Road Coloring Problem for Completely Reachable Automata

Yinfeng Zhu
Institute of Natural Sciences and Mathematics
Ural Federal University
Ekaterinburg, Russia
yinfeng.zhu7@gmail.com

Trahtman (2008) prove the Road Coloring Theorem which state that the edges of any given aperiodic directed multi-digraph with a constant out-degree d can be colored by d colors such that the resulting automaton admits a reset word. Bondar and Volkov (2016) asked an analogous question: characterize multi-digraphs that can be colored such that the resulting automaton is completely reachable. We show a necessary and sufficient condition for the multi-digraphs admit a completely reachable coloring. Moreover, this condition can be determine in polynomail time. For a fixed integer d, we show that to determine a given multi-digraph whether or not admit a completely reachable coloring within d colors is NP-complete. We also obtain a criterion for the multi-digraphs whose road colorings are all completely reachable.

1 Introduction

1.1 Automata, digraphs and road colorings

An *deterministic finite automaton* (DFA) is a triple $\mathscr{A} = (Q, \Sigma, \delta)$ where Q and Σ are finite non-empty sets and δ is a function from $Q \times \Sigma$ to Q. The elements of Q is called *states*; the elements of Σ is called *letters*; and the function δ is called the *transition function* of $\mathscr{A}$. Finite sequences (including the empty sequence denoted by ε) over Σ are called *words* over Σ. Write Σ^* for the set of all words over Σ.

The transition function δ extends to a function $Q \times \Sigma^* \to Q$ (still denoted by δ) via the recursion: for each $q \in Q$, $a \in \Sigma$, $w \in \Sigma^*$, set $\delta(q, \varepsilon) = q$ and $\delta(q, wa) = \delta(\delta(q, w))$. For a subset $P \subseteq Q$ and $w \in \Sigma^*$, write $P.w$ for the set $\{p.w : p \in P\}$.

A subset $P \subseteq Q$ is called *reachable* in $\mathscr{A}$, if there exists a word $w \in \Sigma^*$ such that $P = \delta(Q, w)$. An automaton is called *completely reachable* if every non-empty subset of its state set is reachable. An automaton is called *synchronizing* if there exists a reachable singleton subset of states.

A digraph is a quadruple $G = (V, E, h, t)$ where V, E are non-empty sets and h, t are functions from E to V. The elements in V are called *vertices* of G and the elements of E are called *edges* of G. For an edge $e \in E$, the vertex $h(e)$ is called the *head* of e and the vertex $t(e)$ is called the *tail* of e. When there is no ambiguity, we will use uv or (u, v) to denoted an edge e such that $h(e) = u$ and $t(e) = v$. For a subset $U \in V$, the *out-neighbour* of U is the set $\{h(e) : t(e) \in U, e \in E\}$, denoted $\mathrm{N}_G^+(U)$; the *in-neighbour* of U is the set $\{t(e) : h(e) \in U, e \in E\}$, denoted $\mathrm{N}_G^-(U)$; For a vertex $v \in V$, the *out-degree* of v is the number of edges whose head is v, denoted $\mathrm{d}_+(v)$; the *in-degree* of v is the number of edges whose tail is v, denoted $\mathrm{d}_-(v)$. A digraph is called *aperiodic* if the greatest common divisor of lengths of all directed cycles in the digraph equals to 1.

For a set X, the *power set* of X is the set consisting all subset of X, denoted $\mathscr{P}(X)$.

A *road coloring* of a finite digraph $G = (V, E, h, t)$ over a set Σ is a function $\alpha : E \to \mathscr{P}(\Sigma)$ such that for every vertex $v \in V$, the family of sets $\{\alpha(e)\}_{e \in E}$ forms a partition of Σ. The elements in Σ are called *colors*.

