Probabilistic Modelling and Reasoning № 1

Georgi Tinchev, s1549117, The University of Edinburgh 07/03/2016

1 Question 1: Player skill graphical models

1.a

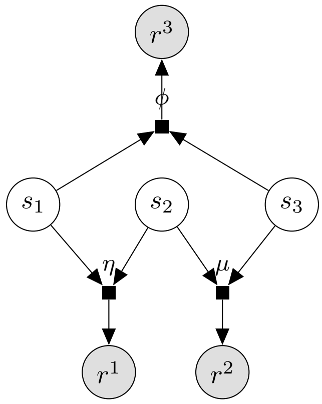

Figure 1: Factor graph of simple player skill model when observing three games between three players. The games are denoted as the factors ϕ, μ and η. Game outcomes are denoted as $r^{(k)}$, where $k \in \{1, 2, 3\}$ being the different games played. The result of the games, denoted $r^{(k)} \in \{0, 1\}$ with 0 indicating a loss and 1 indicating a win. The players' skills are identified as s_1, s_2 and s_3, where the index denotes the ID of the player $\in \{1, 2, 3\}$ respectively.

1.b

1.b.1 $I(s_1, s_2|r^{(2)})$

There are 2 paths leading from s_1 to s_2, namely $s_1 - \eta - s_2$ and $s_1 - \phi - s_3 - \mu - s_2$. The former path is blocked, as η has two incoming edges and neither η, nor $r^{(1)}$ (being all descendants) are in the conditioning set (being $r^{(2)}$). The latter is also blocked, but the case is slightly more complicated. Firstly, μ has two incoming edges, however, its descendant is in the conditioning set ($r^{(2)}$), therefore, the path is not blocked there. Secondly, ϕ also has two incoming edges and neither ϕ, nor $r^{(3)}$ are in the conditioning set, therefore, the path is blocked there. Thus, as the two paths are blocked $=> I(s_1, s_2|r^{(2)})$.

1.b.2 $I(s_1, s_2|r^{(2)}, r^{(3)})$

As identified in Section 1.b.1, there are 2 paths. The path $s_1 - \eta - s_2$ is still blocked, as η, nor any of its descendants are in the conditioning set. The path $s_1 - \phi - s_3 - \mu - s_2$, however, is now unblocked, as for all factors and variables with 2 incoming edges, there is a child that belongs to the conditioning set, namely $r^{(3)}$ is a child of ϕ and $r^{(2)}$ is a child of μ, both of them being in the conditioning set.

1.c

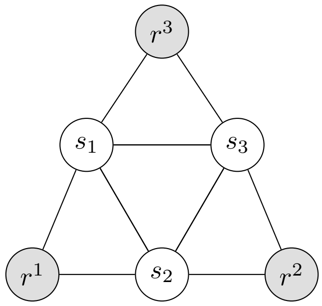

Figure 2: Markov network of simple player skill model when observing three games between three players. The observed and latent variables are denoted in a similar way, as Figure 1. Game outcomes are denoted as $r^{(k)}$, where $k \in \{1, 2, 3\}$ being the different games played. The result of the games, denoted $r^{(k)} \in \{0, 1\}$ with 0 indicating a loss and 1 indicating a win. The players' skills are identified as s_1, s_2 and s_3, where the index denotes the ID of the player $\in \{1, 2, 3\}$ respectively.

1.d

2 Question 2: Gaussian player skill model

2.a

From the figure, one can express the joint probability distribution as follows:

$$p(p_b, p_w, s_b, s_w) = \frac{1}{Z}\phi_1(p_w)\phi_2(s_w, p_w)\phi_3(s_b)\phi_4(s_b, p_b)\phi_5(p_w, p_b, r) \tag{1}$$

