The Power of a Pebble: Exploring and Mapping Directed Graphs

Michael A. Bender* Antonio Fernández† Dana Ron‡ Amit Sahai§ Salil Vadhan¶

Abstract

Exploring and mapping an unknown environment is a fundamental problem that is studied in a variety of contexts. Many works have focused on finding efficient solutions to restricted versions of the problem. In this paper, we consider a model that makes very limited assumptions about the environment and solve the mapping problem in this general setting.

We model the environment by an unknown directed graph G, and consider the problem of a robot exploring and mapping G. We do not assume that the vertices of G are labeled, and thus the robot has no hope of succeeding unless it is given some means of distinguishing between vertices. For this reason we provide the robot with a "pebble" — a device that it can place on a vertex and use to identify the vertex later.

In this paper we show: (1) If the robot knows an upper bound on the number of vertices then it can learn the graph efficiently with only *one* pebble. (2) If the robot does not know an upper bound on the number of vertices n, then $\Theta(\log\log n)$ pebbles are both necessary and sufficient. In both cases our algorithms are deterministic.

1 Introduction.

The problem of exploring and mapping an unknown environment is a fundamental problem with applications ranging from robot navigation to searching the World Wide Web. As such, a large body of work has focused on finding efficient solutions to variants of the problem, with restrictive assumptions on the form of the environment (*cf.* [13, 12, 16, 23, 14, 27, 7, 4, 1].) In this paper, we consider a model that makes very limited assumptions about the environment, and give efficient algorithms to solve the mapping problem in this general setting.

A natural way to model the problem is by a robot exploring a graph $G = (V, E)$. The case in which the graph has both undirected edges and labeled vertices can be solved in time linear in the number of edges by depth first search. Other search techniques [22] improve on this bound by a constant factor. Unfortunately, many exploration and mapping problems do not satisfy these constraints. For instance, if the graph represents a city (having one-way streets) or the Internet, it contains directed edges. This alone does not make the problem substantially more difficult, since the problem with directed edges and labeled vertices can be solved by a greedy search algorithm in time $O(|V| \cdot |E|)$. More sophisticated techniques [16, 1] yield improved running times.

Regardless of whether there are directed edges, a more daunting difficulty arises if vertices are not uniquely labeled. This situation could arise from the limited sensory capabilities of a robot or from the changing appearance of vertices. If no assumptions are made on the labeling of the vertices (so that all vertices may appear the same), then we need a way to mark vertices in order to have any hope of mapping the environment. In this paper, we model a marking device by a pebble, which can be dropped at a vertex and later identified and retrieved. This notion of marking is basic and can be simulated in many situations. It can be shown that a robot provided with a pebble can map an *undirected* graph with unlabeled vertices in time $O(|V| \cdot |E|)$, by repeatedly marking nodes and backtracking.[1] However, without the assumptions that the edges are undirected and the vertices are labeled, the existence of an efficient algorithm has remained open.

The main contribution of this paper is a general mapping algorithm. This algorithm efficiently solves the mapping problem without assuming unique labelings of the vertices and while allowing directed edges.

The problem. Let G be a strongly-connected directed graph over n vertices, where the vertices have no labels. The outdegree of each vertex is d,[2] and the outgoing edges at each vertex are numbered from '1' to 'd'.[3] The vertices' indegrees

*Division of Engineering and Applied Sciences, Harvard University, Cambridge MA 02138. Email: bender@deas.harvard.edu. Supported by NSF grants CCR-95-04436 and CCR-93-13775.

†Laboratory for Computer Science, MIT. Current address: Dpto de Arquitectura y Tecnología de Computadores, Universidad Politécnica de Madrid. Email: anto@eui.upm.es. Supported by the Spanish Ministry of Education, Army grant DAAH04-95-1-0607, and ARPA contract N00014-95-1-1246.

‡Laboratory for Computer Science, MIT, 545 Technology Square, Cambridge, MA 02139. Email: danar@theory.lcs.mit.edu. Supported by a Bunting fellowship.

§Laboratory for Computer Science, MIT, 545 Technology Square, Cambridge, MA 02139. Email: amits@theory.lcs.mit.edu. Supported by a DOD NDSEG doctoral fellowship and partially by DARPA grant DABT63-96-C-0018.

¶Laboratory for Computer Science, MIT, 545 Technology Square, Cambridge, MA 02139. Email: salil@math.mit.edu. Supported by a DOD NDSEG doctoral fellowship and partially by DARPA grant DABT63-96-C-0018.

[1] In addition to undirected edges and labeled vertices, other simplifying assumptions that can be made about the environment include geometric structure (such as planarity) or random access (as on the World Wide Web).

[2] In fact, identical outdegrees is the worst case, and with minor modifications, our algorithms work for graphs having arbitrary degrees.

[3] Without some way of distinguishing edges it is not clear how to reach one vertex from another even given a map of the graph. The assumption that the edges emanating from a vertex are numbered is a *local* (and weak) assumption, as opposed to a *global* assumption

are not assumed to be seen. The robot is placed at an arbitrary starting vertex in G, and at each step it traverses one of the edges emanating from its current vertex. The robot's task is to explore and map G efficiently. That is, after walking a polynomial number of steps (in the size of the graph), it should output a graph $\widehat{G}$ isomorphic to G. However, unless the robot has a tool to help it distinguish vertices, it is condemned to failure as a cartographer. For example, a robot traveling alone cannot decide whether G consists of a single vertex or many vertices. A basic tool for the robot is a *pebble*. Now, as the robot explores G, it can *mark* a vertex by dropping the pebble, and it can *identify* the vertex if it finds the pebble later. Upon finding the pebble, the robot can pick it up. However, because the graph is directed, the robot cannot retrace its steps to retrieve the pebble.

Bender and Slonim [7] show that a robot given a pebble can explore and map any graph in *exponential* time. However, they prove that a robot cannot map graphs in *polynomial* time using a constant number of pebbles, when it does not know a bound on n. This lower bound motivates two questions: (1) How many pebbles are needed to learn graphs efficiently if n is known? (2) How many pebbles are in fact needed if n is unknown?

In this paper we demonstrate that surprisingly few pebbles are needed in both cases. We show that

- If the robot knows n (or an upper bound $\hat{n}$ on n), it can learn the graph with only *one* pebble in time polynomial in n (respectively, $\hat{n}$).
- If the robot does not know n (or $\hat{n}$), then $\Theta(\log \log n)$ pebbles are both necessary and sufficient.

In both cases our algorithms are deterministic. The lower bound of $\Omega(\log \log n)$ for the case of unknown n holds even for probabilistic algorithms.

Intuition. To understand the difficulties facing the exploring robot, consider the problem of *traversing* a graph (i.e., visiting all vertices and edges). Certainly, in order to map a graph, the robot must traverse it. One standard technique that comes to mind is *random walks*. Unfortunately, for *directed* graphs, the expected time until a random walk visits all vertices may be exponential in n and random walks are therefore ineffective for traversing. (For undirected graphs the expected time is polynomial in n.)

Consider, for example, the graph in Figure 1. This graph is called a *combination lock graph*, because in order to reach the rightmost node v_n, the robot must discover the unique sequence of edge labels (the combination) extending from v_1 to v_n. Notice that in polynomial time, with very high probability, a random walk only visits a logarithmic number of vertices in the combination lock. More generally, for any polynomial time (randomized) algorithm that does not mark vertices, there exists a combination lock graph that (with high probability) the algorithm will not fully explore.

We now return to the problem of learning with a pebble. Although one (pebbleless) robot cannot traverse combination locks (efficiently), a robot with a pebble can learn them using random walks [7].[4] However, consider the graph shown in Figure 2. This graph consists of two combination locks, where the end of one combination lock leads into the beginning of the other. If the robot ever drops its pebble in the top lock and travels into the bottom lock, then it is doomed. The robot will be stuck in the bottom combination lock without its pebble, and cannot traverse this lock to learn it. Notice that once the robot has lost its pebble, knowing the size of the graph trapping it is not helpful.