Submitted to:
NCMA 2014

Let $\alpha : \Sigma \to \mathscr{P}(E)$ be a road coloring of a digraph $G = (V, E, h, t)$. A DFA is obtained from G and α. Define $\mathscr{A}(G, \alpha)$ to be the automaton (V, Σ, δ) such that for every $v \in V$ and $a \in \Sigma$, we have $v.a = t(e)$ where e is the unique edge such that $a \in \alpha(e)$. We say that α is a *synchronzing coloring* of G if $\mathscr{A}(G, \alpha)$ is synchronizing; α is a *completely reachable coloring* if $\mathscr{A}(G, \alpha)$ is a completely reachable automaton.

The Road Coloring Theorem can be stated as follows.

Theorem 1 (Trahtman [5])**.** *Let $G = (V, E, h, t)$ be a strongly connected digraph and $d = \max\{\mathrm{d}_+(v), v \in V\}$. The following are equivalent.*

(1) The digraph G admits a synchronizing coloring.

(2) The digraph G admits a synchronizing coloring with d colors.

(3) The digraph G is aperiodic.

As it was conjectured by Adler, Goodwyn, and Weiss [1], the Road Coloring Theorem provides the necessary and sufficient condition for a strongly connected digraph G admitting a synchronizing coloring with d colors.

1.2 Main results

Bondar and Volkov [2] ask an analogous question for completely reachable colorings. They found an interesting and non-parallel phenomena that there are some digraphs that have no completely reachable coloring with 2 letters but admit a completely reachable coloring with 3 letters. In this article, we study on Bondar-Volkov's question.

If we have no restrictions on the number of colors, we can obtain the necessary and sufficient condition for a digraph admitting a completely reachabel coloring.

Theorem 2. *A digraph $G = (V, E, i, t)$ admits a completely reachable coloring if and only if*

(1) G is strongly connected,

(2) G is aperiodic,

(3) for every subset $U \subseteq V$, $|U| \leq |\mathrm{N}_-(U)|$.

These conditions can be determined in polynomial-time. This is obivious for the first and second conditions. The third condition is equivalent to whether a derived bipartite graph (of polynomial size) has a perfect matching.

However, for a fixed integer $d \geq 2$, it is hard to determine whether a given digraph whether or not admit a completely reachabel coloring with d colors.

Theorem 3. *Let $d \geq 2$ be an integer. To determine a given digraph whether or not it admits a completely reachable with d colors is NP-complete.*

We obtain a criterion for determining whether a digraph whose road colorings are all completely reachable.

The remaining of this article will proceed as follows. In Section 2, using two fundamental graph theoretical results, we present a proof of Theorem 2. In Section 3, based on a reduction of Plesńik [4], we prove Theorem 3. In Section 4, we show a criterion for the digraphs whose road colorings are all completely reachable.

2 Proof of Theorem 2

Let us recall two fundamental graph theoretical results which will be used in the proof of Theorem 2.

The *period* of a strongly connected digraph G is the greatest common divisor of the lengths of its cycles, denoted $p(G)$. The following result is well-known.

Theorem 4. *Let G be a strongly connected digraph of period p. The vertex set can be partition into p sets $\{C_i : i \in \mathbb{Z}_p\}$ such that $N_G^+(C_i) = C_{i+1}$ for every $i \in \mathbb{Z}_p$. Moreover, for each vertex v and $i \in \mathbb{Z}_p$, there exists an integer $k \geq 0$ such that*

$$\underbrace{N_G^+ \cdots N_G^+}_{k}(v) = C_i.$$

A *bipartite graph* $H = (X,Y,E)$ is a triple, where X,Y are two nonempty set and $E \subseteq X \times Y$. The elements in $X \cup Y$ are *vertices* and the elements in E are *edges*. An *X-perfect matching* of H is a matching, a set of disjoint edges. which covers every vertex in X. For $U \subseteq X$, the *neighborhood* of U is the set $\{v : (u,v) \in E, u \in U\}$, denoted $\mathrm{N}_H(U)$.

Theorem 5 (Hall's Marriage Theorem)**.** *Let $H = (X,Y,E)$ be a bipartite graph. There exists an X-perfect matching in H if and only if for every subset $U \subseteq X$, we have $|U| \leq \mathrm{N}_H(U)$.*

Next, we prove Theorem 2.

