

Tianxiao Zhao

tzh@kth.se

## Assignment 4

### (i) Gradient Check

First, we check the analytical gradient of the RNN networks compared with the numerical one using two different error criteria:  $\frac{\max |g_a - g_n|}{\max (eps, |g_a| + |g_n|)}$  and  $\frac{\sum |g_a - g_n|}{\max (eps, |g_a| + |g_n|)}$ . The number of hidden nodes is set to 100, and delta is set to  $1e-4$ , which is used for the numerical gradient calculation. Relative errors between analytical and numerical gradients with different batch size are shown below:

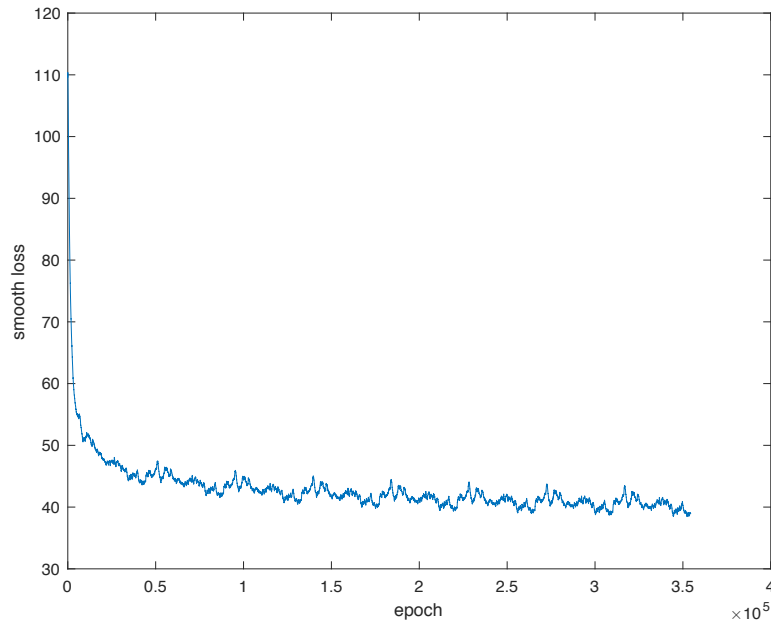
**Tab. 1: Relative errors with respect to different parameters**

Batch size		5	10	15	25
RNN.b	max error	7.136e-11	4.8926e-11	5.6127e-11	8.1869e-11
	sum error	1.6632e-09	1.6555e-09	1.7504e-09	2.0155e-09
RNN.c	max error	6.552e-12	7.8116e-12	8.4527e-12	9.0078e-12
	sum error	4.151e-10	4.2927e-10	4.6423e-10	5.0471e-10
RNN.U	max error	2.3707e-11	1.0603e-11	1.3284e-11	1.1344e-11
	sum error	1.7802e-09	2.0604e-09	2.6012e-09	3.7259e-09
RNN.W	max error	1.801e-11	2.6535e-11	3.1059e-11	4.9581e-11
	sum error	3.6539e-08	5.1613e-08	6.537e-08	1.0142e-07
RNN.V	max error	4.1359e-12	6.3295e-12	7.0474e-12	1.0649e-11
	sum error	7.68e-09	9.9979e-09	1.1924e-08	1.8035e-08

Since the relative errors above are small enough (below a relative error level of  $1e-06$ ), I believe that my analytical gradients are bug-free.

### (ii) Loss Function Plots

In this section, the graph of the smooth loss function for a longish training run (8 epochs) is presented below, see Fig. 1. Here, the parameter settings are mainly based on the instruction given on the KTH Social website:  $m = 100$ ,  $\eta = 0.1$ ,  $\text{seq\_length} = 25$ ,  $\text{sig} = 0.01$ ,  $n\_epochs = 8$ .



**Fig. 1: Evolution of the smooth loss function as epoch increases**

### (iii) Synthesized Text

Here follows the synthesized texts generated by RNN during training. The length of each text is constraint to 200 characters long. From the printed results, we can see that some words gradually appear in the right form after some epochs. But the meaning of the printed text still keeps unclear to us. More training epochs are desirably needed to synthesize a more sensible sentence.

