

# DD2424 - Assignment 4

Timo Nicolai

April 11, 2019

To check my gradient calculations I followed the steps outlined in the assignment description: For an RNN with a five-dimensional hidden state vector I calculated the parameters gradients for the input sequence **HARRY POTTER AND THE GOBL** and the corresponding output sequence **ARRY POTTER AND THE GOBL** (one hot encoded) both analytically and numerically. Figure 1 shows the element-wise maximum relative differences between analytical and numerical gradients for the different network parameters. All of them are very small so I assume my implementation is correct.

Figure 2 shows the loss function for a 10.000 updates training run, the dotted red line indicates the point in the loss curve at which the lowest loss was achieved, the final model parameters were set to be equal to those at this point in the training run.

Listing 1 shows snippets of length 200 (linewrapped to improve readability) that were generated by the partially trained RNN at different points during this training run (**iteration 0** means just before the first gradient update, **iteration 1000 means** means just after the 1000th update and so on).

Listing 2 shows a sequence of 1000 characters generated by the network corresponding to the dotted red line in Figure 2 obtained by initializing the hidden state to zero and choosing a random dummy input character from the set of all possible input characters.

Parameter	$\epsilon_{rel,max}$
$U$	7.12e-5
$V$	5.02e-6
$W$	5.02e-4
$b$	1.31e-5
$b$	2.47e-5

Figure 1: Maximum relative gradient error between analytical and numerical gradient. Numerical gradient obtained via simple forward difference with  $h = 1e - 4$ .

```

iteration 0/100000:
Gitzx;2L)dSüDRFEkmo/üQaHyP"Q:muVQzSqdywEsNqKyEFB;EOüüWAJ?P
9cnh2"MNEHyjTQ90')EhFkT7mAfQ0oYPuRHLdr`X    OYN?    exH    oRksD?
WEbGC_?ü67EEJBn NTD sxhBh_
wk0SO`~}U_MIYX_GmhaF4MrXXD,aYo0qg}'DwüQhQKJfs7Mmyi_NU-}CC0

iteration 10000/100000:
rsurof tarplorisn!" shad, in d Mr seat?" Ha drecsy the qugels ofg amuinp ap was
Cwothaln aroid he doked Boll, wate wai: Bived. . "Ap leccetinghe abon't Hind
Draverill, y be, ofbecal mav'igg To the

iteration 20000/100000:
iciint astlore't Hacry wass hit wee pretemee wonqenory got was sibked Hermiot
thas nosthoke heren ferimaide's - waicce of the Gromittouts jusuy stranh es,
obostinon," "He peale Harnyon hes - you lack

iteration 30000/100000:
ir - in the has mud a faistels ar Flinks as takeinge well comblens dis ham
Chothe looking hir! . . . . Reeked to knou the ouse Dicep conost, dood the ould
fue lyon harre fild his pEin in he ss. "Bege

iteration 40000/100000:
wer, myting the me ons, Vliat uSiled now nat the cracter threm, dammaiffors asee
on so not noth logkent. Atly wans naldsing stawn mentlys.. "Yearmell fouby whe
of'th. Voldewer he lappeatti- andoled h

iteration 50000/100000:
rey. "You chald to and anement beinged onettmered Magll, and Pangr, they hrout
grow fraike yot Portanffed youtind a his wastair, Mngored stateddness ther ao
happeely- ops ght him. "Yead was and at Vem

iteration 60000/100000:
" Foral for soupe." Hars bugalimasenl though as redver thoagh from the stary a
Srlitcey a dood the maten he ceppent and his bascrang heer. " Aubbetare
fromedion wh is to and bead?" Row Ron oustatai

iteration 70000/100000:
cmooce was he had fon Terem a that Furpiringe," saighboslays. "The could
benasted blosmookingly and with spettorted Mode at the Baving her nair!" yound
apto on you Ron havs, . . you to wuren wollad tu

iteration 80000/100000:
are Wecagaclad coidisten't sredit and a Prougrode Nerred if resned, lloke at
enlameloblinse anst dother pold Squexe soppaiserth?, alls. Mr.. . nomaperin; to
acimef loom?" said Hermione hagh, wat opeen

iteration 90000/100000:
t, the Fakenors same Chertappinet to Rour nir oke. Harry be yourmblen which and
warne cotilf had might." Vo he kileed fehts to mud Harry was Fond we healde
fayter ois he he said Harryeser wisher his

iteration 100000/100000:
hat had hatcherould the cound been a who do moth ofter, was to kit day flookry?"
they dobdech shouthly the ficess for ligherping Cridk eyere asp banfing are she
andstedsech Muggly, "Thearnd the dilie

```

Listing 1: Snippets produced by the RNN at intermediate points of the training run, contains occurrences of “Harry”, “Ron” and “Hermione” so it seems that the training was successful.

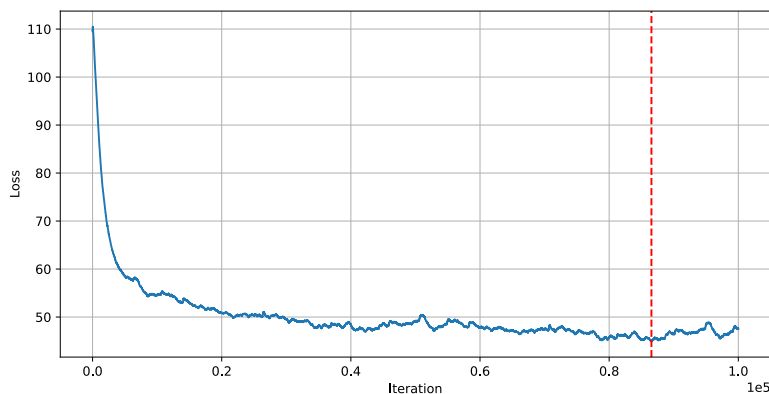


Figure 2: Loss curve for RNN with hidden state size 100, trained on sequences of length 25 with  $\eta = 0.1$ .

ermofuwn Dushey?"

Deem and sel onay gipfhouleed. "So bally. Hanging the Hermbact. Hizard, ald. . . Ceround Cas in Sparvisased. Cupresed. Dumbled of the core. I madine mactez of knees ting as you to sof ?" Do its it mownarviring crowg, who deapparine, asading foke had said of extrese. Harry motghing woll, witl at poturint he was ahear wemestriss head . . .T. I cad, Sfowhars. I noned. . . . "Thou scestore thin a vadempibeving Vor Chers moldnord, it rook ever in the deer you dodase sace toody hes Wisceding all winking it. Alle." He wand of Hog firting Won from then by alaming. . . . OREKr I slear, in sare, deadny, be had han Dumum hised hoursed his fall his was mad Cabby," soad there nook be him to to ant a the you arrix into way," ha! I to intid fuldning so Karkmofge's on ppecchir mantede him hakning in it arbeading and had aren to theme grorg migut to to I was elfed ins." Houmpinaind, stilled me welled. Shay he saping, all to ke do thous Pookad he hathed werrougs wand sume

Listing 2: *You're a Hizard Harry* - Prof. Dumbled of the core