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

List of Symbols

χ	state space of random variable X
$X(t)$	value of random variable X at time t
$X^{[0,T]}$	discrete valued trajectory of random variable X in time interval $[0, T]$
$\mathbf{X}^T$	transpose of matrix/vector $\mathbf{X}$

List of Figures

1. Introduction

1.1. Motivation

1.2. Related Work

1.3. Contributions

1.4. Structure of the Thesis

2. Foundations

This chapter presents the theory applied in this thesis. First, the details of the communication problem is described briefly to put the theory into perspective, and then the mathematical theory of the frameworks used to model this problem is introduced.

2.1. Problem Formulation

The communication model considered in this thesis is given in Figure 2.1. The parent nodes, X_1 and X_2, emit messages which carry information about their states. These messages are translated by an observation model, ψ, and agent node, X_3 makes a decision based on this translated message, y. The main objective is to infer the observation model, given a set of trajectories of nodes.
The transition models of the nodes and the dependencies between them are modelled as continuous-time Bayesian network (CTBN), denoted by $\mathbf{X}$. The network $\mathbf{X}$ represents a stochastic process over a structured multivariate state space $\mathcal{X} = [\mathcal{X}_1, ..., \mathcal{X}_n]$.
The messages that are emitted by the parent nodes X_1 and X_2 are modelled as independent homogeneous continuous-time Markov processes $X_i(t)$, with state space $\mathcal{X}_i = \{x_1, x_2, ..., x_m\}$ for $i \in \{1, 2\}$.
The agent node X_3 does not have direct access to the messages but observes a translation of them. The observation model is defined as the likelihood of a translation given the parent messages.

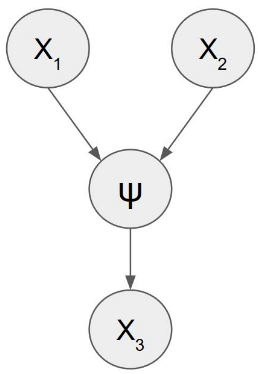

Figure 2.1.: Communication model.

$$\psi(x_1, x_2) := p(y(t) \mid X_1(t) = x_1, X_2(t) = x_2) \tag{2.1}$$

The agent X_3 is modelled as inhomogeneous continuous-time Markov process with state space $\mathcal{X}_3 = \{x_1, x_2, ..., x_m\}$ and set of actions $a \in \{a_0, a_1, ..., a_k\}$ to choose from.
Given the observation, the agent forms a belief over the parent states, $b(x_1, x_2; t)$, that summarizes the past observations. The policy of the agent, $\pi(a \mid b)$, is assumed to be shaped by evolution (close) to optimality. Based on the belief state, the agent takes an action, which in the setting described above corresponds to change its internal dynamics.

2.2. Continuous-Time Bayesian Networks

Consider a directed acyclic graph denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is a set of edges such that $\mathcal{E} = \{(m, n) : m, n \in \mathcal{V}\}$. In this graph, the parent nodes of node n is defined as the set of nodes that feed into it and denoted by $Par_{\mathcal{G}}(n) = \{m \in \mathcal{V} : (m, n) \in \mathcal{E}\}$. A directed acyclic graph is characterized as a Bayesian network where each node respresents a random variable such that $\mathcal{V} = \{X_1, X_2, ..., X_N\}$ and the joint distribution $p(X_1, X_2, ..., X_N)$ factors as

$$p(X_1, X_2, ..., X_N) = \prod_{i=1}^{N} p(X_i \mid Par_{\mathcal{G}}(X_i)). \tag{2.2}$$

A continuous-time Bayesian network (CTBN) is a Bayesian network with graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ that represents a collection of random variables whose values evolve continuously over time. In the CTBN framework, through a directed graph, the dependencies of a set of Markov processes (MPs) can be modelled efficiently, relying on two assumptions. The first assumption is that only one node can transition at a time. Secondly, the instantaneous dynamics of each node depends only on its parent nodes. [1, 2]

2.2.1. Continuous-Time Markov Processes

A continuous-time Markov process (CTMP) is a continuous-time stochastic process which satisfies Markov property, namely, the probability distribution over the states at a later time is conditionally independent of the past states given the current state.[1]
Consider a CTMP X(t) over a single variable with state space $\mathcal{X}$. Then the Markov property can be written as

$$\Pr\left(X(t_k) = x_{t_k} \mid X(t_{k-1}) = x_{t_{k-1}}, \dots, X(t_0) = x_{t_0}\right) = \Pr\left(X(t_k) = x_{t_k} \mid X(t_{k-1}) = x_{t_{k-1}}\right) \tag{2.3}$$

where X(t) denotes the state of the variable at time t such that $X(t) = x_t \in \mathcal{X}, t \geq 0$, and $t_0 < t_1 < ... < t_k$.
A CTMP is represented by its transition intensity matrix, $\mathbf{Q} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. In this matrix, the intensity q_i represents the instantaneous probability of leaving state x_i and $q_{i,j}$ represents the instantaneous probability of switching from state x_i to x_j, where $x_i, x_j \in \mathcal{X}$.

$$\mathbf{Q} = \begin{bmatrix} -q_1 & q_{1,2} & \cdots & q_{1,n} \\ q_{2,1} & -q_2 & \cdots & q_{2,n} \\ \vdots & \vdots & \ddots & \cdots \\ q_{n,1} & q_{n,2} & \cdots & -q_n \end{bmatrix} \tag{2.4}$$

where $q_i = \sum_{i \neq j} q_{i,j}$.[2]

2.2.1.1. Homogenous Continuous-Time Markov Processes

A continuous-time Markov process is time-homogenous when the transition intensities do not depend on time. Let X(t) be a homogenous CTMP, with state sppace $\mathcal{X}$ and transition intensity matrix $\mathbf{Q}$. Infinitesimal transition probability from state x_i to x_j in terms of the transition intensities $q_{i,j}$ can be written as [1]:

$$p_{i,j}(h) = \delta_{ij} + q_{i,j}h + o(h) \tag{2.5}$$

where $p_{i,j}(h) \equiv \Pr(X(t+h) = x_j \mid X(t) = x_i)$ are Markov transition functions, $\delta_{i,j} = \delta(x_i, x_j)$ is Kronecker delta and $o(\cdot)$ is a function decaying to zero faster than its argument such that $\lim_{h\to 0} \frac{o(h)}{h} = 0$.
The *Chapman-Kolmogorov-* or *master-equation* is then derived as follows:

$$\begin{aligned}
p_j(t) &= \Pr(X(t) = x_j) \\
&= \sum_{\forall i} p_{i,j}(h) p_i(t-h) \\
\lim_{h\to 0} p_j(t) &= \lim_{h\to 0} \sum_{\forall i} \left[\delta_{ij} + q_{i,j}h + o(h)\right] p_i(t-h) \\
&= \lim_{h\to 0} p_j(t-h) + \lim_{h\to 0} h \sum_{\forall i} q_{i,j} p_i(t-h) \\
\lim_{h\to 0} \frac{p_j(t) - p_j(t-h)}{h} &= \lim_{h\to 0} \sum_{\forall i} q_{i,j} p_i(t-h) \\
\frac{d}{dt} p_j(t) &= \sum_{\forall i} q_{i,j} p_i(t) \qquad (2.6)
\end{aligned}$$

Equation 2.6 can be written in matrix form:

$$\frac{d}{dt} p(t) = p(t)\mathbf{Q} \tag{2.7}$$

where the time-dependent probability distribution $p(t)$ is a row vector with entries $\{p_i(t)\}_{x_i \in \mathcal{X}}$. The solution of the system of ordinary differential equations (ODEs) is,

$$p(t) = p(0)\exp(t\mathbf{Q}) \tag{2.8}$$

with initial distribution $p(0)$.
The amount of time staying in a state x_i is exponentially distributed with parameter q_i. The probability density function f and cumulative distribution function F for staying in the state x_i [2]:

$$f(t) = q_i \exp(-q_i t), t \geq 0 \tag{2.9}$$
$$F(t) = 1 - \exp(-q_i t), t \geq 0 \tag{2.10}$$

Given the transitioning from state x_i, the probability of landing on state x_j is $q_{i,j}/q_i$.