Proof of Theorem 2. "⇒": Let α be a completely reachable coloring of G and its color set is Σ. The corresponding autoaton is $\mathscr{A}(G,\alpha) = (V,\Sigma,\delta)$. For arbitrary vertices u and v, by completely reachability, there exists a word w such that $V.w = u$ and then $v.w = u$. Then there exists a walk in G from v to u. Hence G is strongly connected.

By Theorem 4, the vertex set V can be partitioned into $p = p(G)$ sets $\{C_i : i \in \mathbb{Z}_p\}$ such that $N_G^+(C_i) = C_{i+1}$ for every $i \in \mathbb{Z}_p$. Then for any words $w \in \Sigma^*$ and $i \in \mathbb{Z}_p$, $V.w \cap C_i \neq \emptyset$. Since $\mathscr{A}(G,\alpha)$ is completely reachable, every singleton set is reachable. Hence $p = 1$ which is equivalent to G is aperiodic.

For a non-empty subset $U \subseteq V$, take a word $w = w'a \in \Sigma^*$ such that $V.w = U$. Let W be the set $V.w'$. Then

$$|U| \leq |W| \leq |\mathrm{N}_G^-(U)|.$$

"⇐": Define H to be the bipartite graph $H = (V_1,V_2,E_H)$ such that $V_1 = V_2 = V$ and $E_H = \{(h(e),t(e)) : e \in E\}$. Then for every non-empty subset $U \subseteq V_1$, then $|U| \leq |\mathrm{N}_H(U)|$. Let W be a non-empty subset of V_1. Let H' be the induced subgraph of H on $W \cup \mathrm{N}_H(W)$. By the Theorem 5, there exists a W-perfect matching M in H'. Now we can define a function $f_W : V_2 \to V_1$ as following:

(1) for $x \in V_2$ which is covered by an edge $\{x,y\} \in M$, set $f_W(x) = y$;

(2) for $x \in \mathrm{N}_H(W)$ which is not covered by the matching M, set $f_W(x)$ to be an arbitrary vertex in $W \cap \mathrm{N}_G^+(x)$.

(3) for $x \in V_2 \setminus \mathrm{N}_H(W)$, set $f_W(x)$ to be an arbitrary vertex in $\mathrm{N}_G^+(x)$.

It is clear that $W = f_W(\mathrm{N}_G^-(W))$.

Now we construct a road coloring $\alpha : E \to \mathscr{P}(\Sigma)$, where $\Sigma = \mathscr{P}(V) \setminus \{\emptyset\}$ by setting

$$\alpha(e) = \{U : f_U(h(e)) = t(e), \emptyset \neq U \subseteq V\}.$$

Let $\mathscr{A} = (V,\Sigma,\delta) = \mathscr{A}(G,\alpha)$. Note that for every non-empty subset U, we have

$$\delta(N_G^-(U),U) = U.$$

Let U_0 be an arbitrary non-empty subset of V, define $U_i = \mathrm{N}_G^-(U_{i-1})$ for all positive integer i. Since G is strongly connected and aperiodic, there exists an integer k such that $U_k = V$. Then U_0 is reachable via the word $U_{k-1}U_{k-2}\cdots U_1U_0$. Hence $\mathscr{A}$ is completely reachable and G admits a completely reachable coloring. □

Remark 6. *Using the notations of the proof of Theorem 2, observe that G satisfies the third condition in Theorem 2 if and only if H has a perfect matching. Then there exists a polynomial-time algorithm (see [3]) to determine whether G satisfies the third condition in Theorem 2.*

3 Proof of Theorem 3

Let $\mathscr{A} = (V, \Sigma, \delta)$ be a DFA. For a letter $a \in \Sigma$, the *defect* of a is the integer $|V| - |V.w|$, denoted $\mathrm{defect}(a)$.

