TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Elektrotechnik und Informationstechnik
Bioinspired Communication Systems

Bayesian Inference of Information Transfer in Graph-Based Continuous-Time Multi-Agent Systems

Master-Thesis
Elektro- und Informationstechnik

Eingereicht von

Gizem Ekinci

am
07.07.2020

1. Gutachten: Prof. Dr. techn. Heinz Koeppl
2. Gutachten: Dominik Linzner

Erklärung zur Abschlussarbeit gemäß §22 Abs. 7 und §23 Abs. 7 APB TU Darmstadt

Hiermit versichere ich, Gizem Ekinci, die vorliegende Arbeit gemäß §22 Abs. 7 APB der TU Darmstadt ohne Hilfe Dritter und nur mit den angegebenen Quellen und Hilfsmitteln angefertigt zu haben. Alle Stellen, die Quellen entnommen wurden, sind als solche kenntlich gemacht worden. Diese Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen. Mir ist bekannt, dass im Falle eines Plagiats (§38 Abs.2 APB) ein Täuschungsversuch vorliegt, der dazu führt, dass die Arbeit mit 5,0 bewertet und damit ein Prüfungsversuch verbraucht wird. Abschlussarbeiten dürfen nur einmal wiederholt werden. Bei der abgegebenen Arbeit stimmen die schriftliche und die zur Archivierung eingereichte elektronische Fassung gemäß §23 Abs. 7 APB überein.

English translation for information purposes only:

Thesis statement pursuant to §22 paragraph 7 and §23 paragraph 7 of APB TU Darmstadt:
I herewith formally declare that I, Gizem Ekinci, have written the submitted thesis independently pursuant to §22 paragraph 7 of APB TU Darmstadt. I did not use any outside support except for the quoted literature and other sources mentioned in the paper. I clearly marked and separately listed all of the literature and all of the other sources which I employed when producing this academic work, either literally or in content. This thesis has not been handed in or published before in the same or similar form. I am aware, that in case of an attempt at deception based on plagiarism (§38 Abs. 2 APB), the thesis would be graded with 5,0 and counted as one failed examination attempt. The thesis may only be repeated once. In the submitted thesis the written copies and the electronic version for archiving are pursuant to § 23 paragraph 7 of APB identical in content.

Darmstadt, den 07.07.2020

(Gizem Ekinci)

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Contents

1 Introduction

1.1 Motivation

1.2 Related Work

1.3 Contributions

1.4 Structure of the Thesis

2 Foundations

2.1 Problem Formulation

Problem: Agent making decisions based on incoming messages, but observes only a summary of them
Objective: To infer this observation model from agent's behaviour
Assuming that the behaviour of agent has been shaped by evolution (close) to optimality X_1 and X_2 homogenous continuous-time Markov processes with $\mathbf{Q}_1$ and $\mathbf{Q}_2$ transition intensity matrices

$$\mathbf{Q}_i \sim Gam(\alpha_i, \beta_i), \quad i \in \{1, 2\} \tag{2.1}$$

X_3 inhomogeneous continuous-time Markov process with set of actions $a \in \{a_0, a_1\}$ and set of transition intensity matrices $\boldsymbol{Q}_3 = \{\mathbf{Q}_{a_0}, \mathbf{Q}_{a_1}\}$

$$\mathbf{Q}_a \sim Gam(\alpha_a, \beta_a) \tag{2.2}$$

$\mathcal{X}_i = \{0, 1\}$ $\psi := p(y(t) \mid X_1(t), X_2(t))$ observation model
$b(x_1, x_2; t)$: belief state $\pi(a \mid b)$: optimal policy of X_3 $X_i^{[0,T]}$: discrete valued trajectory in time interval $[0, T]$

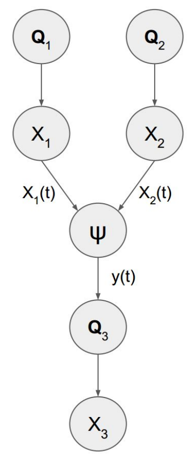

Figure 2.1: Graphical model.