This example illustrates the dilemma facing the robot as it explores the unknown graph G. The robot *must* drop the pebble in order to learn new terrain, but when the robot drops the pebble, it runs the risk of *losing* it.

Closed paths. To avoid losing its pebble, the robot must know how to return to it. Thus, before dropping the pebble at a vertex, the robot should know a closed path containing this vertex. However, such a path may be difficult to obtain. When n is unknown, the robot can only identify a closed path by dropping the pebble and finding it again. Thus, we encounter a chicken-and-egg situation. In order to safely drop the pebble, the robot must find a closed path. But in order to find a closed path, the robot must drop its pebble.

Now we recognize the tangible benefit of knowing n. By repeating the same pattern of edges n times, the robot can enter a closed path *without* dropping its pebble. For example, if the robot repeatedly follows edges labeled '1', it enters a cycle after at most n moves. Once the robot knows a closed path, it can map the subgraph visited by the path using the pebble. However, it is not clear how to harness this additional power. By repeating one pattern of edges, the robot enters a closed path and can map one subgraph. Later, the robot may repeat a different pattern of edges, enter another closed path, and map a second subgraph. Thus, the robot can map many subgraphs, but it is not obvious how to piece these maps together. This is because the robot has little information about how the subgraphs overlap and interconnect. As a result, finding closed paths permits the robot to drop the pebble, map a (small) portion of the graph and retrieve the pebble, but does not solve the mapping problem.

In order to solve the mapping problem, we use an algorithmic tool that we call an *orienting procedure*. An orienting procedure allows our algorithms to construct a limited number of maps. Instead of trying to piece these maps together, the algorithm expands them separately until one maps all of G. This expansion is possible because by executing the orienting procedure, the robot can recognize particular vertices in the graph that are associated with the maps.

Orienting procedures. Intuitively, an orienting procedure for a graph G leads the robot around the graph and ultimately leaves the robot at a vertex it "recognizes". The robot recognizes this vertex by observing the output produced by the procedure. More precisely, if the robot sees the same output in two different executions of the procedure, then both times it ends up at the same vertex.[5] The notion of orienting procedures is analogous to the notion of (*adaptive*) *homing sequences* in automata theory [20], and it is closely related to the notion of *two-robot homing sequences* introduced by Bender and Slonim [7]. In the context of learning, homing sequences were first applied by Rivest and Schapire [27, 26]; they were used for learning environments modeled by finite automata.

We show that given an orienting procedure, the robot can build maps of subgraphs containing the ending vertices of the procedure. Since the robot is not provided with

that the vertices are labeled.

[4] More generally, graphs having high *conductance* can be learned efficiently [7].

[5] Actually, the robot may be at vertices equivalent under automorphism, but we avoid this issue in the introduction.

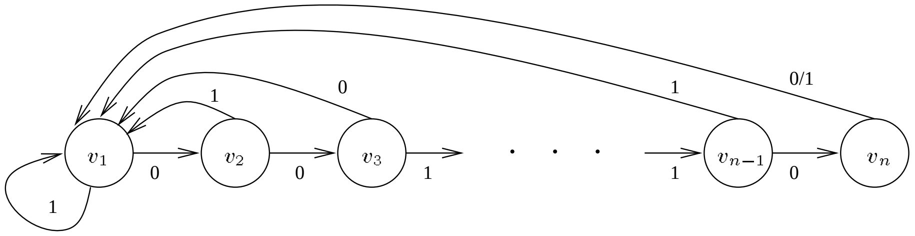

Figure 1: A combination lock graph.

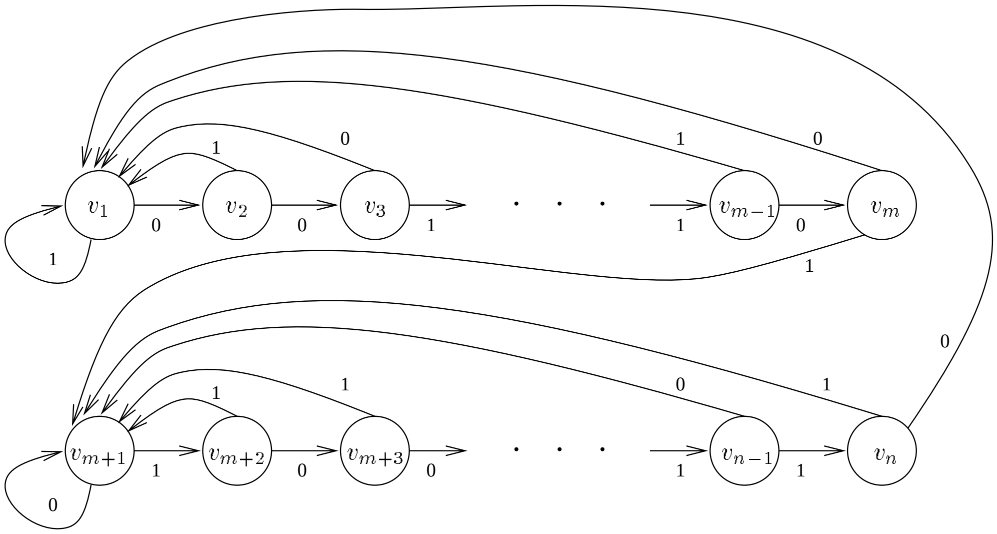

Figure 2: A Graph consisting of Two Combination Locks.

an orienting procedure, it builds maps using a partially-constructed orienting procedure, which it gradually improves. Each map is associated with a different output of the procedure. There is a difficulty, however, in using a partial orienting procedure. Namely, the underlying graph may look different from what the map associated with the procedure's output suggests. As a result, the robot could become disoriented and lose the pebble.

A central idea in our algorithms is how to avoid losing the pebble while using misleading information about the graph. The algorithms employ a two-tiered structure of the cycling technique mentioned above. At the lower level, the robot uses the cycling technique to verify safely whether the underlying graph is consistent with its map. If verification fails the robot is able to improve the partial orienting procedure. At the higher level, the robot uses a generalization of the cycling technique to arbitrary deterministic procedures (instead of edge-label patterns). This generalized cycling technique allows the robot to find closed paths that visit increasingly large portions of G, until all of G is visited and mapped.

Extensions. Our results generalize to the case in which the observed labeling of the edges at the robot's current vertex is a function of the robot's previous vertex. This models the situation that arises when navigating in a city, where the relative location of the streets exiting an intersection is determined by the direction from which the intersection was entered. Some intuition is given in Section 5.

Related work. Our work is most directly related to the work of Bender and Slonim [7]. Bender and Slonim show that two cooperating robots can explore and map unknown directed graphs with unlabeled vertices in polynomial time. The robots do not require any prior knowledge of the size of the graph. Bender and Slonim demonstrate that two robots are strictly more powerful than one robot with O(1) pebbles. They prove that one robot with a constant number of pebbles cannot (efficiently) learn arbitrary directed graphs without knowing an upper bound on the number n of vertices. They conjecture that the same holds when n is known; our results disprove this conjecture. Our $O(\log\log n)$-pebble algorithm (for unknown n) can be simulated by two robots. This yields a deterministic alternative to Bender and Slonim's randomized two-robot algorithm.[6]

Most early work on graph exploration assumed that the robot is a finite automaton. Rabin [24] first proposed the idea of providing the automaton with pebbles to help it explore. This led to a body of work examining the number of pebbles needed to explore various environments [29, 13, 12, 3, 25]. For a survey on automata exploring labyrinths, see [21]. Deng and Papadimitriou [16] propose and study the problem of exploring an unknown directed graph having labeled vertices. Albers and Henzinger [1] give improved al-

[6] In light of our results and those of Bender and Slonim, we see that a friend is only worth $\log\log n$ pebbles.

gorithms for this problem. These works study exploration from the perspective of competitive analysis. The results are stated in terms of the *deficiency* of the graph (i.e., the minimum number of edges to be added to make the graph Eulerian). Betke, Rivest, and Singh [9] and together with Awerbuch [4] study the problem of *piecemeal* learning undirected labeled graphs. In the piecemeal learning problem the robot is required to return to its starting position periodically.

