TUHH
Hamburg
University of
Technology

UKE

Comprehensive Guidelines and Templates for Thesis Writing

Master's Thesis

of

Author

Date of issue:
Date of submission:
Examiner:
Supervisor:

Statutory Declaration

I, Author, hereby affirm that the following Master's thesis has been elaborated solely by myself. No other means and sources except those stated, referenced and acknowledged have been used.

(Author)

Abstract

Contents

Acronyms Index

TUHH:	Hamburg University of Technology
SOFA:	Sequential Organ Failure Assessment
qSOFA:	Quick Sequential Organ Failure Assessment
ICU:	Intensive Care Unit
EHR:	Electronic Health Record
YAIB:	Yet Another ICU Benchmark
FSQ:	Finite Scalar Quantization
RL:	Reinforcement Learning

1 List of Notes

2 Notes

2.1 Base ODE-System

$$\dot{\varphi}_i^1 = \omega^1 - \frac{1}{N}\sum_{j=1}^{N}\{(a_{ij}^1 + \kappa_{ij}^1)\sin(\varphi_i^1 - \varphi_j^1 + \alpha^{11})\} - \sigma\sin(\varphi_i^1 - \varphi_i^2 + \alpha^{12}) \tag{1.1}$$

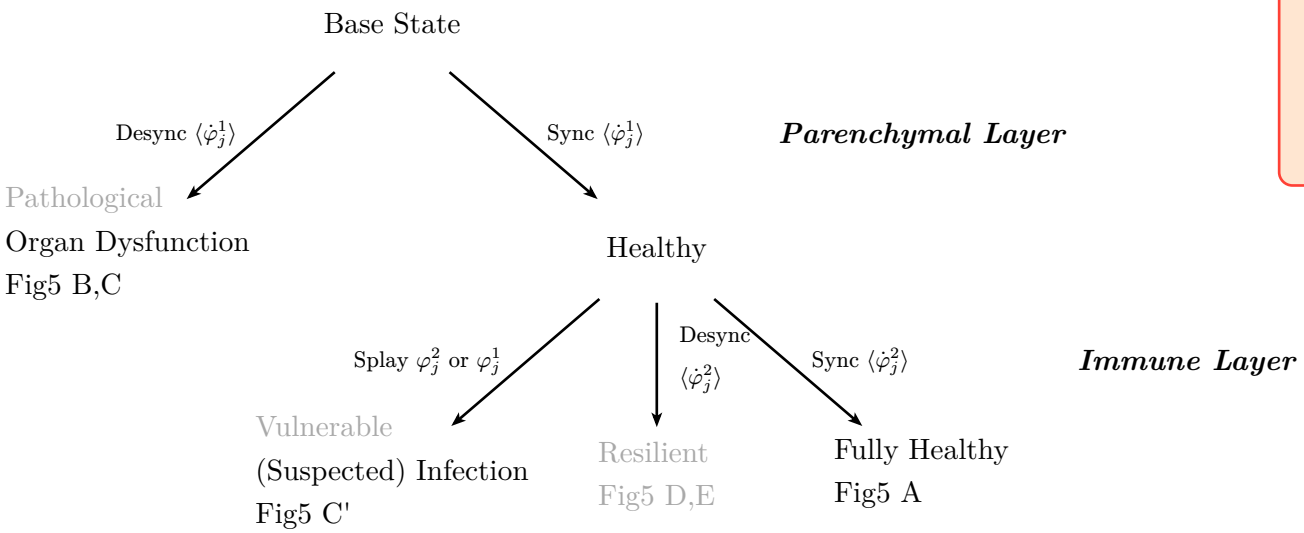
$$\dot{\kappa}_{ij}^1 = -\varepsilon^1(\kappa_{ij}^1 + \sin(\varphi_i^1 - \varphi_j^1 - \beta)) \tag{1.2}$$

$$\dot{\varphi}_i^2 = \omega^2 - \frac{1}{N}\sum_{j=1}^{N}\kappa_{ij}^2\sin(\varphi_i^2 - \varphi_j^2 + \alpha^{22}) - \sigma\sin(\varphi_i^2 - \varphi_i^1 + \alpha^{21}) \tag{1.3}$$

$$\dot{\kappa}_{ij}^2 = -\varepsilon^2(\kappa_{ij}^2 + \sin(\varphi_i^2 - \varphi_j^2 - \beta)) \tag{1.4}$$