Likelihood Function Consider a single transition denoted as $d = < x_i, x_j, t >$, where the transition occurs from state x_i to x_j after spending t amount of time at state x_i. The likelihood of this transition is the product of the probability of having remained at state x_i for duration t from Equation 2.9, and the probability of transitioning to x_j.

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

The likelihood of a trajectory sampled from a homogenous CTMC, denoted by $X^{[0,T]}$, can be decomposed as the product of the likelihood of single transitions. The sufficient statistics summarizing this trajectory can be written as $T[x_i]$, the total amount of time spent in state x_i, $M[x_i, x_j]$ the total number of transitions from state x_i to x_j.

$$M[x_i, x_j] = \sum_{d \in X^{[0,T]}} \mathbb{1}(X(t) = x_i)\mathbb{1}(X(t+h) = x_j) \tag{2.12}$$

$$T[x_i] = \sum_{d \in X^{[0,T]}} \mathbb{1}(X(t) = x_i) \tag{2.13}$$

where $\mathbb{1}(\cdot)$ is the indicator function. Then the likelihood of a trajectory $X^{[0,T]}$ can be written as:

$$\begin{aligned} p(X^{[0,T]} \mid \mathbf{Q}) &= \prod_{d \in X^{[0,T]}} p(d \mid \mathbf{Q}) \\ &= \left(\prod_i q_i^{M[x_i]} \exp\left(-q_i T[x_i]\right) \right) \left(\prod_i \prod_{j \neq i} \left(\frac{q_{i,j}}{q_i} \right)^{M[x_i, x_j]} \right) \\ &= \prod_{j \neq i} \exp(-q_{i,j} T[x_i]) \; q_{i,j}^{M[x_i, x_j]} \end{aligned} \tag{2.14}$$

where $M[x_i] = \sum_{j \neq i} M[x_i, x_j]$ is the total number transitions leaving state x_i.

2.2.1.2. Conditional Markov Processes

A continuous-time Markov process is *time-inhomogenous* when the transition intensities change over time. In a CTBN, while every node is a Markov process, the leaf nodes are characterized as *conditional* Markov processes, a type of inhomogeneous MP, where the intensities change over time, but not as a function of time rather as a function of parent states. [2]
Let X be a conditional Markov process in a graph $\mathcal{G}$, with a set of parents $U = Par_{\mathcal{G}}(X)$. Its *conditional intensity matrix* (CIM) $\mathbf{Q}_{X \mid U}$ can be viewed as a set of homogenous intensity matrices $\mathbf{Q}_{X \mid u}$, with entries $q_{i,j|u}$ (similar to Equation 2.4), for each instantiation of parent nodes $U(t) = u$ such that $u \in \mathcal{U} = \bigtimes_{X_m \in Par_{\mathcal{G}(X)}} \chi_m$, where $\bigtimes$ denotes Cartesian product.[2]
As a result, given a trajectory of parent nodes, X has a trajectory of intensity matrix as

$$\mathbf{Q}^{[0,T]} = [\mathbf{Q}_{X \mid U(t_0)}, \mathbf{Q}_{X \mid U(t_1)}, ..., \mathbf{Q}_{X \mid U(t_N)}], \; 0 < t_0 < ... < t_N \leq T. \tag{2.15}$$

Markov transition function for a conditional MP can be written as follows:

$$\Pr(X(t+h) = x_j \mid X(t) = x_i, U(t) = u, \mathbf{Q}_{X \mid u}) = \delta(i,j) + q_{i,j|u} h + o(h) \tag{2.16}$$

Likelihood Function Given the instantiation of its parents, the complete information on the dynamics of X is obtained. Then the likelihood of a trajectory drawn from a conditional MP

X can be written similar to Equation 2.14,

$$p(X^{[0,T]} \mid \mathbf{Q}_{X\,|\,U}) = \left(\prod_{u}\prod_{i} q_{i|u}^{M[x_i|u]} \exp\left(-q_{i|u}T[x_i \mid u]\right)\right) \left(\prod_{u}\prod_{i}\prod_{j\neq i}\left(\frac{q_{i,j|u}}{q_{i|u}}\right)^{M[x_i,x_j|u]}\right)$$
$$= \prod_{u}\prod_{j\neq i} \exp(-q_{i,j|u}T[x_i \mid u])\; q_{i,j|u}^{M[x_i,x_j|u]} \tag{2.17}$$

with the sufficient statistics introduced in Section 2.2.1.1 are also conditioned on parent nodes.

2.2.2. The CTBN Model

Evidently, a homogenouos CTMP can be considered as a conditional MP whose set of parents is empty. Thus, a CTBN can be formed as a set of conditional Markov processes.
Let S be a CTBN with graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ local variables $\mathcal{V} = \{X_1, X_2, ..., X_N\}$, each with a state space $\mathcal{X}_n$. This results in a factorizing state space for S such that $\mathcal{S} = [\mathcal{X}_1, \mathcal{X}_2, ..., \mathcal{X}_N]$. The joint states of the variables are denoted by $s = (x_1, x_2, ..., x_N) \in \mathcal{S}$ where $x_1 \in \mathcal{X}_1, ..., x_N \in \mathcal{X}_N$. The dependencies of each variable are defined as a set of its parents $U_n = Par_{\mathcal{G}}(X_n)$ with values $U_n(t) = u_n$ such that $u_n \in \mathcal{U}_n = \bigtimes_{X_m \in Par_{\mathcal{G}(X_n)}} \mathcal{X}_m$. In the following, the set of all conditional transition intensity matrices are denoted as $\boldsymbol{Q} = \left\{Q_{X_1\,|\,U_1}, ..., Q_{X_N\,|\,U_N}\right\}$.
Consider a trajectory drawn from S, such that $S^{[0,T]} = \left\{X_1^{[0,T]}, X_2^{[0,T]}, ..., X_N^{[0,T]}\right\}$. Following Equation 2.17, the likelihood of this trajectory can be written as follows.

$$p(S \mid \boldsymbol{Q}) = \prod_{n=1}^{N}\prod_{u\in\mathcal{U}_n}\prod_{x_i\in\mathcal{X}_n}\prod_{x_j\in\mathcal{X}_n\setminus x_i} \exp\left(q_{i,j|u}^{n} T_n[x_i \mid u]\right)(q_{i,j|u}^{n})^{M_n[x_i,x_j|u]} \tag{2.18}$$

where $T_n[\cdot]$ and $M_n[\cdot]$ indicates the sufficient statistics for X_n.
It should be noted that a CTBN can also be represented by a single conditional intensity matrix, through *amalgamation* operation. [2]