Rivest and Schapire [27, 26] study the problem of learning environments modeled by finite automata. Here, an environment is represented by a directed graph, in which each vertex has one of two (or constant) possible labelings. The robot has learned the environment (automaton) when it can predict the label of any vertex (state) reached on an arbitrary walk. Hence, if the automaton is irreducible, then the robot actually learns the topology of the underlying graph. Their algorithms (with the exception of one, for permutation automata) rely on a teacher. The teacher supplies counterexamples to the robot's hypotheses. Variants of this problem that do not rely on a teacher are studied in the following works [14, 18, 28, 17]. We note that Dean et. al. [14] apply a cycling technique related to ours, but for different purposes.

Exploring and navigating in geometric environments is studied extensively. A sample of papers includes [5, 23, 15, 11, 6, 10, 8, 19, 2].

2 Preliminaries

Let $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ be the unknown directed graph the robot has to explore and map. Suppose that the graph is *strongly connected* and that all the vertices of G are unlabeled and have (the same) outdegree d. Let the edges emanating from each vertex be labeled by distinct indices in $\{1, \ldots, d\}$ and denote an edge from u to v with label σ by $\langle u, \sigma, v\rangle$. Let $n = |\mathrm{V}|$ and let $\hat{n}$ be an upper bound on n.

The exploring robot starts at an arbitrary vertex of G. The robot has a single pebble.[7] At each time step, the robot may traverse any outgoing edge from the vertex it is at. In addition, the robot may *drop the pebble* at the vertex. or *pick up the pebble* that it has previously dropped at the vertex.

We say that a graph $\mathrm{M} = (\mathrm{V_M}, \mathrm{E_M})$ is *isomorphic* to G (denoted, $\mathrm{M} \cong \mathrm{G}$) if there exists an isomorphism between the two graphs that preserves edge labels. Namely, there exists a one-to-one and onto mapping $f : \mathrm{V_M} \to \mathrm{V}$, such that the following holds: For every two vertices w and z in $\mathrm{V_M}$, there is an edge labeled σ from w to z in M, if and only if there is an edge labeled σ from $f(w)$ to $f(z)$ in G. Let w_0 and v_0 be distinguished vertices in M and G, respectively. We use the notation $(\mathrm{M}, w_0) \cong (\mathrm{G}, v_0)$ to say that there exists an isomorphism f between M and G such that $f(w_0) = v_0$. We say that (M, w_0) is *consistent* with (G, v_0) if there exists a subgraph G' of G containing v_0, such that $(\mathrm{M}, w_0) \cong (\mathrm{G}', v_0)$.

We say that the robot at vertex v in G has *learned* the graph G when it outputs a graph $\widehat{\mathrm{G}}$ together with a vertex $\hat{v}$ in $\widehat{\mathrm{G}}$ such that $(\widehat{\mathrm{G}}, \hat{v}) \cong (\mathrm{G}, v)$. Since in each time step the robot traverses a single edge, the *running time* of the algorithm is the number of moves the robot makes. Though computation time is ignored in this definition, we note that the total computation time of our algorithms is polynomial in the size of the graph.

[7] In Section 4 we consider a robot having a source of pebbles.

3 Learning with a Single Pebble

In this section we present our algorithm for learning efficiently any graph using a single pebble and knowledge of $\hat{n}$. We start (in Section 3.1) by describing an important subroutine of our algorithm, which we call *path compression*. The robot executes this subroutine (using the pebble) to map subgraphs of G that are visited by closed paths known to the robot. In Section 3.2 we show that the robot can learn G if we assume the robot has access to a *return-path oracle* for G. The robot can query this oracle from any vertex in the graph and receive a sequence of edges that leads it back to its start vertex. In the following sections we progressively weaken this assumption. In Section 3.3 we formally define an *orienting procedure* and describe how to devise such a procedure based on procedures for *distinguishing* between vertices. In Section 3.4 we replace the assumption that the robot has access to a return-path oracle with the assumption that it knows an orienting procedure for G. Finally, in Section 3.5 we show how the robot can use knowledge of $\hat{n}$ to explore and learn the graph while building an orienting procedure on its own.

3.1 Compressing Closed Paths

Here we present an essential building block of our algorithms. Let the robot be at vertex v in G, and assume the robot knows a closed path in G that starts (and ends) at v. The path visits a subgraph $\mathrm{G}_{\mathtt{path}}$ of G. Namely, $\mathrm{G}_{\mathtt{path}}$ consists of all vertices and edges traversed along the path. Since the path may visit the same vertices several times, $\mathrm{G}_{\mathtt{path}}$ is not necessarily a simple cycle. In the path compression procedure the robot uses the pebble to identify repeated vertices on the path and construct a graph M isomorphic to $\mathrm{G}_{\mathtt{path}}$.

More precisely, let $\mathtt{path} = \sigma_1, \ldots, \sigma_k$ be a sequence of edge labels corresponding to a closed path starting (and ending) at v. Let $u_0, u_1, \ldots, u_k$ be the vertices in G visited along the path, where $u_0 = u_k = v$. The robot maintains a list of length $k+1$ where ultimately the i-th entry in the list will identify the i-th vertex occurring on the path in G. Initially the list is $(w_0, \Lambda, \ldots, \Lambda, w_0)$, where Λ means "unidentified." The goal of the robot is to replace all "unidentified" entries with vertex names.

The algorithm proceeds in at most n stages, each starting and ending with the robot and the pebble at v. In the first stage, the robot drops the pebble at vertex v and follows the entire closed path; for each i such that the robot observes the pebble after i steps (i.e., at the vertex reached by traversing $\sigma_1, \ldots, \sigma_i$), the robot replaces the i-th entry in the list with w_0. In the j-th stage, let t be the smallest index such that the t-th entry in the list is Λ. The robot traverses $\sigma_1, \ldots, \sigma_t$, and after the t-th step drops the pebble at the vertex reached. Then it replaces the t-th entry with w_{j-1} (i.e., a new vertex name). As in the first stage, it traverses the rest of the closed path (and returns to v). For each i such that the robot observes the pebble after i steps (counting steps from when it left v), the robot replaces the i-th entry in the list (which must be a Λ) with w_{j-1}. After returning to v, the robot follows $\mathtt{path}$ once more to pick up the pebble.

The algorithm maintains the property that the same label w_j appears in places k and k' in the list if and only if the k-th and k'-th vertices on the closed path in G are the same. When the list is completed, the robot constructs a map M in accordance with the list and the edge labels in $\mathtt{path}$. Namely, the vertices of M are the vertices in the list, and if w_j and $w_{j'}$ appear in places i and $i+1$ in the list, then there is an edge $\langle w_j, \sigma_{i+1}, w_{j'} \rangle$ in M.

Lemma 1 *Let v be a vertex in G and* path *be a sequence of edge labels that corresponds to a closed path in* G *starting and ending at v. Let $\mathrm{G}_{\mathtt{path}}$ be the subgraph of* G *visited by* path*. The path compression procedure runs in time $O(n \cdot |\mathrm{path}|)$ and outputs a graph* M *such that $(\mathrm{M}, w_0) \cong (\mathrm{G}_{\mathtt{path}}, v)$.*

3.2 Learning with a Return-Path Oracle

In this section, we assume that the robot is given access to a *return-path oracle*. Namely, at any time step it can query the oracle and receive a sequence of edge labels that returns the robot to a particular vertex v_0.

