Dynamic Chain Event Graphs

Lorna Barclay

University of Warwick

Graphical Bayes Research Group

Example

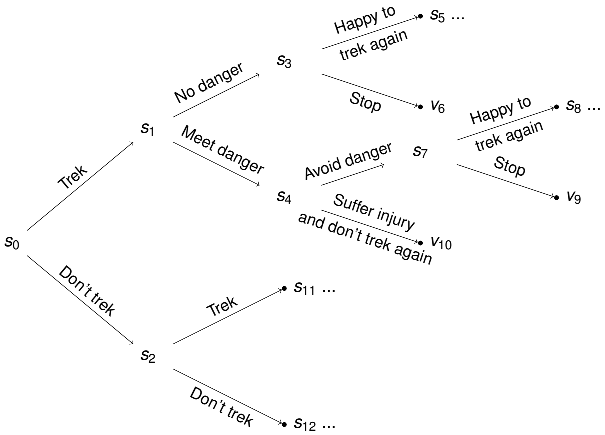

Figure: Trekking example: the beginning of the infinite tree, $\mathcal{T}$

Example

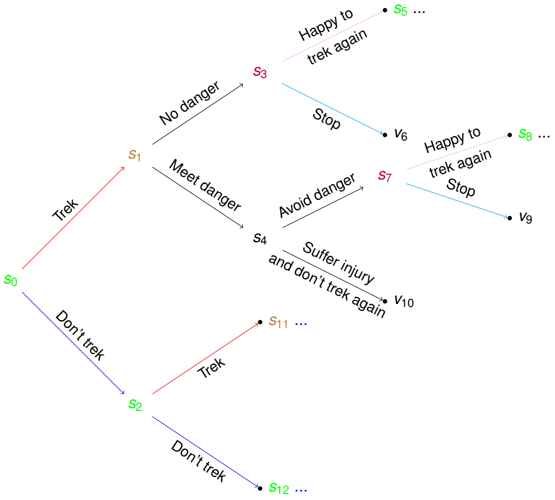

Figure: Trekking example: the beginning of the infinite staged tree, $\mathcal{T}$

Example of a DCEG

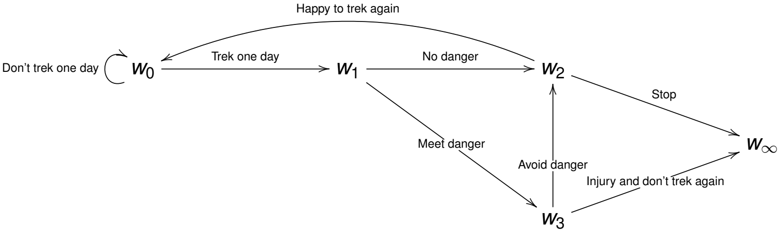

Figure: DCEG of $\mathcal{T}$

The Dynamic Chain Event Graph

Definition

A *Dynamic Chain Event Graph* (DCEG) $\mathcal{D} = (V(\mathcal{D}), E(\mathcal{D}))$ of a staged tree $\mathcal{T}$ is a directed coloured graph with vertex set $V(\mathcal{D}) = W$, the set of positions of the staged tree $\mathcal{T}$, together with a single sink vertex, w_∞, comprising the leaf nodes of T, if these exist.

The edge set $E(\mathcal{D})$ is given as follows: Let $v \in w$ be a single representative vertex of the position w. Then there is an edge from w to a position $w' \in W$ for each child $v' \in ch(v), v' \in w'$ in the tree $\mathcal{T}$. When two positions are also in the same stage then these and their edges are coloured in the same colour as the corresponding vertices and edges in the tree $\mathcal{T}$.

DGEGs and Markov Processes

Let $\{X_n : n \in \mathbb{N}\}$ be a discrete-time Markov process on the state space $\{a, b, c\}$ with transition matrix P given by

$$P = \begin{pmatrix} 0.2 & 0.3 & 0.5 \\ 0.5 & 0.3 & 0.2 \\ 0.5 & 0.3 & 0.2 \end{pmatrix},$$

and with initial distribution $\alpha = (0.4, 0.4, 0.2)$.