2.2 Continuous Time Bayesian Networks

2.2.1 Time-Homogenous Continuous Time Markov Processes

The messages that are emitted by the parent nodes X_1 and X_2 are modelled as independent time-homogeneous continuous-time Markov processes (CTMP). These processes are defined by transition intensity matrices Q_{X_1} and Q_{X2}, whose intensities do not depend on time. In this matrix, the intensity q_i represents the instantaneous probability of leaving state i and $q_{i,j}$ represents the instantaneous probability of switching from state i to j.

Infinitesimal transition probabilities in terms of the entries of transition matrices q_{ij} can be written as [1]:

$$p_{i,j}(h) = \delta_{ij} + q_{i,j}h + o(h) \tag{2.3}$$

where $p_{i,j}(t) \equiv Pr(X^{(t+s)} = j \mid X^{(s)} = i)$ are Markov transition functions and o(.) is a function decaying to zero faster than its argument.

The *forward* or *master equation* is then derived as follows:

$$p_j(t) = \sum_{\forall i} p_{i,j}(h) p_i(t-h) \tag{2.4}$$

$$\begin{aligned} \lim_{h\to 0} p_j(t) &= \lim_{h\to 0} \sum_{\forall i} \left[\delta_{ij} + q_{i,j}h + o(h)\right] p_i(t-h) \\ &= \lim_{h\to 0} p_j(t-h) + \lim_{h\to 0} h \sum_{\forall i} q_{i,j} p_i(t-h) \end{aligned} \tag{2.5}$$

$$\lim_{h\to 0} \frac{p_j(t) - p_j(t-h)}{h} = \lim_{h\to 0} \sum_{\forall i} q_{i,j} p_i(t-h) \tag{2.6}$$

$$\begin{aligned} \frac{d}{dt} p_j(t) &= \sum_{\forall i} q_{i,j} p_i(t) \\ &= \sum_{\forall i \neq j} \left[q_{i,j} p_i(t) - q_{j,i} p_j(t)\right] \end{aligned} \tag{2.7}$$

Eq.2.7 can be written in matrix form:

$$\frac{d}{dt} p = p\mathbf{Q} \tag{2.8}$$

The solution to ODE, the time-dependent probability distribution $p(t)$ is,

$$p(t) = p(0) \exp(t\mathbf{Q}) \tag{2.9}$$

with initial distribution $p(0)$.

The amount of time staying in a state i is exponentially distributed with parameter q_i. The probability density function f for staying in the state i:

$$f(t) = q_i \exp\left(-q_i t\right), t \geq 0 \tag{2.10}$$

2.2.1.1 Likelihood Function

To write down the likelihood of a trajectory sampled from a homogenous CTMC X, first let us consider one transition $d =< i, j, t >$, where transition happens after time t from state i to j. The likelihood of this transition is:

$$L_X(d \mid \mathbf{Q}) = \left(q_i \exp\left(-q_i t\right)\right) \left(\frac{q_{i,j}}{q_i}\right) \tag{2.11}$$

We define sufficient statistics over a dataset D as $T[i]$, total amount of time spent in state i, $M[i,j]$ total number of transitions from state i to j, we can write down the likelihood of a trajectory $X^{[0,T]}$,

$$\begin{aligned} L_X(\mathbf{Q}:D) &= \prod_{d \in D} L(d \mid \mathbf{Q}) \\ &= \left(\prod_{\forall i} q_i^{M[i]} \exp\left(-q_i T[i]\right)\right)\left(\prod_{\forall i}\prod_{\forall j \neq i}\left(\frac{q_{i,j}}{q_i}\right)^{M[i,j]}\right) \end{aligned} \tag{2.12}$$

with $M[i] = \sum_{\forall j} M[i,j]$.