Introduced in [1] and slightly adapted in [2].

table for parameters



2.2 Computational Saving

For parts in the form of $\sin(\theta_l - \theta_m)$ following [3] one can calculate and cache the terms $\sin(\theta_l), \sin(\theta_m), \cos(\theta_l), \cos(\theta_m)$ in advance:

$$\sin(\theta_l - \theta_m) = \sin(\theta_l)\cos(\theta_m) - \cos(\theta_l)\sin(\theta_m) \qquad \forall l, m \in \{1, ..., N\} \tag{2}$$

so the computational cost for the left-hand side for N oscillators can be reduced from $N(N-1)$ to $4N$ trigonometric function evaluations, positively impacting the computational efficiency of the whole ODE-system significantly.

2.3 Implementation Differences

- Actual Mean Phase Velocity instead of averaged difference over time.
- +They Calculate the difference wrong since the phase difference should be mod$[2\pi]$
- Different solver accuracy, but very similar
- They are not batching the computation

2.4 Metrics

2.4.1 Kuramoto Parameter

Kuramoto Order Parameter

cite

$$R_2^\mu = \frac{1}{N}\left|\sum_j^N e^{i\cdot\varphi_j(t)}\right| \quad \text{with} \quad 0 \leq R_2^\mu \leq 1 \tag{3}$$

$R_2^\mu = 0$ splay-state and $R_2^\mu = 1$ is fully synchronized.

Mean Phase Velocities

$$\langle\varphi^\mu\rangle = \frac{1}{N}\sum_j^N \varphi_j^\mu \tag{4}$$

2.5 Data

Entropy, Splay Ratio, MPV Std, Cluster Ratio

MIMIC-3

https://github.com/alistairewj/sepsis3-mimic/tree/v1.0.0°

2.6 Model

For a general model setup, the latent space $z = \left(a^1, \sigma, \alpha, \beta, \omega^1, \omega^2, \frac{C}{N}, \varepsilon^1, \varepsilon^2\right)$ represents the parameter of the dynamic network model, so we have

$$z \in \mathbb{R}^d \quad \text{with} \quad d = 9 \tag{5}$$

As shown in the supplemental material of [2], for example, the parameter α exhibits a π-periodicity, allowing to reduce the effective parameter space by constraining certain parameters with upper and lower bounds. These bounds are omitted here for simplicity but are included in . To further reduce the latent space z, the we keep $a^1, \omega^1, \omega^2, \frac{C}{N}, \varepsilon^1$ and ε^2 fixed. The reduced latent space $z' = (\sigma, \alpha, \beta)$:

table

$$z' \in \mathbb{R}^{d'} \quad \text{with} \quad d' = 3 \tag{6}$$

where both alpha and beta exhibit a periodic behavior Each point in the latent space z_j can be categorized as either of *healthy*, *vulnerable* or *pathological*.

We relate high-dimensional physiological observations (e.g. samples from the MIMIC-III database) to the latent space via:

$$x_j = f(z_j) + \varepsilon \tag{7}$$

where f is unknown an unknown function and ε the measurement noise. Note that different observations x_j can be mapped to the same classification, as for the latent space. We define two the two class mappings Q and R:

$$Q(x_j) = c_j = R(z_j) \quad \text{where} \quad x_j = f(z_j) + \varepsilon \tag{8}$$

mapping observations and the latent representation to a shared class label c. To make things more complicated, R does not directly act on z, but rather the metrics derived from the solution to a dynamical system (initial value problem) (Equation 1) parameterized by z. The metrics are detailed in Section 2.4.

In the setting of structured latent variational learning we want to approximate an encoder $q(z|x)$ to infer the latent variables from observed data X and the class.

How to structure the latent space? Binary classification (sepsis, no sepsis) may not provide enough information to accurately structure the latent space. The options:

- Add more classes like resilient/vulnerable… maybe even the full spectrum?
  - need to be modeled by R
- Introduce the time/action component as additional information (like the Reinforcement Learning (RL) environment?)

