Why does Kruskal's algorithm work?

$T^* - e$ has two components, A and B, with $x \in A$ and $y \in B$.

Let f be an in P has an endpoint in A and an endpoint in B.

Note that $f \notin E(T^*)$

Remember we showed $w(f) < w(e)$

Think about $T^{**} = T^* - e + f$ (the graph obtained by deleting e from T^* and adding f)