1. Introduction

Dear Tourists to Seoul

Hello everyone and welcome to Seoul! What made you come to Seoul? Well, you might expect to find wonders in Seoul, but unfortunately, the season is wrong. It is in very humid and hot summer\(^1\), so you would really like to find a place that can cool down your temperature. And there is the right place in Seoul: Metro.

Especially, the price to use Seoul Metro is relatively cheaper than other cities (you can imagine Tokyo, for instance). And Seoul Metro interconnects many cities nearby Seoul, so you can visit almost every city in Kyeonggi Province by Seoul Metro - at least the underground of these cities. But most importantly, it has a very good air conditioning system especially in August! Therefore it attracts many

\(^1\)This guide is written in August, 2015. Korea is known to have 4 seasons, but nowadays, it only has two seasons, summer and winter.
metro manias, mostly older people. Of course, there are some young manias for Seoul Metro, including Su-Hyeon.

Su-Hyeon, who was and is an everyday Seoul Metro user, really likes to take a subway. So she once asked the following question in her mind.

**Question 1.1. What is the longest path in Seoul Metro?**

If there is no additional condition in this question, then you can repeat any path infinitely many times. So the length of the longest path in Seoul Metro can be infinite. This is too boring. But with giving one additional condition, she made this question much more interesting.

There are many lines in Seoul Metro and in each line there is a line segment between consecutive stations. Let us call each of these an edge. For instance, between 올림포4가 and 동대문역사문화공원, there are two edges from Line 2 and 5, respectively. The below is the reformulation of Question 1.1 with an additional condition, and we are going to mostly think about this question.

**Question 1.2. What is the longest path in Seoul Metro with no repeated edges?**

Surprisingly (or not?), Graph theory can help us to find such a path. In the following sections, we will see how and why graph theory can solve this question.

### 2. Basic Graph Theory and its Application

#### 2.1. Useful Definitions from Graph Theory

Graph Theory can make us analyze the real world in a more simple way, so it can also help to resolve Question 1.2. But before using that, we should agree with the following basic definitions from graph theory.

**Definition 2.1 (Graphs).** A graph $G$ is a pair of sets $(V, E)$ such that $E$ is a multiset of subsets of $V$ of size 2. $V$ is called the set of vertices, and $E$ is called the set of edges.

For Seoul Metro, vertices are the stations, and edges coincide what we defined in Section 1. And a multiset means that we can have multiple edges between two stations like edges between 올림포4가 and 동대문역사문화공원. Edges are defined as sets, but we can use a simple notation for them. For instance, if $e = \{v, w\}$ is an edge for vertices $v, w \in V$, then we just denote $e$ as $vw$.

**Definition 2.2 (Neighborhood and Degree).** Let a graph $G = (V, E)$ be given.

(i) If $vw \in E$, we say that $v$ and $w$ are adjacent.

(ii) For a vertex $v$, a neighborhood of $v$ (or $N(v)$) is the set of vertices which are adjacent to $v$.

(iii) Degree of a vertex $v$ (or $deg(v)$) is the size of $N(v)$.

For instance in Seoul Metro, $deg(망우) = 4$ whereas $deg(신도림) = 5$.

That’s all we need for now. We first see a very famous example using these simple concepts which was made in nearly 300 years ago.

#### 2.2. an Example: Seven Bridges of Könisberg

Leonhard Euler is one of the greatest mathematicians of all the time. In his time, there was a famous but unsolved problem concerning seven bridges of Könisberg which he solved in 1736. The problem is stated as follows.

---

2By the definition of edges and adjacency of vertices, $v$ is not an element of $N(v)$. 
Question 2.3. Is it possible to cross all the 7 bridges of Königsberg without crossing the same bridge more than once?

Euler formulated this problem into an abstract setting which is now known to be a graph. He considered every bridge as an edge and shrunk the region that is reachable without crossing bridges into a vertex. After investigations, Euler concluded that the answer of Question 2.3 is NO because of the following celebrated observation.