2.3. Belief State in Partially Observable Markov Decision Processes

Partially observable Markov decision process (POMDP) framework provides a model of an agent which interacts with its environment but is unable to obtain certain information about its state. Instead, the agent gets an observation which is a function of the true state, e.g. noisy observations, translation. The main goal, similar to Markov decision processes (MDPs), is to learn a policy solving a task by optimizing a reward function. The problem of decision making under uncertainty can be decomposed into two parts for the agent. The first is to keep a belief state which summarizes past experiences, and the second is to optimize a policy which will give an action based on the belief state. [3, 4]
The belief state, if represented as a probability distribution over states, provides a summary over the agent's past experiences.
In the problem considered in this thesis, the agent node X_3 cannot observe the incoming

messages directly, rather a summary of them. This presents a POMDP problem. However, since the optimal policy of the agent is assumed to be given, the theory for policy optimization is skipped.
In the following, update methods for belief state are introduced, where belief state refers to the posterior probability distribution over the environment states.

2.3.1. Exact Belief State Update

Consider a POMDP problem, with discrete state space S, action space A, observation space Ω. In a scenario where a compact representation of the *transition model*, $T(s,a,s')$, and *observation model*, $O(s',a,o)$, is available, the belief state update can be obtained as [3]:

$$\begin{aligned} b'(s') &= \Pr(s' \mid o,a,b) &(2.19)\\ &= \frac{\Pr(o \mid s',a,b)\Pr(s' \mid a,b)}{\Pr(o \mid a,b)} &(2.20)\\ &= \frac{\Pr(o \mid s',a)\sum_{s\in\mathcal{S}}\Pr(s' \mid a,b,s)\Pr(s \mid a,b)}{\Pr(o \mid a,b)} &(2.21)\\ &= \frac{O(s',a,o)\sum_{s\in\mathcal{S}}T(s,a,s')\,b(s)}{\Pr(o \mid a,b)} &(2.22)\end{aligned}$$

For the derivation above, from Equation 2.19 to Equation 2.20, Bayes' theorem, and from Equation 2.20 to Equation 2.21, law of total probability is applied. It should also be noted that the denominator in Equation 2.22 is in the following form,

$$\Pr(o \mid a,b) = \sum_{\forall s'\in\mathcal{S}} O(o \mid s',a) \sum_{\forall s\in\mathcal{S}} T(s' \mid s,a)b(s) \tag{2.23}$$

which is computationally expensive in the case of continuous state space.

2.3.2. Filtering for CTMP

Equation 2.22 is discrete-time solution of belief state. However, since in the model described in Section 2.1, the parent nodes are modelled as CTMPs, thus the environment state for the agent is the state of a CTMP, the belief state should be solved in continuous-time. This is achieved by the inference of the posterior probability of CTMP. [5]
Filtering problem in statistical context, as opposed to deterministic digital filtering, refers to the inference of the conditional probability of the true state of the system at some point in time, given the history of observations. [6]
Let X be a CTMP with transition intensity matrix $\mathbf{Q}$. Assume discrete-time observations denoted by $y_1 = y(t_1), ..., y_N = y(t_N)$. The belief state can be written as:

$$b(x_i;t_N) = \Pr(X(t_N) = x_i \mid y_1, ..., y_N) \tag{2.24}$$

From the master equation given in Equation 2.6, it follows that:

$$\frac{d}{dt}b(x_j;t) = \sum_{\forall i} q_{i,j} \cdot b(x_i;t) \tag{2.25}$$

The time-dependent belief state $b(t)$ is a row vector with $\{b(x_i;t)\}_{x_i \in \mathcal{X}}$. This posterior probability can be described by a system of ODEs:

$$\frac{db(t)}{dt} = b(t)\mathbf{Q} \tag{2.26}$$

where the initial condition $b(0)$ is row vector with $\{b(x_i;t)\}_{x_i \in \mathcal{X}}$ [5]. The solution to this ODE is

$$b(t) = b(0)\exp(t\mathbf{Q}). \tag{2.27}$$

The belief state update at discrete times of observation y_t is derived as

$$\begin{aligned}
b(x_i;t_N) &= \Pr(X(t_N) = x_i, \mid y_1, ..., y_N) \\
&= \frac{\Pr(y_1, ..., y_N, X(t_N) = x_i)}{\Pr(y_1, ..., y_N)} \\
&= \frac{\Pr(y_N \mid y_1, ..., y_{N-1}, X(t_N) = x_i)}{\Pr(y_N \mid y_1, ..., y_{N-1})} \frac{\Pr(y_1, ..., y_{N-1}, X(t_N) = x_i)}{\Pr(y_1, ..., y_{N-1})} \\
&= Z_N^{-1} \; \Pr(y_N \mid X(t_N) = x_i) \; \Pr(X(t_N) = x_i \mid y_1, ..., y_{N-1}) \\
&= Z_N^{-1} \; p(y_N \mid x_i) \; b(x_i; t_N^-)
\end{aligned} \tag{2.28}$$

where $Z_N = \sum_{x_i \in \mathcal{X}} p(y_N \mid x_i) \; b(x_i; t_N^-)$ is the normalization factor [5].

2.3.3. Belief State Update using Particle Filter

In a more realistic scenario, the exact update of belief state may not be feasible for several reasons. The computation of Bayes belief update is expensive for large state spaces as can be seen from Equation 2.23. Moreover, a problem with continuous state spaces requires a belief state represented as probability distributions over infinite state space rather than a collection of probabilities as given in Section 2.3.1.[7] Another reason could be the lack of a compact representation of transition or observation models. Under such circumstances, the belief state is obtained using sample-based approximation methods.[7]
It should be noted that since the belief state is nothing but the conditional probability of true states given the observations, the problem at hand poses a filtering problem as described in Section 2.3.2.

2.3.3.1. Particle Filtering

Particle filtering is one of the most commonly used Sequential Monte Carlo (SMC) algorithms. The popularity of this method thrives from the fact that, unlike other approximation methods such as Kalman Filter, it does not assume a linear Gaussian model. This advantage offers great flexibility and finds application in a wide range of areas.[8]
The key idea in particle filtering is to approximate a target distribution $p(x)$ by a set of samples, i.e. particles, drawn from that distribution. This is achieved by sequentially updating the particles through two steps. The first step is *importance sampling.* Since the target

distribution is not available, the particles are generated from a *proposal distribution* $q(x)$ and weighted according to the difference between target and proposal distributions. The second step is to resample the particles using these weights with replacement.[6]
Consider a problem to calculate the expectation $\widehat{f(x)} = \mathbb{E}[f(x)] = \int f(x)p(x)dx$, and suppose p(x) is an intractable density function from which the particles cannot be sampled. Instead they are drawn from a proposal distribution q(x), which yields an empirical approximation such that

$$x^{(i)} \sim q(x)$$
$$q(x) \approx \frac{1}{N}\sum_{i=1}^{N} \delta_{x^{(i)}}(x)$$

where $\delta_{x^{(i)}}(x)$ is Dirac delta. The expectation in question can be written as

$$\int f(x)p(x)dx = \int f(x)\frac{p(x)}{q(x)}q(x)dx$$
$$\int f(x)\frac{p(x)}{q(x)}\left(\frac{1}{N}\sum_{i=1}^{N} \delta_{x^{(i)}}(x)\right)dx = \frac{1}{N}\sum_{i=1}^{N}\frac{p(x^{(i)})}{q(x^{(i)})}f(x^{(i)})$$

where $w(x^{(i)}) = \frac{p(x^{(i)})}{q(x^{(i)})}$ is defined as *importance weight* of a particle. Then the particles are resampled using the importance weights with replacement, which concludes one iteration of sequential updating.[6]
In this application, the particles to represent the belief state are drawn from marginalized CTBN. The algorithm for belief state update through particle filtering and marginal process is given in the following chapter.