Lemma 7. *Let Σ be a k-element set. Assume that $G = (V, E, h, t)$ is k-out-regular digraph and α is a road coloring of G with the color set Σ. Let $\mathscr{A} = \mathscr{A}(G, \alpha)$. Then*

$$\sum_{v \in V} \max(0, (k - \mathrm{d}_G^-(v))) \le \sum_{a \in \Sigma} \mathrm{defect}(a).$$

Proof. For each $a \in \Sigma$, let G_a be the sub-digraph of G containing all vertices of G and edges which color a. It is clear that $\mathrm{defect}(a)$ equals the number of vertices whose in-degree is zero in G_a. Since $|\Sigma| = k$ and G is k-out-regular, $\alpha(e)$ is a singleton set for every $e \in E$. Then, for every vertex v,

$$\left|\{a \in \Sigma : \mathrm{d}_{G_a}^-(v) = 0\}\right| \ge k - \mathrm{d}_G^-(v).$$

Using double counting argument, we have

$$\sum_{a \in \Sigma} \mathrm{defect}(a) \ge \sum_{v \in V} \max\left(0, k - \mathrm{d}_G^-(v)\right).$$

□

Lemma 8. *Let $k \ge 2$. Let $G = (V, E, h, t)$ be a digraph satisfying the following conditions.*

(1) $|V| = n$ is a prime number and $|V| > k$.

(2) For every vertex v, $\mathrm{d}_G^+(v) = k$.

(3) There exists a vertices x such that there are at least $k-2$ edges from v to x for every $v \in V$.

(4) Let G' be a digraph which is obtained from G by deleting $k-2$ edge from v to x for every $v \in V$. In the digraph G', we have

- $\mathrm{d}_{G'}^-(x) = 1$,
- *there exists a vertex y such that $\mathrm{d}_{G'}^-(y) = 3$,*
- *for each $z \in V \setminus \{x, y\}$, $\mathrm{d}_{G'}^-(z) = 2$.*

Then G admits a completely reachable coloring with k colors if and only if G has a Hamitonian cycle.

Proof. "⇒": Let Σ be a k-element set. Let α be a completely reachable coloring with the color set Σ. Since $n > k$, there exists a color of defect 0, denoted a and a color of defect 1, denoted b.

Claim 9. *For all $c \in \Sigma \setminus \{a, b\}$, $\mathrm{defect}(c) \ge 2$.*

Proof. Assume, for a contradiction, there exists a color $c \in \Sigma$ such that $c \notin \{a,b\}$ and $\text{defect}(c) = 1$. By Lemma 7, we have

$$\sum_{x\in\Sigma} \text{defect}(x) \geq \sum_{v\in V} \max(0, k - \text{d}^-_G(v)) \geq (n-1)(k-2).$$

Meanwhile,

$$\sum_{x\in\Sigma} \text{defect}(x) \leq (n-1)(k-3) + 0 + 1 + 1 = (n-1)(k-2) - (n-3).$$

This is a contradiction. □

Since α is completely reachable, every $n-1$-element subset of V is reachable. Then the action of a is a cyclic permutation. Hence, the edges in G_a form a Hamiltonian cycle.

"$\Leftarrow$": Let C be a hamitonian cycle. Color the edges in C by c. Let D be the set of edges of G' but not belongs to C.

Case 1. Every edge of C is in G'.

Color the edges in D by d. The action of d is clearly 1-defect. We color the other edges such that α is a road coloring. Since n is prime, one can check that α is completely reachable.

Case 2. There exists an edge $e \in C$ such that $e \notin G'$.

Let G'' be the digraph on vertex set V and edge set D. Note that G'' has $n+1$ edges and we have

- $\text{d}^-_{G''}(x) = 2$ and $\text{d}^-_{G''}(z) = 1$ for all $z \neq x$;
- $\text{d}^+_{G''}(y) = 2$ and $\text{d}^-_{G''}(z) = 1$ for all $z \neq y$.

Let e be an edge such that $t(e) = y$. Color the edges in $D - e$ by d. The action of d is clearly 1-defect. We color the other edges such that α is a road coloring. Since n is prime, one can check that α is completely reachable.

□

The problems of determining the existance of Hamiltonian cycles (paths, resp.) for a given digraph is denoted by HCP (HPP, resp.).

