Quadratic and cubic functions defined on the positive integers

by

Stuart Dent

A Dissertation Submitted to the Faculty of the

Graduate Interdisciplinary Program
in Applied Mathematics

In Partial Fulfillment of the Requirements
For the Degree of

Doctor of Philosophy

In the Graduate College

The University of Arizona

2 0 1 0

Get the official approval page
from the Graduate College
before your final defense.

Statement by Author

Dedication

For my mom.

Acknowledgments

First and foremost, thanks go to co-advisors Anna Mehmbuhr and Mbaqanga Uffda. Much of what I have learned in this project has been learned through them; their patience has been endless and their knowledge has been irreplaceable. Susan Kho and Vincenzo Mitti, as well, have given me valuable feedback as I have revised my dissertation this spring.

Summer 2009, fall 2009, and spring 2010 research assistantships were funded by the University of Arizona Department of Mathematics NSF VIGRE (Vertical InteGration in Research and Education) grant. The remaining semesters of my PhD years were funded by teaching assistantships. I thank my college-algebra, trigonometry, and calculus students for helping me to see mathematics from both sides of the classroom.

Table of Contents

List of Figures

List of Tables

Abstract

Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper.

Quadratic and cubic functions defined on the positive integers

Stuart Dent, Ph.D.
The University of Arizona, 2010

Director: Anna Mehmbuhr

Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper. Here is my abstract. It is a summary of the paper.

Chapter 1
Chapter title

1.1 Disclaimers and usage instructions

This is a version of my PhD dissertation which I defended and got through Graduate College format review in spring 2010. I hope it works for you. Please keep in mind, however, that style policies may change. Of course, please consult the style information you received from the Graduate College.

I should acknowledge that I didn't write this from scratch. I got this template from I forget where — somewhere in LPL, I think — and modified to my taste, and to match current formatting requirements. So, this is a current snapshot of collectively developed dissertation style for the University of Arizona.

How to use this template:

- Search around in `stu-dent-dis.tex`, `ua-thesis.cls`, `disack.tex`, etc. for the names "Stuart Dent" and "Anna Mehmbuhr", etc. Change these to your name and the names of your committee members. Also change the title of your dissertation.
- In those same files, look for anything else which needs changing.
- In the first line of `stu-dent-dis.tex`, you can set the style to `draft` or `final`.
- The shell scripts `./dbuild` and `./pbuild`, in the same directory as this file, can be used to create DVI and PDF files, respectively.
- Any LaTeX problem you have is most likely resolvable by a Google search.
- Good luck, and have fun!

— John "Stu Dent" Kerl, April 21, 2010.

1.2 Section title

The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown *fox* jumped over the lazy dogs.

Table 1.1 will be discussed in section 1.3.

x	$f(x)$
0	0
1	1
2	4
3	9
4	16
5	25

TABLE 1.1. Here is the long caption for the body of the document. Note that it's OK for this to have many lines. But the short caption for the list of tables should be short. Hence the difference between square brackets and curly braces in the `caption` command in the LaTeX code. And, you guessed it — this last sentence is here just to make this caption even longer.

1.3 Section title

The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox

x	$g(x)$
0	0
1	1
2	8
3	27
4	64
5	125

TABLE 1.2. Here is another caption.

jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs.

Compare table 1.2 with table 1.1 which was on page 11. We conjecture, but are unable to prove, a relationship between these tabular data and the function of equation (1.3.1):

$$y = x^4 - x^2 + 1. \tag{1.3.1}$$

Chapter 2

Chapter title

The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs.

Please see figure 2.1.