Thereafter, using chain rule, $p(p_b, p_w|s_b, s_w)$ can be expressed as follows:

$$p(p_b, p_w|s_b, s_w) = \frac{p(p_b, p_w, s_b, s_w)}{p(s_b, s_w)} = \frac{p(p_b, p_w, s_b, s_w)}{\int_{p_b,p_w} p(p_b, p_w, s_b, s_w)} = \tag{2}$$

$$\frac{\frac{1}{Z}\phi_1(s_w)\phi_2(s_w, p_w)\phi_3(s_b)\phi_4(s_b, p_b)\phi_5(p_w, p_b, r)}{\int_{p_b,p_w} \frac{1}{Z}\phi_1(s_w)\phi_2(s_w, p_w)\phi_3(s_b)\phi_4(s_b, p_b)\phi_5(p_w, p_b, r)} \tag{3}$$

$$\frac{\frac{1}{Z}\phi_1(p_w)\phi_2(s_w, p_w)\phi_3(p_b)\phi_4(s_b, p_b)\phi_5(p_w, p_b, r)}{\frac{1}{Z}\phi_1(p_w)\phi_3(p_b)\int_{p_b,p_w} \phi_2(s_w, p_w)\phi_4(s_b, p_b)\phi_5(p_w, p_b, r)} \tag{4}$$

$$\frac{\phi_2(s_w, p_w)\phi_4(s_b, p_b)\phi_5(p_w, p_b, r)}{\int_{p_b,p_w} \phi_2(s_w, p_w)\phi_4(s_b, p_b)\phi_5(p_w, p_b, r)} \tag{5}$$

$$\frac{\phi_2(s_w, p_w)\phi_4(s_b, p_b)}{\int_{p_b,p_w} \phi_2(s_w, p_w)\phi_4(s_b, p_b)} \tag{6}$$

$$\phi_2(s_w, p_w)\phi_4(s_b, p_b) \tag{7}$$

2.b

2.b.1 $\mathbb{E}[\theta \mid s_b, s_w]$

$$\mathbb{E}[\theta \mid s_w, s_b] = \frac{1}{\sqrt{2}\beta}\mathbb{E}[(p_w - s_w) + (p_b - s_b) \mid s_w, s_b] \tag{8}$$

$$\frac{1}{\sqrt{2}\beta}\Bigg(\mathbb{E}[p_w + p_b \mid s_w, s_b] - s_w - s_b\Bigg) \tag{9}$$

$$\frac{1}{\sqrt{2}\beta}\Bigg(\mathbb{E}[p_w \mid s_w, s_b] + \mathbb{E}[p_b \mid s_w, s_b] - s_w - s_b\Bigg) \tag{10}$$

$$\frac{1}{\sqrt{2}\beta}\Bigg(s_w + s_b - s_w - s_b\Bigg) = 0 \tag{11}$$

2.b.2 $\mathbb{E}[\psi \mid s_b, s_w]$

$$\mathbb{E}[\psi \mid s_w, s_b] = \frac{1}{\sqrt{2}\beta}\mathbb{E}[(p_w - s_w) - (p_b - s_b) \mid s_w, s_b] \tag{12}$$

$$\frac{1}{\sqrt{2}\beta}\Bigg(\mathbb{E}[p_w - p_b \mid s_w, s_b] - s_w + s_b\Bigg) \tag{13}$$

$$\frac{1}{\sqrt{2}\beta}\Bigg(\mathbb{E}[p_w \mid s_w, s_b] - \mathbb{E}[p_b \mid s_w, s_b] - s_w + s_b\Bigg) \tag{14}$$

$$\frac{1}{\sqrt{2}\beta}\Bigg(s_w - s_b - s_w + s_b\Bigg) = 0 \tag{15}$$

2.b.3 $\mathbb{E}[\theta^2 \mid s_b, s_w]$

$$\mathbb{E}[\theta^2 \mid s_w, s_b] = \frac{1}{2\beta^2}\mathbb{E}[((p_w - s_w) + (p_b - s_b))^2 \mid s_w, s_b] \tag{16}$$

$$\frac{1}{2\beta^2}\Bigg(\underbrace{\mathbb{E}[(p_w - s_w)^2 \mid s_w, s_b]}_{\text{A}} + \underbrace{2\mathbb{E}[(p_w - s_w)(p_b - s_b) \mid s_w, s_b]}_{\text{B}} + \underbrace{\mathbb{E}[(p_b - s_b)^2 \mid s_w, s_b]}_{\text{C}}\Bigg) \tag{17}$$

$$A = \mathbb{E}[p_w^2 \mid s_w, s_b] - 2s_w\mathbb{E}[p_w \mid s_w, s_b] + s_w^2 \tag{18}$$

$$A = \mathrm{Var}(p_w) + \mathbb{E}[p_w \mid s_w, s_b]^2 - 2s_w^2 + s_w^2 = \mathrm{Var}(p_w) = \beta^2 \tag{19}$$