We show how the robot can learn G by querying the oracle and using repeated applications of the path compression procedure. The return-path algorithm proceeds in at most $n \cdot d = |\mathrm{E}|$ stages. In each stage the robot learns at least one new edge in G. In the i-th stage, the robot constructs a strongly connected map M_i having a designated vertex w_0. The initial map, M_0, consists only of the vertex w_0 (and no edges). The final map is the output, $\widehat{\mathrm{G}}$, of the algorithm. The algorithm maintains the invariant that (M_i, w_0) is consistent with (G, v_0) (where consistency is defined in Section 2). The algorithm associates a closed path $\mathtt{path}(\mathrm{M}_i)$ with each map M_i. This path starts and ends at w_0 and passes through all vertices and edges in M_i. Since M_i is strongly connected, the robot can easily compute such a path of length $O(n^2 d)$.

We say that a vertex w in a map M_i is *finished* if it has d outgoing edges in M_i. Otherwise it is *unfinished*. In the $(i+1)$-th stage the algorithm augments the map M_i with a new edge emanating from an unfinished vertex in M_i and perhaps other vertices and edges. This is done as follows. Let w be an unfinished vertex in M_i and let σ be the label of a missing edge from w. Let $\mathtt{explore}(\mathrm{M}_i)$ be a sequence of edge labels connecting w_0 to w, concatenated with σ. The robot performs the walk corresponding to $\mathtt{explore}(\mathrm{M}_i)$ in G starting from v_0. It then queries the return-path oracle. Let the return path that the oracle provides be called $\mathtt{ret}_i$. The robot returns to v_0 using the path $\mathtt{ret}_i$. Then it compresses the closed path $\mathtt{path}_{i+1} = \mathtt{path}(\mathrm{M}_i) \circ \mathtt{explore}(\mathrm{M}_i) \circ \mathtt{ret}_i$. The algorithm lets M_{i+1} be the resulting map. By Lemma 1, we know that $(\mathrm{M}_{i+1}, w_0) \cong (\mathrm{G}_{\mathtt{path}_{i+1}}, v_0)$. Since $\mathtt{path}_{i+1}$ contains $\mathtt{path}(\mathrm{M}_i)$, M_{i+1} contains M_i as a subgraph; by the choice of w and σ, M_{i+1} also contains at least one new edge (the edge labeled σ going out of w).

Note that the time complexity of this algorithm can be improved. However, the above formulation serves as a basis for subsequent algorithms (that do not rely on a return-path oracle). From all the above, we obtain the following lemma.

Lemma 2 *Let ℓ be the length of the longest return path provided by the oracle. The return-path algorithm runs in time $O(n^2 d \cdot (n^2 d + \ell))$ and outputs a map $\widehat{\mathrm{G}}$ isomorphic to* G.

3.3 Orienting Procedures

Intuitively, an *orienting procedure* for a graph G guides the robot around the graph and ultimately leaves the robot at a vertex it "recognizes." We note that an orienting procedure does not lead the robot back to a *particular* vertex. Hence, assuming an orienting procedure is weaker than assuming a return-path oracle. Before we define an orienting procedure formally, we explain the notion of equivalence between vertices. We say that two vertices u and v in G are *equivalent*, denoted $u \equiv v$, if $(\mathrm{G}, u) \cong (\mathrm{G}, v)$, i.e., there exists an automorphism of G mapping u to v.

Definition 1 *An* orienting procedure **op** *for a graph* G *has the following properties.*

1. *It determines the robot's actions (i.e., what edge labels it traverses and when it drops and picks up the pebble).*
2. *The robots starts and ends with its pebble, regardless of the starting vertex.*
3. *The procedure is deterministic.*
4. *The procedure returns an* output*. The output is a function of when the robot sees the pebble.*

   Notice that because the procedure is deterministic, every time the robot executes the orienting procedure starting from any fixed vertex v in G, it returns the same output and finishes at the same final vertex. Thus, an orienting procedure has at most n outputs.
5. *Let* $\mathsf{output}(\mathbf{op}, v)$ *be the output of the procedure* **op** *when started at vertex v, and let* $\mathsf{final}(\mathbf{op}, v)$ *be the final vertex reached. An orienting procedure guarantees that for every u and v in* G $\mathsf{output}(\mathbf{op}, u) = \mathsf{output}(\mathbf{op}, v) \implies \mathsf{final}(\mathbf{op}, u) \equiv \mathsf{final}(\mathbf{op}, v)$.
   Note that the converse is not guaranteed. Namely, the procedure may end at the same vertex with two different outputs.

We show how to build an orienting procedure using *distinguishing procedures* for inequivalent vertices in G.

Definition 2 *Let u and v be two inequivalent vertices in* G*. A* distinguishing procedure $\mathbf{dp}_{u,v}$ *for u and v has the following properties.*

1–4. As in Definition 1.

5. $\mathsf{output}(\mathbf{dp}_{u,v}, u) \neq \mathsf{output}(\mathbf{dp}_{u,v}, v)$.

Notice that a distinguishing procedure differentiates between *starting* vertices whereas an orienting procedure differentiates between *ending* vertices. In addition, a distinguishing procedure differentiates between a single pair of starting vertices whereas an orienting procedure differentiates among all possible ending vertices.

Every orienting procedure **op** that we consider can be viewed as a tree $\mathrm{T}_{\mathbf{op}}$ in the following sense: Each leaf in $\mathrm{T}_{\mathbf{op}}$ corresponds to a different output of **op**. The internal nodes of $\mathrm{T}_{\mathbf{op}}$ are distinguishing procedures. The branches emitting from a node are labeled by the possible outputs of the distinguishing procedure. Leaves are labeled by the sequence of outputs on the branches leading from the root to the leaf. Consider all vertices in G that the robot may end at when **op** terminates with output A at a leaf ζ_{A}; denote this set of vertices by $\mathsf{reach}(\mathrm{A})$. Property 5 dictates that all vertices in $\mathsf{reach}(\mathrm{A})$ are equivalent.

We can build an orienting procedure of the above type in stages, extending the tree in each stage. Initially we let our *candidate* orienting procedure **cop** be the empty procedure, i.e. the robot makes no actions, and the tree $\mathrm{T}_{\mathbf{cop}}$ has a single leaf. Assume inductively that **cop** preserves properties *1–4* and has k possible outputs (so that $\mathrm{T}_{\mathbf{cop}}$ has k leaves). If **cop** is not yet a complete orienting procedure, then for some output A corresponding to leaf ζ_{A} there exist inequivalent vertices u and v in $\mathsf{reach}(\mathrm{A})$. Let $\mathbf{dp}_{u,v}$ be a distinguishing procedure for u and v. We replace the leaf ζ_{A}

with $\mathbf{dp}_{u,v}$. Since $\mathsf{output}(\mathbf{dp}_{u,v}, u) \neq \mathsf{output}(\mathbf{dp}_{u,v}, v)$, the new tree has at least $k+1$ leaves. Therefore, the modified **cop** has at least $k+1$ outputs. Since an orienting procedure has at most n different outputs, we obtain an orienting procedure after at most $n-1$ stages.[8] It can be shown that for every pair of inequivalent vertices there exists a distinguishing procedure with running time $O(n^3 d)$. Hence, every graph has an orienting procedure with running time $O(n^4 d)$. In Section 3.5, we exhibit an algorithm in which the robot devises distinguishing procedures and builds an orienting procedure while exploring the graph.[9]

3.4 Learning with an Orienting Procedure

In this section we assume that the robot has a prespecified *orienting procedure* **op** for the graph G. For ease of the presentation, we assume throughout this section that the graph has no automorphisms (and hence no vertices are equivalent). This assumption can easily be removed here and is not used in later sections.

By the above assumption, for each possible output A, the set $\mathsf{reach}(\mathrm{A})$ (defined in the Section 3.3) contains a single vertex, which we denote v_{A}. With each output A, the algorithm associates a map M(A), which is constructed as the algorithm proceeds. The map M(A) contains a designated vertex $w_0(\mathrm{A})$. The algorithm ensures that each M(A) is strongly connected and maintains the following invariant:

Invariant 1 (orienting procedure): *For every output* A *of* **op**, $(\mathrm{M(A)}, w_0(\mathrm{A}))$ *is consistent with* $(\mathrm{G}, v_{\mathrm{A}})$.

