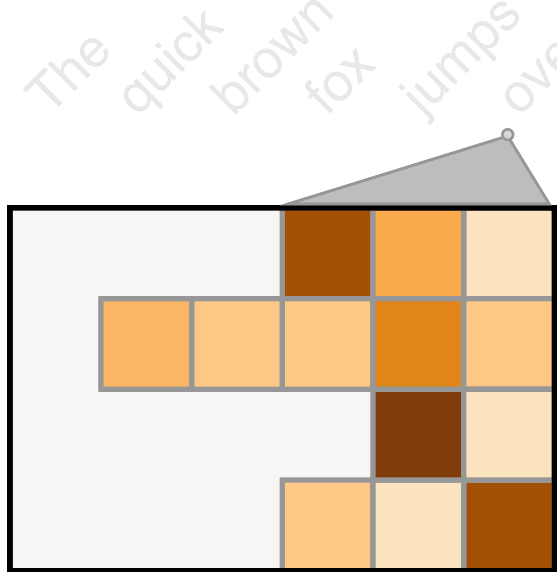
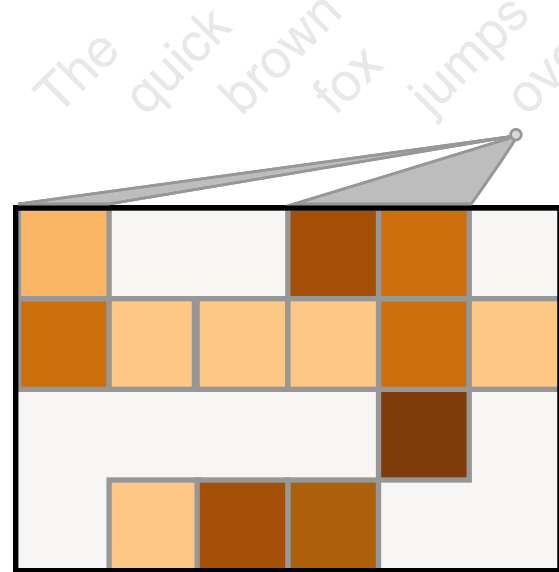


Sparse Transformer



Adaptive Span
Transformer



Adaptively Sparse
Transformer (*Ours*)

head 1

head 2

head 3

head 4