Borne inférieure de circuit : une application des expanders

Simone Bova[1] Florent Capelli[2] Friedrich Slivovski[1]
Stefan Mengel[3]

[1]TU Wien, [2]IMJ-PRG, Paris 7, [3]CRIL, Lens

6 Novembre 2015

JGA 2015

Motivation

We have knowledge on a system, expressed as a list of constraints, a CNF :

$$F = \bigwedge_{i=1}^{n} \bigvee_{j} \ell_j \text{ where } \ell_j \in \{x, \neg x\} \text{ for some variable } x$$

We want to query F many times:

- Is F satisfiable? Is $F[x_1 \leftarrow 0, x_2 \leftarrow 1, x_3 \leftarrow 0]$ still satisfiable?
- How many assignments do satisfy $F[x_1 \leftarrow 0]$?
- etc.

Example: car configuration on the website of Renault.

Motivation

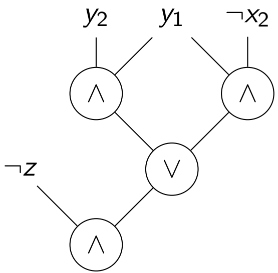
- **Problem**: All these queries are hard (NP or #P complete).
- **Strategy**: Compile F to an optimized data structure that support these queries in polynomial time.
- **Main idea**: Spend time (possibly exponential) only once to optimize and not for each query

Data structure: boolean circuits with good properties.

Which data structure?

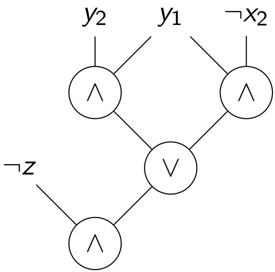
In this talk DNNF: *Decomposable Negation Normal Form*
A DNNF:

- a boolean circuit C with $\vee$ and $\wedge$ gates
- *Negation Normal Form:* inputs are labeled by x or $\neg x$ with x a variable
- *Decomposable:* For α an $\wedge$-gate whose inputs are α_1 et α_2, we have $\mathrm{var}(\alpha_1) \cap \mathrm{var}(\alpha_2) = \emptyset$



Remarks

- DNFs are DNNFs
- Stable by partially assigning variables
- One of the most general family of circuits that still supports interesting queries
  - Satisfiability in linear time
  - Enumeration of satisfying assignments with linear delay
  - Existential quantification of a subset of variables



Questions: upper bounds

Question (Upper bounds)

How can we use the structure of a formula to compile it in FPT*-time?*

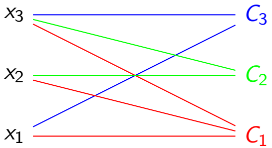


Figure: Incidence graph

- Which parameters are relevent?
- Close to the parametrized complexity of #SAT.

Questions: Lower bounds

In this talk:

Question (Lower bounds)

Can we transform every CNF-formula F into a DNNF *of polynomial size in $|F|$?*

The answer is no:

- A $2^{\Omega(\sqrt{|F|})}$ lower can be deduced from known lower bound on monotone circuits
- **In this talk**: we use expanders to get a $2^{\Omega(|F|)}$ lower bound on an infinite family of CNF.

Graph formula and vertex covers

- Given a graph $G = (V, E)$, define $F_G = \bigwedge_{(x,y)\in E}(x \vee y)$
- Satisfying assignment of F = vertex covers of G
- $S \subseteq V$: $\mathrm{VC}(G, S)$ = vertex covers C of G such that $S \subseteq C$

Key theorem:

Theorem

Let G be a graph of degree d and $\mu_d = (1 + 2^{-d}) > 1$:

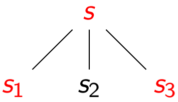
$$\#\mathrm{VC}(G, S) \leq \mu_d^{-|S|} \#\mathrm{VC}(G)$$

$\rightarrow$ if S is big, $\mathrm{VC}(G, S)$ is exponentially smaller than $\mathrm{VC}(G)$

Proof of the key theorem

For $S = \{s\}$, $N_s = \mathrm{neighbors}(s)$, $|N_s| = d$:

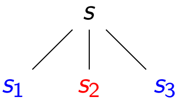
- $\#\mathrm{VC}(G) = \#\mathrm{VC}$ that contain s + $\#\mathrm{VC}$ that do not contain s
- Transform a VC C containing s to one which do not.
  Remember $C \cap N_S$



Proof of the key theorem

For $S = \{s\}$, $N_s = \mathrm{neighbors}(s)$, $|N_s| = d$:

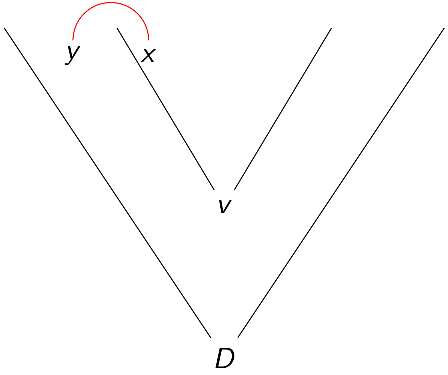
- $\#\mathsf{VC}(G) = \#\mathsf{VC}$ that contain s + $\#\mathsf{VC}$ that do not contain s
- Transform a VC C containing s to one which do not.
  Remember $C \cap N_S$



- From this and $C \cap N_s$, one can reconstruct C
- $\#\mathsf{VC}$ containing $s \leq 2^d \times$ $\#\mathsf{VC}$ that don't
- $(1 + 2^{-d})\#\mathsf{VC}(G, \{s\}) \leq \#\mathsf{VC}(G)$
- For $|S| > 1$, induction.

