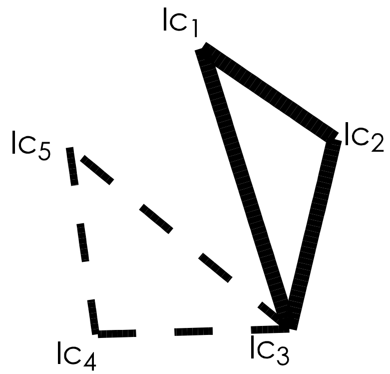
Create the corresponding interval cover set SI



A polynomial time algorithm for 2-IP-DIS-$\{\sqsubset, \between\}$

Create the corresponding interval graph $\Omega(SI)$

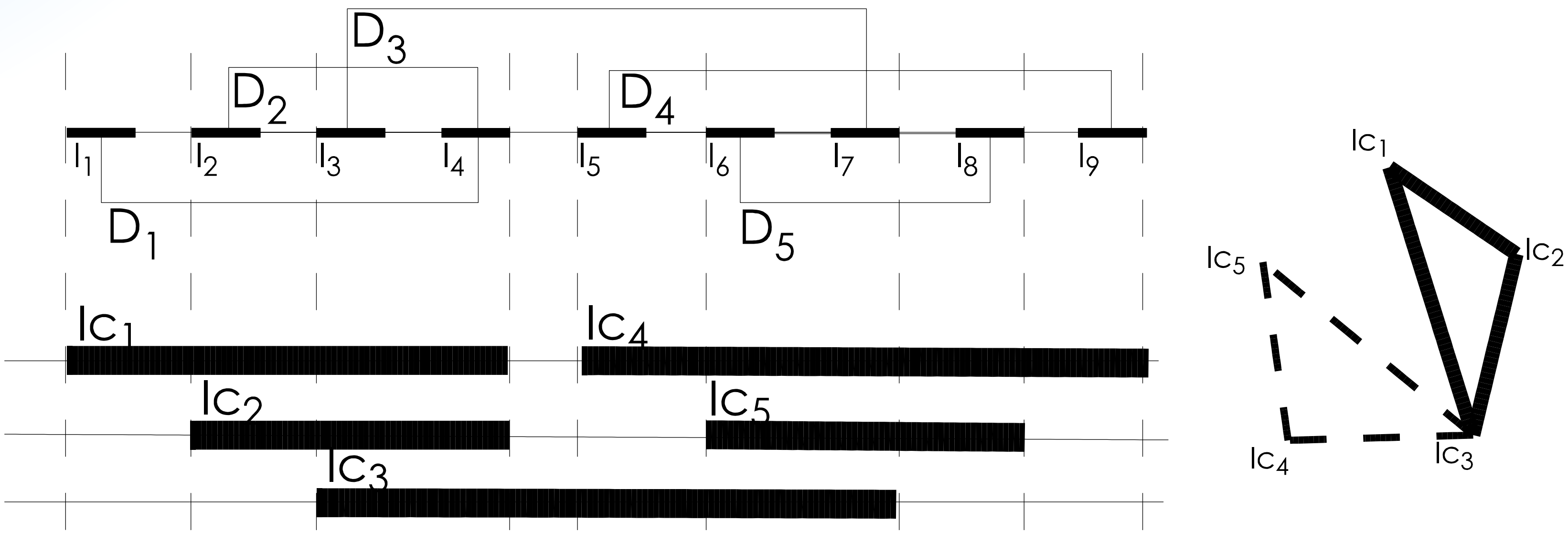




A polynomial time algorithm for 2-IP-DIS-{⊏, ≬}

For each maximal clique C in Ω(SI) (at most n) [Fulkerson et al 65]

Verify the disjonction constraint



Clique	Max Card.
Cl_1	2 (e.g. D_1 D_3)
Cl_2	3 (D_3 D_4 D_5)