2.3.3.2. Marginalized Continuous-Time Markov Processes

Let $\mathbf{X}$ be a CTBN with local variables X_n, $n \in \{1, ..., N\}$, and set of conditional intensity matrices $\boldsymbol{Q}$. In the following, it is assumed that every non-diagonal entry in $\mathbf{Q}_n \mid \mathbf{u}$ is Gamma distributed with shape and rate parameters, $\alpha^n_{i,j|\mathbf{u}}$ and $\beta^n_{i,j|\mathbf{u}}$.
The marginal process description of $\mathbf{X}$ considering a single trajectory in the interval $[0, t)$ is given as follows:

$$\begin{aligned} &\Pr(X_n(t+h) = x_j \mid X_n(t) = x_i, \mathbf{U}_n(t) = \mathbf{u}, \mathbf{X}^{[0,t)}) \\ &\quad = \int \Pr(X_n(t+h) = x_j \mid X_n(t) = x_i, \mathbf{U}_n(t) = \mathbf{u}, Q_{n|\mathbf{u}}, \mathbf{X}^{[0,t)})p(Q_{n|\mathbf{u}})dQ_{n|\mathbf{u}} \\ &\quad = \delta_{i,j} + \mathbb{E}[q^n_{i,j|\mathbf{u}} \mid \mathbf{X}^{[0,t]} = \mathbf{x}^{[0,t]}]\ h + o(h), \end{aligned} \tag{2.29}$$

By integrating out the intensity matrix $Q_{n|\mathbf{u}}$, the parameter is replaced by its expected value given the history of the process. It should be noted that by doing so, the process becomes parameter-free, and thus self-exciting.
The derivation of the conditional expectation for marginal CTBN follows from the Bayes' rule:

$$p\left(\mathbf{Q} \mid \mathbf{X}_{[0,t]}\right) = \frac{p\left(\mathbf{X}_{[0,t]} \mid \mathbf{Q}\right)p(\mathbf{Q})}{p\left(\mathbf{X}_{[0,t]}\right)} \tag{2.30}$$

Equation 2.30, written for single trajectory $\mathbf{X}_{[0,t]}$, can be extended for multiple trajectories. Consider K trajectories drawn from CTBN $\mathbf{X}$, denoted by $\xi_t = \left\{\mathbf{X}^{[0,t],1}, \mathbf{X}^{[0,t],2}, ..., \mathbf{X}^{[0,t],K}\right\}$. Since the trajectories are conditionally independent, given $\mathbf{Q}$, using Equation 2.18 the likelihood of set ξ_t is written as,

$$\Pr(\xi_t \mid \boldsymbol{Q}) = \prod_{n=1}^{N} \prod_{\mathbf{u} \in \mathbf{U}_n} \prod_{x_i \in \mathcal{X}_n} \prod_{x_j \in \mathcal{X}_n \setminus x_i} \exp\left(q^n_{i,j|\mathbf{u}} T_n[x_i \mid \mathbf{u}]\right) (q^n_{i,j|\mathbf{u}})^{M_n[x_i, x_j|\mathbf{u}]} \tag{2.31}$$

where the joint sufficient statistics of X_n over all K trajectories are denoted by $T_n[x_i \mid \mathbf{u}] = \sum_{k=1}^{K} T_n^k[x_i \mid \mathbf{u}]$ and $M_n[x_i, x_j \mid \mathbf{u}] = \sum_{k=1}^{K} M_n^k[x_i, x_j \mid \mathbf{u}]$.
Given independent Gamma-priors on transition intensities, the expectation in Equation 2.29 can be evaluated as follows:

$$\mathbb{E}\left[q^n_{i,j|\mathbf{u}} \mid \xi_t\right] = \frac{\alpha^n_{i,j|\mathbf{u}} + M_n[x_i, x_j \mid \mathbf{u}]}{\beta^n_{i,j|\mathbf{u}} + T_n[x_i \mid \mathbf{u}]} \tag{2.32}$$

2.4. Sampling Algorithms

2.4.1. Gillespie Algorithm for Generative CTBN

Gillespie algorithm is a computer-oriented Monte Carlo simulation procedure that is originally proposed to simulate the reactions of molecules in any spatially homogeneous chemical system. Such systems are regarded as Markov processes and represented via their master equations, which cannot be directly used to obtain realizations of the process. Gillespie algorithm is an efficient tool to overcome this problem. [9]
This algorithm can also be applied to sample *events* from a CTBN given the transition intensity matrices, where an event refers to a transition occurring at a specific point in time. This procedure is introduced as *Generative CTBN* in [2].

Algorithm 1: Generative CTBN

Input : Structure of the network with N local variables $X_1, X_2, ..., X_n$ with state-space $\mathcal{X}_n = \{x_1, ..., x_m\}$
Transition intensity matrices $\mathbf{Q}_n$ with entries $q^n_{i,j}$
T_{max} to terminate simulation
Output : Sample trajectory of the network
Initialize: Initialize node values $X_n(0) = x_i \in \mathcal{X}_n$

while $t < T_{max}$ **do**
    $\tau \sim \exp(\sum_{\forall n} \sum_{\forall i \neq j} q^n_{i,j})$
    transitioning node is randomly drawn with probability $P(X_n) = \frac{q^n_i}{\sum_{\forall n} q^n_i}$
    next state is randomly drawn with probability $P(x_j) = \frac{q^n_{i,j}}{q^n_i}$
    $t \leftarrow t + \tau$
end while

2.4.2. Thinning Algorithm

Thinning algorithm is a method introduced to simulate nonhomogenous Poisson processes. [10] Later, it is adapted to sample from Hawkes processes, a self-exciting process with time-dependent intensity function. [11, 12] This algorithm is used here to simulate inhomogeneous Markov process.

Algorithm 2: Thinning Algorithm

Input : $\lambda(t)$ the intensity function of the inhomogenous process
N number of events to terminate simulation
Output : Sample trajectory of the process
Initialize: Time $t = 0$

```
while i < N do
    the upper bound for intensity, λ*
    transition time τ drawn by  u ~ U(0,1)  and  τ = -ln(u)/λ*
    t ← t + τ
    draw  s ~ U(0,1)
    if  s ≤ λ(t)/λ*  then
        sample accepted and  t_i = t, i = i + 1
    end if
end while
```

3. Experimental Setup

This chapter presents the methodology used in this thesis. First, it is explained how different frameworks introduced in Chapter 2 are put into use. Then, the algorithms used in data generation and inference are given in detail. The results from these experiments are presented in the succeeding chapter.