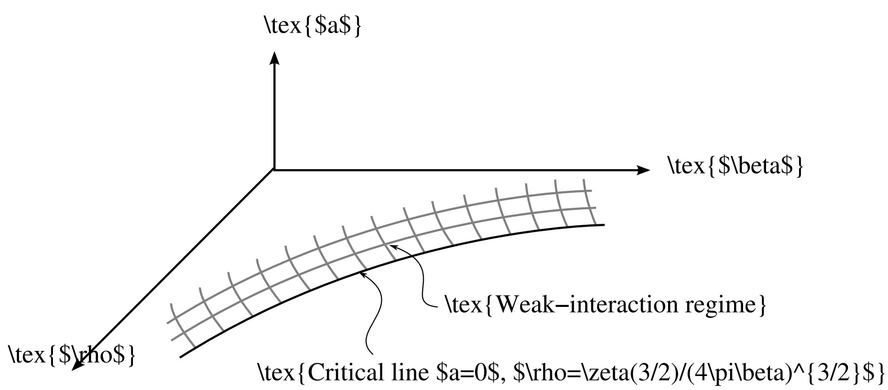

Figure 2.1. Here is an EPS file which I put in here and which surely must be important.

2.1 Section title

The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown *fox* jumped over the lazy *dogs*. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs.

2.2 Section title

The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. The quick brown fox jumped over the lazy dogs. [a4paper]article

gb4e natbib [english]babel [utf8x]inputenc [T1]fontenc

amsmath graphicx [colorinlistoftodos]todonotes [colorlinks=true, allcolors=blue]hyperref fixltx2e Sweave

concordance:cake$_c hapter.tex : cake_c hapter.Rnw : 11821 concordance : cake_c hapter.tex : ./GLOS$

(

Assuming that in the previous chapters the following points are addressed already:

- The nature of glosses has been well-explained (Target audience: CS people without any formal linguistics background):
  - What glosses are: A basic intro of interlinear gloss for non-linguists

  - The golden nature of glosses (encodes NON-LINEAR syntax (i.e. structure parse) and semantics information)
  - The potential of gloss:
    * potential: providing disambiguation, labeling important grammar morphemes in the source language, providing morphological analysis, providing one-to-many and many-to-one relations of source tokens and target tokens.

- A history of machine translation, and a non-mathy description of the methods of doing machine translation. (Target reader: theoretical linguists)

)

2.3 Introduction

The Innovation is to incorporate the gloss information of Interlinear Glossed Text data into machine translation.

In supervised machine learning models, two factors effects the performance of the trained systems kotsiantis2007supervised: a.) the quality of the training data and b.) the choices of the features. The properties of the gloss data as described in *CHAPTERXYZ* make it a better training data than natural language data (Scottish Gaelic in the current case) for the following reasons. First, glosses are more purified that natural language words. The most ideal meaning representation system should be built with one-meaning-to-one-representation mappings; in other words, a meaning is mapped to one and only one representation. Natural languages fail to do so, given that synonyms and ambiguous words/phrases are ubiquitous in natural languages. Glosses provide this one-to-one mapping. Second, the gloss data provides hierarchical (non-linear) syntactic parsing information to some degree. To determine what the gloss of a word is, linguists have to look for hierarchical (non-linear) context information.

Therefore, theoretically incorporation of the gloss data should improve the translation systems. Specifically, I propose the following hypothesis:

Example 2.3.1. Gloss-helps hypothesis: the translation systems trained with the gloss data incorporated should outperform the systems trained with only Gaelic and English sentences pairs (i.e. without gloss data).

The hypothesis can have two versions, strong and weak:

Example 2.3.2. Strong version: Gloss may replace the source natural language totally, and the system outperforms the system trained with source natural language to target language sentence pairs (i.e. the baseline systems).

Example 2.3.3. Weak version: Gloss only increases the performance of the baseline systems, but cannot replace the source language.

The experiments reveal that replacing Gaelic words with glosses doesn't boost up the performance of the translation systems. Thus, the strong version (replacing-Gaelic-with gloss) of the Gloss-helps hypothesis is not attested. However,it is found that if the Gaelic data and the gloss data are combined in a specific way as the training data, the performance of the systems is improved significantly.

This chapter describes the experiments conducted to test the Gloss-helps hypothesis and the results attest the weak version. The rest of the chapter is organized as follows: Section 2.4 describes the constant parameter settings across all the experiments, section 2.5 tests the hypothesis in (2.3.2), section 2.6 tests the hypothesis in (2.3.3),and section 5 concludes the chapter.

