Query enumeration on sparse graphs

Barbizon 2016

Alexandre Vigny

June 7, 2016

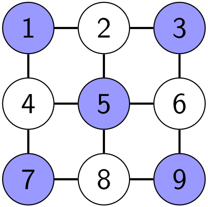
About me

- Alexandre Vigny
- I started my Ph.D in October 2015
- My advisors are :
  - Arnaud Durand (IMJ-PRG, Paris 7)
  - Luc Segoufin (LSV, ENS Cachan)

Introduction

- Query q — *small*
- Database D — *huge*
- Compute $q(D)$ — *gigantic*

Examples :

query q	database D	solutions $q(D)$
$q(x,y) := \exists z (B(x) \wedge E(x,z) \wedge \neg E(y,z))$		$\{(1{,}2)\ (1{,}3)\ (1{,}4)\ (1{,}6)\ (1{,}7) \cdots (3{,}1)\ (3{,}2)\ (3{,}4)\ (3{,}6)\ (3{,}7) \cdots \cdots \}$

Enumeration

Input : $\|D\| := n$ & $\|q\| := k$ $(k \ll n)$

Goal : output solutions one by one

- STEP 1: Preprocessing

  Prepare the enumeration : Database $D \longrightarrow$ Index I

  Preprocessing time : $f(k) \cdot n \rightsquigarrow O(n)$

- STEP 2 : Enumeration

  Enumerate the solutions : Index $I \longrightarrow \overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4}, \cdots$

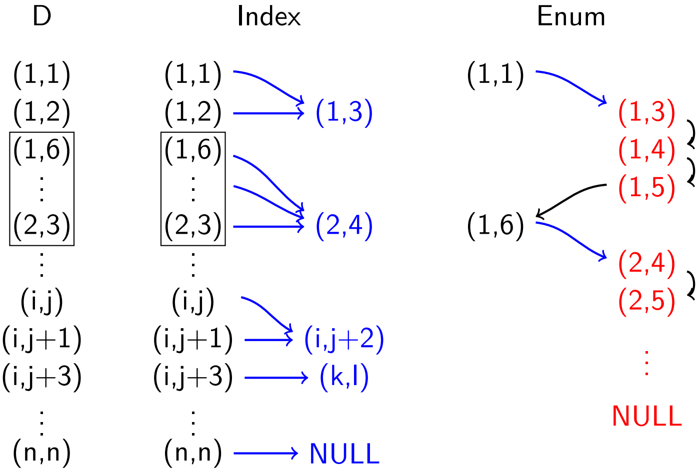
  Delay : $O(f(k)) \rightsquigarrow O(1)$

Constant delay enumeration after linear preprocessing $(CD \circ Lin)$

Example

Database $D := \langle \{1, \cdots, n\}; E \rangle \qquad \|D\| = |E| \quad (E \subseteq D \times D)$

Query $q(x, y) := \neg E(x, y)$



Enumeration VS Model-Checking

Model-Checking	:	Is there a solution ? Yes / No
Constant-delay Enumeration	:	First solution computed in time $O(n)$

Constant-delay Enumeration $\Longrightarrow$ Linear Model-Checking

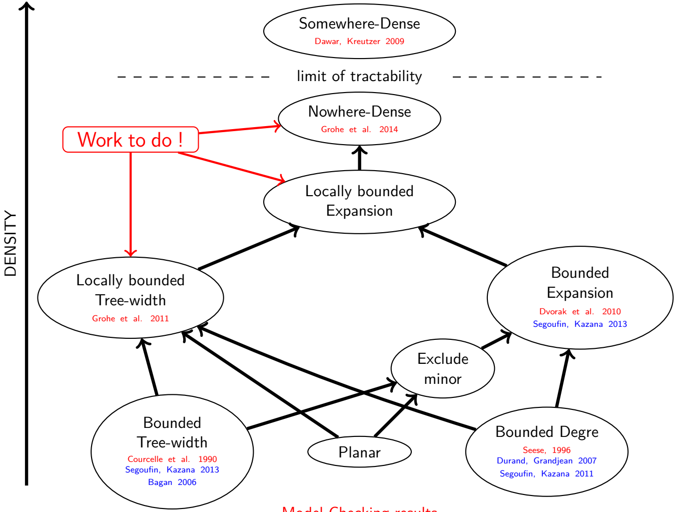
Under some complexity hypothesis, the Model-Checking is not doable in polynomial time.

$\Downarrow$

Restricted databases or/and queries

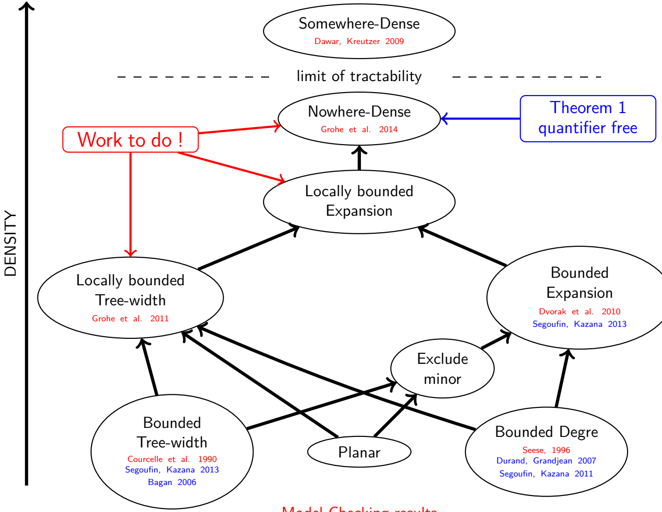
Bonded degree, planar $\cdots$	MSO, quantifier free $\cdots$
Nowhere-dense	**First Order**

Classes of graphs closed under taking sub-graphs



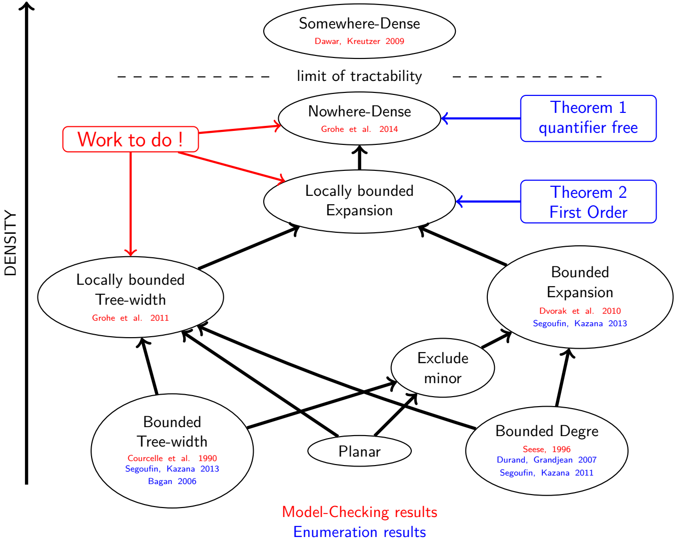
Model-Checking results
Enumeration results

Classes of graphs closed under taking sub-graphs



Model-Checking results
Enumeration results

Classes of graphs closed under taking sub-graphs



First results

Theorem 1

The enumeration of **quantifier free** first-order queries over **nowhere dense** class of graphs is in $CD \circ Lin$.

Theorem 2

The enumeration of **first-order** queries over class of graphs with **locally bounded expansion** is in $CD \circ Lin$.

Future work

- Generalize the Nowhere-Dense case.

- Enumeration with update.

Thank you !