Proof strategy

Let $G = (V, E)$ be a graph $(x, y) \in E$, D a DNNF for F_G, $v \in D$ such that:

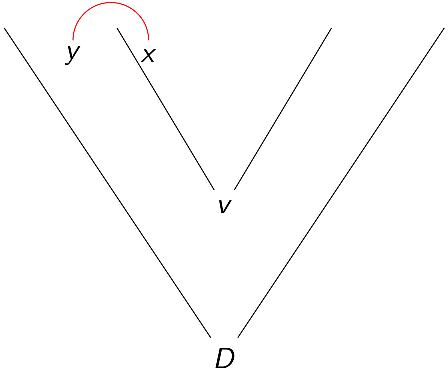


Solutions of D_v and D:

- must assign x or y to 1 (otherwise, not a solution of F)

Proof strategy

Let $G = (V, E)$ be a graph $(x, y) \in E$, D a DNNF for F_G, $v \in D$ such that:

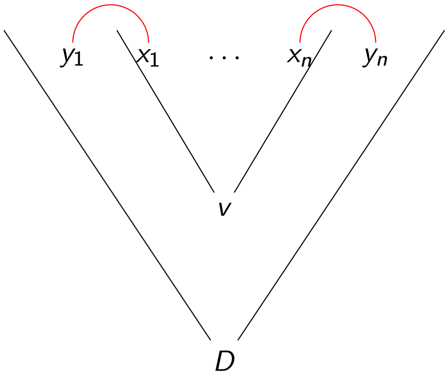


Solutions of D_v and D:

- must assign x or y to 1 (otherwise, not a solution of F)
- Actually they either *all* assign x to 1 or *all* y to 1

How to find such gates

$(x_1, y_1), \dots, (x_n, y_n)$ an induced matching of G and v a gate such that:



One can always find $S \subseteq X \cup Y$ of size n such that each solution of v must contain S

“Proof”

1. Choose G wisely
2. Greedily look for a gate v with enough variables in subcircuit: roughly $|V|/2$
3. Extract large induced matching S from $\text{var}(v)$ to $V \setminus \text{var}(v)$

“Proof”

1. Choose G wisely
2. Greedily look for a gate v with enough variables in subcircuit: roughly $|V|/2$
3. Extract large induced matching S from $\text{var}(v)$ to $V \setminus \text{var}(v)$
4. All solutions of D_v fixe the same value to a large number of variables

“Proof”

1. Choose G wisely
2. Greedily look for a gate v with enough variables in subcircuit: roughly $|V|/2$
3. Extract large induced matching S from $\mathrm{var}(v)$ to $V \setminus \mathrm{var}(v)$
4. All solutions of D_v fixe the same value to a large number of variables
5. Solutions of D_v = exponentially smaller than the solutions of D

"Proof"

1. Choose G wisely
2. Greedily look for a gate v with enough variables in subcircuit: roughly $|V|/2$
3. Extract large induced matching S from $\text{var}(v)$ to $V \setminus \text{var}(v)$
4. All solutions of D_v fixe the same value to a large number of variables
5. Solutions of $D_v =$ exponentially smaller than the solutions of D
6. Disconnect v : it removes a small fraction of solutions
7. Go to 2 until you have removed all gates

“Proof”

1. Choose G wisely
2. Greedily look for a gate v with enough variables in subcircuit: roughly $|V|/2$
3. Extract large induced matching S from $\mathrm{var}(v)$ to $V \setminus \mathrm{var}(v)$
4. All solutions of D_v fixe the same value to a large number of variables
5. Solutions of $D_v =$ exponentially smaller than the solutions of D
6. Disconnect v : it removes a small fraction of solutions
7. Go to 2 until you have removed all gates

Expanders

- **Goal**: ensure that there is always a large induce matching between $W \subseteq V$ of size roughly $|V|/2$ and $(V \setminus W)$ in G
- **Boundary expansion**: $G = (V, E)$ is a (c, d)-expander iff
  - it is of degree d and
  - for each $W \subseteq V$, if $\frac{|V|}{d} \leq |W| \leq \frac{|V|}{2}$ then $\partial W = |N_W \setminus W| \geq c|W|$.
- **Bounded degree + expansion**: one can find large induced matching from subset of variables W of size roughly $|V|/2$ to $V \setminus W$

Theorem

There exists a familly of CNF formulas $(F_n)_{n\in\mathbb{N}}$ such that $|\mathrm{var}(F_n) = n|$ and every DNNF computing F_n is of size $2^{\Omega(n)}$.

Trying to explain old lower bounds

- Known lower bounds of this kind are usually of the form $2^{\Omega(\sqrt{|F|})}$
- Most examples are based on $(n \times n)$ matrices or grids
- In grids, large subsets of variables have a boundary of size roughly $\sqrt{N}$ where $N = n^2$ is the number of variables
- Expander is a way of having a linear size boundary and allows us to lift lower bounds

Conclusion

- We prove a strong exponential lower bound on some family of circuits representing a very restricted class of CNF formulas (2-CNF, monotone, read 3)
- Closes open questions in the domain of knowledge compilation (Marquis, Darwich, 2002)
- Can we find other lower bounds using these kind of techniques?