$$B = 2\mathbb{E}[p_b p_w \mid s_w, s_b] - 2s_b\mathbb{E}[p_w \mid s_w, s_b] - 2s_w\mathbb{E}[p_b \mid s_w, s_b] + 2s_b s_w \tag{20}$$

$$B = 2\Bigg(\mathbb{E}[p_b \mid s_w, s_b]\mathbb{E}[p_w \mid s_w, s_b] + \mathrm{Cov}(p_b, p_w)\Bigg) - 2s_b s_w - 2s_w s_b + 2s_b s_w \tag{21}$$

$$B = 2\left(\mathbb{E}[p_b \mid s_w, s_b]\mathbb{E}[p_w \mid s_w, s_b] + \operatorname{Cov}(p_b, p_w) \right) - 2s_b s_w \tag{22}$$

$$B = 2\left(s_b s_w + \underbrace{\operatorname{Cov}(p_b, p_w)}_{p_w \text{ and } p_b \text{ are CI } => \operatorname{Cov}(p_w, p_b) = 0} \right) - 2s_b s_w = 0 \tag{23}$$

$$C = \mathbb{E}[p_b^2 \mid s_w, s_b] - 2s_b\mathbb{E}[p_b \mid s_w, s_b] + s_b^2 \tag{24}$$

$$C = \operatorname{Var}(p_b) + \mathbb{E}[p_b \mid s_w, s_b]^2 - 2s_b^2 + s_b^2 = \operatorname{Var}(p_b) = \beta^2 \tag{25}$$

From Equation 17:

$$\frac{1}{2\beta^2}(A + B + C) = \frac{\beta^2 + 0 + \beta^2}{2\beta^2} = 1 \tag{26}$$

2.b.4 $\mathbb{E}[\psi^2 \mid s_b, s_w]$

$$\mathbb{E}[\psi^2 \mid s_w, s_b] = \frac{1}{2\beta^2}\mathbb{E}[((p_w - s_w) - (p_b - s_b))^2 \mid s_w, s_b] \tag{27}$$

$$\frac{1}{2\beta^2}\left(\underbrace{\mathbb{E}[(p_w - s_w)^2 \mid s_w, s_b]}_{\text{D}} - \underbrace{2\mathbb{E}[(p_w - s_w)(p_b - s_b) \mid s_w, s_b]}_{\text{E}} + \underbrace{\mathbb{E}[(p_b - s_b)^2 \mid s_w, s_b]}_{\text{F}} \right) \tag{28}$$

$$D = \mathbb{E}[p_w^2 \mid s_w, s_b] - 2s_w\mathbb{E}[p_w \mid s_w, s_b] + s_w^2 \tag{29}$$

$$D = \operatorname{Var}(p_w) + \mathbb{E}[p_w \mid s_w, s_b]^2 - 2s_w^2 + s_w^2 = \operatorname{Var}(p_w) = \beta^2 \tag{30}$$

$$E = 2\mathbb{E}[p_b p_w \mid s_w, s_b] - 2s_b\mathbb{E}[p_w \mid s_w, s_b] - 2s_w\mathbb{E}[p_b \mid s_w, s_b] + 2s_b s_w \tag{31}$$

$$E = 2\left(\mathbb{E}[p_b \mid s_w, s_b]\mathbb{E}[p_w \mid s_w, s_b] + \operatorname{Cov}(p_b, p_w) \right) - 2s_b s_w - 2s_w s_b + 2s_b s_w \tag{32}$$

$$E = 2\left(\mathbb{E}[p_b \mid s_w, s_b]\mathbb{E}[p_w \mid s_w, s_b] + \operatorname{Cov}(p_b, p_w) \right) - 2s_b s_w \tag{33}$$

$$E = 2\left(s_b s_w + \underbrace{\operatorname{Cov}(p_b, p_w)}_{p_w \text{ and } p_b \text{ are CI } => \operatorname{Cov}(p_w, p_b) = 0} \right) - 2s_b s_w = 0 \tag{34}$$

$$F = \mathbb{E}[p_b^2 \mid s_w, s_b] - 2s_b\mathbb{E}[p_b \mid s_w, s_b] + s_b^2 \tag{35}$$

$$F = \operatorname{Var}(p_b) + \mathbb{E}[p_b \mid s_w, s_b]^2 - 2s_b^2 + s_b^2 = \operatorname{Var}(p_b) = \beta^2 \tag{36}$$

