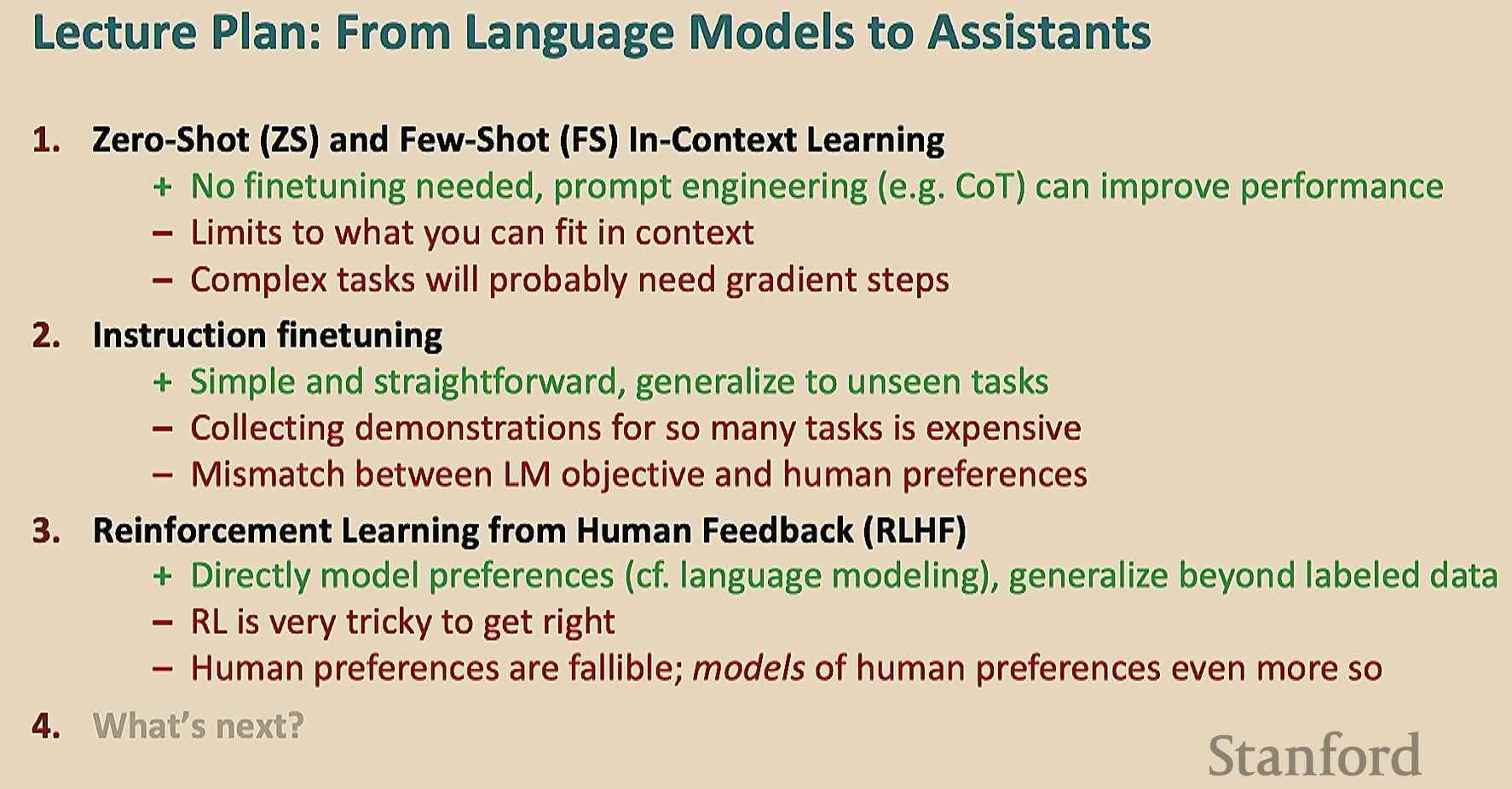
He proposes the idea that Language models *may* be doing rudimentary modelling of *agents, beliefs and actions*. Holy crap. How do we get from language models to multitask assistants like ChatGPT? History of the techniques in chronological order.



1. Zero shot (ZS) and Few Shot (FS) in context learning.
   1. ZS (GPT1 and 2) = ability to do tasks with no examples and no gradient updates = no task specific fine-tuning.
   2. FS (GPT3) More shots > ZS but plateaus. FS is an emergent property of model scale. **Still don’t know why!**
   3. Chain of thought prompting. Also an emergent property of model scale.
   4. Zero-shot chain of thought prompting **The highest ROI technique**. “Let’s think step by step” was the best ***“let’s work this out in a step by step way to be sure we have the right answer”*.**
2. Instruction finetuning (it’s just finetuning!)
   1. Language modelling =/= assisting users. So the objective is to align to human intent.
   2. Super natural instruction dataset: 1.6k and 3m examples. Instruction finetuning is blurring the line with additional pretraining.
3. Reinforcement Learning from Human Feedback (RLHF)
   1. Actual human preferences. More reliable to get humans to compare samples and rate them better than each other rather than absolute scores.
   2. Models of human preferences
4. What’s next
   1. RLHF is very underexplored and fast moving. It works but it is still data expensive. ***Constitutional AI*** by Anthropic AI.