For the cohort extraction and SOFA calculation I use [4] and [5]. The nice thing is we could interpret larger SOFA scores (> 3) as the vulnerable state introduced by [2]. Increases in SOFA score ≥ 2 could then be used as definition for sepsis.

mapping not really clear, which metrics correspond to sofa/infection

YAIB [5] and other resources care about the "onset" of infection and sepsis [6]. For sepsis this isn't really problematic since we could use the "state transitions" as indicators. But for the suspected infection it is problematic, maybe use si_upr and si_lwr provided by [4] (https://eth-mds.github.io/ricu/reference/label_si.html°). These would be 48h - SI - 24h adapted from [7], maybe a bit too much.

3 Introduction

4 Preliminaries

4.1 Sepsis

As the most extreme course of an infectious disease, sepsis poses a serious health threat, with a high mortality rate and frequent long-term consequences for survivors. In 2017, an estimated 48.9 million people worldwide suffered from sepsis and the same year, 11.0 million deaths were associated with sepsis [8], making it the most common cause of in-hospital deaths. Untreated, the disease is always fatal and even with successful treatment, around 75% of those affected suffer long-term consequences. Highlighting the importance of early recognition and treatment of infections in patients with pre-existing health conditions. Overall, untreated septic diseases in particular represent an enormous burden on the global healthcare system.

The triggers for sepsis are varied, but almost half of all sepsis-related deaths occur as a secondary complication of an underlying injury or a non-communicable, also known as chronic disease [9]. Faster recognition of a septic condition significantly increases the chance of survival [10], it urges to develop accurate and robust detection and prediction methods, i.e. reducing the time to receive the appropriate medical attention.

Per definition, sepsis is a "life-threatening organ dysfunction caused by a dysregulated host response to infection" [11]. There are multiple (now historic) more specific definitions available and sometimes blurry terminology used when dealing with the sepsis and septic shocks. The following chapter Section 4.1.1 gives a more detailed introduction to the most commonly used sepsis definition, which is referred to as Sepsis-3, additionally the chapter provides a short explanation of both the pathology and biology of sepsis

4.1.1 Sepsis-3 definition

Out of the need for an update of an outdated and partly misleading sepsis model a task force led by the "Society of Critical Care Medicine and the European Society of Intensive Care Medicine", was formed in 2016. Their resolution, named "Third International Consensus Definitions for Sepsis and Septic Shock" [11], provides until today the most up to date and most widely used sepsis definition and guidance on sepsis identification.

In general sepsis does not classify as a specific illness, rather a condition of "physiologic, pathologic, and biochemical abnormalities" [11], where the original causes are still mostly uncertain. Most commonly the underlying cause of sepsis is diarrhoeal disease, road traffic injury the most common underlying injury and maternal disorders the most common non-communicable disease causing sepsis [8].

The Sepsis-3 defines a condition as septic when a patient exhibits an abnormal host reaction against a documented or suspected infection, i.e. an immune system overreaction to some kind of infection or inflammation harmful to the body and organ system. When infection is first suspected, even modest organ dysfunction is linked to a 10% increase of in-hospital mortality.

In clinical practice any patient prescribed with antibiotics followed by the cultivation of body fluids, or the other way around, is characterized with a suspected infection. The timings of prescription and fluid samplings play a crucial role. If the antibiotics were administered first, then the cultivation has to be done in the first 24h after first prescription, if the cultivation happened first, the antibiotics have to be prescribed in the following 72h [11]. This can be seen in the lower part of figure Figure 2. Regardless which happened first, the earlier of the two times is treated as the time of suspected infection

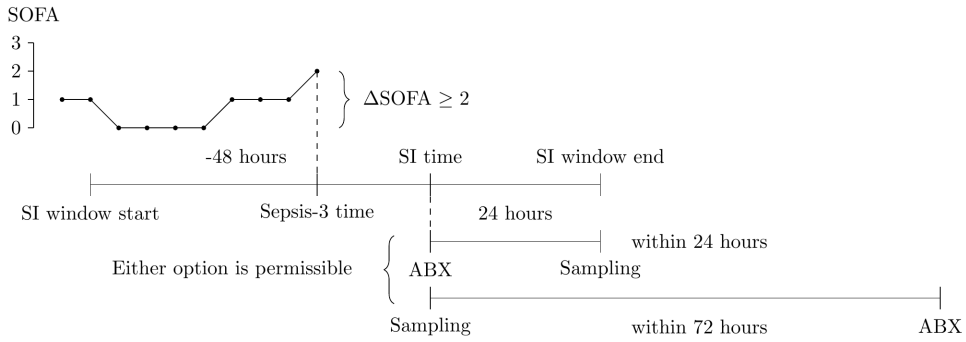


Figure 2: Graphical representation of the timings in the Sepsis-3 definition, taken from [12]

