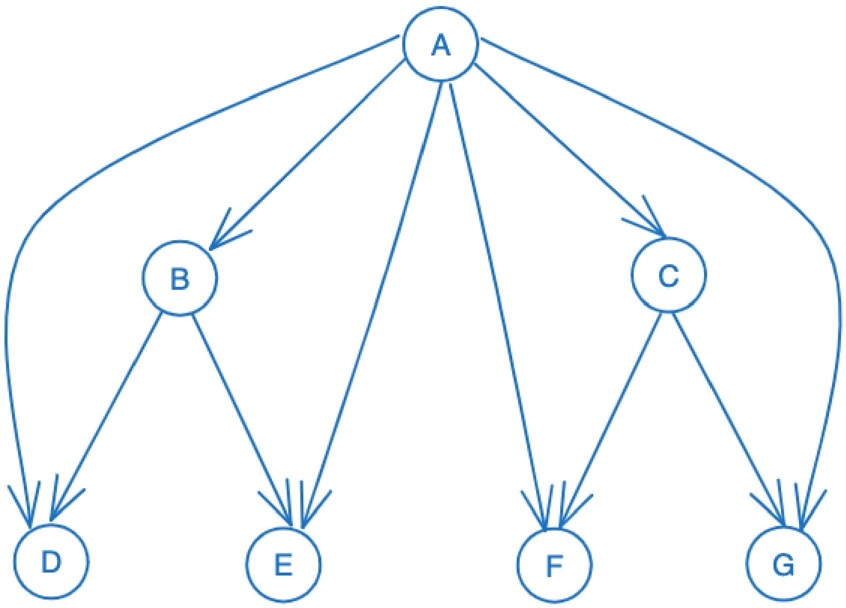
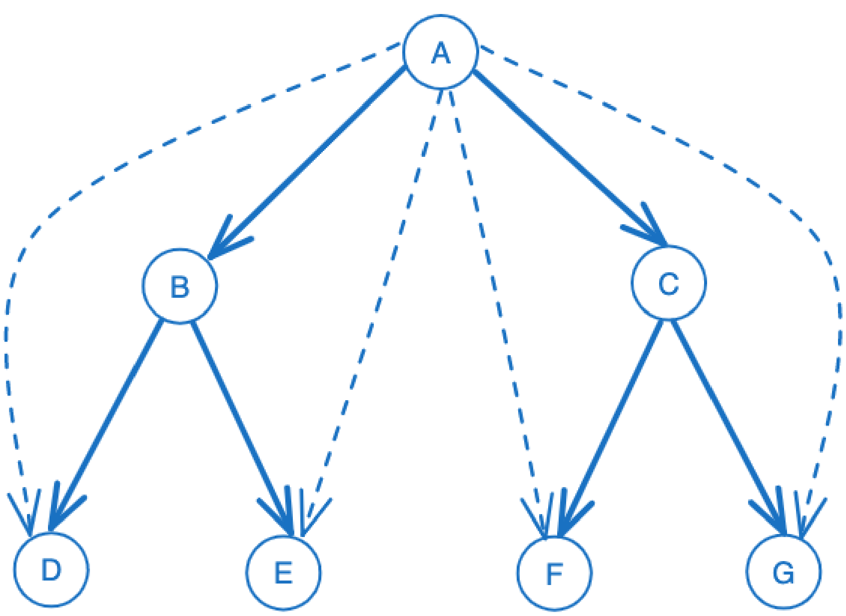
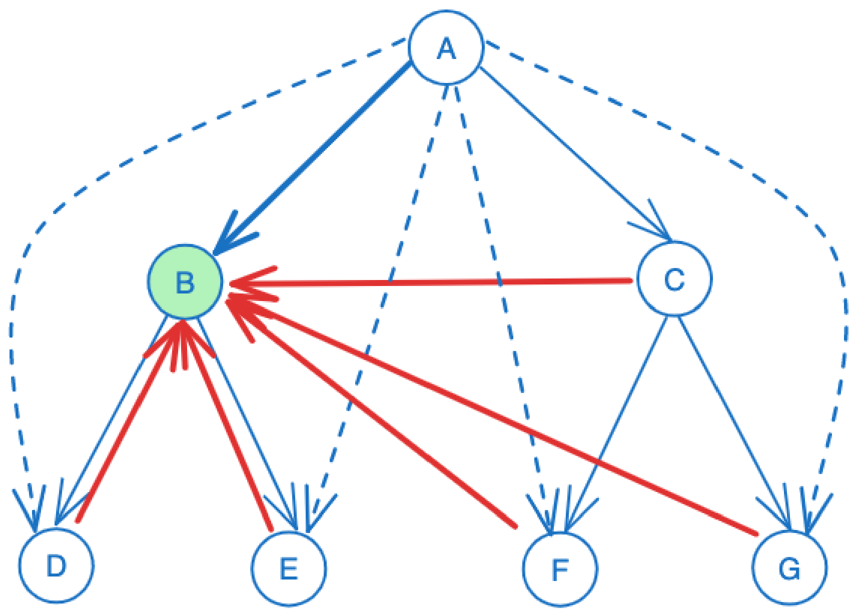
Below is a transitively-closed graph.



