LACOMBE Yoach
yoach.lacombe@gmail.com

1 Question 1

We expect the weights to be the identity matrix and the bias to be zero, in order for $\phi(x)$ to be x. Indeed, then, we will sum over the $\phi(x_i) = x_i$ to get exactly what we want.

2 Question 2

The architectural differences between the lab6 GNN and the DeepSets architecture are quite clear. Indeed, whereas DeepSets is constrained to be $\rho(\sum \phi(x_i))$, with the same ϕ over all of the elements of the set, the GNN is in the form $\rho(\sum \phi_i(x_i))$ with a potentially different ϕ for each node of the graph.
There are no conceptual differences between a set and a graph without edges, both are unordered series of elements.

3 Question 3

The response of this question relies on the usefulness of an embedding layer. Indeed, here, the embedding layer maps the nodes to continuous vector space, which allows numerical comparisons and operations between each nodes. Using a fully-connected layer instead would prevent the network from using the advantages of an embedding, because it would use the nodes id as input of the layer. Yet, there are no order relation in the space of the nodes (contrarily to the space of the embedding), so the fully connected layer would be confused by the apparent order given by the indexing of the nodes. In other words, since the ids are symbolic, the network needs a mapping of the symbols to a vector space in order to be the most efficient.

4 Question 4

It would be a terrible idea to use DeepSets to the above problem. Indeed, the recommendation task highly rests on the input sessions, which themselves highly rest on the order of navigation through the items. In other words, there is an order in the click behavior of the user whereas DeepSets is specialized on unordered ensembles. Using DeepSets would thus be sub-optimal.

References