Homework 3

Victor Croisfelt
victorcroisfelt@gmail.com

January 14, 2022

1 Subgraphs and Order Embeddings

First, we give the definition of a subgraph.

Definition 1.1. (Subgraph) Let $A = (\mathcal{V}_A, \mathcal{E}_A)$ and $B = (\mathcal{V}_B, \mathcal{E}_B)$ denote two different graphs. Without loss of generality, graph A is said to be a **subgraph** of B if and only if (iff) there exists a bijection or one-to-one correspondence $f_{AB'} : \mathcal{V}_A \mapsto \mathcal{V}_{B'}$ such that $(f(u), f_{AB'}(v)) \in \mathcal{E}_B$ for $u, v \in \mathcal{V}_A$ and $(u, v) \in \mathcal{E}_A$; B' is an induced subgraph of B. We denote the subgraph relationship between A and B by overloading the operator $\subseteq$ as $A \subseteq B$.

In the following, we consider that we have an embedding space which preserves the subgraph relationship $\subseteq$ by making all entries of the embedding vector $\mathbf{z}_A \in \mathbb{R}^d_+$ of graph A be smaller than the embedding entries of $\mathbf{z}_B \in \mathbb{R}^d_+$, where d is the dimension of the embedding space. Thus, in the embedding space, a graph A is said to be a subgraph of B iff $\mathbf{z}_A \preccurlyeq \mathbf{z}_B$, where $\preccurlyeq$ evaluates $\leq$ to all d dimensions.

Now, we are good to go :)

Question 2.1 Transitivity

We have three graphs $A = (\mathcal{V}_A, \mathcal{E}_A)$, $B = (\mathcal{V}_B, \mathcal{E}_B)$, and $C = (\mathcal{V}_C, \mathcal{E}_C)$. A is a subgraph of B. B is a subgraph of C. Show that A is also a subgraph of C. We answer this question in two parts.

1) Playing with sets. First, if A is a subgraph of B, the set of nodes $\mathcal{V}_B$ can be decomposed into two disjoint sets:

$$\mathcal{V}_B = \mathcal{V}_{B'} \cup (\mathcal{V}_B \setminus \mathcal{V}_{B'}), \tag{1}$$

where $\mathcal{V}_{B'} = \{f_{AB'}(u) | u \in \mathcal{V}_A\}$ represents the induced subgraph of B which is graph-isomorphic to A. Now, if B is a subgraph of C, we can do something similar and write:

$$\mathcal{V}_C = \mathcal{V}_{C'} \cup (\mathcal{V}_C \setminus \mathcal{V}_{C'}), \tag{2}$$

where $\mathcal{V}_{C'} = \{f_{BC'}(v) | v \in \mathcal{V}_B\}$ is graph-isomorphic to B.

Lets take a closer look at $\mathcal{V}_{C'}$. By using (1), we have that

$$\mathcal{V}_{C'} = \{f_{BC'}(v)|v \in \mathcal{V}_B\} = \{f_{BC'}(v)|v \in (\mathcal{V}_{B'} \cup (\mathcal{V}_B \setminus \mathcal{V}_{B'}))\} \tag{3}$$

$$\stackrel{(a)}{=} \{f_{BC'}(n)|n \in \mathcal{V}_{B'}\} \cup \{f_{BC'}(m)|m(\in \mathcal{V}_B \setminus \mathcal{V}_{B'})\} \tag{4}$$

$$\stackrel{(b)}{=} \{f_{BC'}(n)|n \in \{f_{AB'}(u)|u \in \mathcal{V}_A\}\} \cup \{f_{BC'}(m)|m \in (\mathcal{V}_B \setminus \mathcal{V}_{B'})\} \tag{5}$$

$$\stackrel{(c)}{=} \{f_{BC'}(f_{AB'}(u))|u \in \mathcal{V}_A\} \cup \{f_{BC'}(m)|m \in (\mathcal{V}_B \setminus \mathcal{V}_{B'})\}, \tag{6}$$

where in (a) we used the fact that the sets are disjoint, in (b) we used the definition of $\mathcal{V}_{B'}$ which is graph-isormorphic to $\mathcal{V}_A$, and in (c) we substituted n. From the expressions above, we can conclude that $\mathcal{V}_{C'}$ can be decomposed into two disjoint pats, where one of these parts is graph-isomorphic to A. This conclusion relies on the fact that the sets are disjoint and the bijection of a bijection is also a bijection, meaning that $f_{BC'}(f_{AB'}(u)) : \mathcal{V}_A \mapsto \mathcal{V}_C''$. Since $\mathcal{V}_{C'}$ is a disjoint subset of $\mathcal{V}_C$, this shows that A is also a subgraph of C.

