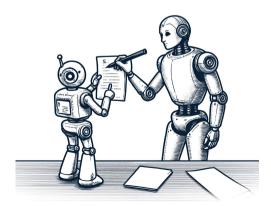
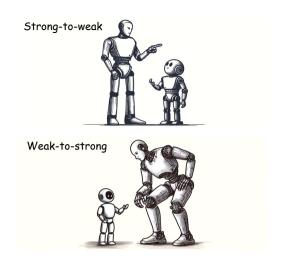


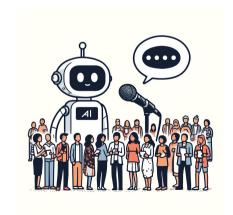
(a) Aligning through inductive bias



(c) Aligning through model feedback



(b) Aligning through behavior imitation



(d) Aligning through environment feedback