The Complexity of Directed Graph Connectivity (Draft Outline)

Bobby (Zelin) Lv

UW-Madison, zlv7@wisc.edu

December 8, 2019

Abstract

We survey the current results of directed graph connectivity. We will focus on completeness reusults, the algorithms of graph connectivity problems, developments in understanding the complexity of the graph connectivity problem in different computational models and remaining open problems.

Contents

1 Introduction

This is the draft outline. It is far from complete and only a beginning outline.

We survey the current results of directed graph connectivity. We will focus on completeness reusults, the algorithms of graph connectivity problems, developments in understanding the complexity of the graph connectivity problem in different computational models and remaining open problems.

For the graph connectivity algorithms, we survey the results about space complexity obtained on graphs with different properties, including planar graphs [Asa+14; CT17], grid graphs, mangroves [AL98; Gar+14], surface-embedded graphs [SV12] and unique-path graphs [KKR08].

For the lower bound side, we approach this problem through circuit computation model and communication computation model. The results for circuit computation cover results on Small Distance Connectivity [Che+16; Ros14; BIP98], and Monotone Circuit, including Span Program [PR17] and Switching Network Model [Pot17]. On the models involving communication complexity, we survey the results on Data Structure Lower Bound for both Dynamic [FS89] and Static [Pa11] setting of graph connectivity problems, as well as Stream Model [SW15].

2 Preliminaries

2.1 Some Notation

Definition 2.1. **L**: Deterministic log space Turing Machine

Definition 2.2. **NL**: Non-Deterministic log space Turing Machine

Definition 2.3. **Parity L**: Non-Deterministic log space Turing Machine with determining if the input if and only if the number of possible computation paths is odd

Definition 2.4. **SL**: Symmetric log-space Turing Machines

Definition 2.5. **SC**: Polylog-space, polynomial-time Turing Machine

Definition 2.6. **USPACE**$(\log n)$: Deterministic log space Turing Machine if, for every input, there exists at most one accepting computation.

Clearly, $DSPACE(\log n) \subseteq USPACE(\log n) \subseteq NSPACE(\log n)$

Definition 2.7. **RUSPACE**$(\log n)$: Deterministic log space Turing Machine if, for every pair of configurations that are reachable from the start configuration, there exists at most one computational path connecting these configurations.

Definition 2.8. **StUSPACE**$(\log n)$: Deterministic log space Turing Machine if, for every pair of configurations, there exits at most one computational path connecting these configurations.

$StUSPACE(\log n) \subseteq RUSPACE(\log n) \subseteq USPACE(\log n)$

2.2 Circuit Complexity

Boolean Circuits are the computation model that take Boolean inputs and compute with a sequence of operations, which can seen as gate. Normally, Boolean Circuits are composed with *de Morgan* Basis, which are $\{\wedge, \vee, \neg\}$. Unlike algorithms, which can take inputs for any arbitrary length, one Boolean Circuit can only handle one input length. Therefore, algorithms or Turing Machines are uniform computation model and Circuits are non-uniform.

2.3 Communication Complexity

Communication Model is a 2-argument function, $f : A \times B \to Z$, In computing, one input $a \in A$ is given to Alice and another input $b \in B$ is given to Bob. They need compute the $f(a, b)$ together to obtain the result $z \in Z$. We measure the communication complexity by how many bit they exchange during the computation.

2.3.1 Data Structure Lower Bound

For data structure lower bound, we use the cell-probe model. In this model, we have one part known as 'the space' and another part as computation. The running time of an operation is measured by the number of cell probes. All computation on already probed cells are free.

2.4 Information Theory

Definition 2.9. Information Entropy is the average rate at which information is produced by a random source of data. For a discrete random variable X whose domain is $\{x_1, x_2, ..., x_n\}$, the information entropy of X is $H(x) = -\Sigma_{i=1}^{n} Pr(X = x_i) \log_2 Pr(X = x_i)$.