In [4], the Plesńik shows polynomial transformations from SAT to HCP (and HPP) for a special class of digraphs. Our proof of Theorem 3 is obtained from Plesńik reduction by some small modifications. For convenience, let us define a digraph operator Replace. Let $G = (V_G, E_G, h_G, t_G)$ and $H = (V_H, E_H, h_H, t_H)$ be two digraphs. Let $v \in V_G$ and $x, y \in V_H$. Define $\text{Replace}(G, v, H, x, y)$ to be the digraph $(V_G \cup V_H \setminus \{v\}, E_G \cup E_H, h, t)$ such that

$$h(e) = \begin{cases} h_G(e) & \text{if } e \in E_G \text{ and } h_G(e) \neq v \\ h_H(e) & \text{if } e \in E_H \\ x_2 & \text{otherwise} \end{cases}$$

and

$$t(e) = \begin{cases} t_G(e) & \text{if } e \in E_G \text{ and } t_G(e) \neq v \\ t_H(e) & \text{if } e \in E_H \\ x_1 & \text{otherwise.} \end{cases}$$

Proof of Theorem 3. For a Boolean formula F, Plesńik constructed a digraph $G = (V, E, h, t)$ such that F is satisfiable if and only if G has a Hamiltonian path. The digraph G satisfies the properties: the size of G is at most a polynomial of the size of F; for every vertex v in G, either $\text{d}^-_G(v) = 1$, $\text{d}^+_G(v) = 2$ or $\text{d}^-_G(v) = 2$, $\text{d}^+_G(v) = 1$.

Write E_1 for the set $\{e \in E : \mathrm{d}^+_G(e) = 1\}$. Define G_1 to be the digraph that adding a copy for each edge in E_1 into G. It is easy to check that for every vertex v of G_1, $\mathrm{d}^+_{G_1} = \mathrm{d}^-_{G_1} = 2$.

Let H be the digraph with vertex set $\{x_1, x_2, x_3\}$ and edge set $\{e_1, e_2, e_3, e_4\}$ such that $h(e_1) = h(e_2) = h(e_3) = x_1$, $h(e_4) = x_2$, $t(e_1) = t(e_2) = x_2$ and $t(e_3) = t(e_4) = x_3$. Choose an arbitrary vertex v of G_1. Define $G_2 = \mathrm{Replace}(G_1, v, H, x_1, x_2)$.

For a positive integer n, define P_n to be the digraph with vertex set $\{x_1, \ldots, x_{n+1}\}$ and having two edge from x_i to x_{i+1} for all $1 \le i \le n$. Choose a vertex u of G_2 such that $\mathrm{d}^+_{G_2}(u) = \mathrm{d}^-_{G_2}(u) = 2$. Define $G_3 = \mathrm{Replace}(G_2, u, P_m, x_1, x_{n+1})$, where m is the least positive integer such that $m + |V_{G_2}|$ is a prime. It is routine to check

- G_3 has a Hamitonian path if and only G has a Hamitonian path;
- there exist vertices x and y such that $\mathrm{d}^-_{G_3}(x) = 1$ and $\mathrm{d}^-_{G_3}(y) = 3$;
- for each $z \in V \setminus \{x, y\}$, $\mathrm{d}^-_{G_3}(z) = 2$;
- for each $z \in V$, $\mathrm{d}^+_{G_3}(z) = 2$;

Define G_4 to be the digraph which is obtained from G_3 by adding $k-2$ edge from v to x for every vertex v of G_3. Observe that G_4 has a Hamitonian cycle if and only if G_3 has a Hamitonian path. Note that G_4 fulfills all properties in Lemma 8. Hence, we polynomially transform the SAT problem to the problem of determining the existance of completely reachable coloring for a digraph. □

4 Digraphs satisfying all road colorings are completely reachable

Let $\mathscr{ARC}$ be the family of digraphs consisting the digraphs whose road colorings are all completely reachable.