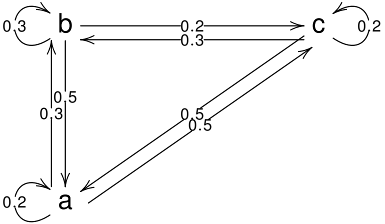

Figure: Example 2: State-transition diagram of a Markov process

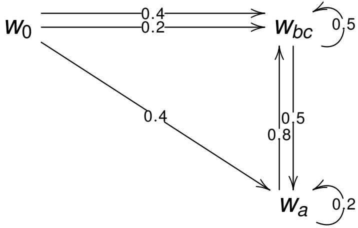

Figure: Example 2: DCEG representation of a Markov process

DGEGs and Markov Processes

A coin is tossed independently, with probability $P(H) = \lambda$ of throwing heads and probability $P(T) = 1 - \lambda = \bar{\lambda}$ of throwing tails. The coin is tossed until three heads have appeared when the game terminates.

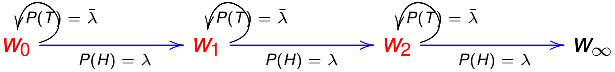

Figure: Example 3: DCEG representation of coin tossing example

DCEG with Holding Times: Extended DCEG

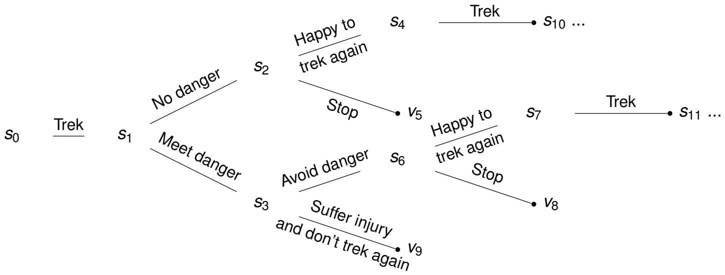

Figure: Variant of trekking example: infinite tree $\mathcal{T}^*$

DCEG with Holding Times: Extended DCEG

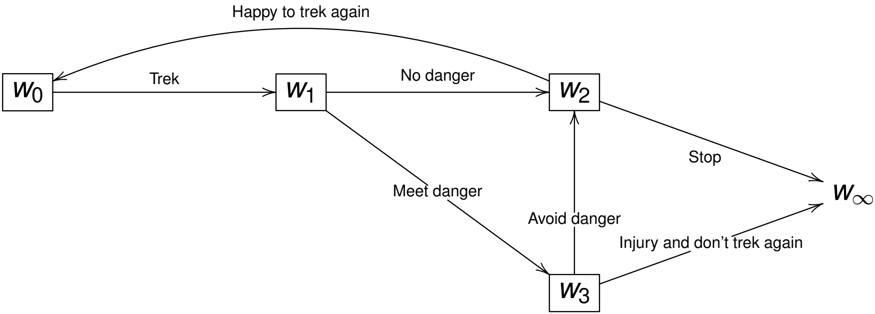

Figure: DCEG of $\mathcal{T}^*$

Extended DCEG

Definition

An *Extended DCEG* $\mathcal{D} = (V(\mathcal{D}), E(\mathcal{D}))$ is a DCEG with no loops from a position into itself and with conditional holding time distributions conditioned on the current stage, u, and the next edge, e_{uj}, to be passed through:

$$F_{uj}(h) = P(H_{uj} \leq h), h \geq 0, \forall u \in U, j = 1, \ldots m_u. \quad (1)$$

Hence $F_{uj}(h)$ describes the time an individual stays in a position $w \in u$ before moving along the j_{th} edge, e_{wj}.