3 Completeness Results

Theorem 3.1. *STCONN is complete for* ***NL****.*

So as we have shown in the section of Savitch's Theorem, we have $STCONN \in \text{DSPACE}(log^2 n)$

Theorem 3.2. *USTCONN is complete for* ***SL****.*

Theorem 3.3. *CYCLE is complete for* ***L****.*

4 Directed Graphs Connectivity Problems Algorithms

The first problem in this area is can we improve Savitch's Bound. [Sav70] proved that the reachability problem over directed graph can be deterministically solved in space of $O(\log^2 n)$. This result implies for polynomial space bounds, nondeterminism does not add any additional power to determinism. This is bound is still the current best upper bound known. In order to improve this result, previous researches have focused on graph with different properties.

4.1 Planar Graphs

A planar graph is a graph with the property that we can draw it on the plane such that its edges only interest at their endpoints.

For reachability problem in planar graph, we have the following result:

Theorem 4.1. *[Asa+14] There exists an algorithm that decides directed planar graph reachability in polynomial time and $\widetilde{O}(\sqrt{n})$ space.*

Furthermore, for graph directed layered planar graphs, we have the following result:

Theorem 4.2. *[CT17] For every $\epsilon > 0$, there is a polynomial time and $O(n^\epsilon)$ space algorithm that decides reachability in directed layered planar graphs.*

4.2 Grid Graphs

Theorem 4.3. *There exists an algorithm that computes the grid graph reachability in polynomial-time and $\widetilde{O}(n^{1/3})$ space.*

There is one work in arxiv claims they have obtained $O(n^{1/4+\epsilon})$.

4.3 Mangroves

G is *unambiguous* if there is at most one path from 1 to n as $N(1,n) \leq 1$. G is called *strongly unambiguous* or a *Mangrove* if for any pair (x,y) of nodes there is at most one path leading from x to y, as $\forall_{x,y \in B}, N(x,y) \leq 1$. Although a mangrove does not need to be a tree, for each x the subgraph of G induced by $T(x)$ is indeed a tree and the same is true for the set of all nodes from which x can be reached. G is called *reach-unambiguous* if $N(1,i) \leq 1$ for each i.

Theorem 4.4. *[AL98] $RUSPACE(\log n) \subseteq DSPACE(\log^2 n / \log\log n)$.*

The computation class ReachFewL is the class of computational tasks that are decidable by nondeterministic log-space machines, where for every input, there are at most polynomially many computation paths from the start configuration to any other configuration.

Theorem 4.5. *[Gar+14] ReachFewL=UL.*

4.4 Surface-Embedded Graphs

In this section, we focus on designing reachability algorithms that improve Savitch's Bound for directed graphs with some topological structure, which are graphs that are embedded on topological surfaces.

Theorem 4.6. *[SV12] There is a log-space reduction that, given an instance $< G, s, t >$ (presented as a combinatorial embedding) where $G \in \Gamma(m, g)$ and s, t are vertices of G, outputs an instance $< G', s', t' >$ where G' is a directed graph and s', t' vertices of G', so that*

- *there is a directed path from s to t in G if and only if there is a directed path from s' to t' in G',*
- *G' has $O(m + g)$ vertices.*

4.5 Unique-Path Graphs

Unique-Path Graphs are graphs such that to a source vertex s if there is at most one simple path from s to any vertex $v \in V(G)$.

Theorem 4.7. *[KKR08] For any $\epsilon \in (0, 1]$, STCON in unique-path graphs is solvable with $O(\frac{n^\epsilon}{\epsilon})$ space in $n^{O(\frac{1}{\epsilon})}$ time.*

5 Circuit Complexity

Theorem 5.1. *uniform* $\mathsf{NC}^1 \subseteq \mathsf{L}$

Theorem 5.2. $\mathsf{NL} \subseteq$ *uniform* $\mathsf{AC}^1 \subseteq$ *uniform* NC^2

5.1 Small Distance Connectivity

Most of new results on proving the lower bound of graph connectivity problem study the small distance connectivity problem. Small distance connectivity problems determine whether an n-vertex graph has an s-to-t path of length at most k. The current best result of this problem is near-optimal [Che+16]. [BIP98] and [Ros14] also gave some results on small distance connectivity problem.