3.1. The Model

A detailed graphical model explored in this thesis is given in the Figure 3.1. This model presents an intersection of continuous-time Bayesian network and partially observable Markov decision process frameworks.

- The transition models of the nodes X_1, X_2 and X_3, and the dependencies between them are modelled as CTBN.
- The interaction of agent node X_3 and its environment is modelled as POMDP.

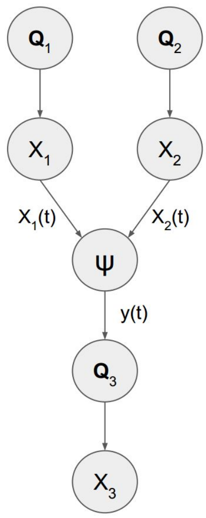

Figure 3.1.: Hierarchical model.

3.1.1. CTBN Model

The transition models of the nodes and the dependencies between them are modelled as continuous-time Bayesian network (CTBN), denoted by S with graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{X_1, X_2, X_3\}$ and $\mathcal{E} = \{(X_1, X_3), (X_2, X_3)\}$. The network S represents a stohastic process over a structured factorising state space $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3$.
The parent nodes X_1 and X_2 emit their states as messages. The dynamics of these nodes are modelled as independent homogeneous continuous-time Markov processes $X_i(t)$, with binary-valued states $\mathcal{X}_i = \{0, 1\}$ for $i \in \{1, 2\}$. These processes are defined by transition intensity matrices $\mathbf{Q}_i$, which are assumed to be Gamma distributed with shape and rate parameters $\alpha = [\alpha_0, \alpha_1]$ and $\beta = [\beta_0, \beta_1]$, re-

spectively, and are in the following forms.

$$\mathbf{Q}_i = \begin{bmatrix} -q_0^i & q_0^i \\ q_1^i & -q_1^i \end{bmatrix} \tag{3.1}$$

$$\mathbf{Q}_i \sim Gam(\alpha^i, \beta^i) \quad for\ i \in \{1, 2\} \tag{3.2}$$

It should be noted that in Equation 3.1, the suffixes are simplified using the fact that $q_i = \sum_{i \neq j} q_{i,j}$.
The agent X_3 is modelled as inhomogenouos continuous-time Markov process with binary states $\mathcal{X}_3 = \{0, 1\}$ and set of actions $a \in \{a_0, a_1\}$, and set of transition intensity matrices which contains one matrix corresponding to each action, $\boldsymbol{Q}_{3|a} = \left\{\mathbf{Q}_{3|a_0}, \mathbf{Q}_{3|a_1}\right\}$.
The dependencies are represented by set of parents for each node $U_n = Par_{\mathcal{G}}(X_n)$ and for the model shown in Figure 3.1 can be written as follows:

$$U_1, U_2 = \emptyset$$
$$U_3 = \{X_1, X_2\}$$

In order to have a compact representation of parent messages, a subsystem of S consisting of only the parent nodes, X_1 and X_2 can be considered as a single system. These two processes can be represented as a *joint* process, X_P, with factorising state space $\mathcal{X}_P = \mathcal{X}_1 \times \mathcal{X}_2$. The transition intensity matrix of the new joint system, $\mathbf{Q}_P$ is obtained by amalgamation operation between $\mathbf{Q}_1$ and $\mathbf{Q}_2$ (see Appendix A) [2].

$$\mathbf{Q}_P = \mathbf{Q}_1 * \mathbf{Q}_2 \tag{3.3}$$

3.1.2. POMDP Model

In a conventional POMDP scenario, there are two problems to be addressed, one is belief state update and the other is policy optimization. As mentioned in Section 2.3, in the problem at hand, the policy of agent X_3 is assumed to be optimal and given. Thus, the POMDP model of the agent only consists of belief state update. A detailed view of the agents interaction with its environment from POMDP framework perspective is given in the Figure 3.2.

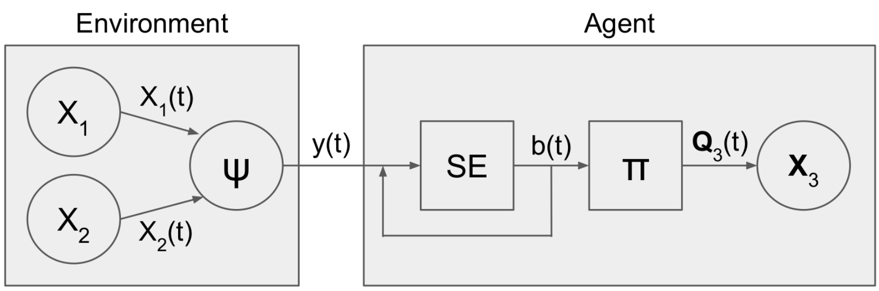

Figure 3.2.: Closer look to agent-environment interaction from the perspective of POMDP framework.

It should be noted that, the interaction in Figure 3.2 is only one-sided, the state or action of the agent does not affect the environment.

3.1.2.1. Observation Model

The messages sent by the parent nodes are translated by the observation model. The agent node X_3 does not have a direct access to the messages, but observes a translation of them. The observation is denoted by $y(t) = y_t$ such that $y_t \in \mathcal{Y}$ where $\mathcal{Y}$ is the observation space. The observation model defines a probability distribution over the observation for each combination of parent messages.

$$\psi(x_1, x_2) = p(y(t) \mid X_1(t) = x_1, X_2(t) = x_2) \tag{3.4}$$

where $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$. As explained in Section 3.1.1, using the joint process X_P for the sake of conciseness, Equation 3.4 can be written as

$$\psi(x_p) = p(y(t) \mid X_P(t) = x_p) \tag{3.5}$$

where $x_p \in \mathcal{X}_P$.
$\psi(x_p)$ is defined as categorical distribution over the observation space $\mathcal{Y}$. ψ denotes the matrix with rows $\left\{\psi(x_p)\right\}_{x_p \in \mathcal{X}_P}$.

3.1.2.2. Belief State

The belief state provides a summary over agents past experiences and allows the agent to take its own uncertainty into account. The belief state is formed by *state estimator* (labelled as *SE* in Figure 3.2) over the parent states, denoted by $b(x_1, x_2; t)$.

$$b(x_p; t) = \Pr(X_P(t) = x_p \mid y_1, ..., y_t) \tag{3.6}$$

Exact Belief State Update As discussed in Section 2.3.2, given the transition intensity matrices of parent nodes, $\mathbf{Q}_1$ and $\mathbf{Q}_2$, the continuous-time belief state update poses a filtering problem for CTMPs. This problem can be formulated according to the joint process of parents.

$$b(x_p; t) = \Pr(X_P(t) = x_p \mid y_1, ..., y_t) \tag{3.7}$$

Consider discrete-time observations from this process, denoted by $y_1 = y(t_1), ..., y_N = y(t_N)$ and time-dependent belief state $b(t)$ as a row vector with $\left\{b(x_p; t)\right\}_{x_p \in \mathcal{X}_P}$. Following Equation 2.27 and Equation 2.28, the belief state update is evaluated as

$$b(t) = b(0) \exp(t\mathbf{Q}_P) \tag{3.8}$$

with the initial condition $b(0)$. The update at discrete times of observation y_t is

$$b(x_p; t_N) = Z_N^{-1} \Pr(y_N \mid X_P(t_N) = x_p)\ b(x_p; t_N^-) \tag{3.9}$$

$$= Z_N^{-1}\ \psi(x_p)\ b(x_p; t_N^-) \tag{3.10}$$

where $Z_N = \sum_{x_p \in \mathcal{X}_P} \psi(x_p)\ b(x_p; t_N^-)$ is the normalization factor.