A polynomial time algorithm for 2-IP-DIS-$\{\sqsubset, \between\}$

- Complexity analysis:
  - Construction of the interval cover set in **O(n)** time
  - Construction of the interval graph in **$O(n^2)$** time
  - Finding all the maximal clique (at most n) in **O(m+n)** time
    - Verify the disjonction constraint in **$O(n\sqrt{n})$**
  - Global complexity of : $O(n)+O(n^2)+O(m+n)+ n.O(n\sqrt{n})$ -> **$O(n^2\sqrt{n})$**

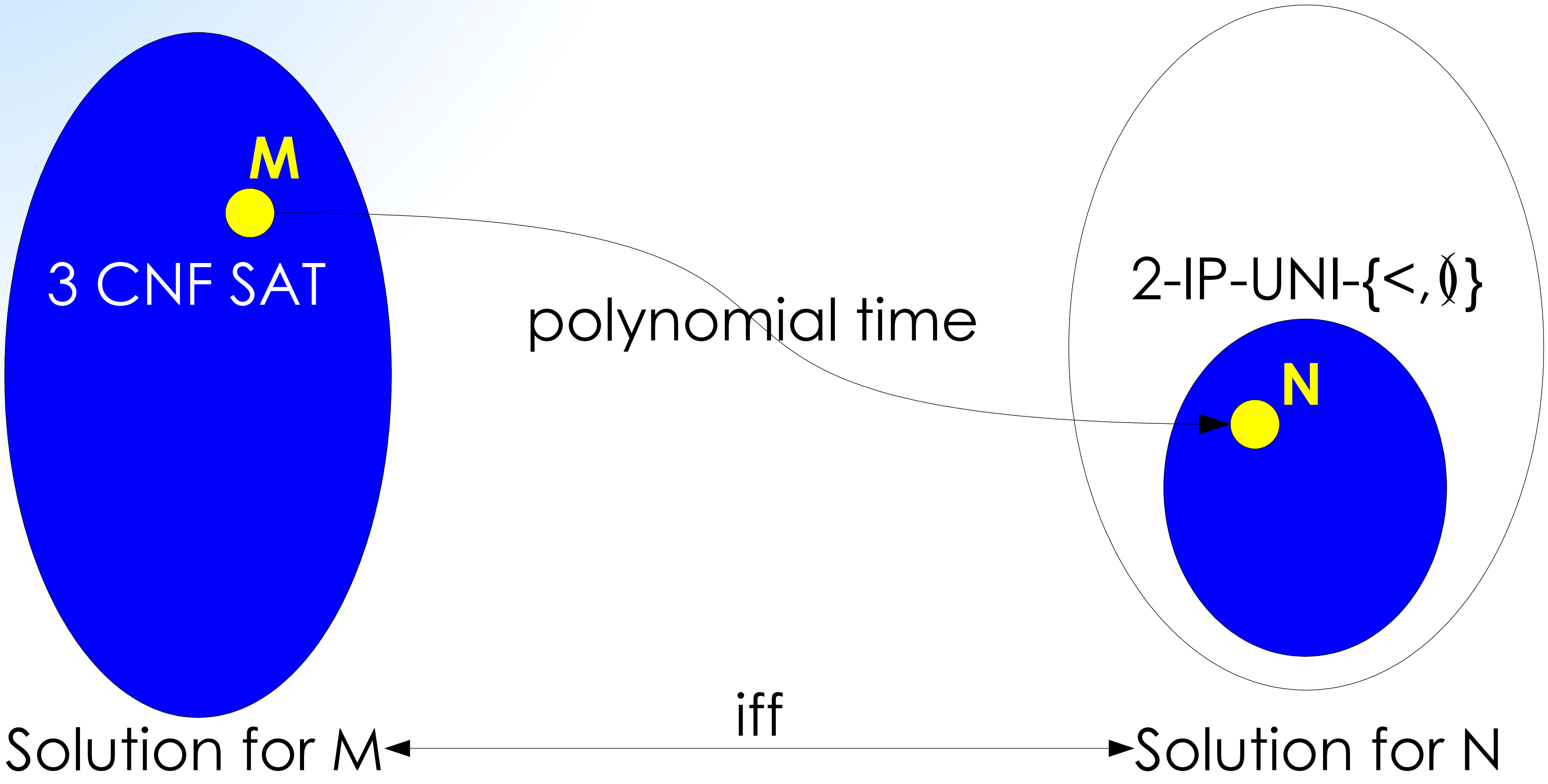
Outline

- Preliminaries
- Two polynomial time algorithms
- Np-completeness proof
  - 2-IP-UNI-{<, ≬} is NP-complete
  - 2-IP-UNI-{<, ≬} is fixed-parameter tractable

2-IP-UNI-$\{<, \between\}$ is NP-complete