2.2.1.2 Marginalized Likelihood Function

X_1 and X_2 are independent homogenous Markov processes, with state space $Val(X_{1,2}) = \{0,1\}$. The transition intensity matrices Q_1 and Q_2 can be written in the following form for convenience,

$$\begin{pmatrix} -q_0 & q_0 \\ q_1 & -q_1 \end{pmatrix}$$

where the transition intensities q_0 and q_1 are gamma-distributed with parameters α_0, β_0 and α_1, β_1, respectively. The marginal likelihood of a sample trajectory from binary-valued homogenous Markov process X with transition intensity matrix Q can be written as follows:

$$P(X^{[0,T]}) = \int P(X^{[0,T]} \mid Q)P(Q)dQ \tag{2.13}$$

$$= \int_0^\infty \left(\prod_x \exp(-q_x T_x) \prod_{x'} q_{xx'}^{M[x,x']}\right) \frac{\beta_{xx'}^{\alpha_{xx'}}\, q_{xx'}^{\alpha_{xx'}-1}\, \exp(-\beta_{xx'} q_{xx'})}{\Gamma(\alpha_{xx'})}\, dq_{xx'} \tag{2.14}$$

$$= \prod_{x \in 0,1} \int_0^\infty q_x^{M_x}\ \exp(-q_x T_x)\ \frac{\beta_x^{\alpha_x}\ q_x^{\alpha_x - 1}\ \exp(-\beta_x q_x)}{\Gamma(\alpha_x)}\, dq_x \tag{2.15}$$

$$= \prod_{x \in 0,1} \frac{\beta_x^{\alpha_x}}{\Gamma(\alpha_x)} \int_0^\infty q_x^{M_x + \alpha_x - 1}\ \exp(-q_x(T_x + \beta_x))\, dq_x \tag{2.16}$$

$$= \prod_{x \in 0,1} \frac{\beta_x^{\alpha_x}}{\Gamma(\alpha_x)} \left(-(T_x + \beta_x)^{M_x + \alpha_x}\ \Gamma(M_x + \alpha_x,\ q_x(T_x + \beta_x))\right)\Big|_0^\infty \tag{2.17}$$

$$= \prod_{x \in 0,1} \frac{\beta_x^{\alpha_x}}{\Gamma(\alpha_x)} \left((T_x + \beta_x)^{M_x + \alpha_x}\ \Gamma(M_x + \alpha_x)\right) \tag{2.18}$$

where T_x, the amount of time spent in state x, $M[x,x']$ the number of transitions from state x to x' and $M[x] = \sum_{x \neq x'} M[x,x']$.

2.2.2 Time-inhomogeneous continuous-time Markov Processes

In an conventional CTBN, while every node is a Markov process itself, the leaf nodes are *conditional* Markov processes, a type of inhomogeneous Markov process, where the intensities change over time, but not as a function of time rather as a function of parent states. [2]

For inhomogeneous Markov processes, Eq.2.10 becomes:

$$f(t) = q_i(t) \exp\left(-\int_0^t q_i(u)du\right) \tag{2.19}$$

2.2.2.1 The Likelihood Function

Let X be an inhomogeneous Markov process, and $X^{[0,T]}$ is a trajectory sampled from this process. We define m number of transitions, with $0 = t_0 < t_1 < ... < t_m$ are the times where transition occurred, and $x_0, x_1, ..., x_m$ are the observed states. The likelihood of trajectory $X^{[0,T]}$ is as follows:

$$L(\mathbf{Q}_X : X^{[0,T]}) = \prod_{k=1}^{m} \left[q_{x_{k-1}}(t_k) \exp\left(-\int_{t_{k-1}}^{t_k} q_{x_{k-1}}(u)du\right) \frac{q_{x_{k-1},x_k}(t_k)}{q_{x_{k-1}}(t_k)} \right] \tag{2.20}$$

2.3 Belief State in Partially Observable Markov Decision Processes

In our problem, cell Z does not have a direct access to its parents' states, rather it observes a summary of them. This is different than observing the states through noisy binary channels, and presents a POMDP problem. Even though it is assumed that the policy of Z is optimal given the true states of parents, the problem of inferring the observational model remains unsolved and to be addressed.