A responses dysregulation is measured by the change of organ functionality over time. Capturing the Sequential Organ Failure Assessment (SOFA) score [13] is regularly used to evaluate the severity of an illness and helps to guide treatment decisions and predict the risk of mortality outside of a sepsis context. The SOFA score is calculated at least every 24 hours and assess six different organ systems and assigns a score from 0 (normal function) to 4 (high degree of dysfunction) each, as stated in Table 1. While the magnitude or baseline of a patients initials SOFA score captures preexisting organ dysfunction, an increase in SOFA score ≥ 2 between measurements indicates an acute organ dysfunction and a drastic worsening in the patients condition.

Category	Indicator	1	2	3	4
Respiration	PaO_2/FiO_2 [mmHg]	< 400	< 300	< 200	< 100
	Mechanical Ventilation			yes	yes
Coagulation	Platelets $[\times \frac{10^3}{mm^3}]$	< 150	< 100	< 50	< 20
Liver	Bilirubin $[\frac{mg}{dl}]$	1.2-1.9	2.0-5.9	6.0-11.9	> 12.0
Cardiovascular[1]	MAP [mmHg]	< 70			
	or Dopamine		≤ 5	> 5	> 15

[1] Adrenergica agents administered for at least 1h (doses given are in [µg/kg · min]

Category	Indicator	1	2	3	4
	or Dobutamine		any dose		
	or Epinephrine			≤ 0.1	> 0.1
	or Noepinephrine			≤ 0.1	> 0.1
Central Nervous System	Glasgow Coma Score	13-14	10-12	6-9	< 6
Renal	Creatinine $[\frac{\text{mg}}{\text{dl}}]$	1.2-1.9	2.0-3.4	3.5-4.9	> 5.0
	or Urine Output $[\frac{\text{ml}}{\text{day}}]$			< 500	< 200

Also newly introduced in [13] a bedside clinical score termed Quick Sequential Organ Failure Assessment (qSOFA):

label

- Respiratory rate >= 22/min
- Altered mentation
- Systolic blood pressure <= 100 mm Hg

If a patient fulfills at least two of these criteria have an increased risk of organ failure, but it is not as accurate as the SOFA score and is designed as a fast patient screening tool.

risk increase by classification

4.2 Biology and Cytokine Storms

4.3 The need for sepsis prediction

4.4 Maybe Treatment

The Parenchymal (Equation 1.1° and Equation 1.2°) and Immune (Equation 1.3° and Equation 1.4°) layer and their respective states of the dynamical system naturally are consistent with the two cornerstones of the Sepsis-3 definition [11], i.e. SOFA score and suspicion of an infection.

dynamic

Healthy $\rightarrow$ sync $\langle \dot{\varphi}_j^1 \rangle$ and $\langle \dot{\varphi}_j^1 \rangle$

SOFA $\rightarrow$ desync $\langle \dot{\varphi}_j^1 \rangle$

Suspected infection $\rightarrow$ splay/desync $\langle \dot{\varphi}_j^2 \rangle$

septic $\rightarrow$ desync $\langle \dot{\varphi}_j^1 \rangle$ and $\langle \dot{\varphi}_j^1 \rangle$

5 The Data and Task problems

6 Problem definition

In [14], a survey among clinicians regarding AI-assistance in healthcare, one participant emphasizes that specific vitals signs might not be to be of less importance, rather the change/trend of a patients trajectory. Another piece of finding of the same study was the preference of trajectories over plain event predictions.