- Rather technical proof using a reduction from the **Exact 3-CNF SAT problem.** (coming soon in the full version)
- Instance :
  - Conjunction of disjunction of boolean variables
    $BF = (\overline{x1} \vee x2 \vee x3) \wedge (x1 \vee \overline{x2} \vee x4) \wedge (x1 \vee x2 \vee \overline{x3})$
- Question:
  - Is there an assignment of the variables that sastifies BF ?

2-IP-UNI-{<, ≬} is NP-complete

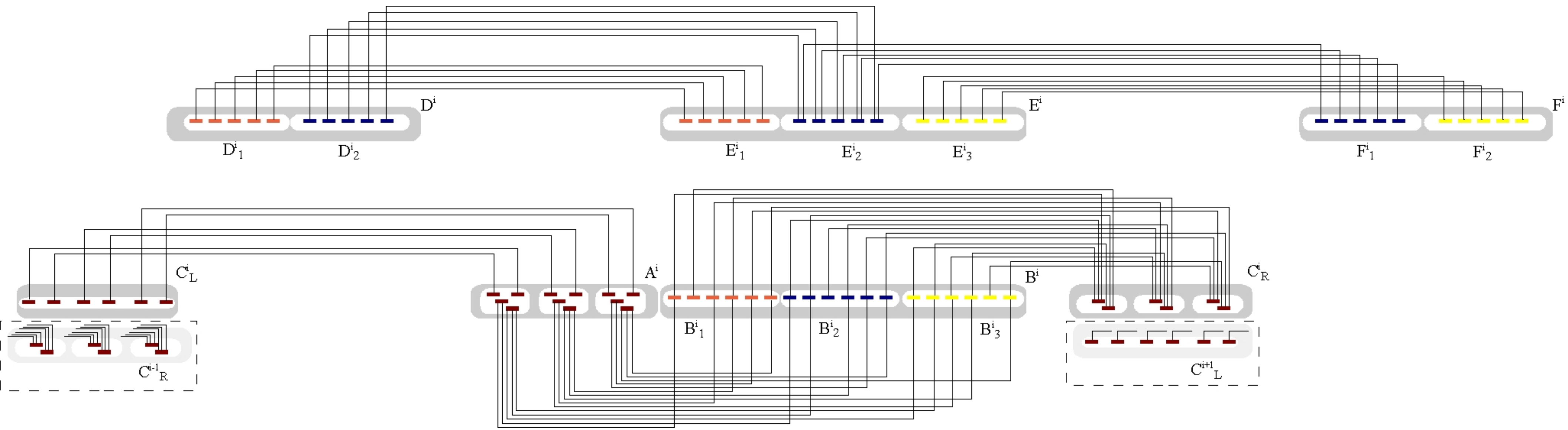


Known to be a ***Difficult problem*** → It is at least as difficult

2-IP-UNI-{<, ≬} is NP-complete

- **Exact 3-CNF SAT problem.**
  - BF=C1 $\wedge$ C2 $\wedge$ C3 … Cq
  - V={x1,x2, … xn}

