Notes, 7(c)

ECE 606

Optimal Substructure, Revisited

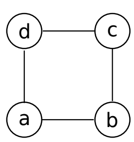
Optimal substructure is quite an important notion to appreciate, as it can tell us how to design an algorithm. Similar to the greedy choice property. What is optimal substructure?

Informally, it is: within an optimal solution is an optimal solution to a subproblem. We have seen several examples of problems that possess optimal substructure.

Are there problems that appear to not demonstrate optimal substructure? The answer is: yes, there appear to be problems that do not. But we cannot necessarily prove this, as if we could, we would resolve arguably the most famous open problem on the planet. Which we will discuss a bit later in the course.

We have observed that shortest simple path in a graph demonstrates optimal substructure: a sub-path of a shortest path is a shortest-path. How about longest simple path?

~~It appears that~~ longest simple path does not demonstrate the same kind of optimal substructure as shortest simple path. Consider the following graph.



A longest simple path $a \rightsquigarrow c$ is $a \to d \to c$. However, the subpath $d \to c$ is not a longest simple path $d \rightsquigarrow c$. Rather, $d \to a \to b \to c$ is.

You'll notice that the above graph has a cycle. What if we are given an acyclic graph? Then it turns out indeed that longest simple path does possess optimal substructure exactly in the manner that shortest simple path does.

Another example of a problem that does not appear to possess optimal substructure is maximum-sized clique.

Given an undirected graph $G = \langle V, E \rangle$, a *clique* in G is subset of the vertices $C \subseteq V$ such that $u, v \in C, u \neq v \implies \langle u, v \rangle \in E$. That is, the subgraph of G that is induced by C is complete, i.e., has all possible edges. The maximum-sized clique problem is, given undirected G, to find a clique of maximum size.

Consider the graph from the previous page. The size of a maximum-sized clique is 2. Such a clique is $\{a, b\}$. We have three other cliques of that size: $\{d, c\}$, $\{c, b\}$ and $\{a, d\}$.

Maximum-sized clique does not possess the following optimal substructure. Suppose our graph is $G = \langle V, E \rangle$ and C is a maxium-sized clique in it. Now, if I pick a vertex $u \in C$ and remove: (i) u from C, (ii) u from V, and, (iii) every edge in G that is incident on u, then I do not necessarily end up with a maximum-sized clique in the resultant graph.

We can see this from the above example. If I remove the vertex a from both $\{a, b\}$ and the graph G, I still have a maximum-sized clique of size 2, e.g., $\{d, c\}$.

One may argue that maximum-sized clique may possess some other kind of optimal substructure. At this point, all I will say is that it is unlikely to. We will revisit this once we have started to discuss computational complexity, in particular, the class **NP**.