2.4 Technical Settings of the Machine Translation Experiments

The experiments are conduced by using OpenNMT 2017opennmt, which implements the state-of-the-art neural net machine translation algorithms cho2014properties, cho2014learning, bahdanau2014neural. The following default hyper-parameter settings of OpenNMT[1] are used across all models so that the only independent variable is the type of the training data:

- Word vector size: 500
- Type of recurrent cell: Long Short Term Memory
- Number of recurrent layers of the encoder and decoder: 2
- Number of epochs: 13
- Size of mini batches: 64

The settings of the hyper-parameters do have effects on the performances of the trained models. A common practice to find the optimal settings of the hyper-parameters is to hold out a subset of the training dataset as the developing dataset, then test the models on the developing data to see what settings are optimal, then merge the developing dataset and training dataset as a new training set, and then train on this new training set using the found optimal hyper-parameters. However, finding the optimal settings of the hyper-parameters is not relevant to our research and causing unnecessary complications. So, the process of optimizing the settings of the hyper-parameters is not implemented, and I simply adopt OpenNMT's default settings. So, the employed settings of the hyper-parameters may be viewed as arbitrarily chosen. Critically, these settings are viewed as constants, so that we can

[1] See their documentation for the complete default hyper-parameter settings: `http://opennmt.net/OpenNMT-py/`.

focus on the effects of different treatments on the source sequences in the translation experiments.

The data and the scripts will be accessible on GitHub[2], so that the results can be reproduced.

2.5 Gloss Representation Solely Does NOT Outperform Gaelic Sentences

This section tests the strong version of Gloss-helps hypothesis in (2.3.2). Given the assumption that gloss may be better than any natural language in terms of representing meanings, it is expected that for neural net machine translation systems it is easier to learn how to translate from the glosses of Scottish Gaelic to English than to learn how to translate from Scottish Gaelic to English. However, the results show that there is no significance difference between the two types of data (i.e. GLOSS $\rightarrow$ English and Gaelic $\rightarrow$ English).

2.5.1 Procedure of the Experiments

I use repeated random sub-sampling validation to compare the performances of the two type of models.

Totally we have 8,388 indexed 3-tuples of Gaelic sentence, a gloss line and an English translation. In the interlinear glossed text example below, each line is an argument of a 3-tuple sample.

Example 2.5.1. Tha a athair nas sine na a mhàthair.
be.pres 3sm.poss father comp old.cmpr comp 3sm.poss mother
'His father is older than his mother.'

The 3-tuple above is:

[2] `https://github.com/lucien0410/Scottish_Gaelic`

Example 2.5.2. ¡“Tha a athair nas sine na a mhàthair.”, “be.pres 3sm.poss father comp old.cmpr comp 3sm.poss mother”, “His father is older than his mother.”¿

First, the samples (i.e. the 3-tuples) are randomly split into three datasets: training set (N=6,388), validation set (N=1,000), and test set (N=1,000).

Example 2.5.3. Definitions of datasets:
Let:

Example 2.5.4. IndexTrain, IndexValidation, and IndexTest be sets of random indexes from 0 to 8,387.

Example 2.5.5. IndexTrain ∩ IndexValidation ∩ IndexTest = ∅

Example 2.5.6. —IndexTrain— = 6,388; —IndexValidation— = 1,000; —IndexTest— = 1,000. The step above just randomly splits the indexes of the 3-tuples into three distinct sets: IndexTrain, IndexValidation, and IndexTest. Based on the indexes, we generate the sets of samples. For each index, the 3-tuple is split into two pairs: ¡gloss, English¿, ¡Gaelic, English¿, so that later we can compare the different effects of gloss lines and Gaelic sentences. For each pair, the first item is the source sequence, and the second item is the target sequence. The systems learns how to map the source sequence to the target sequence.