Learning proceeds in at most $n^2 d$ phases. In each phase, some map M(A) is augmented with at least one new edge. We say that a map is *finished* if all its vertices are finished. The algorithm terminates when some map M(A) is finished, in which case it outputs M(A). We use the shorthand `path`(A) to represent `path`(M(A)) and `explore`(A) to represent `explore`(M(A)), where `path`(·) and `explore`(·) were defined in Section 3.2. Let $\mathrm{G}_{\mathtt{path}(\mathrm{A})}$ be the subgraph of G visited by `path`(A) when starting from v_{A}. In each phase the algorithm uses the orienting procedure to find a closed path satisfying the following:

1. For some output A, the path starts and ends at v_{A}.
2. The path visits all of $\mathrm{G}_{\mathtt{path}(\mathrm{A})}$ and at least one additional edge.

The robot compresses this closed path and replaces M(A) with the resulting map.

To find a closed path satisfying the above properties the robot does the following. Starting from its current vertex, it executes the orienting procedure, observes its output A_1, and follows $\mathtt{path}(\mathrm{A}_1) \circ \mathtt{explore}(\mathrm{A}_1)$. It then executes the orienting procedure again, observes its output A_2, and follows $\mathtt{path}(\mathrm{A}_2) \circ \mathtt{explore}(\mathrm{A}_2)$. The robot repeats the above until it observes an output A_j that it has previously seen (i.e., $\mathrm{A}_j = \mathrm{A}_i$ for $i < j$). Note that some output must reappear after at most $n+1$ repetitions (though the robot need not know n). At this point the robot has discovered a closed path that starts and ends at v_{A_j}. Furthermore, this closed path starts with $\mathtt{path}(\mathrm{A}_i) \circ \mathtt{explore}(\mathrm{A}_i)$, and hence visits all of $\mathrm{G}_{\mathtt{path}(\mathrm{A}_i)}$ and at least one additional edge. Informally, since the robot does not know to which vertex it will return, it "prepares" all vertices v_{A_i} for the possibility. It does so by following $\mathtt{path}(\mathrm{A}_i) \circ \mathtt{explore}(\mathrm{A}_i)$ from each v_{A_i}.

Let T(**op**) be the running time of **op**. Since for every map M(A), $|\mathtt{path}(\mathrm{A})| = O(n^2 d)$, and $|\mathtt{explore}(\mathrm{A})| \leq n-1$, the length of the closed path found is $O(n \cdot (\mathrm{T}(\mathbf{op}) + n^2 d))$. By Lemma 1, the closed path can be compressed in time $O(n^2 \cdot (\mathrm{T}(\mathbf{op}) + n^2 d))$. We obtain the following lemma.

Lemma 3 *A robot with a single pebble can learn any strongly connected graph* G *using an orienting procedure* **op** *for* G *in time* $O(n^4 d \cdot (\mathrm{T}(\mathbf{op}) + n^2 d))$.

3.5 Learning the Graph while Building an Orienting Procedure

In this section we show that a robot having a single pebble can efficiently explore and map any strongly-connected directed graph if it knows an upper bound $\hat{n}$ on the size of the graph. Recall that if the robot does not know $\hat{n}$ then this task is impossible. The structure of the algorithm presented here is similar to the structure of the algorithm described in Section 3.4. Since the robot does not have a *real* orienting procedure it uses a *candidate* orienting procedure **cop**. In each phase, for some output A of **cop** the algorithm either (1) replaces M(A) with a new, larger M(A) or (2) discovers a distinguishing procedure $\mathbf{dp}_{u,v}$ for some inequivalent vertices u and v in $\mathsf{reach}(\mathrm{A})$. In the latter case it improves **cop** using $\mathbf{dp}_{u,v}$ (as described in Section 3.3). Since the improved **cop** will never again output A, the algorithm discards M(A). The algorithm terminates when some M(A) is finished, in which case it outputs M(A). We show that the algorithm maintains the following invariant, which is a relaxation of Invariant 1.

Invariant 2 (candidate orienting procedure): *For every output* A *of* **cop** *there exists a vertex* $u \in \mathsf{reach}(\mathrm{A})$ *such that* $(\mathrm{M(A)}, w_0(\mathrm{A}))$ *is consistent with* (G, u).

In particular this invariant ensures that the finished map is isomorphic to G.

In Section 3.4 we had the property that $\mathsf{reach}(\mathrm{A})$ consisted of a single vertex v_{A}. This provided a method for the robot to identify closed paths that start and end at some v_{A}. Here, this method does not work since $\mathsf{reach}(\mathrm{A})$ may contain several vertices (equivalent or inequivalent). Therefore, the robot could observe output A twice *without* being on a closed path. The robot's knowledge of $\hat{n}$ combined with the following observation suggests a remedy for this problem — that is, how to find a closed path that starts and ends at a vertex u in some $\mathsf{reach}(\mathrm{A})$.

Observation 1 *Let* $f : \mathrm{V} \rightarrow \mathrm{V}$ *be any deterministic function. Then for every vertex* $v \in \mathrm{V}$, *the sequence* $v, f(v), f(f(v)), \ldots$ *becomes cyclic within the first* n *applications of* f.

Suppose the robot repeats the following: it executes **cop**, observes its output A, and follows $\mathtt{path}(\mathrm{A}) \circ \mathtt{explore}(\mathrm{A})$. Then after at most $\hat{n}$ repetitions it has entered a cycle. We later show how after another $2\hat{n}$ repetitions it can find a closed path that starts and ends at a vertex u in $\mathsf{reach}(\mathrm{A})$, for some output A.

Suppose the robot runs the algorithm from the previous section with the enhancement above. The robot can now

[8] For the purposes of this construction, it actually suffices to relax the definition of a distinguishing procedure to allow *either* $\mathsf{output}(\mathbf{dp}_{u,v}, u) \neq \mathsf{output}(\mathbf{dp}_{u,v}, v)$ *or* $\mathsf{final}(\mathbf{dp}_{u,v}, u) \equiv \mathsf{final}(\mathbf{dp}_{u,v}, v)$.

[9] However, our algorithm may terminate (correctly) before the orienting procedure is complete.

find closed paths, but the algorithm still has a serious flaw. Consider a map M(A) that results from compressing a closed path that starts and ends at $u \in \mathsf{reach}(A)$. Assume that in a subsequent stage in the algorithm, the robot obtains a new M(A) by compressing a closed path that starts and ends at $u' \in \mathsf{reach}(A)$. If $u' \equiv u$ then the argument that the new M(A) is larger than the old M(A) holds as before. However, if $u' \not\equiv u$ then we can claim nothing about the size or structure of the new M(A). This is because (old $M(A), w_0(A)$) may not be consistent with (G, u'). Hence, the argument that the new M(A) is bigger than the old M(A) is no longer valid. This motivates the need for a *map verification procedure*.

Map Verification. Suppose the robot is at a vertex v in some $\mathsf{reach}(A)$. We would like a procedure to verify that $(M(A), w_0(A))$ is consistent with (G, v). This is not difficult if we allow the robot to lose its pebble. In particular the robot hypothesizes that path(A) corresponds to a closed path in G starting at v. Then the robot attempts to compress path(A). If path(A) is not a closed path starting from v and the robot loses the pebble, then clearly $(M(A), w_0(A))$ is not consistent with (G, v). Otherwise, the robot compares M(A) to the map resulting from compressing the closed path.

Since we cannot allow the robot to lose the pebble (or else it will not be able to learn the graph), we must modify the above procedure. The new procedure, described below, performs a weaker form of verification. We later show that it nonetheless meets the needs of the algorithm.

1. The robot starts from v and follows path(A) $\hat{n}$ times. Clearly, if $(M(A), w_0(A))$ is consistent with (G, v), then the robot ends at v. However, even if $(M(A), w_0(A))$ is not consistent with (G, v) then by Observation 1 we know that the robot has entered a cycle.
2. Next the robot drops the pebble at its current vertex v' and follows path(A) once.
   - If the pebble is not at the vertex reached, then verification fails. To retrieve the pebble, the robot continues repeating path(A) until it finds the pebble.
   - Otherwise, the robot compresses path(A), which it has now identified as a closed path, starting from v'. If the resulting map differs from M(A) then verification fails. Otherwise verification passes.