Extended DCEG

Definition

An Extended DCEG is semi-Markov if

$$P(E_n, H_n \leq h | W_0, W_1, \ldots, W_n, E_0, E_1, \ldots, E_{n-1}, H_0, H_1, \ldots, H_{n-1}) = P(E_n, H_n \leq h | W_n).$$

Hence, the joint probability of the n_{th} holding time and the n_{th} edge passed along depends only on the current position of the individual.

As we are assuming a time-homogeneous DCEG we further have that

$$P(H_n \leq h | W_n = w, E_n = e_{wj}) = F_{uj}(h), w \in u,$$

and also

$$P(E_n = e_{wj} | W_n = w) = \pi_{uj}, w \in u,$$

Therefore,

$$P(E_n = e_{wj}, H_n \leq h | W_n = w) = P(E_n = e_{wj} | W_n = w) P(H_n \leq h | W_n = w, E_n = e_{wj}) = \pi_{uj} F_{uj}(h).$$

Learning the Parameters of a DCEG

- We have associated with each stage u a CPV $\pi_u = (\pi_{u1}, \pi_{u2}, .., \pi_{m_u})$ and we denote the full set of CPVs by π.
- We can further attach a vector of conditional holding time distributions $(F_{u1}, F_{u2}, \ldots, F_{um_u})$ to each stage u with parameters $\lambda_u = (\lambda_1, \lambda_2, \ldots, \lambda_{m_u})$. We call the full set of parameters λ.
- We record the number of times the individuals pass along a position $w \in u$ and go along the j_{th} edge, which we denote by N_{uj}.
- We let $\mathbf{h}_{uj}$ be the vector of conditional holding times for the individuals which arrive at stage u and move along the j_{th} edge next and we let h_{uji} be the holding time of the i_{th} pass along this edge.

Learning the Parameters of a DCEG

Then, immediately from the definition of a time-homogeneous and semi-Markov Extended DCEG $\mathcal{D}$, the likelihood $L(\pi, \lambda | \mathbf{N}, \mathbf{h}, \mathcal{D})$ of this random sample separates

$$L(\pi, \lambda | \mathbf{N}, \mathbf{h}, \mathcal{D}) = L_1(\pi | \mathbf{N}, \mathcal{D}) L_2(\lambda | \mathbf{h}, \mathbf{N}, \mathcal{D}).$$

$$\prod_{u \in U} L_u(\pi_u | \mathbf{N}_u, \mathcal{D}) = \prod_{u \in U} \prod_{j=1}^{m_u} \pi_{uj}^{Nuj},$$

$$\prod_{u \in U} L_{uj}(\lambda_{uj} | \mathbf{h}_{uj}, \mathbf{N}_{uj}, \mathcal{D}) = \prod_{u \in U} \prod_{i=1}^{N_{uj}} \frac{1}{\lambda_{uj}} exp(-\frac{1}{\lambda_{uj}} h_{uji}).$$

Learning the Parameters of a DCEG

It is immediate that if λ and π are believed to be a priori independent so that

$$p(\pi, \lambda|\mathcal{D}) = p_1(\pi|\mathcal{D})p_2(\lambda|\mathcal{D}),$$

then the posterior density $p(\pi, \lambda|\mathbf{h}, \mathbf{N}, \mathcal{D})$ separates into

$$p(\pi, \lambda|\mathbf{h}, \mathbf{N}, \mathcal{D}) = p_1(\pi|\mathbf{N}, \mathcal{D})p_2(\lambda|\mathbf{h}, \mathbf{N}, \mathcal{D})$$

and we can perform the updating of the CPVs, π, and the holding time parameters, λ, without reference to the other.