Example 2.5.7. Gloss to English

Example 2.5.8. GLOSStoENTrain = $\{< gloss_i, En_i >| \; i \in IndexTrain\}$

Example 2.5.9. GLOSStoENValidation = $\{< gloss_i, En_i >| \; i \in IndexValidation\}$

Example 2.5.10. GLOSStoENTest = $\{< gloss_i, En_i >| \; i \in IndexTest\}$

Example 2.5.11. Example: ¡"be.pres 3sm.poss father comp old.cmpr comp 3sm.poss mother", "His father is older than his mother."¿

Example 2.5.12. Gaelic to English

Example 2.5.13. GDtoENTrain $= \{< GD_i, En_i >| \; i \in IndexTrain\}$

Example 2.5.14. GDtoENValidation $= \{< GD_i, En_i >| \; i \in IndexValidation\}$

Example 2.5.15. GDtoENTest $= \{< GD_i, En_i >| \; i \in IndexTest\}$

Example 2.5.16. Example: ¡"Tha a athair nas sine na a mhàthair.", "His father is older than his mother."¿ The models are trained with the training set and validation set (i.e. the model learns how to map the source sequence to the target sequence). Both training set and validation set are known information for the models[3]. Specifically, the neural net system learns how to maps gloss lines to English sentences from samples in (2.5.8) and (2.5.9), and another neural net system learns how to maps Gaelic sentences to English sentences from from samples in (2.5.13) and (2.5.14).

Example 2.5.17. Models:

Example 2.5.18. ModelGLOSStoEN = Model trained with GLOSStoENTrain in (2.5.8) and GLOSStoENValidation in (2.5.9)

Example 2.5.19. ModelGDtoEN = Model trained with GDtoENTrain in (2.5.13) and GDtoENValidation in (2.5.14) The two trained models (gloss-to-English and Gaelic-to-English) then take the right source sequences of the test sets (i.e. glossing

[3] Technically speaking, the validation set is part of the training data in terms of machine learning. The presence of the validation set is a special requirement of neural net machine learning, which uses the validation set to evaluate the convergence of the training.

lines and Gaelic sentences for ModelGLOSStoEN and ModelGDoEN respectively) as inputs and then generate the predicted target sequences (i.e. English sentences).

Example 2.5.20. Predictions:

Example 2.5.21. PredictionsGLOSStoEN = A list of English sequences that ModelGLOSStoEN maps to from the gloss sequences in (2.5.10)

Example 2.5.22. PredictionsGDtoEN = A list of English sequences that ModelGDtoEN maps to from the Gaelic sentences in (2.5.15)

To evaluate the model, the predicted target sequences are checked against the target sequences of the test set (i.e. the gold standard/human-translated English sentences). Specifically, the BLEU (bilingual evaluation understudy)[4] score metric bleu of each prediction is calculated using the multi-bleu.perl[5] script, a public implementation of Moses moses. The BLEU score calculation is an automatic evaluation of how similar two copora are. In the current experiments we are comparing the predicted target sequences with the gold standard. The BLEU score of 100 means the two copora are identical, and the BLEU score of 0 means the two copora are completely distinct from each other.

Example 2.5.23. Gold-Standard = English sentences in (2.5.10) = English sentences in (2.5.15) Note that the gold-standard is the same because they are the same English sentences in the 3-tuples samples. Then the two sets of predicted English sentences are evaluated, yielding two BLEU scores.

Example 2.5.24. Scores:

[4]There are other automatic machine translation evaluation algorithms available, such as translation edit rate Snover06astudy and DamerauLevenshtein distance damerau1964technique, levenshtein1966binary. BLEU is chosen for the current experiments because it is the most widely used evaluation algorithm, and the correlation between the BLUE score evaluation and human judgment evaluation is also well-acknowledged.

[5]The script can be downloaded from: `https://github.com/moses-smt/mosesdecoder/blob/master/scripts/generic/multi-bleu.perl`

Example 2.5.25. ScoreGLOSStoEN = BLEU(Gold-Standard, PredictionsGLOSStoEN)

Example 2.5.26. ScoreGDtoEN = BLEU(Gold-Standard, PredictionsGDtoEN)
This procedure of splitting the data into three sub-sets, training the models, and evaluating the models is executed for ten times.