Theorem 5.3. *[Che+16] For any $k(n) \leq n^{1/5}$ and any $d = d(n)$, any depth-d circuit computing STCONN(k(n)) must have size $n^{\Omega(k^(1/d)/d)}$. Furthermore, for any $k(n) \leq n$ and any $d = d(n)$, any depth-d circuit computing STCONN(k(n)) must have size $n^{\Omega(k^{1/5d}/d)}$.*

[Ros14] shows a tight lower bound on the size of bounded-depth formulas solving distance $k(n)$ connectivity.

Theorem 5.4. *[Ros14] Formulas of depth $\log n/(\log\log n)^{O(1)}$ solving STCONN($k(n)$) have size $n^{\Omega(\log k)}$ for all $k(n) \leq \log\log n$.*

Corollary 5.5. *Polynomial-size circuits solving STCONN($k(n)$) require depth $\Omega(\log k)$ for all $k(n) \leq \log\log n$.*

5.2 Monotone Circuit

5.2.1 Span Program

[PR17] shows the current best lower bound of size of monotone span program. Furthermore, lower bound for st-connectivity implies a quasipolynomial separation between $m\mathsf{NC}^2$ and monotone span programs, since st-connectivity is well-known to be computable.

Theorem 5.6. *[PR17] The st-connectivity function requires $n^{\Omega(\log n)}$ size monotone span programs over $\mathbb{R}$.*

The proof of this theorem also requires Communication Complexity.

5.2.2 Switching Network Model

[Pot17] analyzes the monotone space of complexity of directed connectivity for a large class of input graphs G using the switching network model. This works generalizes and improves the results obtained in the previous works in switching network model.

Theorem 5.7. *[Pot17] If z and m are constants such that $m \leq \frac{n}{2000z^4}$ and G is a directed acyclic input graph such that*

- *There is no path of length at most 2^{z-1} from s to t.*

- *For any vertex $v \in V(G)$, there are at most m vertices $w \in V(G)$ such that either there is a path of length at most 2^{z-2} from v to w in G or there is a path of length at most 2^{z-2} from w to v in G, then*

$$m(G) \geq \frac{(9mn)^{\frac{1}{4}}}{20|E(G)|(z+1)\sqrt{2^z z!}(\frac{n}{9m})^{\frac{z}{4}}}$$

6 Communication Complexity

6.1 Data Structure Lower Bound

6.1.1 Static Connectivity in Sparse Directed Graphs

Communication complexity can be used to prove that such data structure for directed graphs doesn't exist [Pa11]:

Theorem 6.1. *[Pa11] In any static data structure solving the directed graph connectivity problem on graphs with at most nw edges and word size w, we must have*

$$t \geq \Omega(\frac{\log n}{\log \frac{sw \log n}{n \log w}})$$

If $t = O(1)$ and $w = O(\log n)$ then $s = n^{1+\Omega(1)}$.

6.1.2 Lower bound for Dynamic Graph Connectivity

The dynamic data structure of graph connectivity supports addition of edges and connectivity queries. It maintains a graph on the vertex set $[n]$. [FS89] shows the following lower bound for dynamic graph connectivity:

Theorem 6.2. *[FS89] Any data structure solving the graph connectivity problem with error at most $1/3$ for $m \geq n+1$ operations satisfies*

$$t_q \times \log t_u w \geq \omega(\log n)$$

6.2 Streams Model

In streams model, [SW15] shows tight $\Omega(n \log n)$ space lower bounds for randomized algorithms which succeed with constant probability in a stream of edge insertions for a number of graph problems.

7 Open Problems

8 Acknowledgements

This is the draft outline of Bobby (Zelin) Lv's Senior Honors Thesis, under guidance of Professor Dieter van Melkebeek.

References

[AL98] Eric Allender and klaus-joern Lange. “RUSPACE(log n) $\subseteq DSPACE(log2n/loglogn)$”. In: *Theory of Computing Systems* 31 (Oct. 1998). DOI: 10.1007/s002240000102.

[Asa+14] Tetsuo Asano, David C. Kirkpatrick, Kotaro Nakagawa, and Osamu Watanabe. “Õ (n)-Space and Polymomial-time Algorithm for the Planar Directed Graph Reachability Problem”. In: 2014.

