Inductive bias (learning bias) is the set of assumptions of the learner (model) to predict outputs. E.g., the parameters learned on the training data as well as the prior assumptions.

Fertility is the propensity for a word to be translated as a consistent number of words in the other language

REINFORCE is an algorithm for gradient estimation in stochastic computation graphs (Williams 1992)

A **Lesioning experiment** is used to quantify the contributions of layers towards the network performance

References

Williams, Ronald J. 1992. "Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning." In *Reinforcement Learning*, 5–32. Springer.