However, we are still not satisfied. The above conclusion is based on the fact that a bijection of a bijection is also bijection. How do we know that?

2) Is a Bijection of a Bijection, a Bijection?.[1] To show this[2], let us forget about subgraphs and graph-isomorphism a little. Thus, we have that: $f : \mathcal{A} \mapsto \mathcal{B}$ and $g : \mathcal{B} \mapsto \mathcal{C}$ are bijections. Proof that the composition $h = g \circ f$ is also a bijection for $h : \mathcal{A} \mapsto \mathcal{C}$. To do so, let us recall the definition of a bijective function.

Definition 1.2. A function $f : \mathcal{X} \mapsto \mathcal{Y}$ is said to be a bijection if the following conditions are satisfied:

- every element of $\mathcal{X}$ must be paired with exactly one element of $\mathcal{Y}$, that is, $\mathcal{X}$ is the domain of f;
- each element of $\mathcal{Y}$ must be paired with at least one element of $\mathcal{X}$, that is, the codomain and the image of f are the same thing – f is a *surjection*;
- no element of $\mathcal{Y}$ must be paired with more than two elements of $\mathcal{X}$, that is, f is a *injection* (one-to-one).

From these conditions, we can see that a bijective function is simultaneously a subjection and an injection.

Proof. We now proof that the composition of bijections is a bijection. Note that $f(\mathcal{A}) = \mathcal{B}$ and $g(\mathcal{B}) = \mathcal{C}$.

First, we show that h is surjective. To do so, we evaluate what happens when h is

[1] Is someone missing a bijection anywhere?

[2] We will refer as "this", because... you know.

evaluated over its domain, yielding in:

$$
\begin{aligned}
h(\mathcal{A}) &= (g \circ f)(\mathcal{A}) && (7)\\
&\overset{(a)}{=} \{c \in C | (g \circ f)(a),\ \text{for } a \in \mathcal{A}\} && (8)\\
&\overset{(b)}{=} \{c \in C | g(f(a)) = c,\ \text{for } a \in \mathcal{A}\} && (9)\\
&\overset{(c)}{=} \{c \in C | g(b) = c,\ \text{for } b \in f(\mathcal{A})\} && (10)\\
&\overset{(d)}{=} g(f(\mathcal{A})) && (11)\\
&\overset{(e)}{=} g(\mathcal{B}) && (12)\\
&\overset{(f)}{=} C, && (13)
\end{aligned}
$$

where in (a) we use the definition of the function h (evaluate it over all domain $a \in \mathcal{A}$ to obtain an element in the co-domain C), (b) we use the definition of composition, (c) we use the fact that, if $f(a)$ is surjective, $f(a)$ is mapped to $b \in \mathcal{B}$, in (d) we use the definition of g (evaluate it over the domain), in (e) we explicitly use the fact that f is a surjection, and we do the same in (f) for g. So, h is surjective, since going from its domain $\mathcal{A}$, all elements of its co-domain C are mapped at least to one element from $\mathcal{A}$.

Second, we show that h is injective by evaluating if $h(a) = h(a')$ implies that $a = a'$ for arbitrary $a, a' \in \mathcal{A}$. Thus, assuming that this equation is true,

$$
\begin{aligned}
h(a) &= h(a') && (14)\\
(g \circ f)(a) &= (g \circ f)(a') && (15)\\
g(f(a)) &= g(f(a')). && (16)
\end{aligned}
$$

The above holds true since $g(b) = g(b')$, since g is a bijection; the same applies for f. Then, h is an injection.