2.5.2 Result

After ten rounds of repeated random sub-sampling validation, ten pairs of scores of the two models are generated, as shown in the following table. The average score of

Round	Gaelic (Baseline)	GLOSS
0	17.29	18.39
1	16.42	18.00
2	15.29	16.02
3	15.97	20.22
4	17.79	19.02
5	16.73	15.53
6	17.11	18.00
7	16.37	20.08
8	15.93	15.82
9	16.99	15.93
Mean	16.59	17.70

TABLE 2.1. BLEU scores of ModelGDtoEN and ModelGLOSStoEn

the ModelsGLOSStoEN is only sightly higher than the average score of the Models-GDtoEN. Also, after doing a paired T-test, the difference between the two types of models is not attested (MGDToEn=16.59, SDGDToEn=0.74; MGLOSStoEN=17.70, SDGLOSStoEN=1.78,; t(9)=1.97, p=0.080)

2.5.3 Summary

The ultimate practical goal of the dissertation is to use glossing data to develop better machine translation systems. Here *better* means to be better than a baseline system, which is the machine translation system trained with Gaelic-to-English translation samples. The models in (2.5.19) are the baseline systems, and their scores are in the Gaelic column of table (**??**). These are the target scores that we aims to outperform. The experiment above is the first attempt to improve that scores by using the *gloss treatment*, in which the Gaelic sentences are replaced with gloss lines. However, the result shows that this *gloss treatment* is not effective as the scores of the gloss models are not statistically higher than the baseline Gaelic-to-English models.

2.5.4 Discussion

It is assumed that the performances of the machine translation systems are correlated with the quality of the representation of meanings in the source sequences. Better representations of meanings yield better machine translation systems. Given the results in (2.5.2) that the gloss models are not better than the Gaelic models, it is concluded that glosses and natural languages are about the same in terms of representing meanings. The strong version of the Gloss-helps hypothesis does not hold.

We may now combine Gaelic and Gloss sentences as the training data to test the weak version of the Gloss-helps hypothesis. The experiments and results are reported in the next section.

2.6 Combining Gaelic words with Glosses

In the previous section, we attempt to build a system by using the *gloss treatment* to outperform the baseline system. It turns that using gloss line solely is not effective enough to improve the system. However, this result does not falsify the gloss-helps hypothesis; instead, it indicates that combination of the gloss line data and the Gaelic sentence data is necessary. This section reports various ways of combining the gloss line data and the Gaelic sentence data, and the experiments and their results using these different treatments. Critically, a specific way of combining Gloss data and Gaelic date (termed as '*ParaPart*' treatment) boosts the performance significantly. The model trained with this specially arranged training data also significantly outperforms Google's Gaelic-to-English translation system.

In this section, I will first describe the most effective treatment, termed as '*ParaPart*' treatment, and the results, and then I will report other experiments attempting other relevant logical treatments.

2.6.1 The 'ParaPart' Treatment Outperforms Any Other Treatments and The Baseline Significantly

Round	Gaelic (Baseline)	ParaPart
0	17.29	32.64
1	16.42	32.28
2	15.29	29.94
3	15.97	31.18
4	17.79	32.83
5	16.73	31.11
6	17.11	32.19
7	16.37	33.52
8	15.93	30.93
9	16.99	34.35
Mean	16.59	32.10

Table 2.2. BLEU scores of ModelGDtoEN and ModelParaParttoEn

(MGDToEn=16.59, SDGDToEn=0.74; MParaPart=32.10, SDParaPart=1.33,; t(9)=48.95, p=0.000)

2.6.2 Other Logical Treatments

Para (MGDToEn=16.59, SDGDToEn=0.74; MPara=24.51, SDPara=1.84,; t(9)=17.50, p=0.000)

Round	Gaelic (Baseline)	Para
0	17.29	25.42
1	16.42	25.32
2	15.29	20.72
3	15.97	22.22
4	17.79	24.27
5	16.73	24.55
6	17.11	27.03
7	16.37	25.34
8	15.93	24.24
9	16.99	25.96
Mean	16.59	24.51