iter = 1, smooth\_loss = 109.5599:

```
AkAXQBe' nloh /PHWBGdG6'my9rHP,A"Hd"PGJuD-Pbwn'rX!GdGNPbG:RN
RTDEE2W_dpZpQRGV9PCTHFUcW_u"0JACK:Zu 眉 9vG,fG?RyNDB/prxI_
HAG/GBEoYNdGNVC;sVeFyxw"63/?NARH G'B/GE(FZPm)GK 眉 oJEg"KRGGGw(GG4E(PQzgMOus 眉 GO}
```

iter = 10000, smooth\_loss = 52.791:

```
ean. . Cnaight Sagks betercare bon atrad of noblu harr in ffore her Oare a gore ther
Cincertores var brit mid fu't tho doull follet, sowrla mot non yoned cons am tow ko, o
poagetterise htarry a vea
```

iter = 20000, smooth\_loss = 49.4631:

```
e stogal W it. Arest if his the in, ther noblofe I breeding should githt said Harry wiwh the
sig the keat wariy wingibe did. The roh to Wevery hairig his lipcht the miondenss a it
surding tile him Fl
```

iter = 30000, smooth\_loss = 48.0622:

```
e ipmeiding the cansed she," sable thasces acow," so dornter lown foog betun tarty him at at
alod bith, Him walle, say a fame momeen someter aparns and wobcy'ccer lest dortiome bros..
stires stiresnec
```

iter = 40000, smooth\_loss = 47.1597:

o conis eleating farkis VAtowh griekno brich knet and for have Dith muft Vot correl ttan  
Mestries araich bow rilch out hengoreagigg perent not laort in there any that the gurdssire -  
Vover frourbed th

=====  
iter = 50000, smooth\_loss = 50.7797:

's find and tlet arved --lyo raocting peging Brelde corle sie shomaldvees d!" Ron we the, wass  
highthi nes shache!"

-The might tice. serm ther. fo.

"Thad loubuim. 's me cuigall starkant, beongs bughe

=====  
iter = 60000, smooth\_loss = 49.2246:

iw sleaven had inaer Ron be thing - he stionf undey loter coudou. HV'midey rarnguld oly had  
did vevely wead ther, hif ol tle gof Weak him shiggore arous!" said astiound Plagled shot a  
kin of''s atch

=====  
iter = 70000, smooth\_loss = 48.4393:

ewdeten wilked ho stoak duve wowe theor reonge nead youg, and knius," said Pnoinst caxap wing  
pink gobnch grand offinks whanwers. Bas der attaicis wond sailing was had the blemures at  
untold weve.

=====  
iter = 80000, smooth\_loss = 46.4801:

om tofirnare Kalking but his a doow ame had look." saght

Harry arared ots Berkaites you kiiming wat mutted, sture yous sisten shat he were Kwum wy dny  
- dudor. Jote at rowires wank and for awrile had

=====  
iter = 90000, smooth\_loss = 46.9418:

. ... Forn, sime his out ettreet, it his Hald word-castas.

Nowers mont wauthan keet.

Rox lik.... Stasce bo look. Hagrill, acaid knee ons they spoft him illow.

"Eaminow soits, over his, oldout.

Wo

=====  
iter = 100000, smooth\_loss = 47.6936:

e'ving deother "of overed easleons,

Hirk now sored afwead IBed scall grinc dravor veese Dard yeiger. spoided inary unletted -"

"It witale the atoreenm aboulederioid armaich of the wing coldiboont, Aob

=====  
iter = 110000, smooth\_loss = 45.7502:

. He wat what be blike, Harry age thing and Heride minks, withencoutdnored Harry Heaving a  
doncang with uikeling bettery, pling bight," hear on shobensbous of tilker.

Bot dading misthed the a dere tin

=====  
iter = 120000, smooth\_loss = 45.8332:

oggacked now ien't tuitd lave as blake bain soby nothing as you chatk thlaving wourder's the  
a sire sudding he sos Kturk his infented and, that of thak hay arrate all. The youts out ofter  
leek,"What,

=====  
iter = 130000, smooth\_loss = 44.6294:

e coming jely all flilins," said Dumbleoan't rit tiving.

"Seme heams did gare the pilour the blen a diges a yould Mong. . . . Faid, pol it and as tilks to Hurimed hess fil same on this from for his lo

#### (iv) Best Result

Now from the trained model, we pick out the best one with the lowest smooth loss, and print out its synthesized text with a length of 1000 characters. The best result is shown below. It seems that more vocabularies are spelled correctly compared to the previous cases. But still, the sentences are kind of incomplete and in chaos.

```
iter = 352353, smooth_loss = 38.4635:
```

aronnthe jow up unding to me toumned time he was the got was I what sirnittilyt he Harry was this?" I and do you plail seidine, you hin were spokice you voice thing it rour, he was had let as them's conrcing to for Trus happove, list, him mable of long to memory toicularow the off experiund had got a with of the somer ichid wime. That's to was this and quies the heating fambing the could work as I in bith ust."

Bat I geter Snyowing me around a deepled was I do his; to his gred see of he have was filace, shoutune you fude thungeon to pulb again indrest to to the laining night." Dely a sire!" said whel.

"Course who not lyice and all to the tow," said Fred mal Dumbled he remaked away. Harry scarifate Ighing a dalledficts midate. Harry going time, voate the there boftandausery unvond there was founded bwast, wizard town to black, wish how fillouch, usich his reburt."

Ron anyoge how reasined Smappe from - who was in the very lears you exalle was now. And Unstick of Perrous and used I