Since h is simultaneously a surjection and an injection, h is a bijection. □

This ends our way of showing that **Transitivity** holds for the Subgraph definition: $(A \subseteq B) \wedge (B \subseteq C) \implies A \subseteq C$, for graphs A, B, and C[3].

Question 2.2 Anti-symmetry

The anti-symmetry property reads as follows: if A is a subgraph of B, and B is a subgraph of A, then A and B are graph-isomorphic. To show that, we evaluate each of the conditions individually.

1) If $A \subseteq B$, there exists a subgraph B' of B defined by the nodes $\mathcal{V}_{B'} = \{f_{AB'}(v) | v \in \mathcal{A}\}$. Note that $|\mathcal{V}_{B'}| = |\mathcal{V}_A|$.

2) If $B \subseteq A$, there exists a subgraph A' of A defined by the nodes $\mathcal{V}_{A'} = \{f_{BA'}(u) | u \in \mathcal{B}\}$. Note that $|\mathcal{V}_{A'}| = |\mathcal{V}_B|$.

Therefore, if we want that 1) and 2) occur together, we have that:

$$|\mathcal{V}_{B'}| = |\mathcal{V}_A| \wedge |\mathcal{V}_{A'}| = |\mathcal{V}_B|.$$

[3] $\wedge$ stands for logical and.

Since $|\mathcal{V}_B| \geq |\mathcal{V}_{B'}|$ and $|\mathcal{V}_A| \geq |\mathcal{V}_{A'}|$, we can rewrite the condition as:

$$|\mathcal{V}_B| \geq |\mathcal{V}_A| \wedge |\mathcal{V}_A| \geq |\mathcal{V}_B|,$$

both can only hold true together if $|\mathcal{V}_A| = |\mathcal{V}_B|$. Since the number of nodes in both graphs are the same, $\mathcal{V}_A = \mathcal{V}_{A'}$ and $\mathcal{V}_B = \mathcal{V}_{B'}$. This means that $f_{AB} : \mathcal{V}_A \mapsto \mathcal{V}_B$, while $f_{BA} : \mathcal{V}_B \mapsto \mathcal{V}_A$ is its inverse. Due to the respective edge mappings coming from the subgraph definition, we conclude that A and B are graph-isomorphic.

Question 2.3 Common subgraph

Let A have embedding $\mathbf{z}_A \in \mathbb{R}^2_+ = [x_A, y_A]^T$, B have embedding $\mathbf{z}_B \in \mathbb{R}^2_+ = [x_B, y_B]^T$, and X have embedding $\mathbf{z}_X \in \mathbb{R}^2_+ = [x_X, y_X]^T$. We have the following conditions for X be a common subgraph of A and B.
1) If X is a subgraph of A, $\mathbf{z}_X \preccurlyeq \mathbf{z}_A$ or $x_X \leq x_A \wedge y_X \leq y_A$.
2) If X is a subgraph of B, $\mathbf{z}_X \preccurlyeq \mathbf{z}_B$ or $x_X \leq x_B \wedge y_X \leq y_B$.

Putting both together and evaluating per coordinates:

$$\text{for x-dimension, } x_X \leq x_A \wedge x_X \leq x_B \tag{17}$$
$$\wedge \tag{18}$$
$$\text{for y-dimension, } y_X \leq y_A \wedge y_X \leq y_B. \tag{19}$$

Alternatively, this can be written as $x_X \leq \min\{x_A, x_B\} \wedge y_X \leq \min\{y_A, y_B\}$. Following the independence of the dimensions, we have that X is simultaneously a subset of A and B iff:

$$\mathbf{z}_X \leq [\min\{x_A, x_B\}, \min\{x_A, x_B\}]^T \preccurlyeq \min\{\mathbf{z}_A, \mathbf{z}_B\}. \tag{20}$$

Q2.4 Order embeddings for graphs that are not subgraphs of each other

Given that $\mathbf{z}_A[0] > \mathbf{z}_B[0] > \mathbf{z}_C[0]$,[4] we need to look at the second dimension in order to avoid that one of the graphs is graph-isomorphic or a subgraph of another graph. Formally, the conditions are
a) Non-isomorphic:

$$\mathbf{z}_A \neq \mathbf{z}_B \neq \mathbf{z}_C, \tag{21}$$

where $\neq$ evaluates all dimensions simultaneously.
b) Non-subgraph: the observations below are based on $\mathbf{z}_A[0] > \mathbf{z}_B[0] > \mathbf{z}_C[0]$.