Theorem 2.4 (Euler, 1736). Given a graph $G = (V, E)$, there is a path using every edge exactly once if and only if there are only two vertices of odd degree or there are no vertices of odd degree.

Look at Figure 2. Then you can see that there are 4 vertices of odd degree. So we can easily conclude that Euler’s answer is TRUE. To celebrate his work, if there is a path satisfying the condition in Theorem 2.4, such a path (and even the graph itself) is called an Eulerian circuit.

So, how can one prove Theorem 2.4? Well, $\Rightarrow$ direction is easy. When we follow the given Eulerian circuit, then at each step, it contributes 2 to every middle vertices of Eulerian circuit. So you just need to think about the starting and ending vertices. And there are 2 cases: (i) The starting vertex is the same as the ending vertex. (ii) The starting vertex is different from the ending vertex. One can easily see that (i) is the case when there are no vertices of odd degree while (ii) is the case when there are two vertices of odd degree.

For $\Leftarrow$, you can use induction. If the given graph $G$ satisfies the condition in the righthand side, you can find a smaller path $P$ than $G$ with no repeated edges such that starting and ending vertices are the same (Why?). After deleting edges of $P$ from $G$, the resulting graph also satisfies the condition on the righthand side. So we can use mathematical induction.

2.3. Application to our main problem. Now we go back to Question 1.2. Actually, I want to restrict Question 1.2 into a smaller problem like below, since two questions are the same in principle, but the original one is more involved in computation.

Question 2.5. What is the maximum Eulerian circuit which is a subgraph of Seoul Metro such that the starting vertex and the ending vertex are the same?

Here, a subgraph of a given graph $G$ means the smaller graph contained in $G$. And by maximum, we mean that the given graph is maximum by the number of edges. From now on, we will call an subgraph of $G$ satisfying the condition in Question 2.5 an Eulerian subgraph of $G$. 
Now, we imagine that there is an Eulerian subgraph $H$ of $G$. And let $S$ be a subgraph of $G$ obtained by deleting all the edges of $H$ from $G$. Note that the degree of every vertex of $H$ is even. Hence, $v$ is a vertex of odd degree in $G$ if and only if $v$ is a vertex of odd degree in $S$. We denote the set of vertices of odd degree in $G$ (and also in $S$) as $O$. But how many vertices of odd degree?

**Theorem 2.6.** Given a graph $G$, the number of vertices of odd degree is even.

*Proof.* We count $n$ which is the number of pairs $(v, e) \in V \times E$ where $v$ is in $e$. Then, $\sum_{v \in V} \deg(v) = n = 2|E|$. This implies that there must be even number of vertices of odd degree. \qed

So, at each component\(^3\) of $S$, we can find an even number of vertices of odd degree. Choosing an arbitrary pair $v$ and $w$ in the same component among them, we can find a path $P$ between them. After deleting edges of $P$ in $G$, $v$ and $w$ becomes vertices of even degree and the parities of other vertices remain the same. In this way, we can find $k := |O|/2$ paths $P_1, P_2, \ldots, P_k$. Note that those paths are *edge-disjoint*, that is, share no common edges. If $S$ were not the same as $\bigcup_{j=1}^k P_j$, we could choose $\bigcup_{j=1}^k P_j$ rather than $S$ to get bigger $H$. And less edges in $\bigcup_{j=1}^k P_j$ we have, more edges in its complement which is an Eulerian subgraph. So, to find a maximum Eulerian subgraph of $G$, it is sufficient to find $k$ edge-disjoint paths with minimum total number of edges.

And due to the following theorem, we can delete the condition of edge-disjointness.

**Theorem 2.7.** Let $O$ be the set of vertices of odd degree in the given graph $G$. Let’s assume that $|O| = 2k$ for a fixed a natural number $k$. And let

$\mathcal{P} := \{\{L_1, \ldots, L_k\} : L_j$’s are paths between vertices in $O$ and if $i \neq j$, $L_i$ and $L_j$ share no ending points.$\}$

If $\{P_1, \ldots, P_k\}$ is the minimum by the number of edges in its union among elements in $\mathcal{P}$, then $P_j$’s are *edge-disjoint*.

