Formalizing Randomized Matching Algorithms

Dai Tri Man Lê and Stephen Cook

Department of Computer Science
University of Toronto
Canada

LICS 2011

Two Aspects of Proof Complexity

1. Propositional Proof Complexity *(Pitassi's invited talk)*
   - the lengths of proofs of tautologies in various proof systems
2. Bounded Arithmetic
   - the power of weak formal systems to prove theorems of interest in computer science

- (1) and (2) are related by "propositional translations"
  - a proof in theory $T \rightsquigarrow$ uniform short proofs in propositional system P_T
  - bounded arithmetic = uniform version of propositional proof complexity
- "bounded": induction axioms are restricted to bounded formulas

1. Propositional Proof Complexity *(Pitassi's invited talk)*
   - the lengths of proofs of tautologies in various proof systems
2. Bounded Arithmetic
   - the power of weak formal systems to prove theorems of interest in computer science

- (1) and (2) are related by "propositional translations"
  - a proof in theory $T \rightsquigarrow$ uniform short proofs in propositional system P_T
  - bounded arithmetic $=$ uniform version of propositional proof complexity
- "bounded": induction axioms are restricted to bounded formulas

Complexity Theory	Bounded Arithmetic
Classify problems according to complexity classes	Classify theorems according to the computational complexity of concepts needed to prove them. "Bounded Reverse Mathematics" [Cook-Nguyen '10]
Separate (or collapse) complexity classes	Separate (or collapse) formal theories for various complexity classes

Complexity Theory	Bounded Arithmetic
Classify problems according to complexity classes	Classify theorems according to the computational complexity of concepts needed to prove them. "Bounded Reverse Mathematics" [Cook-Nguyen '10]
Separate (or collapse) complexity classes	Separate (or collapse) formal theories for various complexity classes

Feasible reasoning with VPV

The VPV theory

- associated with complexity class P (polytime)
- universal theory based on Cook's theory *PV* ('75)
- with symbols for all polytime functions and their defining axioms based on Cobham's Theorem ('65).
- Induction on polytime predicates: a derived result via binary search.
- Proposition translation: polynomial size extended Frege proofs

The VPV theory

- associated with complexity class P (polytime)
- universal theory based on Cook's theory *PV* ('75)
- with symbols for all polytime functions and their defining axioms based on Cobham's Theorem ('65).
- Induction on polytime predicates: a derived result via binary search.
- Proposition translation: polynomial size extended Frege proofs

Proofs in *VPV* are **feasibly constructive**.

- Given a proof in *VPV* for the formula $\forall X \exists Y \varphi(X, Y)$, where φ represents a polytime predicate, we can extract a polytime function $F(X)$ and a correctness proof in *VPV* of $\forall X \varphi(X, F(X))$.
- Induction is restricted to polytime "concepts".

Feasible proofs

Polytime algorithms usually have feasible correctness proofs, e.g.,

- the "augmenting-path" algorithm: finding a maximum matching
- the Hungarian algorithm: finding a minimum-weight matching
- ...

(formalized in *VPV*, see the full version on our websites)

Feasible proofs

Polytime algorithms usually have feasible correctness proofs, e.g.,

- the "augmenting-path" algorithm: finding a maximum matching
- the Hungarian algorithm: finding a minimum-weight matching
- ...

(formalized in *VPV*, see the full version on our websites)

Main Question

How about randomized algorithms and probabilistic reasoning?

"Formalizing Randomized Matching Algorithms"

Two fundamental randomized matching algorithms

1. RNC^2 algorithm for testing if a bipartite graph has a perfect matching (Lovász '79)
2. RNC^2 algorithm for finding a perfect matching of a bipartite graph (Mulmuley-Vazirani-Vazirani '87)

Recall that:

$$\boxed{\begin{array}{c}\text{Log-Space} \subseteq \mathsf{NC}^2 \subseteq \mathsf{P} \\ \mathsf{RNC}^2 \subseteq \mathsf{RP}\end{array}}$$

Important Remark

The two algorithms above also work for general undirected graphs, but we only consider bipartite graphs.