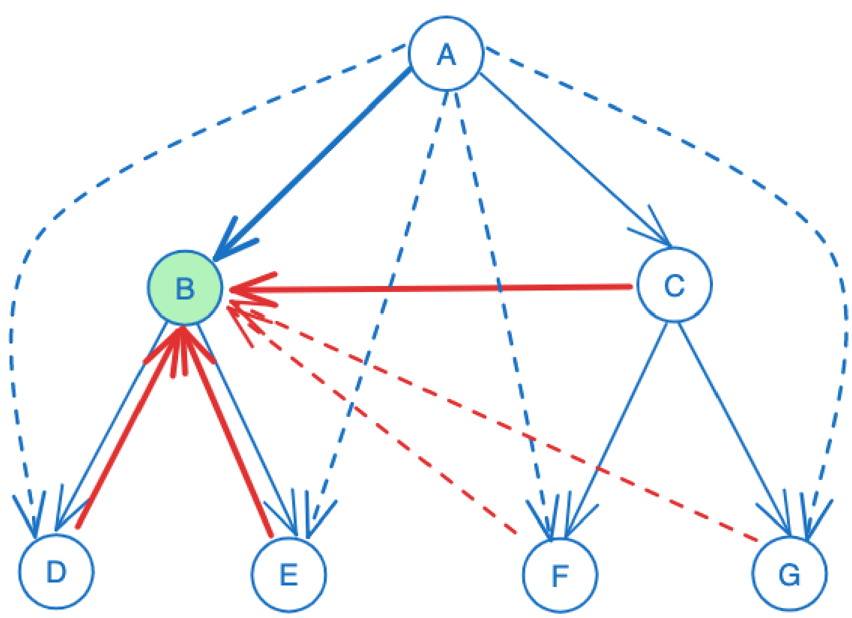
It is well known that we can use edges in the transitive reduction and recover those edges in the transitive closure, and for a model with transitive bias we do not need to explicitly train on the dotted edges below.



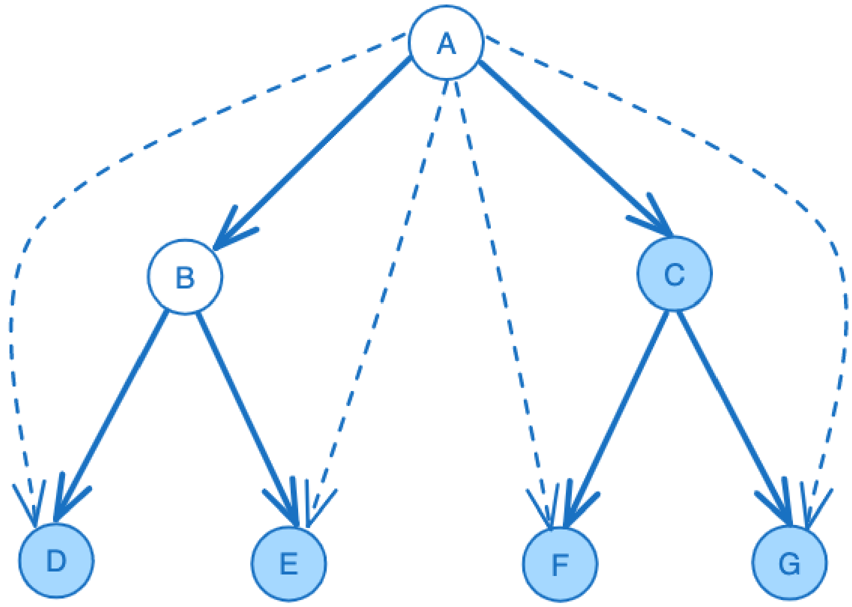
However, when training, we also must specify some negative edges. Consider the positive edge (A, B). The negative edges ending at B are $(C, B), (D, B), (E, B), (F, B), (G, B)$, indicated below in red.



Due to the transitivity property, however, we can omit explicitly mentioning the edges (F, B) and (G, B), as they are implied to not exist once we know (C, B) does not exist.



We can connect this to the notion of the minimal reachability-equivalent subset $(V \setminus \mathrm{An}^{\cup}(B))^{mres}$. The set $(V \setminus \mathrm{An}^{\cup}(B))$ is shown below shaded in blue.



The minimal reachability-equivalent subset $(V \setminus \mathrm{An}^{\cup}(B))^{mres}$ excludes F and G from the shaded nodes $(V \setminus \mathrm{An}^{\cup}(B))$ in the figure above, leaving only nodes C, D, and E as the heads of the “hierarchy-aware” negative edges ending at B.