

TITLE OF YOUR THESIS

A Dissertation
Presented to
The Academic Faculty

By

George P. Burdell

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Fiction

Georgia Institute of Technology

January 1927

Copyright © George P. Burdell 1927

TITLE OF YOUR THESIS

Approved by:

Dr. Burdell, Advisor
School of Myths
Georgia Institute of Technology

Dr. Two
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Three
School of Electrical Engineering
Georgia Institute of Technology

Dr. Four
School of Computer Science
Georgia Institute of Technology

Dr. Five
School of Public Policy
Georgia Institute of Technology

Dr. Six
School of Nuclear Engineering
Georgia Institute of Technology

Date Approved: January 11, 2000

A great quote to start the thesis

George P. Burdell

A great dedication goes here.

ACKNOWLEDGEMENTS

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

TABLE OF CONTENTS

Acknowledgments	v
List of Tables	viii
List of Figures	ix
Chapter 1: Practical Considerations Concerning Tensorflow	1
1.1 Why Not Use Gradient Tape and TensorFlow-1-Style Code?	1
1.2 Shared Weights Between Layers	1
1.3 A Note on Custom Gradients	2
1.4 Updating TensorFlow Variables After Applying Gradients	3
1.5 Other Considerations	4
Chapter 2: Technical Approach	5
2.1 Section	5
2.1.1 Example Subsection	5
Chapter 3: Results	7
Chapter 4: Discussion	8
Chapter 5: Conclusion	9

Appendix A: Experimental Equipment	11
Appendix B: Data Processing	12
References	13
Vita	14

LIST OF TABLES

LIST OF FIGURES

2.1	This is another example Figure, rotated to landscape orientation.	6
-----	---	---

SUMMARY

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

CHAPTER 1

PRACTICAL CONSIDERATIONS CONCERNING TENSORFLOW

Most of the computations for my research rely on TensorFlow version 2.3.1 [1], a Python library for machine learning specializing in building models with differentiable, parameterizable composite functions and learning model parameters using gradient descent or other gradient-based optimization methods. TensorFlow is a common platform for researchers and developers working on artificial neural networks, and there are many tutorials and examples freely available online, so I will not replicate that work here. This chapter assumes the reader already has some familiarity with TensorFlow and Keras [2] (a high-level library inside TensorFlow). The goal of this chapter is to provide the reader with the tools and workarounds to be able to replicate my work without resorting to hacking things together with gradient tape and/or TensorFlow-1-style code.

Why Not Use Gradient Tape and TensorFlow-1-Style Code?

Keras offers a high-level environment. Code written in Keras's framework is easier to integrate with other work. Gradient tape is great for hacking something together or debugging, but promotes styles of coding that are less readable, less maintainable, and less portable. Keras also has a lower learning curve than the broader TensorFlow library.

Shared Weights Between Layers

Trainable TensorFlow variables declared outside of any Keras layer will not be automatically added to a Keras model's list of trainable variables. In most cases, this limitation is not a problem; it is intuitive to declare a layer's weights inside that layer. However, sometimes the same variable is needed in multiple distinct layers. To include a variable in the

model's trainable variables, it is sufficient to declare the variable in one layer and pass the variable (or the layer it was initialized in) as an input argument to the `__init__` function of the other layers that share that variable. This will work even if the Keras model does not use the layer that declared the variable.¹

A Note on Custom Gradients

TensorFlow offers a well-documented means of replacing TensorFlow's gradient computations of an operation with specified custom gradient computations. However, if the operation involves multiple tensors that are inputs or trainable variables, the standard approach replaces all the gradients with custom gradients. If TensorFlow's gradient computations are sufficient for some tensors but not others, a workaround is necessary. This workaround is best explained by example.