From Equation 28:

$$\frac{1}{2\beta^2}(D - E + F) = \frac{\beta^2 - 0 + \beta^2}{2\beta^2} = 1 \tag{37}$$

2.b.5 $\mathbb{E}[\psi\theta \mid s_b, s_w]$

$$\mathbb{E}[\psi\theta \mid s_w, s_b] = \frac{1}{2\beta^2}\mathbb{E}[(p_w - s_w)^2 - (p_w - s_w)(p_b - s_b) + (p_w - s_w)(p_b - s_b) - (p_b - s_b)^2 \mid s_w, s_b] \tag{38}$$

$$\mathbb{E}[\psi\theta \mid s_w, s_b] = \frac{1}{2\beta^2}\mathbb{E}[(p_w - s_w)^2 - (p_w - s_w)(p_b - s_b) + (p_w - s_w)(p_b - s_b) - (p_b - s_b)^2 \mid s_w, s_b] \tag{39}$$

$$\mathbb{E}[\psi\theta \mid s_w, s_b] = \frac{1}{2\beta^2}\left(\underbrace{\mathbb{E}[(p_w - s_w)^2 \mid s_w, s_b]}_{\text{G}} - \underbrace{\mathbb{E}[(p_b - s_b)^2 \mid s_w, s_b]}_{\text{H}}\right) \tag{40}$$

$$G = \mathbb{E}[p_w^2 \mid s_w, s_b] - 2s_w\mathbb{E}[p_w \mid s_w, s_b] + s_w^2 \tag{41}$$

$$G = \mathrm{Var}(p_w) + \mathbb{E}[p_w \mid s_w, s_b]^2 - 2s_w^2 + s_w^2 = \mathrm{Var}(p_w) = \beta^2 \tag{42}$$

$$H = \mathbb{E}[p_b^2 \mid s_w, s_b] - 2s_b\mathbb{E}[p_b \mid s_w, s_b] + s_b^2 \tag{43}$$

$$H = \mathrm{Var}(p_b) + \mathbb{E}[p_b \mid s_w, s_b]^2 - 2s_b^2 + s_b^2 = \mathrm{Var}(p_b) = \beta^2 \tag{44}$$

From Equation 40:

$$G - H = \frac{\beta^2 - \beta^2}{2\beta^2} = 0 \tag{45}$$

2.b.6 $\mathbb{p}[\theta, \psi \mid s_b, s_w]$

$\mathbb{p}(\theta|s_b, s_w)$ is a distribution, however, as it is composed of p_w, p_b, s_w, s_b, that are all Gaussian distributions, $\mathbb{p}(\theta|s_b, s_w)$ is also a Gaussian distribution. Its mean($\mathbb{E}[\theta|s_b, s_w]$) is 0 and its variance is 1 ($\mathbb{E}[\theta^2|s_b, s_w]$), therefore, $\mathbb{p}(\theta|s_b, s_w)$ is a standard normal Gaussian distribution.

$\mathbb{p}(\psi|s_b, s_w)$ is a distribution, however, as it is composed of p_w, p_b, s_w, s_b, that are all Gaussian distributions, $\mathbb{p}(\psi|s_b, s_w)$ is also a Gaussian distribution. Its mean($\mathbb{E}[\psi|s_b, s_w]$) is 0 and its variance is 1 ($\mathbb{E}[\psi^2|s_b, s_w]$), therefore, $\mathbb{p}(\psi|s_b, s_w)$ is a standard normal Gaussian distribution.

Based on the 2 aforementioned statements, the joint distribution $\mathbb{p}(\theta, \psi|s_b, s_w)$ is a joint standard multivariate Gaussian distribution, expressed as follows:

$$\mathbb{p}(\theta, \psi|s_b, s_w) = \frac{1}{\sqrt{(2\pi)^2 \mid \Sigma \mid}} exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \tag{46}$$

With the following x, μ and Σ:

$$x = \begin{bmatrix}\theta \\ \psi\end{bmatrix} \tag{47}$$

$$\mu = \begin{bmatrix}\mathbb{E}[\theta \mid s_b, s_w] \\ \mathbb{E}[\psi \mid s_b, s_w]\end{bmatrix} = \begin{bmatrix}0 \\ 0\end{bmatrix} \tag{48}$$

$$\Sigma = \begin{bmatrix} \mathrm{Var}(\theta) & \mathrm{Cov}(\theta,\psi) \\ \mathrm{Cov}(\theta,\psi) & \mathrm{Var}(\psi) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{49}$$

The $\mathrm{Cov}[\theta, \psi \mid s_b, s_w] = 0$ as $\mathbb{E}[\psi\theta \mid s_b, s_w] = 0$, $\mathbb{E}[\psi \mid s_b, s_w] = 0$ and $\mathbb{E}[\theta \mid s_b, s_w] = 0$, meaning that θ and ψ are independent.