Example

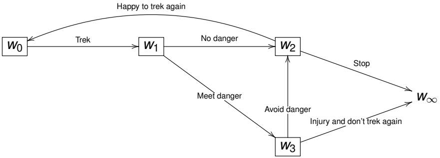

Figure: DCEG of trekking example

Description	Holding time distribution	Prior
Hrs until trek	$H_{u_0 1} \sim Exp(\lambda_0)$	$\lambda_0 \sim IG(10\frac{1}{3}, 9\frac{1}{3})$
Meet danger	$N_{u_1} \sim Mult(\pi_{u_1})$	$\pi_{u_1} \sim Dir(10\frac{1}{3}, 10\frac{1}{3})$
Hrs until danger met	$H_{u_1 1} \sim Weibull(\lambda_1, k_1)$	$\lambda_1^{k_1} \sim IG(10\frac{1}{3}, 9\frac{1}{3})$
Hrs until finished trek/no danger met	$H_{u_1 2} \sim Weibull(\lambda_3, k_3)$	$\lambda_3^{k_3} \sim IG(10\frac{1}{3}, 9\frac{1}{3})$
Avoid danger	$N_{u_3} \sim Mult(\pi_{u_3})$	$\pi_{u_3} \sim Dir(4, 4)$
Hrs until avoids danger	$H_{u_3 1} \sim Weibull(\lambda_2, k_2)$	$\lambda_2^{k_2} \sim IG(4, 3)$
Hrs until injury	$H_{u_3 2} \sim Exp(\lambda_6)$	$\lambda_6 \sim IG(4, 3)$
Happy to trek again	$N_{u_2} \sim Mult(\pi_{u_2})$	$\pi_{u_2} \sim Dir(6\frac{2}{3}, 6\frac{2}{3})$
Days until happy to trek	$H_{u_2 1} \sim Exp(\lambda_4)$	$\lambda_4 \sim IG(6\frac{2}{3}, 5\frac{2}{3})$
Hrs until decides to stop	$H_{u_2 2} \sim Exp(\lambda_5)$	$\lambda_5 \sim IG(6\frac{2}{3}, 5\frac{2}{3})$

Table: Prior distributions on CPVs and conditional holding times

Example

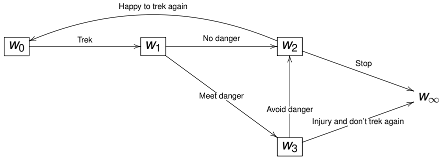

Figure: DCEG of trekking example

- The paths are simulated by assuming that the individual meets danger with probability $\frac{3}{4}$ and avoids danger, if danger is met, with probability $\frac{2}{3}$. The probability of trekking again is also $\frac{3}{4}$.
- We simulate from $Exp(5)$ to describe the time until trekking
- We simulate from a $Weibull(3, 2)$ for the time until meeting danger, from a $Weibull(5, \frac{2}{3})$ for the time until the trek is finished and from a $Weibull(2, \frac{1}{2})$ to describe the time until danger is avoided.
- We choose $Exp(10)$, $Exp(3)$ and $Exp(2)$ for the days until trekking again again, the hours until deciding to stop and the hours until injury.