We refer to this procedure as **ver**(A).

Note 2 *There are two situations in which* **ver**(A) *passes:*

1. $(M(A), w_0(A))$ *is consistent with* (G, v)*, or*
2. $(M(A), w_0(A))$ *is not consistent with* (G, v)*, but* $(M(A), w_0(A))$ *is consistent with* (G, v')*.*

If verification fails, then because of Invariant 2 **ver**(A) is a distinguishing procedure. This procedure distinguishes between v and the vertex u in $\mathsf{reach}(A)$ such that $(M(A), w_0(A))$ is consistent with (G, u). Since for every map M(A), the length of path(M(A)) is $O(n^2 d)$, the running time of **ver**(A) is $O(\hat{n} \cdot n^2 d)$. We are now ready to describe the algorithm.

The Algorithm. The algorithm proceeds in at most $2n^2 d$ phases. Initially, its candidate orienting procedure **cop** is the empty procedure (as described in Section 3.3). In each phase:

1. To enter a closed path, the robot repeats the following $\hat{n}$ times.

   (∗) The robot executes **cop** and obtains an output A. If this is the first appearance of output A then the algorithm creates a new map M(A) consisting of a single vertex $w_0(A)$. Next the robot executes **ver**(A) to verify the map M(A).

   - If **ver**(A) fails, then **ver**(A) is a distinguishing procedure between a pair of vertices in $\mathsf{reach}(A)$. The robot uses this distinguishing procedure, which outputs PASS or FAIL, to improve **cop** (as described in Section 3.3). Thus, the output of **cop** is in $\{\text{PASS}, \text{FAIL}\}^*$. Because of the extension to **cop**, **cop** will never again output A, so the robot discards M(A). The robot stops repeating (∗), skips Stages 2–4, and goes to the next phase with the improved **cop**.
   - Otherwise (**ver**(A) passes), the robot follows explore(A). Note that by definition of **ver**(A), the robot follows explore(A) starting from a vertex u such that $(M(A), w_0(A))$ is consistent with (G, u).

   The subroutine (∗) can be viewed as a function taking the vertex at which the robot starts to the vertex at which it finishes. By Observation 1, we know that after $\hat{n}$ repetitions of (∗), the robot enters a closed path consisting of some number of executions of (∗).
2. The aim of this stage is to determine the closed path the robot has entered.[10] To determine this closed path, the robot repeats (∗) another $2\hat{n}$ times. Let the sequence S of outputs observed be $A_1, \ldots, A_{2\hat{n}}$. The robot finds the smallest p such that S is composed entirely of periodic repetitions of the last p outputs of S. More precisely, for all i, $A_{2\hat{n}-i} = A_{2\hat{n}-(i \bmod p)}$. Let seq be the sequence of edge labels traversed in $\mathbf{cop} \circ \mathbf{ver}(A_{2\hat{n}-p+1}) \circ \mathtt{explore}(A_{2\hat{n}-p+1}) \circ \ldots \circ \mathbf{cop} \circ \mathbf{ver}(A_{2\hat{n}}) \circ \mathtt{explore}(A_{2\hat{n}})$. By the minimality of p, the closed path consists of one or more repetitions of seq. To determine the closed path, the robot drops the pebble and repeatedly traverses seq until it finds the pebble at the end of one of its traversals of seq. It then retrieves the pebble for future use.
3. The robot proceeds along the closed path found above until it reaches the end of any execution of **cop**, say with output A. The robot then compresses the closed path and replaces M(A) with the resulting map.
4. If the new M(A) is finished then the algorithm outputs (the new) M(A) and terminates.

As noted above, if **ver** ever fails in Stage 1, the robot can improve **cop**. If all verifications pass, by Lemma 1 we know that in each phase (new $M(A), w_0(A)$) is consistent with (G, u) for some $u \in \mathsf{reach}(A)$, and thus Invariant 2 is preserved. Because **ver**(A) is part of the closed path and by Note 2, the new M(A) contains the old M(A) as a subgraph. Because explore(A) is part of the closed path (and is followed from u) the new M(A) also contains at least one new edge.

The algorithm terminates after at most $2n^2 d$ phases because in each phase the algorithm can either improve the

[10] Note that the robot cannot simply drop the pebble and repeat (∗) until it sees the pebble again because the robot needs the pebble to execute (∗).

candidate orienting procedure or enlarge a map. More precisely, since the candidate orienting procedure can be improved at most $n-1$ times, at most $n-1$ maps are discarded. At any time the algorithm maintains at most n maps, and so the algorithm builds at most $2n-1$ maps. Since each map contains at most $n \cdot d$ edges, the bound on the number of phases follows. Note that the algorithm may terminate before completing the orienting procedure.[11]

The running time of each phase is the sum of (1) the time to find a closed path, and (2) the running time of the compression procedure. Item (1) is $O(\hat{n})$ times the sum of (a) the running time of the candidate orienting procedure, (b) the running time of the verification procedure, and (c) the length of the exploration sequence (which is at most n). Recall that the running time of the verification procedures is $O(\hat{n} \cdot n^2 d)$. Also recall that verification procedures (that fail) are distinguishing procedures for improving the candidate orienting procedure. Therefore, we can bound the running time of any candidate orienting procedure by $O(\hat{n} \cdot n^3 d)$. Thus, Item (1) amounts to $O(\hat{n}^2 n^3 d)$. By Lemma 1, Item (2) is bounded by $O(\hat{n}^2 n^4 d)$. Since there are at most $2n^2 d$ phases, we obtain the following Theorem.

Theorem 1 *A robot having a single pebble can learn any strongly connected graph given an upper bound $\hat{n}$ on the size of the graph in time $O(\hat{n}^2 n^6 d^2)$.*

4 Learning without an Upper Bound on n

In this section we state our results concerning the number of pebbles needed to learn graphs efficiently if the graph size is unknown. We use the algorithm of Section 3.5 as a subroutine to show that for any $c > 0$, $\lceil c \log\log n \rceil$ pebbles are sufficient. The resulting algorithm is deterministic. In addition, we prove a matching lower bound demonstrating that $\Omega(\log\log n)$ pebbles are necessary. The lower bound applies to any randomized algorithm that uses an expected polynomial number of moves. We note that in our upper bound the total computation time to decide on moves is polynomial, whereas the lower bound applies even when the robot is computationally unbounded. Furthermore, our upper bound holds even when the pebbles used by the robot are indistinguishable from each other, while the lower bound holds for distinguishable pebbles.

We want to study how the number of pebbles needed grows with the size of the unknown graph. We denote the *expected number of pebbles* a (probabilistic) robot A uses on graphs of size n, by $p_A(n)$. Namely,

$$p_A(n) \stackrel{\text{def}}{=} \max_{G \in \mathcal{G}_n} E[\# \text{ of pebbles that A uses on G}],$$

where $\mathcal{G}_n$ is the set of all graphs on n vertices. The *expected running time* of A is defined analogously. (Recall that in each time step the robot makes a single move, and hence the running time of the algorithm is the number of moves the robot makes.)

Theorem 2 *For every constant $c > 0$, there exists a (deterministic) algorithm that learns graphs of size n in polynomial-time using at most $\lceil c \log\log n \rceil$ pebbles, without knowledge of n.*

[11] In fact, our algorithm as a whole can be viewed as an orienting procedure that outputs a completed map and a designated vertex.