Belief State Update Using Particle Filter The assumption that full information of parent dynamics being available is unrealistic. In an environment as described above, the agent most probably doesn't have access to the parameters $\mathbf{Q}_1$ and $\mathbf{Q}_2$, rather may have some prior beliefs over them. Thus, in order to simulate a more realistic model and be able to marginalize out these parameters from inference problem, the joint process (introduced in previous section) is replaced with its marginalized counterpart. Using the Gamma-priors over $\mathbf{Q}_1$ and $\mathbf{Q}_2$ (Equation 3.2) and sufficient statistics over the particle history, the particles are drawn from this marginalized process. With every new observation, the particles are propogated through the process, while the sufficient statistics are updated and the parameters are re-estimated after each particle using the Equation 2.32. The belief state then obtained as the distribution of states over the particles,

$$b(x_p;t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{p_i(t),x_p} \tag{3.11}$$

where N is the number of particles, $p_i \in \mathbf{p}$ is the set of particles, and δ is the Kronecker delta.

Algorithm 3: Marginal particle filter for belief state update

Input : Observation y_k at time t_k, set of particles $\mathbf{p}^{k-1}$, estimated $\hat{Q}$
Output: New set of particles $\mathbf{p}^k$, $\mathbf{b}^{[t_{k-1},t_k]}$

for $p_m \in \mathbf{p}^{k-1}$ **do**
    $p_m = \left\{x_m, \hat{Q}\right\} \leftarrow$ *Propagate particle through marginal process from* t_{k-1} *to* t_k
    $\hat{Q} \leftarrow$ *sufficient statistics added from* $p_m[t_{k-1}, t_k]$
    `// observation likelihood assigned as particle weight`
    $w_m \leftarrow p(y_k \mid X_P(t_k) = x_m)$
end for
`// belief state from` t_{k-1} `to` t_k
$\mathbf{b}^{[t_{k-1},t_k]} \leftarrow \left\{\frac{1}{N}\sum_{i=1}^{N} \delta_{p_i^{[t_{k-1},t_k]},x_p}\right\}_{x_p \in \mathcal{X}_P}$
`// normalize weights`
$w_m \leftarrow \frac{w_m}{\sum_m w_m}$
`// resample particles`
for $p_m \in \mathbf{p}_k$ **do**
    $p_m \leftarrow$ *Sample from* p_k *with probabilities* w_m *with replacement*
end for

3.1.2.3. Optimal Policy

The optimal policy is defined using a polynomial function of belief state.

$$\pi(b) = \begin{cases} a_0 & \text{if } \mathbf{w}b^\top > 0.5 \\ a_1 & \text{otherwise} \end{cases} \tag{3.12}$$

where $\mathbf{w}$ is a row vector of weights.
Given the optimal policy, $\pi(b)$, the agent takes an action based on the belief state. In the setting

described above, taking an action means to change its internal dynamics to the transition intensity matrix corresponding to that action.

$$a(t) = \pi(b(t)) \tag{3.13}$$

$$\mathbf{Q}_3(t) = \begin{cases} \mathbf{Q}_{3|a_0} & \text{if } a(t) = a_0 \\ \mathbf{Q}_{3|a_1} & \text{otherwise} \end{cases} \tag{3.14}$$

3.2. Inference of Observation Model

Inference problem is considered for deterministic observation models. Considering the number of states of parents and the observations, there are a number of possible observation models. Given the parent trajectories, the belief state and the resulting $\mathbf{Q}_3$ trajectory is computed for each observation model. Then the likelihood of X_3 trajectory given these $\mathbf{Q}_3$ trajectories are compared for maximum likelihood estimation.

$$\hat{\psi} = argmax \Pr(X_3^{[0,T]} \mid Q_3^{[0,T]}) \tag{3.15}$$

3.2.1. Likelihood Model

Consider a trajectory in the dataset, denoted by $\mathbf{X}^{[0,T]} = \left\{X_1^{[0,T]}, X_2^{[0,T]}, X_3^{[0,T]}\right\}$. The set of parameters to the system, as introduced before, is written as $\theta = \left\{\mathbf{Q}_1, \mathbf{Q}_2, \pi, \psi\right\}$. Then likelihood of the sample trajectory $\mathbf{X}^{[0,T]}$ can be written as:

$$\begin{aligned} p(\mathbf{X}^{[0,T]} \mid \theta) &= p(X_1^{[0,T]}, X_2^{[0,T]}, X_3^{[0,T]} \mid \mathbf{Q}_1, \mathbf{Q}_2, \pi, \psi) \\ &= p(X_3^{[0,T]} \mid X_1^{[0,T]}, X_2^{[0,T]}, \mathbf{Q}_1, \mathbf{Q}_2, \pi, \psi)\ p(X_1^{[0,T]} \mid \mathbf{Q}_1)\ p(X_2^{[0,T]} \mid \mathbf{Q}_2) \\ &= p(X_3^{[0,T]} \mid X_1^{[0,T]}, X_2^{[0,T]}, \pi, \psi)\ p(X_1^{[0,T]} \mid \mathbf{Q}_1)\ p(X_2^{[0,T]} \mid \mathbf{Q}_2) \\ &= p(X_3^{[0,T]} \mid \mathbf{Q}_3^{[0,T]})\ p(X_1^{[0,T]} \mid \mathbf{Q}_1)\ p(X_2^{[0,T]} \mid \mathbf{Q}_2) \end{aligned} \tag{3.16}$$

As mentioned before, it is plausible to marginalize out the parameters $\mathbf{Q}_1$ and $\mathbf{Q}_2$, for a more realistic model and inference. Noting that in case the belief state is updated using filtering of CTMPs (See Section 3.1.2.2), $\mathbf{Q}_3^{[0,T]}$ becomes a deterministic function of all the parameters including $\mathbf{Q}_1$ and $\mathbf{Q}_2$, the marginalization cannot be carried out analytically on Equation 3.16. On the other hand, particle filtering removes this dependency on $\mathbf{Q}_1$ and $\mathbf{Q}_2$ by using marginalized counterpart of CTMPs (See Section 3.1.2.2), leaving it straightforward to marginalize out the parameters on Equation 3.16.