Ci= ($\overline{x1}$ v x2 v x3)



Outline

- Preliminaries
- Two polynomial time algorithms
- Np-completeness proof
  - 2-IP-UNI-$\{<, \between\}$ is NP-complete
  - 2-IP-UNI-$\{<, \between\}$ is fixed-parameter tractable

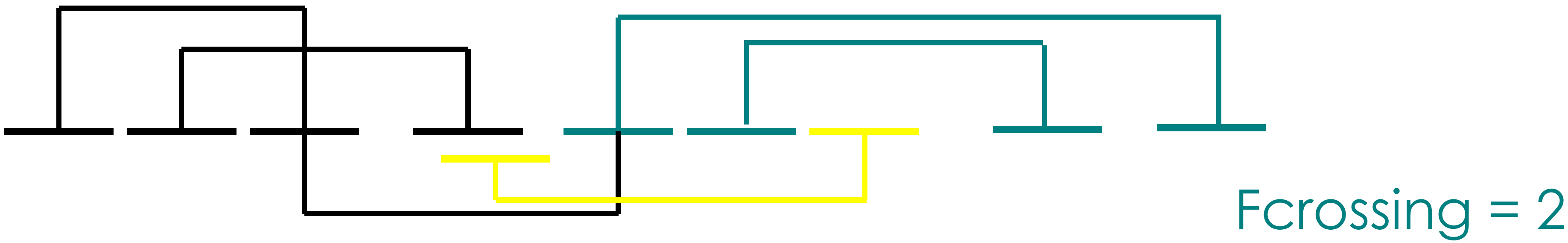
A fixed-parameter algorithm for 2-IP-UNI-{<, $\between$}

- An exact algorithm with strong emphasis on the crossing structure
- Solved for any support by means of dynamic programming in a complexity exponential only in a fixed parameter

A fixed-parameter algorithm for 2-IP-UNI-{<,≬}

- Forward Crossing Number of *D:*

$$FCrossing(D) = max_{D_i \in D} \left| D_j : D_i \between D_j \right|$$



A fixed-parameter algorithm for 2-IP-UNI-{<,≬}

- The problem can be solved by means of dynamic programming in $\mathbf{O(n^2.FCrossing(D).2^{FCrossing(D)}(\log(n)+FCrossing(D)))}$
- We expect Fcrossing(D) to be small compared to |D|

Conclusion

2-INTERVAL PATTERN PROBLEM			
	SUPPORT		
MODEL	UNLIMITED	UNITARY	DISJOINT
$\{<, \sqsubset, \between\}$	**NP**-complete		$O(n\sqrt{n})$[MV80]
$\{\sqsubset, \between\}$	**NP**-complete		$O(n^2\sqrt{n})$ ⋆
$\{<, \sqsubset\}$	$O(n^2)$		
$\{<, \between\}$	**NP**-complete ⋆		?
$\{<\}$	$O(n \log n)$		
$\{\sqsubset\}$	$O(n \log n)$ ⋆ •		
$\{\between\}$	$O(n^2 \log n)$		

Open Complexity

Is it fixed-parameter tractable with respect to parameter lower than FCrossing(D)?

Depth(D)=max {|D'| :D' is a $\{\between\}$-comparable subset of D}

CPM2004
Istanbul, Turkey

Questions on *New results for the 2-interval pattern problem*

Guillaume Blin[1], Guillaume Fertin[1] and Stéphane Vialette[2]

[1] LINA, FRE CNRS 2729
Université de Nantes,
2 rue de la houssinière
BP 92208 44322 Nantes Cedex 3
FRANCE
{blin,fertin}@lina.univ-nantes

[2] LRI, UMR CNRS 8623
Université Paris-Sud
Faculté des Sciences d'Orsay,
Bât 490, 91405 Orsay Cedex
FRANCE
vialette@lri.fr

lina LABORATOIRE D'INFORMATIQUE DE NANTES ATLANTIQUE