Suppose the operation is the following:

$$z = f(x, y)$$

for which the standard TensorFlow gradient computations of f are desired in respect to x , but custom gradient computations are desired in respect to y is $g(\nabla_z \text{LOSS})$. This can be rewritten as the following:

```
@tf.custom_gradient
def h(z, y):
    def grad_fun(grad):
        return (tf.identity(grad), g(grad))
    return z, grad_fun

z = f(x, tf.stop_gradient(y))
```

¹One could instead declare the variable outside any layers, pass it into the `__init__` functions of all the variables that depend on it, and then manually add the variable to the model's list of trainable variables, but I do not recommend this approach. The resulting code will be less readable and much less maintainable.

```
z = h(z, y)
```

The function h does nothing on the forward pass, but in the backward pass computes the custom gradient in respect to y as intended.

Updating TensorFlow Variables After Applying Gradients

To update TensorFlow Variables after applying gradients, it is necessary to track which variables are affected and what their corresponding update functions are. To accomplish this, I store the update functions in a Python dictionary using variable names as the dictionary keys. This Python dictionary needs to be widely accessible so that layers can add update functions when they are initialized; a simple way to do this is to make the update function Python dictionary a class attribute. The keys need to be unique, but TensorFlow variable names can conflict. It is easy to avoid this problem by checking for conflicts before adding a new update function.

```
class PostProcess:
    update = {}

    def add_update(varName, update_fun):
        assert varName not in PostProcess.update
        PostProcess.update[varName] = update_fun
```

In the standard Keras training paradigm, models are trained using the fit function, a method in the Keras model object. The fit function calls the function `train_step`, where gradients are applied. To update TensorFlow Variables after gradients are applied, `train_step` is the function to modify. The only change that needs to be made is adding a function call to all update functions that correspond to the model's list of trainable variables.

```
class Model_subclass(tf.keras.Model):
    def train_step(self, data):
```

```
trainStepOutputs = tf.keras.Model.train_step(self, data)
update_ops = []
for tv in self.trainable_variables:
    if tv.name in PostProcess.update:
        PostProcess.update[tv.name]()
return trainStepOutputs
```

Changes to Tensorflow variables in the update function must use the assign command (or its variants: assign_add, assign_sub, ect). Otherwise, TensorFlow will detect that computations lie outside of its computational graph and throw an error. Note that using the assign command on Python variables that are not TensorFlow variables will produce some very cryptic error messages, so be sure to use the assign command correctly. If the value change of one TensorFlow variable depends on the value of another TensorFlow variable value pre-update, it may be necessary to use the TensorFlow control_dependencies command to get TensorFlow to track that dependency. TensorFlow has a useful tool called TensorBoard that helps visualize TensorFlow's dependencies, but a workaround is required to use TensorBoard on update functions that are called after applying gradients. To use TensorBoard to visualize dependencies in an update function, temporarily call the update function in the layer's call method, use TensorBoard to verify all necessary dependencies are being tracked, then remove the update function call from the layer's call method.

Other Considerations

The TensorFlow Probability version 0.11.1 [3] is an extension of TensorFlow mostly used for probabilistic models. The library contains a Cholesky update function, but the function does not properly handle complex inputs. To compute Cholesky updates for complex inputs, users should either write their own implementation or use my code (included in supplementary material).

CHAPTER 2

TECHNICAL APPROACH

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Section

This is a section in Chapter 2.

Example Subsection

This is a subsection in Chapter 2.

Example Subsubsection

This is a subsubsection in Chapter 2.

