num\_steps = 35, hidden\_size = 512, max\_epoch = 6, max\_max\_epoch = 39, keep\_prob = 0.5, lr\_decay = 0.8, batch\_size = 20, vocab\_size = 10000, optimizer = GradDesc, loss\_function = sequence\_loss\_by\_example init\_scale = 0.05, learning\_rate = 0.5 init\_scale = 0.05, learning\_rate = 1.0 init\_scale = 0.5, learning\_rate = 0.5 init scale = 0.5, learning rate = 1.0 Perplexity Steps

 $max_grad_norm = 5.0$ ,  $num_layers = 2$ ,