[BIP98] P. Beame, Russell Impagliazzo, and Toniann Pitassi. “Improved depth lower bounds for small distance connectivity”. In: *Computational Complexity* 7 (Jan. 1998), pp. 325–345. DOI: 10.1007/s000370050014.

[Che+16] Xi Chen, Igor C. Oliveira, Rocco A. Servedio, and Li-Yang Tan. “Near-optimal Small-depth Lower Bounds for Small Distance Connectivity”. In: *Proceedings of the Forty-eighth Annual ACM Symposium on Theory of Computing*. STOC ’16. Cambridge, MA, USA: ACM, 2016, pp. 612–625. ISBN: 978-1-4503-4132-5. DOI: 10.1145/2897518.2897534.

[CT17] Diptarka Chakraborty and Raghunath Tewari. “An O(N&Epsi;) Space and Polynomial Time Algorithm for Reachability in Directed Layered Planar Graphs”. In: *ACM Trans. Comput. Theory* 9.4 (Dec. 2017), 19:1–19:11. ISSN: 1942-3454. DOI: 10.1145/3154857.

[FS89] M. Fredman and M. Saks. “The Cell Probe Complexity of Dynamic Data Structures”. In: *Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing*. STOC ’89. Seattle, Washington, USA: ACM, 1989, pp. 345–354. ISBN: 0-89791-307-8. DOI: 10.1145/73007.73040.

[Gar+14] Brady Garvin, Derrick Stolee, Raghunath Tewari, and N. V. Vinodchandran. “ReachFewL = ReachUL”. In: *Comput. Complex.* 23.1 (Mar. 2014), pp. 85–98. ISSN: 1016-3328. DOI: 10.1007/s00037-012-0050-8.

[KKR08] Sampath Kannan, Sanjeev Khanna, and Sudeepa Roy. “STCON in directed unique-path graphs”. In: *Leibniz International Proceedings in Informatics, LIPIcs* 2 (Jan. 2008). DOI: 10.4230/LIPIcs.FSTTCS.2008.1758.

[Pa11] Mihai Patraşcu. “Unifying the Landscape of Cell-Probe Lower Bounds”. In: *SIAM J. Comput.* 40.3 (June 2011), pp. 827–847. ISSN: 0097-5397. DOI: 10.1137/09075336X.

[Pot17] Aaron Potechin. “Bounds on Monotone Switching Networks for Directed Connectivity”. In: *J. ACM* 64.4 (Aug. 2017), 29:1–29:48. ISSN: 0004-5411. DOI: 10.1145/3080520.

[PR17] Toniann Pitassi and Robert Robere. “Strongly Exponential Lower Bounds for Monotone Computation”. In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2017. Montreal, Canada: ACM, 2017, pp. 1246–1255. ISBN: 978-1-4503-4528-6. DOI: 10.1145/3055399.3055478.

[Ros14] Benjamin Rossman. “Formulas vs. Circuits for Small Distance Connectivity”. In: *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*. STOC ’14. New York, New York: ACM, 2014, pp. 203–212. ISBN: 978-1-4503-2710-7. DOI: 10.1145/2591796.2591828.

[Sav70] Walter J. Savitch. “Relationships Between Nondeterministic and Deterministic Tape Complexities”. In: *J. Comput. Syst. Sci.* 4.2 (Apr. 1970), pp. 177–192. ISSN: 0022-0000. DOI: 10.1016/S0022-0000(70)80006-X.

[SV12] D. Stolee and N. V. Vinodchandran. "Space-Efficient Algorithms for Reachability in Surface-Embedded Graphs". In: *2012 IEEE 27th Conference on Computational Complexity*. June 2012, pp. 326–333. DOI: 10.1109/CCC.2012.15.

[SW15] Xiaoming Sun and David P. Woodruff. "Tight Bounds for Graph Problems in Insertion Streams". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2015)*. Ed. by Naveen Garg, Klaus Jansen, Anup Rao, and José D. P. Rolim. Vol. 40. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2015, pp. 435–448. ISBN: 978-3-939897-89-7. DOI: 10.4230/LIPIcs.APPROX-RANDOM.2015.435.