Leading to the following form:

$$\mathbb{p}(\theta, \psi | s_b, s_w) = \frac{1}{2\pi} exp\Big(-\frac{1}{2}(\theta^2 + \psi^2)\Big) \tag{50}$$

2.c

For $r = \mathbb{1}[p_b > p_w]$

We can express p_b and p_w in terms of θ and ψ.

$$p_b = -\sqrt{2}\beta\psi + p_w - s_w + s_b \tag{51}$$

$$p_w = \sqrt{2}\beta\theta + s_w - p_b + s_b \tag{52}$$

Substituting Equation 51 in Equation 52 results in:

$$p_w = \sqrt{2}\beta\theta + s_w + \sqrt{2}\beta\psi - p_w + s_w - s_b + s_b \tag{53}$$

$$p_w = \frac{\sqrt{2}\beta(\theta + \psi) + 2s_w}{2} \tag{54}$$

Substituting Equation 52 in Equation 51 results in:

$$p_b = -\sqrt{2}\beta\psi + \sqrt{2}\beta\theta + s_w - p_b + s_b - s_w + s_b \tag{55}$$

$$p_b = \frac{\sqrt{2}\beta(\theta - \psi) + 2s_b}{2} \tag{56}$$

Substituting Equation 54 and 56 in $r = \mathbb{1}[p_b > p_w]$:

$$\frac{\sqrt{2}\beta(\theta - \psi) + 2s_b}{2} > \frac{\sqrt{2}\beta(\theta + \psi) + 2s_w}{2} \tag{57}$$

$$\sqrt{2}\beta(\theta - \psi) + 2s_b > \sqrt{2}\beta(\theta + \psi) + 2s_w \tag{58}$$

$$-\sqrt{2}\beta\psi + 2s_b > \sqrt{2}\beta\theta + 2s_w \tag{59}$$

$$2s_b + 2s_w > 2\sqrt{2}\beta\psi \tag{60}$$

$$\mathbb{P}(r = 1 | \theta, \psi, s_b, s_w) = \mathbb{1}\left[\frac{s_b + s_w}{\sqrt{2}\beta} > \psi\right] \tag{61}$$

2.d

$$\mathbb{P}(r = 1|s_b, s_w) = \frac{\mathbb{P}(r = 1, s_b, s_w)}{\mathbb{P}(s_b, s_w)} \tag{62}$$

$$\mathbb{P}(r = 1|s_b, s_w) = \frac{\int_{\theta,\psi} \mathbb{P}(r = 1, \theta, \psi, s_b, s_w)}{\mathbb{P}(s_b, s_w)} \tag{63}$$

$$\mathbb{P}(r = 1|s_b, s_w) = \frac{\int_{\theta,\psi} \mathbb{P}(r = 1|\theta, \psi, s_b, s_w)\mathbb{P}(\theta, \psi, s_b, s_w)}{\mathbb{P}(s_b, s_w)} \tag{64}$$

$$\mathbb{P}(r = 1|s_b, s_w) = \frac{\int_{\theta,\psi} \mathbb{P}(r = 1|\theta, \psi, s_b, s_w)\mathbb{P}(\theta, \psi|s_b, s_w)\mathbb{P}(s_b, s_w)}{\mathbb{P}(s_b, s_w)} \tag{65}$$

Substituting Equation 61 and 50 and using x, μ and Σ from Equation 47, 48 and 49:

$$\mathbb{P}(r = 1|s_b, s_w) = \int_{\theta,\psi} \mathbb{1}\left[\frac{s_b + s_w}{\sqrt{2}\beta} > \psi\right]\mathcal{N}\Big(x|\mu, \Sigma\Big) = \Phi\left[\frac{s_b - s_w}{\sqrt{2}\beta}\right] \tag{66}$$

2.e

3