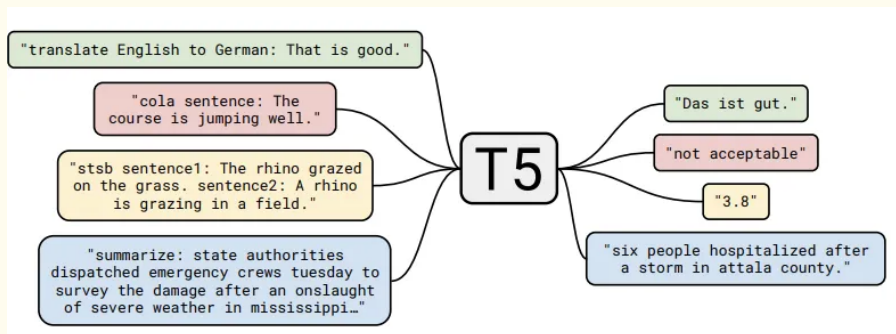
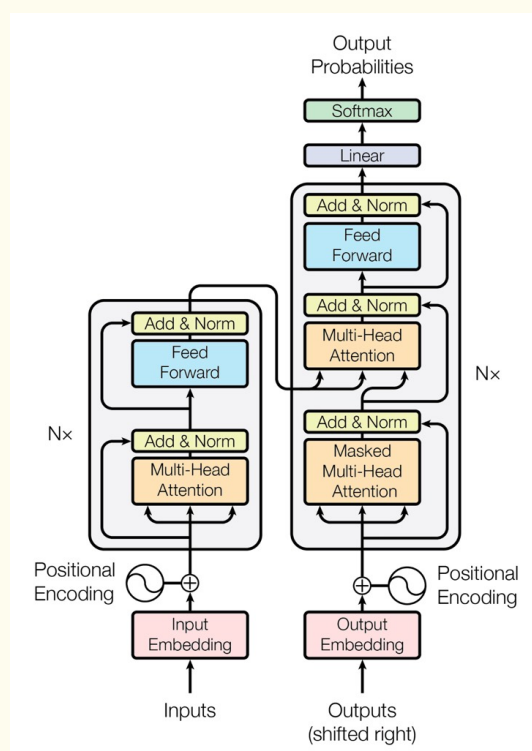




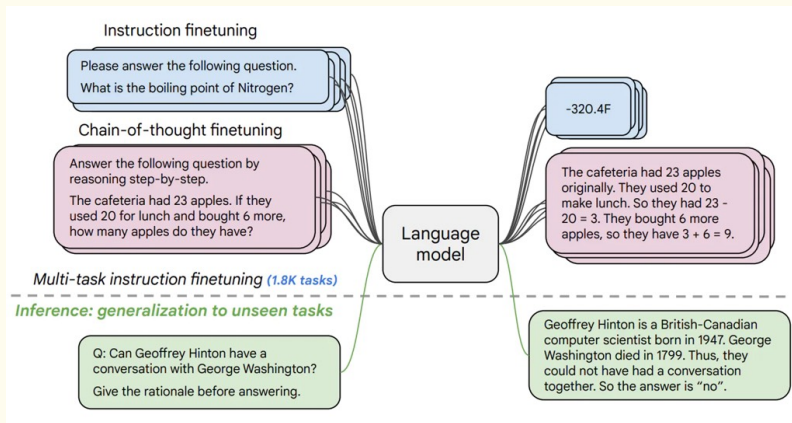
T5 and FlanT5:



- Text to Text Transfer Transformer
- capable of doing multiple task -
 - Question-answering
 - summarization
 - classification, regression
 - machine translation
- T5 is trained on 750 GB C4 (colossal clean crawled corpus) dataset.



Flan T5:



→ Finetuned Language Net

→ Better than T5. Flan T5 is finetuned for 1000 additional tasks with same # no. of parameters, covering more languages.

→ Flan T5 is an encoder-decoder model that has been pre-trained on a multi-task mixture of unsupervised and supervised task.

→ Input: Text
Output: Text

Google's Flan-T5 is available via 5 pre-trained checkpoints:

- Flan-T5 small
- Flan-T5-base
- Flan-T5-large
- Flan-T5-XL
- Flan-T5 XXL

Model size variants

Model	Parameters	# layers	d_{model}	d_{ff}	d_{kv}	# heads
Small	60M	6	512	2048	64	8
Base	220M	12	768	3072	64	12
Large	770M	24	1024	4096	64	16
3B	3B	24	1024	16384	128	32
11B	11B	24	1024	65536	128	128

FLAN-T5 variants with their parameters and memory usage

