Faculty of Science
Graduate School of Informatics

APPROVAL OF INDIVIDUAL PROGRAMME CONTENT MSC LOGIC

The Examinations Board of the MSc Logic accepts the individual programme content enumerated in Appendix I and the proposed Thesis plan given in Appendix II as sufficient for obtaining the MSc Logic degree, and gives permission to start with the research for the MSc Logic thesis as outlined in Appendix II.

Name	Daan van Stigt	Address	Amsteldijk 75
Student number	10255141	Postal code/City	1074 JA, Amsterdam
MSc	Logic	Phone	06 493 07 365
Track	Logic & Computation	Email	daan_douwe@hotmail.com

	Obligatory	Advised	Not permitted
Basic Logic:	☐	☐	☐

Total amount of credits already rewarded for the MSc Logic programme: 99 ECTS
(max. 18 ECTS left open, none of which are core elements: see graduation plan)

Appendices:

- Request for approval of the individual programme content (Appendix I)
- Thesis plan (Appendix II)

Approval grantedon (date): , Amsterdam

	Date	**Signature**
Student		
Thesis Supervisor		
Programme Manager		
Programme Director		
Chair Examinations Board		

This form, when completed and signed, will be kept with the Education Service Centre (together with the appendices 'Request for approval of the individual programme content' and 'Thesis plan').

Request for approval of the individual programme content

Obligatory elements

(including one research project of 6 EC)

Courses	Code	EC
Logic, Language and Computation	NIS 5314LOLC3Y	3
Research Project Master of Logic 1 (Imperative Programming with Python)	NIS 53141RPL6Y	6
Computational Complexity	NIS 5314COCO6Y	6
Introduction to Modal Logic	NWIS 5122INML6Y	6
Information Theory	NIS 5314INTH6Y	6

for research projects please use the official title as registered (or to be registered) in SIS. To see the title in SIS, go to 'main menu' > 'self service' > 'student center' > 'my study details', click on the project and then click on 'individual course info'.

Non-obligatory elements

Courses	Code	EC
Research Project Master of Logic 2 (Genetic Algorithms in Haskell)	NIS 53142RPL6Y	6
Combinatorics with Computer Science Applications	NIS 5314CWCS6Y	6
Game Theory	NIS 5314GATH6Y	6
Machine Learning 1	NII 52041MAL6Y	6
Natural Language Processing 1	NII 52041NLP6Y	6
Complex Networks	NWIS 5374CONE8Y	8
Research Project Master of Logic 3 (Probabilistic Modelling with PCFGs)	NIS 53143RPL6Y	6
Natural Language Processing 2	NII 52042NLP6Y	6
Machine Learning 2	NII 52042MAL6Y	6
Quantum Computing	NWIS 5334QUCO8Y	8
Parallel Algorithms	NWIS 53348PAA8Y	8

for research projects please use the official title as registered (or to be registered in SIS).

Total amount of credits for the MSc Logic

Obligatory Elements	27	EC
Non-Obligatory Elements	72	EC
Master Thesis	30	EC
Total	129	EC

Thesis plan

Name student: Daan van Stigt

Subject of thesis project:

Deep generative models
for unsupervised
problems in NLP

Thesis supervisors:
(at least one has to be a staff-member of ILLC)

1. Wilker Aziz
2.

Starting date of thesis project: 01/02/2018

Supervision frequency: ☐ Weekly ☐ Biweekly ☐ Other, namely

Planned defense date: 30/08/2018

Planned date for handing thesis to committee: 7/08/2018

Courses yet to complete

Course	EC	Semester and year of study	(planned) end date
Total			

Please give a description of your thesis project on page 4

Description and Planning of Thesis Project

The subject of this thesis is the emerging synthesis between probabilistic graphical models and deep neural networks: deep generative models. In particular we are concerned with the application of these models to unsupervised problems in natural language processing. We have identified three possible applications: unsupervised word and speech segmentation (Goldwater et al., 2009; Kamper et al., 2016); language modelling with (recurrent) switching linear dynamical systems (Belanger, 2015; Linderman, 2016); and language models combining (recurrent) neural networks with topic models (Dieng et al., 2017; Srivastava & Sutton, 2017; Miao et al. 2017). These are different applications that are nevertheless amenable to similar modelling methods within the deep generative framework, and that, most importantly, are connected by a similar interest: to extract interpretable latent structure from text (speech) in an unsupervised manner.