7 State of the Art

7.1 Model Based Methods

7.2 Data Based Methods

7.2.1 Selected Works

8 Dynamic Network Model (DNM)

8.1 Description

8.1.1 Functional Models

8.1.2 Parenchymal

8.1.3 Immune System

8.1.4 Kuramoto

8.2 Implementation

8.2.1 Standard

8.2.2 Lie

9 Latent Dynamics Model

9.1 Task - Definition of Ins and Outs

9.2 Data

9.2.1 MIMIC-III/IV

9.2.2 YAIB + (Further) Preprocessing

9.2.2.1 ricu-Concepts

9.3 Latent Dynamics Model (LDM)

9.3.1 The high level ideas

9.3.1.1 Representation Learning and Latent Spaces

9.3.1.2 Semantics

9.3.1.3 Autoregressive Prediction

9.3.2 The Lookup (FSQ)

9.3.3 Encoder

9.3.4 Decoder

9.3.5 Introducing time

9.3.6 Combining the building blocks

9.4 Metrics (How to validate performance?)

10 Experimental Results

10.1 Metrics

10.2 Further Experiments

10.2.1 Custom Latent Space

10.2.2 SOFA vs Infection

11 Conclusion

Bibliography

[1] J. Sawicki, R. Berner, T. Löser, and E. Schöll, "Modeling Tumor Disease and Sepsis by Networks of Adaptively Coupled Phase Oscillators," *Frontiers in Network Physiology*, vol. 1, 2022, doi: 10.3389/fnetp.2021.730385°.

[2] R. Berner, J. Sawicki, M. Thiele, T. Löser, and E. Schöll, "Critical Parameters in Dynamic Network Modeling of Sepsis," *Frontiers in Network Physiology*, vol. 2, 2022, doi: 10.3389/fnetp.2022.904480°.

[3] T. Böhle, C. Kuehn, and M. Thalhammer, "On the reliable and efficient numerical integration of the Kuramoto model and related dynamical systems on graphs," *International Journal of Computer Mathematics*, vol. 99, no. 1, pp. 31–57, 2022, doi: 10.1080/00207160.2021.1952997°.

[4] N. Bennett, D. Plečko, I.-F. Ukor, N. Meinshausen, and P. Bühlmann, "ricu: R's interface to intensive care data," *GigaScience*, vol. 12, p. giad41, 2023.

[5] R. van de Water, H. N. A. Schmidt, P. Elbers, P. Thoral, B. Arnrich, and P. Rockenschaub, "Yet Another ICU Benchmark: A Flexible Multi-Center Framework for Clinical ML," in *The Twelfth International Conference on Learning Representations*, Oct. 2024.

[6] M. Moor, B. Rieck, M. Horn, C. R. Jutzeler, and K. Borgwardt, "Early Prediction of Sepsis in the ICU Using Machine Learning: A Systematic Review," *Frontiers in Medicine*, 2021, doi: 10.3389/fmed.2021.607952°.

[7] C. W. Seymour *et al.*, "Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)," *JAMA*, vol. 315, no. 8, pp. 762–774, 2016, doi: 10.1001/jama.2016.0288°.

[8] K. E. Rudd and et al., "Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study," *The Lancet*, vol. 395, no. 10219, pp. 200–211, 2020, doi: 10.1016/S0140-6736(19)32989-7°.

[9] C. Fleischmann-Struzek, D. Schwarzkopf, and K. Reinhart, "Inzidenz der Sepsis in Deutschland und weltweit: Aktueller Wissensstand und Limitationen der Erhebung in Abrechnungsdaten," *Medizinische Klinik - Intensivmedizin und Notfallmedizin*, vol. 117, no. 4, pp. 264–268, May 2022, doi: 10.1007/s00063-021-00777-5°.

[10] C. W. Seymour *et al.*, "Time to Treatment and Mortality during Mandated Emergency Care for Sepsis," *The New England Journal of Medicine*, vol. 376, pp. 2235–2244, Jun. 2017, doi: 10.1056/NEJMoa1703058°.

[11] M. Singer *et al.*, "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)," *JAMA*, vol. 315, no. 8, pp. 801–810, Feb. 2016, doi: 10.1001/jama.2016.0287°.

[12] N. Bennett, D. Plěcko, and I.-F. Ukor, "Sepsis 3 label — \textttsep3." 2025.

[13] J. L. Vincent *et al.*, "The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine," *Intensive care medicine*, vol. 22, no. 7, pp. 707–710, 1996.

[14] B. Eini-Porat, O. Amir, D. Eytan, and U. Shalit, "Tell me something interesting: Clinical utility of machine learning prediction models in the ICU," *Journal of Biomedical Informatics*, vol. 132, p. 104107, 2022, doi: 10.1016/j.jbi.2022.104107°.