Theorem 3 *Consider any algorithm A that, with probability greater than 1/2, learns any graph in expected polynomial time without knowing the size of the graph. Then $p_A(n) = \Omega(\log\log n)$.*[12]

The algorithm for Theorem 2 uses the one-pebble algorithm of Section 3.5 combined with a variant of the standard "guess and double" technique; namely, instead of doubling, it takes the k'th power for a suitably chosen k. In particular, setting $k = \lceil 2^{1/c} \rceil$, it runs the one-pebble algorithm trying $\hat{n} = 2^k$, $\hat{n} = 2^{k^2}$, $\hat{n} = 2^{k^3}$, until one of these executions succeeds. The algorithm succeeds after at most $\lceil c \log\log n \rceil$ iterations, and each iteration uses at most one pebble. The details of the proof will appear in the full version of this paper. We note that this algorithm can be deterministically simulated by two robots, giving a deterministic alternative to Bender and Slonim's randomized algorithm [7].

Proof (of Theorem 3): In order to prove the theorem, we analyze the behavior of any algorithm on two types of graphs of outdegree 2: *cycles* and *combination locks with tails*. Formally, the *cycle* of n nodes is the labeled, directed graph C_n on vertex set $\{w_0, \cdots, w_{n-1}\}$, where there are two directed edges labeled 0 and 1 going from w_i to $w_{(i+1) \bmod n}$. A combination lock with tail has the following structure (see Figure 3). Let $\alpha = \alpha_1 \alpha_2 \cdots \alpha_\ell \in \{0,1\}^\ell$ be any string and let $m \geq 0$ be an integer. The combination lock with combination α and tail m is the graph $L_{\alpha,m}$ on vertex set $\{u_1, u_2, \ldots, u_m, v_1, \ldots, v_{\ell+1}\}$ with the following edges: For each $1 \leq i \leq m-1$, there are two edges labeled 0 and 1 from u_i to u_{i+1}; there are two edges labeled 0 and 1 from u_m to v_1; for each $1 \leq i \leq \ell$, there is an edge labeled α_i from v_i to v_{i+1} and an edge labeled $\bar{\alpha}_i$ from v_i to v_1; there are two edges labeled 0 and 1 from $v_{\ell+1}$ to u_1. It is important to note that a robot starting at vertex v_1 (i.e., the start of the combination lock) does not reach vertex v_{k+1} unless it executes the consecutive sequence of moves $\alpha_1 \cdots \alpha_k$ at some point. We start by giving the intuition behind the proof.

We analyze any algorithm based on the times it drops pebbles in the case that it does not see previously-dropped pebbles. We show that there must be huge gaps in these pebble-dropping times or else the algorithm uses $\Omega(\log\log n)$ pebbles on sufficiently large cycles of length n. The quantity $\Omega(\log\log n)$ is exactly the threshold below which the gaps between pebble drops become superpolynomial. That is, for any polynomial f there are infinitely many time steps t such that no pebble is dropped between time t and time $f(t)$ with high probability. Then, for one of these big gaps, we can construct a combination lock with tail for which the following holds. With high probability, the algorithm drops no pebble within the combination lock and fails to reach the last few vertices of the lock in its allotted running time. Thus the robot fails to learn the graph. The idea of using combination locks with tails to foil a robot comes from Bender and Slonim's argument that a constant number of pebbles is insufficient [7]. The novel aspect of our proof is the analysis of pebble-dropping times to determine on which sizes of combination locks the algorithm fails.

We now turn to the details of the proof. Suppose, in contradiction to the claim in the theorem, that we have an expected polynomial-time algorithm A which succeeds in learning graphs with probability greater than 1/2, but does not use $\Omega(\log\log n)$ pebbles. Let $q(n) = O(n^k)$ be a

[12] It is easy to see from the proof that the success probability of 1/2 is arbitrary and can be replaced by any constant.

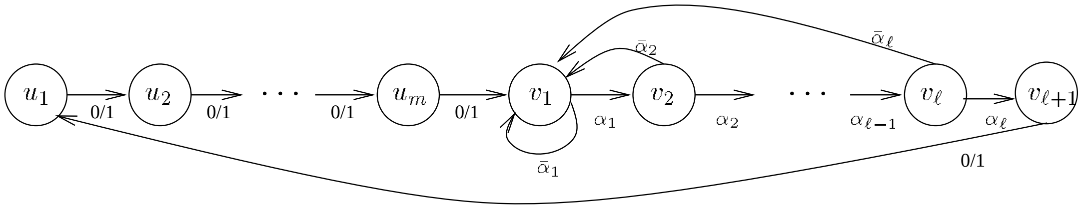

Figure 3: A Combination Lock with a Tail.

polynomial upper bound on the expected running time of the algorithm. In this proof, we use the standard technique of treating the randomized algorithm A as a distribution on deterministic algorithms A_r, *i.e.* for every infinite string $r \in \{0,1\}^{\mathbb{N}}$, A_r is the deterministic algorithm given by A using random coins r. All probabilities and expectations in this proof are taken over the choice of r.

We wish to study how the robot behaves when it doesn't see the pebbles it has dropped previously. To formalize this, we look at the infinite graph I on vertex set $\{w_1, w_2, \ldots\}$ where there are two edges labeled 0 and 1 from w_i to w_{i+1} for every $i \geq 1$. Now consider the behavior of the robot when it is placed at vertex w_1. Notice that when the robot drops a pebble at vertex w_i and moves, it never sees its pebble again. For $t \geq s \geq 1$, let $\mathrm{P}(s,t)$ be the probability that the robot drops at least one pebble between vertices w_s and w_{t-1}, inclusive, and let $\mathrm{E}(s,t)$ be the expected number of pebbles dropped by the robot between vertices w_s and w_{t-1}, so $\mathrm{E}(s,t) \geq \mathrm{P}(s,t)$. Notice that $\mathrm{E}(1,t)$ is a *lower bound* on the expected number of pebbles the robot uses on a cycle C_t of t vertices, because for every r, A_r's behavior in its first $t-1$ moves is the same in C_t as in I. We now use this to show that that there are superpolynomial gaps in the pebble-dropping times.

Claim: For every fixed $c > 0$, there are infinitely many t such that $\mathrm{P}(t, t^c) < 1/8$.

Proof of Claim: Suppose not, *i.e.* there is some t_0 such that for all $t \geq t_0$, $\mathrm{P}(t,t^c) \geq 1/8$. Then for every $\ell \geq 0$, $\mathrm{E}(t_0, t_0^{c^\ell}) = \sum_{j=1}^{\ell} \mathrm{E}(t_0^{c^{j-1}}, t_0^{c^j}) \geq \sum_{j=1}^{\ell} \mathrm{P}(t_0^{c^{j-1}}, t_0^{c^j}) \geq \ell/8$.

For $n \geq t_0$, let ℓ_n be the smallest value of ℓ such that $n < t_0^{c^\ell}$. Then $\log\log n < \log\log t_0 + \ell_n \log c$, so $\ell_n = \Omega(\log\log n)$. We also have

$$\mathrm{E}(1,n) \geq \mathrm{E}(t_0,n) \geq \mathrm{E}(t_0, t_0^{c^{\ell_n - 1}}) \geq \frac{\ell_n - 1}{8} = \Omega(\log\log n).$$

But $\mathrm{E}(1,n)$ is a lower bound on the expected number of pebbles the robot uses on a cycle of length n, so we have a contradiction. $\Rightarrow\Leftarrow$ $\square$

Recall that the expected running time of A is $q(n) = \mathrm{O}(n^k)$. Using the above claim with $c = k+1$, we can find a t with the following properties:

- $\mathrm{P}(t, t^{k+1}) < \frac{1}{8}$.
- $\frac{8q(2t+4)}{2^t} < \frac{1}{8}$.
- $t^{k+1} \geq 8q(2t+4)$.

Consider the random variable W which is a string consisting of the robot's first $8q(2t+4)$ moves in I. There are less than $8q(2t+4)$ contiguous subsequences of length t in W, so there is some string $\alpha \in \{0,1\}^t$ which occurs in W with probability less than $8q(2t+4)/2^t < 1/8$. In other words there is a sequence of moves α of length t which the robot performs with probability less than $1/8$ in its first $8q(2t+4)$ steps in I.