2.3.1 Exact Belief State Update

In a conventional POMDP, given the *transition function*, $T(s,a,s')$ and *observation function*, $O(s',a,o)$, the belief state update is computed as follows [3] [1]:

$$\begin{aligned} b'(s') &= \Pr(s'|o,a,b) = \Pr(s'|o,b) \\ &= \frac{\Pr(o|s',b)\Pr(s'|b)}{\Pr(o|b)} \\ &= \frac{\Pr(o|s')\sum_{s\in\mathcal{S}}\Pr(s'|b,s)\Pr(s|b)}{\Pr(o|b)} \\ &= \frac{O(s',o)\sum_{s\in\mathcal{S}}T(s,s')\,b(s)}{\Pr(o|b)} \end{aligned} \tag{2.21}$$

In Eq.2.21, the transition function is time-independent. However, in our setting, the parent nodes X and Y are modelled as CTMCs with time-dependent transition matrices. Now we can derive the belief state as follows:

$$b(x_1,x_2;t) = P(X_1(t)=x_1, X_2(t)=x_2 \mid y_1,...,y_t) \tag{2.22}$$

Denote $b(t)$, $t \geq 0$, as row vector with $\{b(x_1,x_2;t)_{x_i\in\mathcal{X}_i}\}$. This posterior probability can be described by a system of ODEs

$$\frac{db(t)}{dt} = b(t)\ \mathbf{T} \tag{2.23}$$

where the initial condition $b(0)$ is row vector with $\{P(X_1(0)=x_1, X_2(0)=x_2)_{x_i\in\mathcal{X}_i}\}$ [4]. $\mathbf{T}$ is the joint transition intensity matrix of X_1 and X_2 and given by amalgamation operation between $\mathbf{Q}_1$ and $\mathbf{Q}_2$ [2].

$$\mathbf{T} = \mathbf{Q}_1 * \mathbf{Q}_2 \tag{2.24}$$

The belief update at discrete times of observation y_t

$$\begin{aligned} b(x_1,x_2;t) &= P(X_1(t)=x_1, X_2(t)=x_2 \mid y_1,...,y_t) \\ &= \frac{P(y_1,...,y_t, X_1(t)=x_1, X_2(t)=x_2)}{P(y_1,...,y_t)} \\ &= \frac{P(y_t \mid y_1,...,y_{t-1}, X_1(t)=x_1, X_2(t)=x_2)}{P(y_t \mid y_1,...,y_{t-1})}\,\frac{P(y_1,...,y_{t-1}, X_1(t)=x_1, X_2(t)=x_2)}{P(y_1,...,y_{t-1})} \\ &= Z_t^{-1}\ P(y_t \mid x_1,x_2)\ P(X_1(t)=x_1, X_2(t)=x_2 \mid y_1,...,y_{t-1}) \\ &= Z_t^{-1}\ P(y_t \mid x_1,x_2)\ b(x_1,x_2;t^-) \end{aligned} \tag{2.25}$$

where $Z_t = \sum_{x_1,x_2\in X} P(y_t \mid x_1,x_2)\ b(x_1,x_2;t^-)$ is the normalization factor [4].

[1] Since it is assumed that there is no affect of agent X_3's action on the observation or transition function, a is omitted from the equation.

2.3.2 Belief State Update using Particle Filter

2.3.2.1 Marginalized Continuous Time Bayesian Networks

2.3.2.2 Particle Filter

2.4 Likelihood Model of Communication System (?)

3 Simulation and Experiments

Environment with exact belief update and belief update using particle filter

3.1 Data Generation

3.1.1 Sampling Trajectories

3.1.1.1 Gillespie Algorithm

3.1.1.2 Thinning Algorithm

3.2 Inference of Deterministic Observation Model

Our dataset contains a number of trajectories from all the nodes involved in the communication. $\mathbf{D} = \{D_1, ..., D_N\}$. Every trajectory comprises of state transitions in time interval $[0, T]$, and the times of these transitions.

4 Experimental Results and Evaluation

4.1 Results

4.2 Evaluation

5 Conclusion

5.1 Discussion

5.2 Future Work

Bibliography

[1] I. Cohn, T. El-Hay, N. Friedman, and R. Kupferman, "Mean Field Variational Approximation for Continuous-Time Bayesian Networks," vol. 11, pp. 1–39, 2010.

[2] U. Nodelman, C. R. Shelton, and D. Koller, "Continuous Time Bayesian Networks," 1995.

[3] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Amino Acids*, vol. 40, no. 2, pp. 443–451, 2011.

[4] L. Huang, L. Paulevé, C. Zechner, M. Unger, A. Hansen, and H. Koeppl, "Supporting Information for Reconstructing dynamic molecular states from single-cell time series," 2016.