Two fundamental randomized matching algorithms

1. RNC^2 algorithm for testing if a bipartite graph has a perfect matching (Lovász '79)
2. RNC^2 algorithm for finding a perfect matching of a bipartite graph (Mulmuley-Vazirani-Vazirani '87)

Recall that:

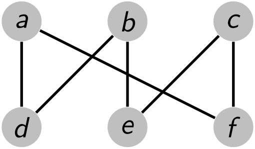
$$\text{Log-Space} \subseteq \mathsf{NC}^2 \subseteq \mathsf{P}$$
$$\mathsf{RNC}^2 \subseteq \mathsf{RP}$$

Important Remark

The two algorithms above also work for general undirected graphs, but we only consider bipartite graphs.

Problem:

Given a bipartite graph G, decide if G has a perfect matching.



$$\begin{array}{c c} & \begin{array}{ccc} d & e & f \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \end{array}$$

replace ones with distinct variables $\rightsquigarrow$

$$M_G = \begin{bmatrix} x_{11} & 0 & x_{13} \\ x_{21} & x_{22} & 0 \\ 0 & x_{32} & x_{33} \end{bmatrix}$$

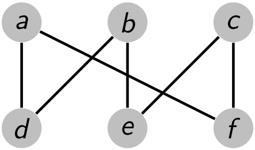
Edmonds' Theorem (provable in VPV)

G has a perfect matching if and only if $\mathrm{Det}(M_G)$ is not identically zero.

Lovász's Algorithm

Problem:

Given a bipartite graph G, decide if G has a perfect matching.



$$\begin{array}{c c} & \begin{array}{ccc} d & e & f \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \end{array} \xrightarrow{\text{replace ones with distinct variables}} M_G = \begin{bmatrix} x_{11} & 0 & x_{13} \\ x_{21} & x_{22} & 0 \\ 0 & x_{32} & x_{33} \end{bmatrix}$$

Edmonds' Theorem (provable in VPV)

G has a perfect matching if and only if $\mathrm{Det}(M_G)$ is not identically zero.

The usual proof is not feasible since...

it uses the formula $\mathrm{Det}(M) = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i=1}^{n} M(i, \sigma(i))$, which has $n!$ terms.

$$\begin{array}{c|ccc} & d & e & f \\ \hline a & 1 & 0 & 1 \\ b & 1 & 1 & 0 \\ c & 0 & 1 & 1 \end{array} \quad \xrightarrow[\text{distinct variables}]{\text{replace ones with}} \quad M_G = \begin{bmatrix} x_{11} & 0 & x_{13} \\ x_{21} & x_{22} & 0 \\ 0 & x_{32} & x_{33} \end{bmatrix}$$

Edmonds' Theorem (provable in VPV)

G has a perfect matching if and only if $\mathrm{Det}(M_G)$ is not identically zero.

$$\begin{array}{c|ccc} & d & e & f \\ \hline a & 1 & 0 & 1 \\ b & 1 & 1 & 0 \\ c & 0 & 1 & 1 \end{array} \quad \overset{\text{replace ones with distinct variables}}{\rightsquigarrow} \quad M_G = \begin{bmatrix} x_{11} & 0 & x_{13} \\ x_{21} & x_{22} & 0 \\ 0 & x_{32} & x_{33} \end{bmatrix}$$

Edmonds' Theorem (provable in VPV)

G has a perfect matching if and only if $\mathrm{Det}(M_G)$ is not identically zero.

- **Observation:** instance of the polynomial identity testing problem
- $\mathrm{Det}(M_G^{n\times n})$ is a polynomial in n^2 variables x_{ij} with degree at most n.
  - $\mathrm{Det}(M_G)$ is called the *Edmonds' polynomial* of G.

Edmonds' Theorem (provable in VPV)

G has a perfect matching if and only if $\mathrm{Det}(M_G)$ is not identically zero.

- **Observation:** instance of the polynomial identity testing problem
- $\mathrm{Det}(M_G^{n\times n})$ is a polynomial in n^2 variables x_{ij} with degree at most n.
  - $\mathrm{Det}(M_G)$ is called the *Edmonds' polynomial* of G.

Lovász's RNC^2 Algorithm

- Pick n^2 random values r_{ij} from $S = \{0, \ldots, 2n\}$
- If $\mathrm{Det}(M_G)(\vec{r}) = 0$ then YES ($\mathrm{Det}(M_G) \equiv 0$) else NO.

Edmonds' Theorem (provable in VPV)

G has a perfect matching if and only if $\mathrm{Det}(M_G)$ is not identically zero.