- **Graph A:** there is no way that A is a subgraph of the other two. Therefore, $\mathbf{z}_A[1]$ is free to be placed over $\mathbb{R}_+$ by now. It depends on the constraints imposed by the following observations.

- **Graph B:** B can be a subgraph of A if $\mathbf{z}_B[1] \leq \mathbf{z}_A[1]$. Therefore, $\mathbf{z}_B[1] > \mathbf{z}_A[1]$. B cannot be a subgraph of C.

[4] [0], [1] index the first and second dimension of a vector.

- **Graph C:** C can be a subgraph of A if $\mathbf{z}_C[1] \leq \mathbf{z}_A[1]$. In addition, C can be a subgraph of B if $\mathbf{z}_C[1] \leq \mathbf{z}_B[1]$. Therefore, $\mathbf{z}_C[1] > \mathbf{z}_A[1]$ and $\mathbf{z}_C[1] > \mathbf{z}_B[1]$.

By putting the three observations for non-subgraph and the non-isomorphic condition together, we have that

$$\mathbf{z}_B[1] > \mathbf{z}_A[1] \wedge \mathbf{z}_C[1] > \mathbf{z}_A[1] \wedge \mathbf{z}_C[1] > \mathbf{z}_B[1]. \tag{22}$$

Therefore,

$$\mathbf{z}_C[1] > \mathbf{z}_B[1] > \mathbf{z}_A[1]. \tag{23}$$

Given the derived relations of the embeddings $\mathbf{z}_A[0] > \mathbf{z}_B[0] > \mathbf{z}_C[0] \wedge \mathbf{z}_C[1] > \mathbf{z}_B[1] > \mathbf{z}_A[1]$ so as to have three non-isomorphic and non-subgraphs graphs, Fig. 1 illustrates the embedding space.

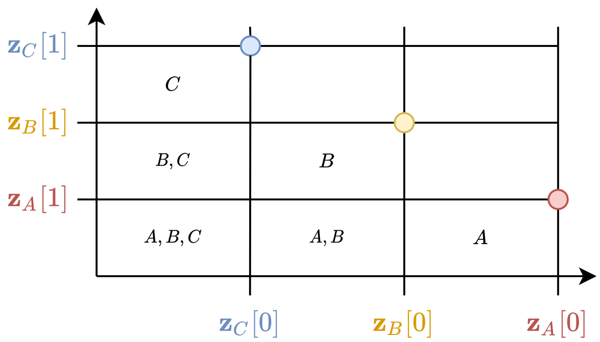

Figure 1: Embedding space considering three non-isomorphic and non-subgraphs graphs: A, B, and C. The labels within the boxes formed by the node embeddings coordinates indicate of which graphs an arbitrary graph would be a subgraph if its embedding was placed within that box.

Q2.5 Limitation of 2-dimensional order embedding space

Fig. 2 visually evaluates the regions in which we can place the subgraphs X, Y, Z so as to have the desired relationships:

- X is a common subgraph of A and B;
- Y is a common subgraph of A and C;
- Z is a subgraph of C.

Note that we cannot have disjoint regions in which Y holds true alone (this is the limitation).

To better evaluate the scenario, we follow the proposed guidelines. Without loss of generality, we assume that $\mathbf{z}_A[0] > \mathbf{z}_B[0] > \mathbf{z}_C[0]$, therefore $\mathbf{z}_C[1] > \mathbf{z}_B[1] > \mathbf{z}_A[1]$ (Q2.4). We construct the subgraphs X, Y, Z to obtain the relationships as defined above. From those, we have:

- $X \subseteq (A, B)$: $\mathbf{z}_X \preccurlyeq \mathbf{z}_A \wedge \mathbf{z}_X \preccurlyeq \mathbf{z}_B$ or $\mathbf{z}_X \preccurlyeq \min\{\mathbf{z}_A, \mathbf{z}_B\}$, which evaluates to $\mathbf{z}_X \preccurlyeq [\mathbf{z}_B[0], \mathbf{z}_A[1]]^T$ given the assumptions;

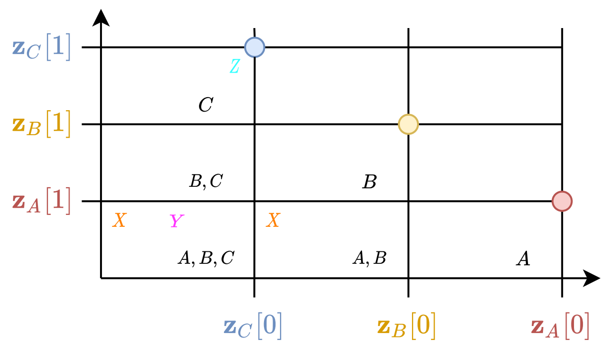

Figure 2: Embedding space considering three non-isomorphic and non-subgraphs graphs: A, B, and C. X is a common subgraph of A and B. Y is a common subgraph of A and C. Z is a subgraph of C.

- $Y \subseteq (A, C)$: $\mathbf{z}_Y \preccurlyeq \mathbf{z}_A \wedge \mathbf{z}_Y \preccurlyeq \mathbf{z}_C$ or $\mathbf{z}_Y \preccurlyeq \min\{\mathbf{z}_A, \mathbf{z}_C\}$, which evaluates to $\mathbf{z}_Y \preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_A[1]]^T$ given the assumptions;
- $Z \subseteq C$: $\mathbf{z}_Z \preccurlyeq \mathbf{z}_C$, which evaluates to $\mathbf{z}_Z \preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_C[1]]^T$ given the assumptions.

As suggested, we want to satisfy $\mathbf{z}_X \preccurlyeq \mathbf{z}_Y \wedge \mathbf{z}_X \preccurlyeq \mathbf{z}_Z$. By splitting the analysis, we have:[5]

1. $\mathbf{z}_X \preccurlyeq \mathbf{z}_Y \implies \mathbf{z}_X \preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_A[1]]^T$
2. $\mathbf{z}_X \preccurlyeq \mathbf{z}_Z \implies \mathbf{z}_X \preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_C[1]]^T$

Therefore, X needs to satisfy three different conditions simultaneously

$$\begin{aligned} \mathbf{z}_X &\preccurlyeq [\mathbf{z}_B[0], \mathbf{z}_A[1]]^T, \text{ from } X \subseteq (A, B) && (24)\\ &\preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_A[1]]^T, \text{ from } \mathbf{z}_X \preccurlyeq \mathbf{z}_Y && (25)\\ &\preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_C[1]]^T, \text{ from } \mathbf{z}_X \preccurlyeq \mathbf{z}_Z. && (26) \end{aligned}$$

By putting those together (taking the minimum over each dimension), we have that $\mathbf{z}_x \preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_A[1]]^T$. Therefore, in the end, we have:

$$\begin{aligned} \mathbf{z}_X &\preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_A[1]]^T, && (27)\\ \mathbf{z}_Y &\preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_A[1]]^T, \text{ and} && (28)\\ \mathbf{z}_Z &\preccurlyeq [\mathbf{z}_C[0], \mathbf{z}_C[1]]^T. && (29) \end{aligned}$$

From these, we can conclude that X is a subgraph of Z, because $\mathbf{z}_C[1] > \mathbf{z}_A[1]$; and that, X is also a subgraph of Y, since $\mathbf{z}_X \preccurlyeq \mathbf{z}_Y$ is an imposed condition. Note that $\mathbf{z}_X$ and $\mathbf{z}_Y$ are confined to the same space.

Conclusion. The limitation is that some relationships cannot be realized depending on the embedding dimension d. For example, for $d = 2$, we do not have a region where a graph is simultaneously a subgraph of A and C without being of B.

[5]Recall to substitute the thinks that we figured out above for $\mathbf{z}_Y$ and $\mathbf{z}_Z$