Marginalizing the likelihood over Q_1 and Q_2:

$$\begin{aligned}
p(\mathbf{X}^{[0,T]} \mid \pi, \psi) &= \int\int p(\mathbf{X}^{[0,T]} \mid \theta)\ p(\mathbf{Q}_1)\ p(\mathbf{Q}_2)\ d\mathbf{Q}_1 d\mathbf{Q}_2 \\
&= \int\int p(X_3^{[0,T]} \mid \mathbf{Q}_3^{[0,T]})\ p(X_1^{[0,T]} \mid \mathbf{Q}_1)\ p(X_2^{[0,T]} \mid \mathbf{Q}_2)\ p(\mathbf{Q}_1)\ p(\mathbf{Q}_2)\ d\mathbf{Q}_1 d\mathbf{Q}_2 \\
&= p(X_3^{[0,T]} \mid \mathbf{Q}_3^{[0,T]}) \int p(X_1^{[0,T]} \mid \mathbf{Q}_1)\ p(\mathbf{Q}_1)\ d\mathbf{Q}_1 \int p(X_2^{[0,T]} \mid \mathbf{Q}_2)\ p(\mathbf{Q}_2)\ d\mathbf{Q}_2
\end{aligned} \tag{3.17}$$

Marginalized likelihood function for binary-valued homogenous CTMP is derived in Appendix B.

Plugging Equation B.3 in Equation 3.17 for both X_1 and X_2:

$$\begin{aligned}
p(\mathbf{X}^{[0,T]} \mid \pi, \Phi) = p(X_3^{[0,T]} \mid Q_3^{[0,T]}) &\prod_{x_1 \in 0,1} \frac{\beta_{x_1}^{\alpha_{x_1}}}{\Gamma(\alpha_{x_1})}\ (T_{x_1} + \beta_{x_1})^{M_{x_1}+\alpha_{x_1}}\ \Gamma(M_{x_1} + \alpha_{x_1}) \\
&\prod_{x_2 \in 0,1} \frac{\beta_{x_2}^{\alpha_{x_2}}}{\Gamma(\alpha_{x_2})}\ (T_{x_2} + \beta_{x_2})^{M_{x_2}+\alpha_{x_2}}\ \Gamma(M_{x_2} + \alpha_{x_2})
\end{aligned} \tag{3.18}$$

3.3. Data Generation

The dataset contains a number of trajectories drawn from CTBN $\mathbf{X}$. Following the notation in Chapter 2, K trajectories in time interval $[0, T]$ are denoted by $\xi_T = \left\{\mathbf{X}^{[0,T],1}, \mathbf{X}^{[0,T],2}, ..., \mathbf{X}^{[0,T],K}\right\}$, where $\mathbf{X}^{[0,T],k} = \left\{X_1^{[0,T],k}, X_2^{[0,T],k}, X_3^{[0,T],k}\right\}$ denotes a single trajectory for all nodes. Every trajectory comprises of state transitions in the interval, and the times of these transitions.

3.3.1. Sampling Algorithm

In order to sample trajectories from CTBN, two sampling algortihms introduced in Section 2.4 are combined. Gillespie algorithm is used to sample from the parent nodes, X_1 and X_2, while thinning algorithm is applied to overcome the challenges that come with conditional intensity matrix of the agent, X_3. It should be noted that Algorithm 1 is applicable to any nodes in a CTBN, both homogenous and conditional MPs. However, since in this setting, the intensity matrix is conditioned on the belief state and the policy, instead of directly on the parent states, a more general algorithm suitable for inhomogenous MPs, thinning algorithm, is preferred. Algorithm 4 describes the procedure to draw samples using particle filtering.

Algorithm 4: Sampling trajectories with particle filtering

Input : Gamma-prior parameters on parents' transition intensity matrices $\alpha^1, \beta^1, \alpha^2, \beta^2$
Set of agent's transition intensity matrices $\boldsymbol{Q}_3$
T_{max} to terminate simulation

Output : Sample trajectory of the network

Initialize: Sample $\mathbf{Q}_1$ and $\mathbf{Q}_2$ from their priors
Initialize nodes uniformly $X_n(0) = x_i \in \mathcal{X}_n$
Initialize particles uniformly $p^i(0) = x_p \in \mathcal{X}_P$
$t = 0$

```
while t < T_max do
    Draw next transition for X_1 and X_2 (τ_parent, x_1 and x_2 using Algorithm 1)
    t_parent ← t + τ_parent // transition time for parents
    y_{t_parent} ~ ψ(x_1, x_2) // new observation at t_parent
    Update particle filter and obtain b^[t, t_parent]
    a^[t, t_parent] ← π(b^[t, t_parent])
    Q_3^[t, t_parent] ← Q_{3|a^[t, t_parent]}
    t_agent ← t
    while t_agent < t_parent do
        the upper bound for intensity, q_3^* [1]
        transition time τ_agent drawn by u ~ U(0,1) and τ_agent = -ln(u) / q_3^*
        t_agent ← t_agent + τ_agent
        draw s ~ U(0,1), accept transition if s ≤ q_3(t_agent) / q_3^*
    end while
    t ← t_parent
end while
```

In math notation:

- $t_{parent} \leftarrow t + \tau_{parent}$
- $y_{t_{parent}} \sim \psi(x_1, x_2)$
- $\mathbf{b}^{[t, t_{parent}]}$, $a^{[t, t_{parent}]} \leftarrow \pi(\mathbf{b}^{[t, t_{parent}]})$
- $\mathbf{Q}_3^{[t, t_{parent}]} \leftarrow \mathbf{Q}_{3|a^{[t, t_{parent}]}}$
- $\tau_{agent} = \frac{-ln(u)}{q_3^*}$, $s \le \frac{q_3(t_{agent})}{q_3^*}$

[1] q is the transition intensity associated with the current state of the agent.

4. Results

The experimental results are presented in this chapter. First, the parameters for the variables introduced in Chapter 3 are given. Then a sample of simulated trajectories are shown as an example. Finally, the inference results are presented.

4.1. Configurations

The configurations given below are used for the results presented in the following sections, if not specified otherwise.

- Gamma priors for parent dynamics such that $\mathbf{Q}_i \sim Gam(\alpha^i, \beta^i)$ for $i \in \{1, 2\}$, and $\alpha = [\alpha_0, \alpha_1]$ and $\beta = [\beta_0, \beta_1]$

$$\alpha^1 = [5, 10] \quad \beta^1 = [5, 20] \tag{4.1}$$

$$\alpha^2 = [10, 10] \quad \beta^2 = [10, 5] \tag{4.2}$$

- Transition intensity matrices of X_1 and X_2 sampled from priors given above

$$\mathbf{Q}_1 = \begin{bmatrix} -1.117 & 1.117 \\ 0.836 & -0.836 \end{bmatrix} \tag{4.3}$$

$$\mathbf{Q}_2 = \begin{bmatrix} -1.1 & 1.1 \\ 2.445 & -2.445 \end{bmatrix} \tag{4.4}$$

- State space, $S = \mathcal{X}_P = \{(0,0), (0,1), (1,0), (1,1)\}$

- Observation space, $\mathcal{Y} = \{0, 1, 2\}$

- Action space, $A = \{a_0, a_1\} = \{0, 1\}$

- The set of transition intensity matrices of X_3

$$\boldsymbol{Q}_3 = \left\{\mathbf{Q}_{3|a_0}, \mathbf{Q}_{3|a_1}\right\} = \left\{\begin{bmatrix} -0.5 & 0.5 \\ 2 & -2 \end{bmatrix}, \begin{bmatrix} -3 & 3 \\ 0.02 & -0.02 \end{bmatrix}\right\} \tag{4.5}$$

- Number of particles, $N = 200$

- Weights of the policy introduced in Equation 3.12, $\mathbf{w} = [0.02, 0.833, 0.778, 0.87]$

4.2. Simulation

4.3. Inference of Observation Model

5. Discussion

Only because the agent can see information (obs model) doesn't mean it needs to use it (policy)
equivalence classes and how they come about

6. Future Work

so what is still missing is the joint inference of obs model and policy

Bibliography

[1] I. Cohn, T. El-Hay, N. Friedman, and R. Kupferman, "Mean Field Variational Approximation for Continuous-Time Bayesian Networks," *Journal of Machine Learning Research*, vol. 11, pp. 2745–2783, 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," *Artificial Intelligence*, vol. 101, no. 1, pp. 99–134, 1998.

