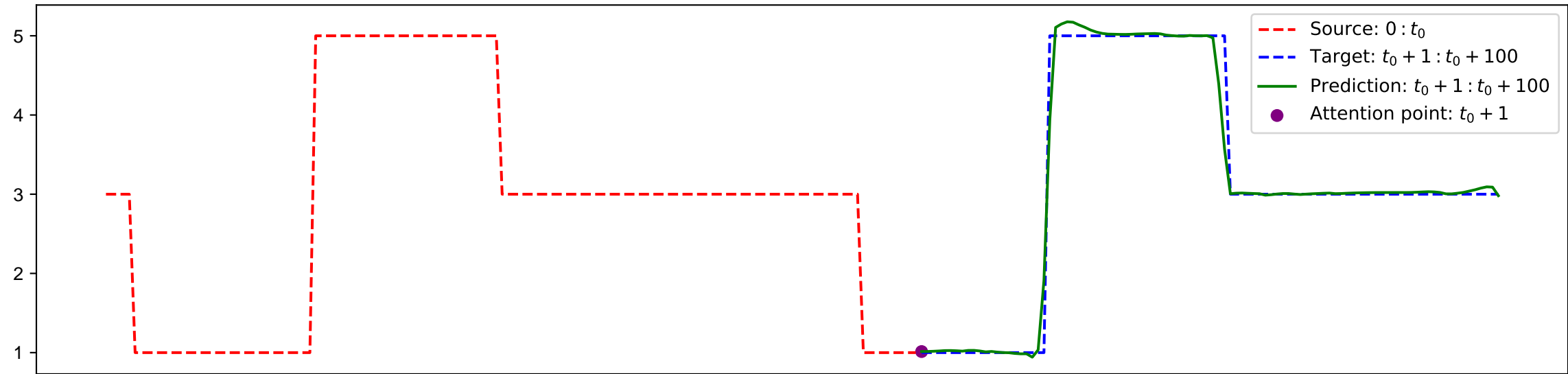


n-steps non-overlapping, total epochs: 300



Attention patterns for $t_0 + 1$