Use $\oplus$ to stand for addition modulo n. Let S be a subset of $(\mathbb{Z}_n, \oplus)$. Let $G(S, n)$ be the digraph with the vertex set $\mathbb{Z}_n$ and the edge set $\{(i, i+1) : i \in \mathbb{Z}_n\} \cup \{(n-1, s) : s \in S\}$.

Theorem 10. *Let G be a digraph with n vertices. The digraph G is in $\mathscr{ARC}$ if and only if there exists $S \subseteq \mathbb{Z}_n$ such that $\langle S, \oplus \rangle = (\mathbb{Z}_n, \oplus)$ and G is isomorphic to $G(S, n)$.*

Lemma 11. *Let G be a digraph with n vertices. If $G = (V, E, h, t) \in \mathscr{ARC}$, then there exists a unique vertex v such that $\mathrm{N}^+_G(v) > 1$. Moreover, G contains a hamitonian cycle.*

Proof. By Theorem 2, G is strongly connected. Then every vertex has at least one out-neighbour. Assume, for contradiction, that there exists two distinct vertices u and v such that $\mathrm{N}^+_G(v) > 1$ and $\mathrm{N}^+_G(u) > 1$.

Let $H = (V_1, V_2, E_H)$ be the bipartite graph corresponding to G which is define in the proof of Theorem 2. By Theorem 2, H has a V_1-perfect matching M. Write $M(v)$ for the vertex such that $(v, M(v)) \in M$. For every edge $e = (x_1, x_2) \in E_H \setminus M$, we can find an edge $(y_1, y_2) \notin M$ and $y_1 \neq x_1$ and let

$$M_e = M \cup \{(x_1, x_2), (y_1, y_2)\} \setminus \{(x_1, M(x_1)), (y_1, M(y_1))\}.$$

Let $\mathscr{M} = \{M\} \cup \{M_e : e \in E_H \setminus M\}$.

For every $e \in E$, write $\overline{e}$ for the edge $(h(e), t(e)) \in E_H$. Define $\alpha : E \to \mathscr{M}$ as the function such that

$$\alpha(e) = \{N \in \mathscr{M} : \overline{e} \in N\}.$$

It is clear that α is a road coloring of G. Let $\mathscr{A} = \mathscr{A}(G, \alpha)$. Note that the action of M is a bijection and for all $e \in E_H$, the defect of the action of M_e is 2. Observe that every $(n-1)$-element subset is not reachable in $\mathscr{A}$ which is a contradiction.

Since G is strongly connected has a unique vertex of out-degree ≥ 1, the digraph G has a hamitonian cycle.

□

Proof of Theorem 10. "$\Leftarrow$": It is trivial.

"$\Rightarrow$": By Lemma 11, we have G is isomorphic to $G(S,n)$ for some subset $S \subseteq \mathbb{Z}_n$. Let $K = \langle S, \oplus \rangle$. It is clear that the period of G equals $\frac{n}{|K|}$. By Theorem 10, $K = (\mathbb{Z}_n, \oplus)$. □

Acknowledgments

I thank Prof. Mikhail V. Volkov for introducing this problem to me and valuable discussions.

References

[1] Roy L Adler, L Wayne Goodwyn & Benjamin Weiss (1977): *Equivalence of topological Markov shifts*. Israel Journal of Mathematics 27, pp. 49–63.

[2] Eugenija A. Bondar & Mikhail V. Volkov (2016): *Completely reachable automata*. In: Descriptional complexity of formal systems, Lecture Notes in Comput. Sci. 9777, Springer, [Cham], pp. 1–17, doi:10.1007/978-3-319-41114-9_1.

[3] Alan Gibbons (1985): *Algorithmic graph theory*. Cambridge University Press, Cambridge.

[4] J. Plesńik (1979): *The np-completeness of the hamiltonian cycle problem in planar diagraphs with degree bound two*. Information Processing Letters 8(4), pp. 199–201, doi:10.1016/0020-0190(79)90023-1.

[5] Avraham N Trahtman (2009): *The road coloring problem*. Israel Journal of Mathematics 172, pp. 51–60.