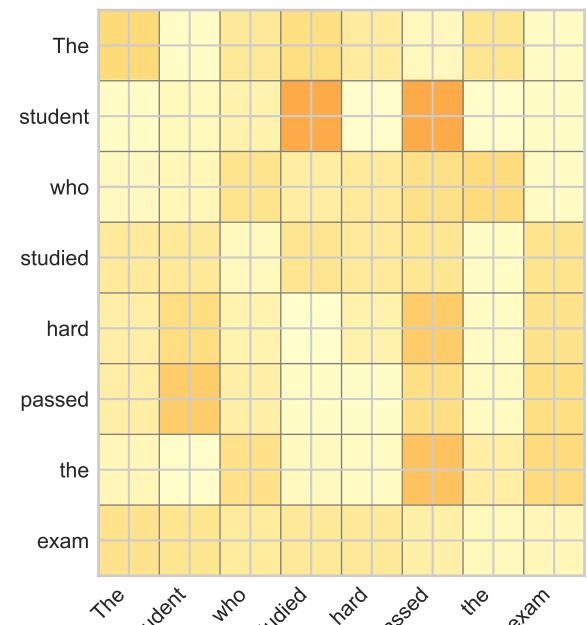
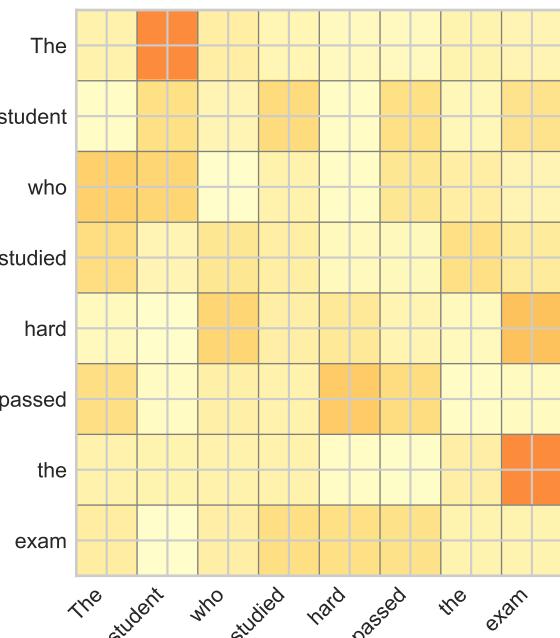


Multi-Head Attention: 8 Different Perspectives on the Same Sentence

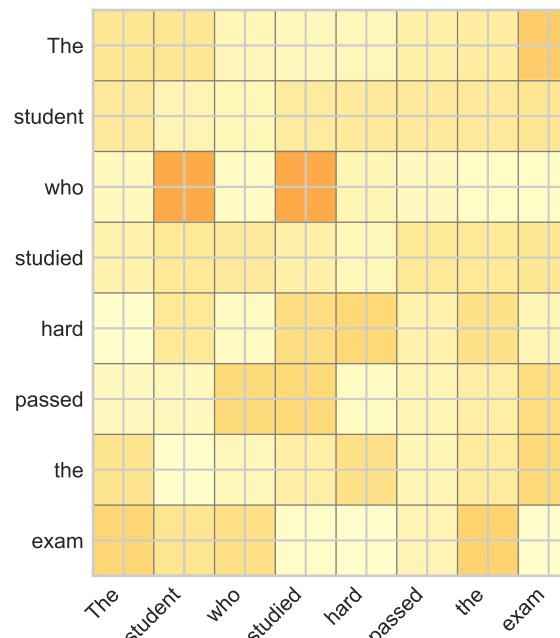
Head 1: Subject-Verb



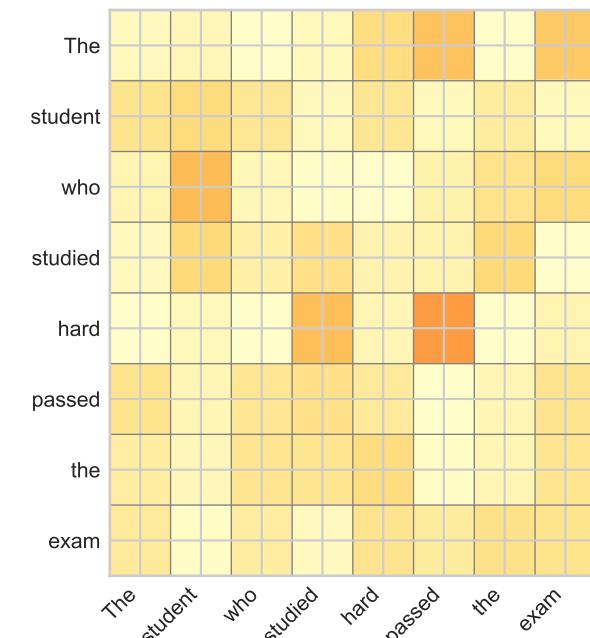
Head 2: Determiner



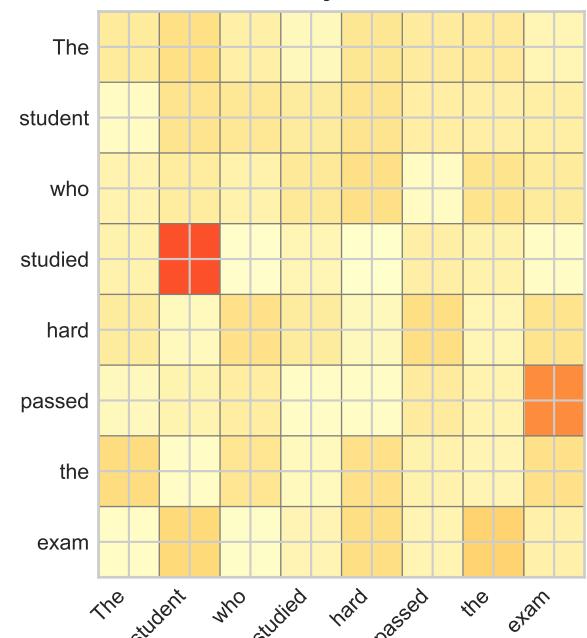
Head 3: Relative Clause



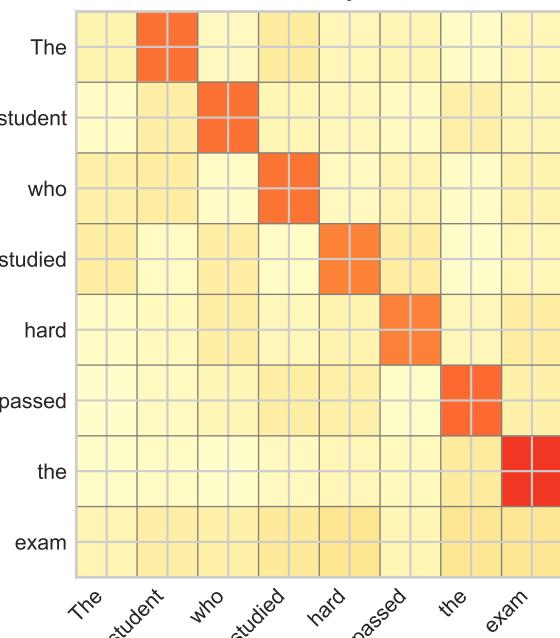
Head 4: Adverb Modifier



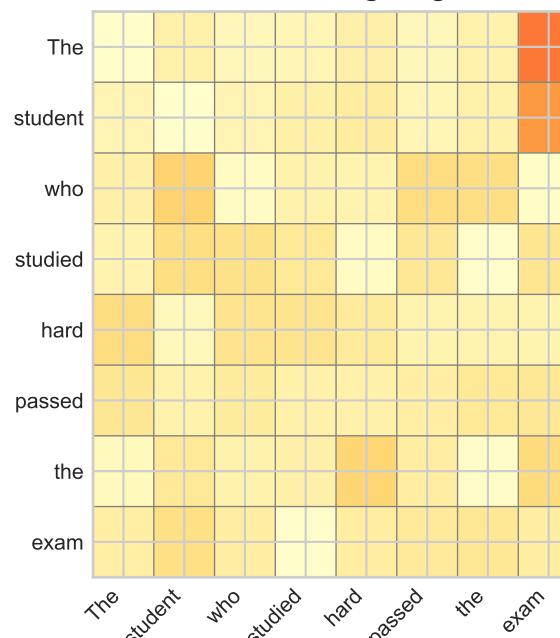
Head 5: Object Relations



Head 6: Sequential



Head 7: Long-range



Head 8: Self-attention

