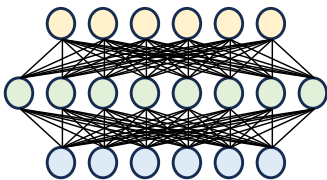
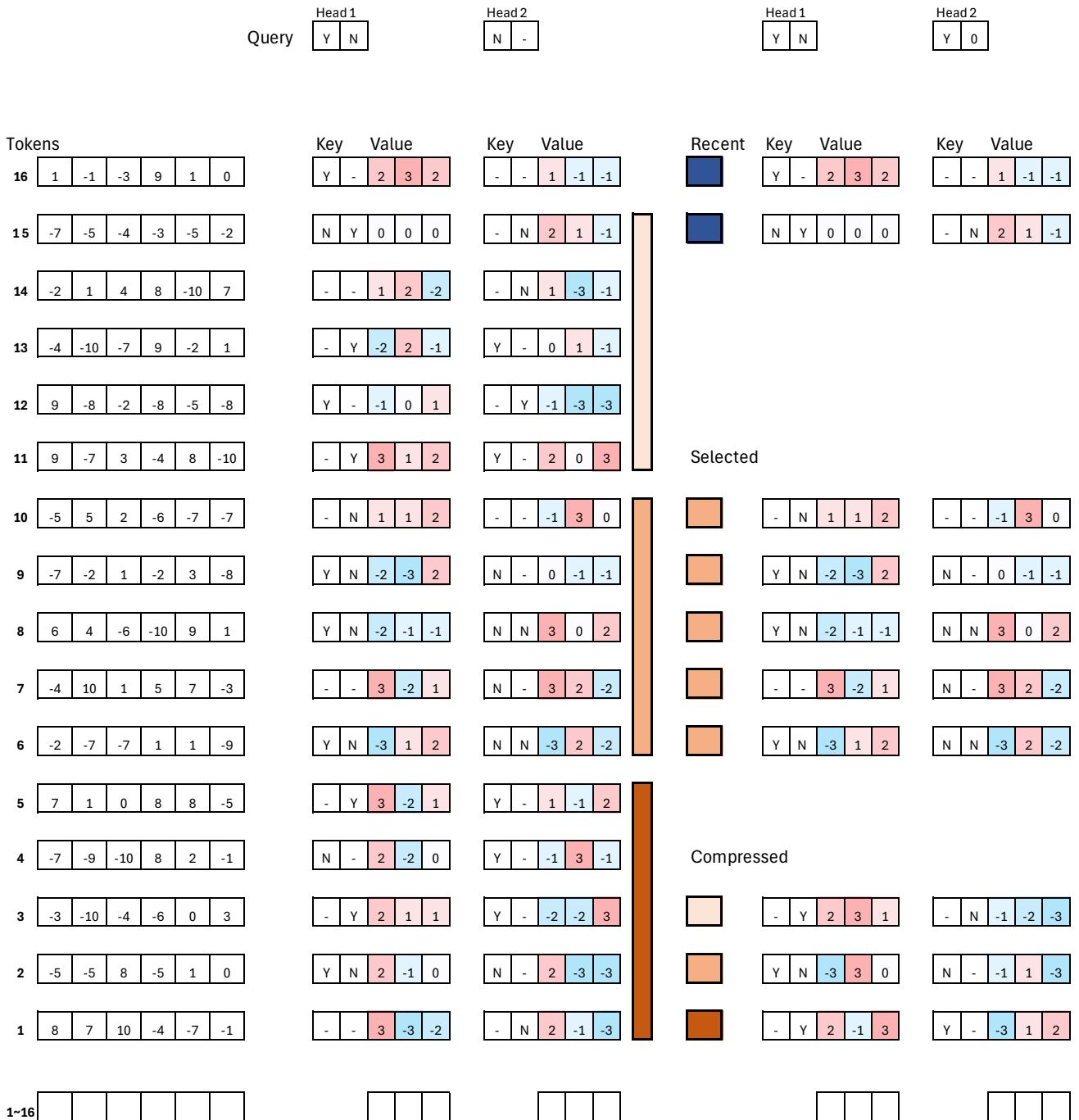


## Multi-head Attention

## Native Sparse Attention



--	--	--	--	--