*Proof.* Suppose $\{P_1, \ldots, P_k\}$ is the minimum among elements in $\mathcal{P}$, and suppose that $P_1 := x_1x_2 \ldots x_l$ and $P_2 := y_1y_2 \ldots y_m$ share an edge $vw$. WLOG, let $x_l = v = y_j$ and $x_{l+1} = w = y_{j+1}$. Then, we can find a new path $Q_1 = x_1 \ldots x_{j-1}y_{j-1}x_{j+1}y_j \ldots y_l$ and $Q_2 = x_{l+1} \ldots x_{j+1}y_{j+2}y_{j+3} \ldots y_m$ in $P_1 \cup P_2$ not using $vw$. Then the union of $\{Q_1, Q_2, P_3, \ldots, P_k\}$ has smaller number of edges than the union of $\{P_1, \ldots, P_k\}$ which leads to a contradiction. \qed

For the simplicity of notations, we will usually call a $k$-set of paths like $\{P_1, \ldots, P_k\}$ as in Theorem 2.7. an *optimal $k$ paths*.

### 3. The real computation

We follow the following steps to compute the optimal paths discussed in Section 2.3.

**STEP 1:** Make a simple\(^4\) graph $M$ from the graph of Seoul Metro.

**STEP 2:** Compute the shortest paths between odd degree vertices of $M$ and their lengths.

**STEP 3:** Find a perfect matching among vertices of odd degree with minimum cost.

\(^3\)A component of a graph $G$ is a maximal connected subgraph of $G$

\(^4\)In the terminology of graph theory, the world *simple* has a different meaning. In this context we assume that this word has the usual meaning as an English word.
It is alright if you don’t understand every word in those steps. They will be introduced in the following subsections.

3.1. Making a simple graph out of the graph of Seoul Metro. Suppose there is a vertex \( v \) of degree 1 in the given graph \( G \). Then \( v \) and the edge \( e \) which is the only edge containing \( v \) cannot be a part of an Eulerian subgraph of \( G \). Hence, we first delete every vertex \( v \) of degree 1 and the only edge incident to \( v \) in the graph of Seoul Metro system until there are no vertices of degree 1. Let \( \tilde{M} \) be the resulting graph.

Now, suppose that there is a vertex \( w \) of degree 2 in \( G \). If an Eulerian subgraph of \( G \) contains \( w \), then it must contain all two edges incident with \( w \), automatically. So, we just delete \( w \) and its incident edges \( xw \) and \( wy \), and add an edge \( xy \). To preserve the information of distance, we define the weight of each edge \( e \) (or \( w(e) \)) as 1 at the beginning, and at each step of deleting and adding edges, we newly define \( w(xy) = w(xw) + w(wy) \) while other weights remain the same. We apply this to \( \tilde{M} \) until there are no vertices of degree 2. And let \( M \) be the resulting graph.

This procedure is to reduce the time of computation. Note that although the edges nearby 응암 is directed, we temporarily forget about this direction in this procedure. For instance, we just think we can go from 응암 to 구산 directly even if it is not possible. The resulting \( M \) consists of 73 vertices and 148 edges and there are 18 vertices of odd degree. And the list of the vertices of odd degree is:

응암, 노원, 광운대, 태평로, 부평구청, 부평, 구로, 까치산, 가좌, 신도림, 정자, 모란, 천호, 강남, 종합운동장, 성수, 신길동, 왕십리.

Figure 3. the simplified Seoul Metro System
3.2. **Compute the shortest paths between odd degree vertices of** $M$ **and their lengths.** In Step 2 and 3, we use mathematical software called SAGE. SAGE has a good library of graph algorithms useful for our purpose. We will mainly concentrate on how to make use of those algorithms rather than considering the principle of them because those are beyond of our scope.