- **Observation:** instance of the polynomial identity testing problem
- $\mathrm{Det}(M_G^{n\times n})$ is a polynomial in n^2 variables x_{ij} with degree at most n.
  - $\mathrm{Det}(M_G)$ is called the *Edmonds' polynomial* of G.

Lovász's RNC^2 Algorithm

- Pick n^2 random values r_{ij} from $S = \{0, \ldots, 2n\}$
- If $\mathrm{Det}(M_G)(\vec{r}) = 0$ then YES ($\mathrm{Det}(M_G) \equiv 0$) else NO.

1. if $\mathrm{Det}(M_G) \equiv 0$, then $\mathrm{Det}(M_G)(\vec{r}) = 0$
2. if $\mathrm{Det}(M_G) \not\equiv 0$, then $\Pr_{\vec{r}\in_R S^{n^2}}\left[\mathrm{Det}(M_G)(\vec{r}) \neq 0\right] \geq 1/2$

((2) follows from the Schwartz-Zippel Lemma)

Obstacle #1 - Talking about probability

- Given a polytime predicate $A(X, R)$,

$$\Pr_{R \in \{0,1\}^n}\big[A(X,R)\big] = \frac{|\{R \in \{0,1\}^n \mid A(X,R)\}|}{2^n}$$

- The function $F(X) := \big|\big\{R \in \{0,1\}^n \mid A(X,R)\big\}\big|$ is in #P.
- #P problems are generally harder than NP problems

Obstacle #1 - Talking about probability

- Given a polytime predicate $A(X, R)$,

$$\Pr_{R\in\{0,1\}^n}\big[A(X,R)\big] = \frac{|\{R \in \{0,1\}^n \mid A(X,R)\}|}{2^n}$$

- The function $F(X) := \big|\big\{R \in \{0,1\}^n \mid A(X,R)\big\}\big|$ is in $\#\mathrm{P}$.
- $\#$P problems are generally harder than NP problems

Solution [Jeřábek '04]

- We want to show $\Pr_{R\in\{0,1\}^n}\big[A(X,R)\big] \geq s/t$, it suffices to show

$$|\{R \in \{0,1\}^n \mid A(X,R)\}| \cdot t \geq 2^n \cdot s$$

- Key idea: construct in *VPV* a polytime surjection

$$G : \{R \in \{0,1\}^n \mid A(X,R)\} \times [t] \twoheadrightarrow \{0,1\}^n \times [s],$$

where $[m] := \{1, \ldots, m\}$.

Definition (Jeřábek 2004 – modified)

Let $\Gamma, \Delta \subseteq \{0,1\}^n$ be polytime definable sets. Define Γ is "larger" than Δ if there exists a polytime surjective function $F : \Gamma \twoheadrightarrow \Delta$.

A bit of history

A series of papers by Jeřábek (2004–2009) justifying and utilizing the above definition

- A very sophisticated framework
- Based on approximate counting techniques
- Related to the theory of derandomization and pseudorandomness
- Application: formalizing probabilistic complexity classes

The Schwartz-Zippel Lemma

Let $P(X_1, \ldots, X_n)$ be a non-zero polynomial of degree D over a field $\mathbb{F}$. Let S be a finite subset of $\mathbb{F}$. Then

$$\Pr_{\vec{R} \in S^n}\left[P(\vec{R}) = 0\right] \leq \frac{D}{|S|}.$$

Obstacle #2

- The usual proof assumes we can rewrite

$$P(X_1, \ldots, X_n) = \sum_{J=0}^{D} X_1^J \cdot P_J(X_2, \ldots, X_n)$$

- This step is not feasible when P is given as arithmetic circuit or symbolic determinant

The Schwartz-Zippel Lemma

Let $P(X_1, \ldots, X_n)$ be a non-zero polynomial of degree D over a field $\mathbb{F}$. Let S be a finite subset of $\mathbb{F}$. Then

$$\Pr_{\vec{R} \in S^n}[P(\vec{R}) = 0] \leq \frac{D}{|S|}.$$

Obstacle #2

- The usual proof assumes we can rewrite

$$P(X_1, \ldots, X_n) = \sum_{J=0}^{D} X_1^J \cdot P_J(X_2, \ldots, X_n)$$

- This step is not feasible when P is given as arithmetic circuit or symbolic determinant

Solution

- Being less ambitious: restrict to the case of Edmonds' polynomials
- Take advantage of the special structure of Edmonds' polynomials