Example

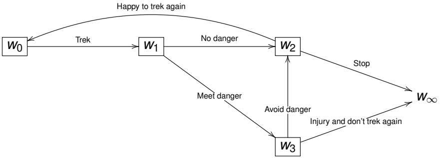

Figure: DCEG of trekking example

Description	Posterior	Mean
Hrs until trek	$\lambda_0 \sim IG(1010\frac{1}{2}, 5018.58)$	$4.97(4.67, 5.29)$
Meet danger	$\pi_{u_1} \sim Dir(761\frac{1}{3}, 259\frac{1}{3})$	$0.75(0.72, 0.77)$
Hrs until danger met	$\lambda_1^{k_1} \sim IG(259\frac{1}{3}, 2056.59)$	$2.50(2.35, 2.66)$
Hrs until finished trek/no danger met	$\lambda_3^{k_3} \sim IG(761\frac{1}{3}, 2128.16)$	$6.23(5.59, 6.92)$
Avoid danger	$\pi_{u_3} \sim Dir(176, 81)$	$0.68(0.63, 0.74)$
Hrs until avoids danger	$\lambda_2^{k_2} \sim IG(176, 251.60)$	$4.16(3.10, 5.60)$
Hrs until injury	$\lambda_6 \sim IG(81, 150.39)$	$1.88(1.51, 2.34)$
Happy to trek again	$\pi_{u_2} \sim Dir(700\frac{2}{3}, 234\frac{2}{3})$	$0.75(0.72, 0.78)$
Days until happy to trek	$\lambda_4 \sim IG(701\frac{2}{3}, 7209.729)$	$10.29(9.56, 11.08)$
Hrs until decides to stop	$\lambda_5 \sim IG(234\frac{2}{3}, 629.64)$	$2.69(2.37, 3.06)$

Table: Posterior distributions on CPVs and conditional holding times

Semi-Markov Processes (Medhi, 1994)

Definition

Let $\{Y_t, t \geq 0\}$ be a process with discrete state space and with transitions occurring at times $t_0, t_1, t_2, \ldots$. Also, let $\{X_n, n \in \mathbb{N}\}$ describe the state of the process at time t_n and let H_n be the holding time before transition to X_n. Hence $Y_t = X_n$ on $t_n \leq t < t_{n+1}$. If

$$P(X_{n+1} = j, H_{n+1} \leq t | X_0, X_1, .., X_n, H_1, .., H_n) = P(X_{n+1} = j, H_{n+1} \leq t | X_n),$$

then $\{X_n, H_n\}$ is called a Markov Renewal process and $\{Y_t, t \geq 0\}$ a semi-Markov process. Also, $\{X_n, n \in \mathbb{N}\}$ is the embedded Markov chain with transition probability matrix $P = (p_{ij})$, where $p_{ij} = P(X_{n+1} = j | X_n = i)$.

Semi-Markov Processes (Medhi, 1994)

The semi-Markov kernel Q has ij_{th} entry

$$Q_{ij}(t) = P(X_{n+1} = j, H_{n+1} \leq t | X_n = i).$$

We assume here that all Markov processes considered are time-homogeneous and hence the above equations do not depend on the index n. We write the semi-Markov kernel as

$$Q_{ij}(t) = p_{ij} F_{ij}(t),$$

where

$$F_{ij}(t) = P(H_{n+1} \leq t | X_{n+1} = j, X_n = i).$$

DCEGs and Semi-Markov Processes

Theorem

Let an Extended DCEG be simple and let no two children lead from the same parent into the same child. Then the DCEG is a semi-Markov process with state space W (the set of positions), with conditional holding time distributions

$$F_{w_i w_j}(t) = P(H_{w_i j} \leq t),$$

whenever $e_{w_i j} = e(w_i, w_j)$ exists and 0 otherwise, and with the entries of the transition probability matrix of the embedded Markov Chain $\{X_n, n \in \mathbb{N}\}$ given by

$$p_{w_i w_j} = P(e_{w_i j} | w_i),$$

if the edge $e_{w_i j} = e(w_i, w_j)$ exists and 0 otherwise. If the position w_0 is a source node then the state-transition diagram of the semi-Markov process omits w_0 and the initial distribution is given by π_{w_0}. Otherwise the initial distribution assigns probability 1 to w_0.

References I

L. M. Barclay, J. L. Hutton, and J. Q. Smith. Refining a Bayesian Network using a Chain Event Graph. *International Journal of Approximate Reasoning*, 2013. 10.1016/j.ijar.2013.05.006.

G. Freeman and J. Q. Smith. Bayesian MAP model selection of Chain Event Graphs. *Journal of Multivariate Analysis*, 102(7): 1152–1165, 2011.

J. Medhi. *Stochastic processes*. New Age International, 1994.

J. Q. Smith and P. E. Anderson. Conditional independence and Chain Event Graphs. *Artificial Intelligence*, 172(1):42–68, 2008.