Graphical models with latent variables posit rich hidden structure that governs observed data. More technically, these models specify a joint distribution over both the observed and latent variables, and the graphical structure specifies how this joint distribution factorises. A variety of distributions (e.g. from the *exponential family*) are typically chosen to model the conditionally independent parts. Learning in these models then becomes a matter of *inference*: conditioned on observations, what are our updated believes about the posited latent structure? This involves computation of the posterior distribution of the latent variables. This computation is generally intractable and approximate methods are needed for this. Luckily, for the mentioned models these methods have been developed successfully and worked to a great deal of generality (Blei et al., 2016).

This class of models lend themselves naturally to unsupervised problems. We can incorporate prior knowledge about the data via the latent structure and the conditional independences, and the posterior distribution over the latent structure can be easy to interpret and very informative. On the other hand, deep learning methods have shown to be very good at automatically learning flexible representations of data. These methods are complimentary:

> "Probabilistic graphical models provide many tools to build structured representations, but often make rigid assumptions and may require significant feature engineering. Alternatively, deep learning methods allow flexible data representations to be learned automatically, but may not directly encode interpretable or tractable probabilistic structure." (Johnson et al., 2016)

A fruitful enterprise is then to combine their complimentary strengths: to use neural networks to learn flexible representations of the observed data and to use graphical models to represent latent structure in terms of those representations. The challenge in combing these is *inference*: the generic approximate inference techniques mentioned above break down in this settings. Much of the recent research on deep generative models has focused on developing approximate inference methods for these settings (Ranganath et al., 2014; Kingma & Welling, 2014; Kucukelbir et al., 2017). The Structured Variational Autoencoder (SVAE) (Johnson et al. 2016) is a proposed general inference method for these types of model. We study the SVAE and its potential application to one of the problems mentioned in the introduction.

References

David Belanger, Sham Kakade (2015). A Linear Dynamical System Model for Text. International Conference of Machine Learning 2015. David M. Blei, Alp Kucukelbir, Jon D. McAuliffe (2017), *Variational Inference: A Review for Statisticians*. Journal of the American Statistical Association, Vol. 112 , Iss. 518, 2017.

Adji B. Dieng, Chong Wang, Jianfeng Gao, John Paisley, (2017) *TopicRNN: A Recurrent Neural Network with Long-Range Semantic Dependency.* Arxiv preprint.

Sharon Goldwater, Thomas L. Griffiths and Mark Johnson (2009). *A Bayesian framework for word segmentation: exploring the effects of context.* Cognition 112 (1), pp. 21–54. 2009.

Matthew Johnson, David Duvenaud, Alex Wiltschko, Bob Datta, Ryan P. Adams (2016). *Composing graphical models with neural networks for structured representations and fast inference.* NIPS 2016.

Herman Kamper, Aren Jansen and Sharon Goldwater (2016). *Unsupervised Word Segmentation and Lexicon Discovery using Acoustic Word Embeddings.* IEEE Transactions on Audio, Speech and Language Processing 24 (4), pp. 669–679. 2016.

Durk P. Kingma, Max Welling (2014). *Auto-Encoding Variational Bayes.* Proceedings of the 2nd International Conference on Learning Representations 2014.

Alp Kucukelbir, Dustin Tran, Rajesh Ranganath, Andrew Gelman, David M. Blei (2017). *Automatic Differentiation Variational Inference.*18(14):1−45, 2017.

Scott W. Linderman, Andrew C. Miller, Ryan P. Adams, David M. Blei, Liam Paninski, Matthew J. Johnson, (2017). *Recurrent switching linear dynamical systems.* Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, PMLR 54:914-922, 2017.

Yishu Miao, Edward Grefenstette and Phil Blunsom. (2017). *Discovering Discrete Latent Topics with Neural Variational Inference.* In the 34th International Conference on Machine Learning (ICML). 2017.

Rajesh Ranganath, Sean Gerrish, David Blei,(2014). *Black Box Variational Inference.* Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics.

Akash Srivastava, Charles Sutton (2017). *Autoencoding Variational Inference For Topic Models.* Proceedings of the 5th International Conference on Learning Representations 2017.