Edmonds' polynomials

$$\begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} \begin{array}{ccc} d & e & f \end{array} \\ \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \end{array} \quad \xrightarrow[\text{distinct variables}]{\text{replace ones with}} \quad M_G = \begin{bmatrix} x_{11} & 0 & x_{13} \\ x_{21} & x_{22} & 0 \\ 0 & x_{32} & x_{33} \end{bmatrix}$$

Edmonds' matrix:

Useful observation:

- Each variable x_{ij} appears at most once in M_G.
- From the above example, by the cofactor expansion,

$$\mathrm{Det}(M_G) = -x_{33} \cdot \mathrm{Det}\begin{pmatrix} x_{11} & 0 \\ x_{21} & x_{22} \end{pmatrix} + \mathrm{Det}\begin{pmatrix} x_{11} & 0 & x_{13} \\ x_{21} & x_{22} & 0 \\ 0 & x_{32} & \mathbf{0} \end{pmatrix}$$

- Thus, we can apply the idea in the original proof.

Theorem (provable in VPV)

Assume the bipartite graph G has a perfect matching.

- Let $S = \{0, \dots, s\}$ be the sample set.
- Let $M_G^{n \times n}$ be the Edmonds' matrix of G.

Then we can construct polytime surjection

$$F : [n] \times S^{n^2-1} \twoheadrightarrow \{\vec{r} \in S^{n^2} \mid \mathrm{Det}(M_G)(\vec{r}) = 0\}.$$

- The degree of the polynomial $\mathrm{Det}(M_G)$ is at most n.
- The surjection F witnesses that

$$\Pr_{\vec{r} \in S^{n^2}}\left[\mathrm{Det}(M_G)(\vec{r}) = 0\right] = \frac{\left|\left\{\vec{r} \in S^{n^2} \mid \mathrm{Det}(A)(\vec{r}) = 0\right\}\right|}{s^{n^2}} \leq \frac{n}{s}$$

The Mulmuley-Vazirani-Vazirani Algorithm

- RNC^2 algorithm for finding a perfect matching of a bipartite graph
- Key idea: reduce to the problem of finding a **unique** min-weight perfect matching using **the isolating lemma**.

Obstacle

The isolating lemma seems too general to give a feasible proof.

Solution

Consider a specialized version of the isolating lemma.

Lemma

Given a bipartite graph G. Assume the family $\mathcal{F}$ of all perfect matchings of G is nonempty. If we assign random weights to the edges, then

$$\Pr\Big[\textit{the min-weight perfect matching is unique}\Big] \textit{ is high.}$$

Summary

Main motivation

Feasible proofs for randomized algorithms and probabilistic reasoning:
"Formalizing Randomized Matching Algorithms"

Summary

Main motivation

Feasible proofs for randomized algorithms and probabilistic reasoning:
"Formalizing Randomized Matching Algorithms"

We demonstrate the techniques through two randomized algorithms:

1. RNC^2 algorithm for testing if a bipartite graph has a perfect matching [Lovász '79]
   - the Schwartz-Zippel Lemma for Edmonds' polynomials
2. RNC^2 algorithm for finding a perfect matching of a bipartite graph [Mulmuley-Vazirani-Vazirani '87]
   - a specialized version of the isolating lemma for bipartite matchings.

Take advantage of special linear-algebraic properties of
Edmonds' matrices and Edmonds' polynomials

Open questions

1. Can we prove in VPV more general version of the Schwartz-Zippel lemma? *(We only considered Edmonds' polynomials.)*
2. Can we do better than VPV, e.g., VNC^2 [Cook & Nguyen '10]?

Open problems and future work

Open questions

1. Can we prove in *VPV* more general version of the Schwartz-Zippel lemma? *(We only considered Edmonds' polynomials.)*
2. Can we do better than *VPV*, e.g., VNC^2 [Cook & Nguyen '10]?

Future work

1. How about RNC^2 matching algorithms for undirected graphs?
   - Use properties of the **pfaffian**
   - Need to generalize results from [Soltys '01] [Soltys & Cook '02] (with Cook and Fontes)
2. Use Jeřábek's techniques to formalize constructive aspects of fundamental theorems that require probabilistic reasoning.
   - Cryptography: the Goldreich-Levin Theorem, construction of pseudorandom generator from one-way functions, etc. (with George)
   - Moser-Tados constructive proof of Lovász Local Lemma (with Filmus)