TABLE 2.3. BLEU scores of ModelGDtoEN and ModelParatoEn

Round	Gaelic (Baseline)	interleavingGdGLOSS
0	17.29	13.67
1	16.42	12.49
2	15.29	11.01
3	15.97	12.33
4	17.79	12.56
5	16.73	12.13
6	17.11	11.55
7	16.37	12.78
8	15.93	12.43
9	16.99	11.65
Mean	16.59	12.26

TABLE 2.4. BLEU scores of ModelGDtoEN and ModelinterleavingGdGLOSStoEn

Interleaving Gaelic Words and Gloss Words (MGDToEn=16.59, SDGDToEn=0.74; MinterleavingGdGLOSS=12.26, SDinterleavingGdGLOSS=0.74,; t(9)=-17.06, p=0.000)

Concating Gaelic Words and Gloss Words (MGDToEn=16.59, SDGDToEn=0.74; MConcatGLOSSGaelic=15.44, SDConcatGLOSSGaelic=1.23,; t(9)=-3.64, p=0.010)

Replacing ambiguous Gaelic Words (MGDToEn=16.59, SDGDToEn=0.74; MReplacingGaelic=9.24, SDReplacingGaelic=0.89,; t(9)=-21.03, p=0.000)

Round	Gaelic (Baseline)	ConcatGLOSSGaelic
0	17.29	15.42
1	16.42	14.31
2	15.29	15.38
3	15.97	14.18
4	17.79	18.63
5	16.73	14.89
6	17.11	15.16
7	16.37	15.20
8	15.93	15.50
9	16.99	15.72
Mean	16.59	15.44

TABLE 2.5. BLEU scores of ModelGDtoENModel and textsubscriptConcatGLOSS-GaelictoEn

Replacing ambiguous GLOSS Items (MGDToEn=16.59, SDGDToEn=0.74; MReplacingGLOSS=15.47, SDReplacingGLOSS=1.03,; t(9)=-3.67, p=0.005)

Complete Tables

2.6.3 literature

what about 2.4 2.8 Linguistics-informed MT: sennrich2016linguistic

Multi-task Sequence to Sequence Learning: luong2015multi

what is Multi-task learning: Overview$_Multi-Task_Learning$

$addccctotargetseq : ccg_target_seq$

$googlezeroshot : google_zero_shot$

Round	Gaelic (Baseline)	ReplacingGaelic
0	17.29	9.44
1	16.42	9.07
2	15.29	7.69
3	15.97	9.12
4	17.79	9.08
5	16.73	10.45
6	17.11	8.62
7	16.37	10.00
8	15.93	10.52
9	16.99	8.46
Mean	16.59	9.24

Table 2.6. BLEU scores of ModelGDtoEN and ModelReplacingGaelictoEn

Round	Gaelic (Baseline)	ReplacingGLOSS
0	17.29	15.95
1	16.42	15.60
2	15.29	14.15
3	15.97	14.72
4	17.79	15.74
5	16.73	14.88
6	17.11	14.45
7	16.37	16.41
8	15.93	15.15
9	16.99	17.61
Mean	16.59	15.47

Table 2.7. BLEU scores of ModelGDtoEN and ModelReplacingGLOSStoEn

Round	GLOSS	Gaelic	ConcatGLOSSGaelic	interleavingGdGLOSS	ReplacingGaelic	Rep
0	18.39	17.29	15.42	13.67	9.44	
1	18.00	16.42	14.31	12.49	9.07	
2	16.02	15.29	15.38	11.01	7.69	
3	20.22	15.97	14.18	12.33	9.12	
4	19.02	17.79	18.63	12.56	9.08	
5	15.53	16.73	14.89	12.13	10.45	
6	18.00	17.11	15.16	11.55	8.62	
7	20.08	16.37	15.20	12.78	10.00	
8	15.82	15.93	15.50	12.43	10.52	
9	15.93	16.99	15.72	11.65	8.46	
Mean	17.70	16.59	15.44	12.26	9.24	

Table 2.8. 'Para' is 'GLOSS -¿ Gaelic; Gloss -¿ En; Gaelic -¿ En', 'ParaPart' is 'Para plus Gaelic word token -¿ Gloss token', 'Over' means oversampling, 'Half' means using half of the training data.