Now, we are given a graph $M$ with weighted edges and let $O$ be the set of vertices of odd degree of $M$. And we want to compute the shortest paths between vertices in $O$. You can use the function `shortest_path(u, v, by_weight=True)` in SAGE for it where $u, v$ are the vertices in $O$.

3.3. **Find a perfect matching among vertices of odd degree with minimum cost.** Let $|O| = 2k$. Our purpose is to find optimal $k$ paths discussed in Section 2. Since $k$ paths are disjoint if they are chosen to be optimal, it is equivalent to find a perfect matching among vertices of odd degree with minimum cost. But what? Wait! What is a perfect matching? I will explain that.

Given a graph $G$, a matching $P$ is a set of edges of $G$ such that two edges $e, f \in P$ share no vertices. And if there is an edge $vw \in P$, then we say that $v$ and $w$ are matched. Intuitively, it is very similar to find a mate between boys and girls: No girls have two boyfriends at the same time, and vice versa. And $P$ is a perfect matching of $G$ if every vertex of $G$ is matched by $P$. So, if there is a perfect matching in $G$, there would be no lonely vertices.

Now, we construct a graph $K$ with $2k$ vertices each of which represents a vertex in $O$ and between every pair of distinct vertices $v$ and $w$ there is an edge with a weight given by their distance in $M$. It is just another representation of paths between vertices in $O$. So finding a perfect matching in $K$ with minimum cost is the same as finding optimal $k$ paths.

And in SAGE, there is a function `matching()` which computes the matching of the given graph with maximum cost. If we flip the sign of every weight of $K$ and add the sufficiently large number in each weight samely, the result of the function `matching()` is exactly the perfect maching of $K$ with minimum cost.

And here is the result which are optimal nine paths $P_1$ to $P_9$ among vertices in $O$.

- $P_1 = 응암 - 디지털미디어시티 - 김포공항 - 가지산$
- $P_2 = 노원 - 태릉입구$
- $P_3 = 광운대 - 상봉 - 회기 - 청량리 - 신설동$
- $P_4 = 부평구청 - 부평$
- $P_5 = 구로 - 신도림$
- $P_6 = 가좌 - 홍대입구 - 공덕 - 용산 - 이촌 - 동작 - 고속터미널 - 교대 - 강남$
- $P_7 = 정자 - 모란$
- $P_8 = 천호 - 잠실 - 종합운동장$
- $P_9 = 성수 - 왕십리$

After deleting these paths, we get the desired Eulerian subgraph of the graph of Seoul Metro like Figure 4. In this resulting graph, the direction nearby 응암 station does not really matter, since we can stop at 응암 and take a subway to the other direction by going to the other platform.

We can even visualize the Eulerian circuit of this subgraph. First we delete six cycles $C_1$ to $C_6$ defined as follows.

- $C_1 = 디지털미디어시티 (경의선) 서울역 (공항철도선) 디지털미디어시티
Figure 4. the maximum Eulerian subgraph in the graph of Seoul Metro

- $C_2 = \text{전대입구 (2호선) 신길동 (1호선) 신도림 (2호선) 대림 (7호선) 전대입구}
- $C_3 = \text{상봉 (경춘) 망우 (경의 중앙) 상봉}
- $C_4 = \text{울지로4가 (2호선) 동대문역사문화공원 (5호선) 울지로4가}
- $C_5 = \text{연신내 (6호선) 별광 (6호선) 음암 (6호선) 연신내}
- $C_6 = \text{까치산 (2호선) 신도림 (2호선) 영등포구청 (5호선) 까치산}

Then the remaining graph contains a big cycle as shown in Figure 5.

Figure 5. the desired Eulerian circuit
Note that every gray region corresponds to a small cycle surrounding it. By adding cycles $C_1$ to $C_6$, we can get the desired Eulerian circuit. And here is one example of such an Eulerian circuit.

\[ \text{E-mail address: prosolver@kaist.ac.kr} \]

\[ \text{Enjoy your trip!} \]