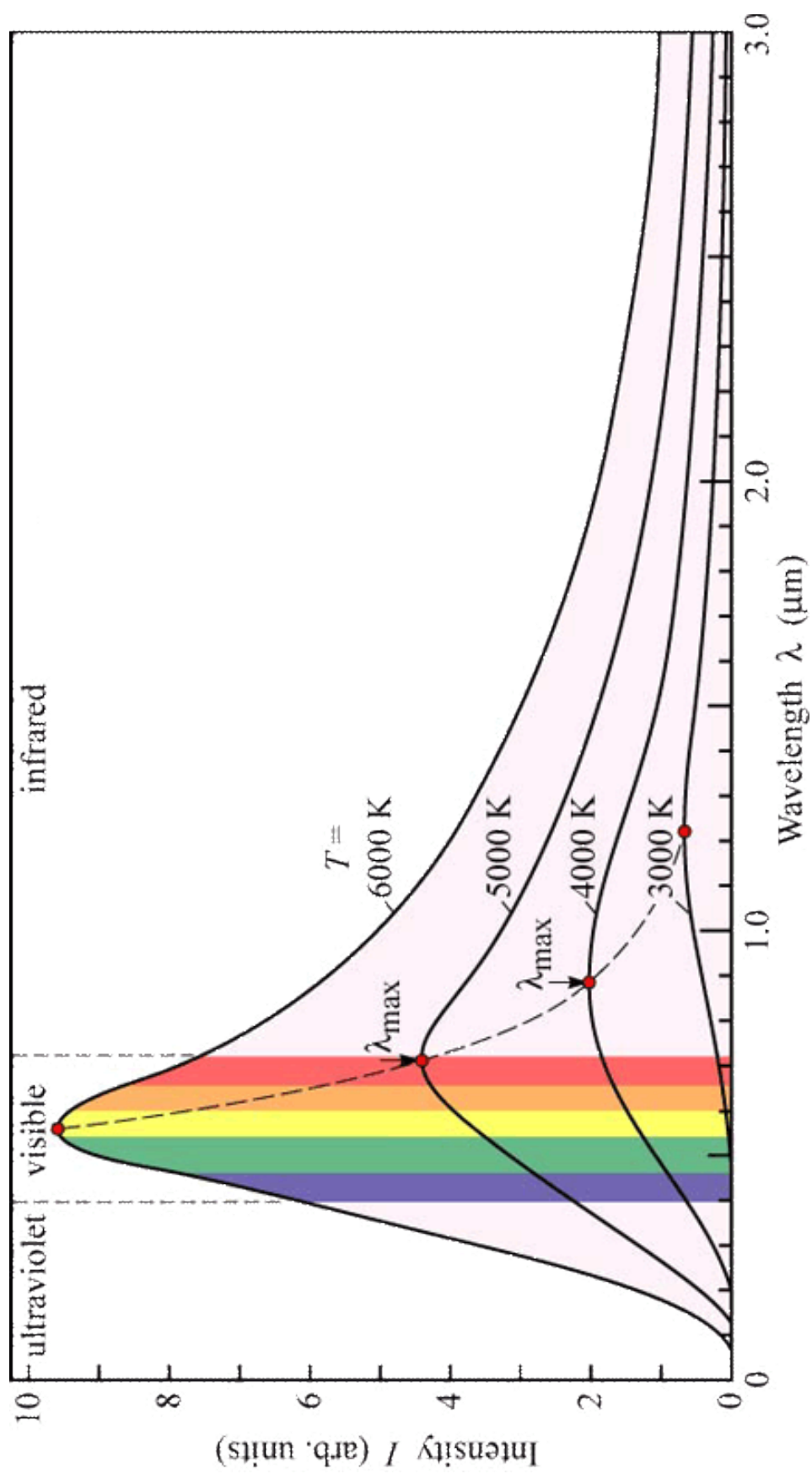


Figure 2.1: This is another example Figure, rotated to landscape orientation.

CHAPTER 3

RESULTS

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

CHAPTER 4

DISCUSSION

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum..

CHAPTER 5

CONCLUSION

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Appendices

APPENDIX A

EXPERIMENTAL EQUIPMENT

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

APPENDIX B

DATA PROCESSING

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

REFERENCES

- [1] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, *Tensorflow: Large-scale machine learning on heterogeneous systems*, Software available from tensorflow.org, 2015.
- [2] F. Chollet *et al.*, *Keras*, <https://keras.io>, 2015.
- [3] J. V. Dillon, I. Langmore, D. Tran, E. Brevdo, S. Vasudevan, D. Moore, B. Patton, A. Alemi, M. Hoffman, and R. A. Saurous, “Tensorflow distributions,” *ArXiv preprint arXiv:1711.10604*, 2017.

VITA

Vita may be provided by doctoral students only. The length of the vita is preferably one page. It may include the place of birth and should be written in third person. This vita is similar to the author biography found on book jackets.