Let β by any binary string of length 4, and consider the behavior of the robot when placed at vertex u_1 in a combination lock G^β with tail $t-1$ and combination $\alpha\beta$ (and vertex set $\{u_1, \ldots, u_{t-1}, v_1, \ldots, v_{t+5}\}$ as above). Since A runs in expected time $q(n)$ and G has $2t+4$ vertices, the probability that A makes more than $8q(2t+4)$ moves in G^β is at most $1/8$.

Let R_1 be the set of random coins r for which A_r would drop a pebble between vertex w_t and $w_{t^{k+1}-1}$ in I. Let R_2 be the set of random coins r for which A_r executes the sequence of moves α at some point during its first $2t+4$ moves in I. Let R_3 be the set of random coins r for which A_r makes more than $8q(2t+4)$ moves in G^β. Let $\mathrm{R} = \mathrm{R}_1 \cup \mathrm{R}_2 \cup \mathrm{R}_3$. We have shown that $\Pr[r \in \mathrm{R}] < 3/8$. Notice that for any $r \notin \mathrm{R}$, the output of A_r on G^β is the same as its output on G^γ for any string γ of length 4 because the robot never sees a pebble that it has dropped and never reaches vertex v_{t+1}. Let S^γ be the set of $r \notin \mathrm{R}$ on which A_r outputs G^γ when placed in G^γ (equivalently, G^β). Then since A has overall success probability at least $1/2$, A must succeed on at least $1/8$ of the $r \notin R$. So $\Pr[r \in \mathrm{S}^\gamma] > 1/8$. But there are 16 S^γ's and they are disjoint. $\Rightarrow\Leftarrow$ $\blacksquare$

5 Extensions

We have generalized our result to the case when the local labelling of the edges at a vertex is a function of the previous vertex in the robot's path. (We assume there is at most one edge between two vertices.) We give only the intuition for the proof here. First, suppose that the robot could drop its pebbles on edges. In this case, one can simply execute our algorithm on the edge-adjacency graph of the original graph. In this new graph, edges of the old graph correspond to vertices, and there is an edge between vertices e_1 and e_2 if there exists vertices in the old graph v_1, v_2, v_3 such that $e_1 = (v_1, v_2)$ and $e_2 = (v_2, v_3)$. The edge (e_1, e_2) in the new graph is given the label that e_2 has in the original graph when arriving at v_2 by way of e_1. Thus, in this new graph, edges will have unique local labellings. One can execute any of our algorithms on this graph, and then reconstruct the actual graph once the algorithm has terminated. With some care, one can remove the need to allow dropping of pebbles on edges without increasing the number of pebbles needed.

References

[1] S. Albers and M. R. Henzinger. Exploring unknown environments. In *Proceedings of the Twenty Ninth Annual*

ACM Symposium on the Theory of Computing, 1997.

[2] D. Angluin, J. Westbrook, and W. Zhu. Robot navigation with range queries. In *Proceedings of the Twenty Eighth Annual ACM Symposium on the Theory of Computing*, pages 469–478, 1996.

[3] V. Anjan. *Doctoral Thesis.* PhD thesis, Mathematical Institute of the Academy of Sciences, Minsk, 1987.

[4] B. Awerbuch, M. Betke, R. L. Rivest, and M. Singh. Piecemeal graph exploration by a mobile robot. In *Proceedings of the Eighth Annual ACM Conference on Computational Learning Theory*, pages 321–328, 1995.

[5] R. Baeza-Yates, J. Culberson, and G. Rawlins. Searching in the plane. *Information and Computation*, pages 234–252, 1993.

[6] E. Bar-Eli, P. Berman, A. Fiat, and P. Yan. On-line navigation in a room. In *Proceedings of the Third Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 8–15, 1992.

[7] M. A. Bender and D. Slonim. The power of team exploration: Two robots can learn unlabeled directed graphs. In *Proceedings of the Thirty Fifth Annual Symposium on Foundations of Computer Science*, pages 75–85, 1994.

[8] P. Berman, A. Blum, A. Fiat, H. Karloff, A. Rosen, and M. Saks. Randomized robot navigation algorithms. In *Proceedings of the Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 74–84, 1996.

[9] M. Betke, R. L. Rivest, and M. Singh. Piecemeal learning of an unknown environment. In *Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory*, pages 277–286, 1993.

[10] A. Blum and P. Chalasani. An on-line algorithm for improving performance in navigation. In *Proceedings of the Thirty Fourth Annual Symposium on Foundations of Computer Science*, pages 2–11, 1993.

[11] A. Blum, P. Raghavan, and B. Schieber. Navigating in unfamiliar geometric terrain. In *Proceedings of the Twenty Third Annual ACM Symposium on Theory of Computing*, pages 494–504, May 1991.

[12] M. Blum and D. Kozen. On the power of the compass (or, why mazes are easier to search than graphs. In *Proceedings of the Nineteenth Annual Symposium on Foundations of Computer Science*, pages 132–142, October 1978.

[13] M. Blum and W. J. Sakoda. On the capability of finite automata in 2 and 3 dimensional space. In *Proceedings of the Eighteenth Annual Symposium on Foundations of Computer Science*, pages 147–161, 1977.

[14] T. Dean, D. Angluin, K. Basye, S. Engelson, L. Kaelbling, E. Kokkevis, and O. Maron. Inferring finite automata with stochastic output functions and an application to map learning. *Machine Learning*, 18(1):81–108, January 1995.

[15] X. Deng, T. Kameda, and C. H. Papadimtriou. How to learn an unknown environment. In *Proceedings of the Thirty Second Annual Symposium on Foundations of Computer Science*, pages 298–303, 1991.

[16] X. Deng and C. H. Papadimtriou. Exploring an unknown graph. In *Proceedings of the Thirty First Annual Symposium on Foundations of Computer Science*, pages 356–361, 1990.

[17] Y. Freund, M. Kearns, Y. Mansour, D. Ron, R. Rubinfeld, and R. E. Schapire. Efficient algorithms for learning to play repeated games against computationally bounded adversaries. In *Proceedings of the Thirty Sixth Annual Symposium on Foundations of Computer Science*, pages 332–341, 1995.

[18] Y. Freund, M. Kearns, D. Ron, R. Rubinfeld, R. Schapire, and L. Sellie. Efficient learning of typical finite automata from random walks. In *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, pages 315–324, 1993.

[19] F. Hoffman, C. Icking, R. Klein, and K. Kriegel. A competitive strategy for learning a polygon. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 166–174, 1997.

[20] Z. Kohavi. *Switching and Finite Automata Theory.* McGraw-Hill, second edition, 1978.

[21] V. B. Kudryavtsev, Sh. Ushchumlich, and G. Kilibarda. On the behavior of automata in labyrinths. *Discrete Math. and Applications*, 3:1–28, 1993.

[22] Petrisor Panaite and Andrzej Pelc. Exploring unknown undirected graphs. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, 1998.

[23] C.H. Papadimitriou and M. Yannakakis. Shortest paths without a map. *Theoretical Computer Science*, 84:127–150, 1991.

[24] M. O. Rabin. Maze threading automata. Seminar Talk presented at the University of California at Berkeley, October 1967.

[25] L. Reyzin. Traversal problems for certain types of deterministic and non-deterministic automata. Unpublished manuscript, 1992.

[26] R. Rivest and R. Schapire. Inference of finite automata using homing sequences. *Information and Computation*, 103(2):299–347, 1993.

[27] R. Rivest and R. Schapire. Diversity-based inference of finite automata. *Journal of the Association for Computing Machinery*, 43(3):555–589, 1994.

[28] D. Ron and R. Rubinfeld. Exactly learning automata of small cover time. *Machine Learning*, 27(1):69–96, 1997.

[29] A.N. Shah. Pebble automata on arrays. *Computer Graphics and Image Processing*, pages 236–246, 1974.

[30] L. Zhang. A survey of the problem of learning an unknown environment. Unpublished manuscript, 1994.