[4] K. P. Murphy, "A survey of POMDP solution techniques," *Environment*, vol. 2, no. September, p. X3, 2000.

[5] 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.

[6] S. Godsill, "PARTICLE FILTERING: THE FIRST 25 YEARS AND BEYOND," pp. 7760–7764, 2019.

[7] S. Thrun, "Monte Carlo POMDPs," vol. 40, no. 10, pp. 117–151, 1904.

[8] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," *Handbook of Nonlinear Filtering*, no. December, pp. 4–6, 2009.

[9] D. T. Gillespie, "A general method for numerically simulating the stochastic time evolution of coupled chemical reactions," *Journal of Computational Physics*, vol. 22, no. 4, pp. 403–434, 1976.

[10] P. A. Lewis and G. S. Shedler, "Simulation of Nonhomogeneous Poisson Processes By Thinning.," *Naval research logistics quarterly*, vol. 26, no. 3, pp. 403–413, 1979.

[11] Y. Ogaata, "On Lewis' Simulation Method for Point Processes," *IEEE Transactions on Information Theory*, vol. 27, no. 1, pp. 23–31, 1981.

[12] M.-A. Rizoiu, Y. Lee, and S. Mishra, "Hawkes processes for events in social media," *Frontiers of Multimedia Research*, pp. 191–218, 2017.

A. Amalgamation Operation

A CTBN with multiple variables can be represented with a single CIM. This is done by amalgamation operation. Amalgamation defines a combining operation over multiple CIMs and produces a single CIM for the entire system. [2]

A.1. Amalgamation of Independent Processes

Consider a CTBN with graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ over two variables such that $\mathcal{V} = \{X_1, X_2\}$. Assume variables X_1 and X_2 with intensity matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$, are both parent nodes, i.e. $\mathcal{E} = \emptyset$ and $Par_{\mathcal{G}}(X_1) = Par_{\mathcal{G}}(X_1) = \emptyset$. This CTBN can be identified as a subsytem of the CTBN model described in Section 3.1.1.
Analogous to Equation 2.5, Markov transition function for the joint process can be derived as

$$\begin{aligned}
\Pr(X_P(t+h) = x'_p \mid X_P(t) = x_p) &= \Pr(X_1(t+h) = x'_1, X_2(t+h) = x_2 \mid X_1(t) = x_1, X_2(t) = x_2) \\
&= \Pr(X_1(t+h) = x'_1 \mid X_1(t) = x_1, X_2(t) = x_2) \\
&\quad \Pr(X_2(t+h) = x_2 \mid X_1(t) = x_1, X_2(t+h) = x_2) \\
&= (\delta_{x'_1,x_1} + hq^1_{x_1,x'_1} + o(h))(1 + hq^2_{x_2,x_2} + o(h)) \\
&= \delta_{x'_1,x_1} + hq^1_{x_1,x'_1} + h\delta_{x'_1,x_1} q^2_{x_2,x_2} + o(h) \qquad \text{(A.1)}
\end{aligned}$$

where $x_1, x'_1 \in \mathcal{X}_1$, $x_2, x'_2 \in \mathcal{X}_2$, $x_p = (x_1, x_2), x'_p = (x'_1, x_2) \in \mathcal{X}_P$.
Suppose the intensity matrices of X_1 and X_2 are in the form

$$\mathbf{Q}_i = \begin{bmatrix} -q^i_0 & q^i_0 \\ q^i_1 & -q^i_1 \end{bmatrix} \quad \text{for } i \in \{1, 2\} \qquad \text{(A.2)}$$

Then the intensity matrix for the joint process X_P with factorising state space $\mathcal{X}_P = \mathcal{X}_1 \times \mathcal{X}_2$ can be written as

$$\mathbf{Q}_P = \begin{bmatrix} -q^2_0 - q^1_0 & q^2_0 & q^1_0 & 0 \\ q^2_1 & -q^2_1 - q^1_0 & 0 & q^1_0 \\ q^1_1 & 0 & -q^1_1 - q^2_0 & q^2_0 \\ 0 & q^1_1 & q^2_1 & -q^1_1 - q^2_1 \end{bmatrix} \quad \text{for } i \in \{1, 2\} \qquad \text{(A.3)}$$

As it can be observed from Equation A.3, the transition intensities which corresponds to state transition in both variables, i.e. anti-diagonal entries, are zero, due to the one of the assumptions in CTBN framework that only one variable can transition at a time, as given in Section 2.2.

B. Marginalized Likelihood Function for Homogenous Continuous Time Markov Processes

Let X be a homogenous CTMP. For convenience, it is assumed to be binary-valued, $\mathcal{X} = \{x_0, x_1\}$. The transition intensity matrix can be written in the following form:

$$\mathbf{Q} = \begin{bmatrix} -q_0 & q_0 \\ q_1 & -q_1 \end{bmatrix} \tag{B.1}$$

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 $X^{[0,T]}$ can be written as follows:

$$\begin{aligned}
P(X^{[0,T]}) &= \int P(X^{[0,T]} \mid Q)P(Q)dQ \\
&= \int_0^\infty = \prod_{j \neq i} \exp(-q_{i,j}T[x_i])\ q_{i,j}^{M[x_i,x_j]} \frac{\beta_{i,j}^{\alpha_{i,j}} q_{i,j}^{\alpha_{i,j}-1} \exp(-\beta_{i,j}q_{i,j})}{\Gamma(\alpha_{i,j})}\ dq_{i,j} \\
&= \prod_{i \in 0,1} \int_0^\infty q_i^{M[x_i]}\ \exp(-q_iT[x_i])\ \frac{\beta_i^{\alpha_i}\ q_i^{\alpha_i-1}\ \exp(-\beta_i q_i)}{\Gamma(\alpha_i)}\ dq_i \\
&= \prod_{i \in 0,1} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \int_0^\infty q_i^{M[x_i]+\alpha_i-1}\ \exp(-q_i(T[x_i]+\beta_i))\ dq_i \qquad \text{(B.2)} \\
&= \prod_{i \in 0,1} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \Big(-(T[x_i]+\beta_i)^{-M[x_i]-\alpha_i}\ \Gamma(M[x_i]+\alpha_i,\ q_i(T[x_i]+\beta_i))\Big)\Big|_0^\infty \\
&= \prod_{i \in 0,1} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} \Big((T[x_i]+\beta_i)^{-M[x_i]-\alpha_i}\ \Gamma(M[x_i]+\alpha_i)\Big) \qquad \text{(B.3)}
\end{aligned}$$

In Equation B.2, the integral is solved using computer algebra system WolframAlpha as follows:

$$\int x^a\ \exp(-xb)\ dx = -b^{-a-1}\ \Gamma(a+1,\ bx) + C \tag{B.4}$$