

TS: pretrained encoder / decoder model
seq 2 seq

↳ generate text, unlike BERT

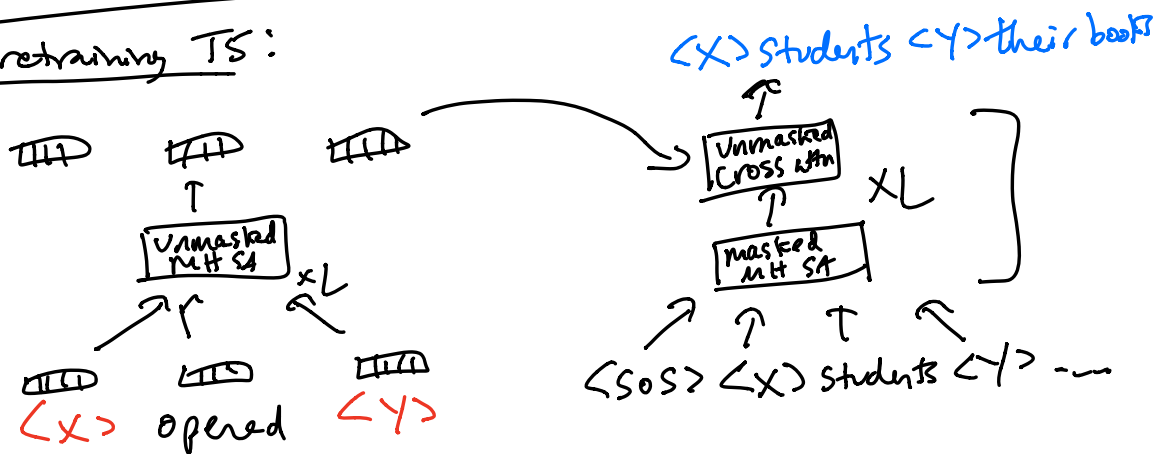
↳ "span denoising" objective

students opened their books



<X> opened <Y>

pretraining TS:



fine-tuning TS:

