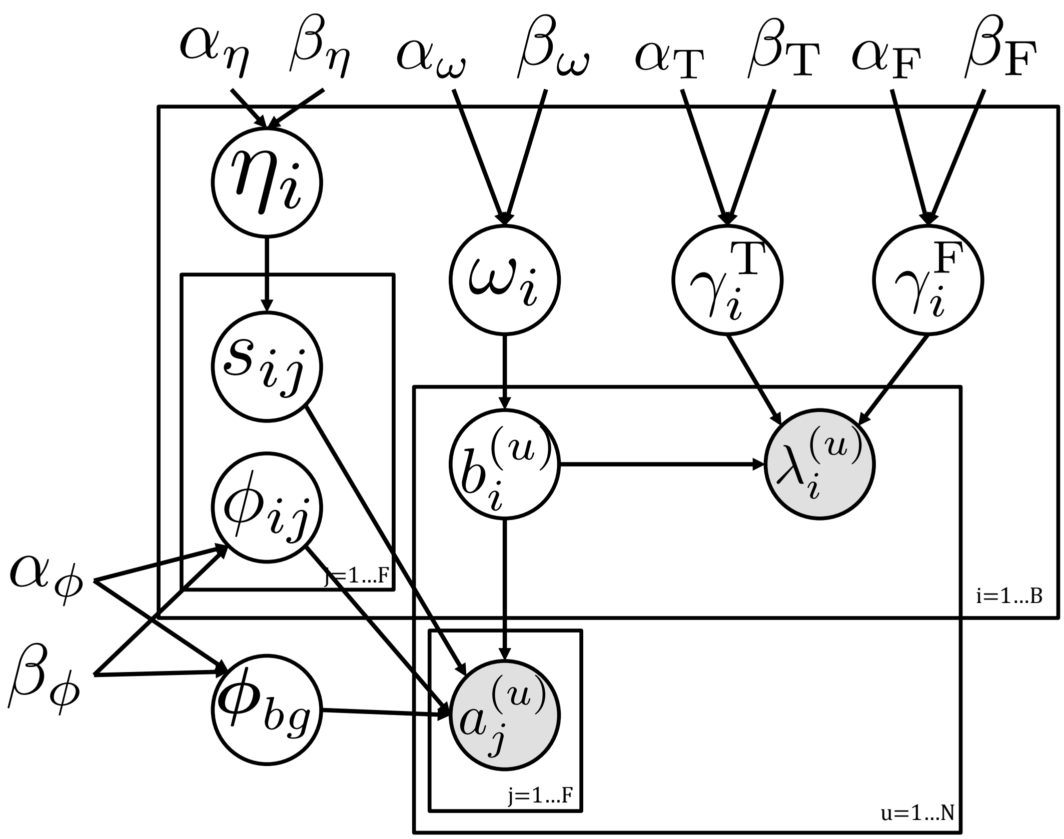
Transparent User Models for Personalization: Appendix

Khalid El-Arini
Ralf Herbrich
Blaise Agüera y Arcas
Ulrich Paquet
Jurgen Van Gael

1 Sampler Derivation



$$\omega_i \sim Beta(\alpha_\omega, \beta_\omega) \tag{1}$$
$$\gamma_i^T \sim Beta(\alpha_T, \beta_T) \tag{2}$$
$$\gamma_i^F \sim Beta(\alpha_F, \beta_F) \tag{3}$$
$$b_i^{(u)} \sim Bernoulli(\omega_i) \tag{4}$$
$$\lambda_i^{(u)} \sim Bernoulli(b_i^{(u)}\gamma_i^T + (1 - b_i^{(u)})\gamma_i^F) \tag{5}$$
$$\eta_i \sim Beta(\alpha_\eta, \beta_\eta) \tag{6}$$
$$s_{ij} \sim Bernoulli(\eta_i) \tag{7}$$
$$\phi_{ij} \sim Beta(\alpha_\phi, \beta_\phi) \tag{8}$$
$$\phi_{bg,j} \sim Beta(\alpha_\phi, \beta_\phi) \tag{9}$$
$$a_j^{(u)} \sim Bernoulli(1 - (1 - \phi_{bg,j}) \prod_{i:b_i^{(u)}=1} (1 - \phi_{ij}s_{ij})) \tag{10}$$

We only sample $b_i^{(u)}$, $\phi_{bg,j}$, ϕ_{ij} and s_{ij}, while collapsing everything else. We assume that $a_j^{(u)}$ and $\lambda_i^{(u)}$ are observed.

1.1 Sampling $b_i^{(u)}$

$$P(b_i^{(u)}|\mathbf{a}^{(u)}, \boldsymbol{\lambda}_i, \mathbf{b}_i^{(-u)}, \boldsymbol{\phi}, \mathbf{s}, \mathbf{b}_{-i}^{(u)}) \tag{11}$$
$$\propto P(\mathbf{a}^{(u)}, \boldsymbol{\lambda}_i|\mathbf{b}_i, \mathbf{b}_{-i}^{(u)}, \boldsymbol{\phi}, \mathbf{s})P(b_i^{(u)}|\mathbf{b}_i^{(-u)}) \tag{12}$$
$$= P(\mathbf{a}^{(u)}|\mathbf{b}_i, \mathbf{b}_{-i}^{(u)}, \boldsymbol{\phi}, \mathbf{s})P(\boldsymbol{\lambda}_i|\mathbf{a}^{(u)}, \mathbf{b}_i, \mathbf{b}_{-i}^{(u)}, \boldsymbol{\phi}, \mathbf{s})P(b_i^{(u)}|\mathbf{b}_i^{(-u)}) \tag{13}$$
$$= \underbrace{P(\mathbf{a}^{(u)}|\mathbf{b}^{(u)}, \boldsymbol{\phi}, \mathbf{s})}_{A}\underbrace{P(\boldsymbol{\lambda}_i|\mathbf{b}_i)}_{B}\underbrace{P(b_i^{(u)}|\mathbf{b}_i^{(-u)})}_{C} \tag{14}$$

On line 14, A and B correspond to the likelihood contributions of $\mathbf{a}^{(u)}$ and $\boldsymbol{\lambda}_i$, respectively. C corresponds to the prior term on $b_i^{(u)}$. We take these each in turn.

A:

$$P(\mathbf{a}^{(u)}|\mathbf{b}^{(u)},\boldsymbol{\phi},\mathbf{s}) = \prod_{j=1}^{F} P(a_j^{(u)}|\mathbf{b}^{(u)},\boldsymbol{\phi}_{\bullet j},\mathbf{s}_{\bullet j}) \tag{15}$$

$$\propto \prod_{j:s_{ij}=1} P(a_j^{(u)}|\mathbf{b}^{(u)},\boldsymbol{\phi}_{\bullet j},\mathbf{s}_{\bullet j}) \tag{16}$$

$$= \prod_{j:s_{ij}=1} Bernoulli(a_j^{(u)}; 1-(1-\phi_{bg,j}) \prod_{k:b_k^{(u)}=1} (1-\phi_{kj}s_{kj})) \tag{17}$$

B:

$$P(\boldsymbol{\lambda}_i|\mathbf{b}_i) = \int_{\gamma_i^T}\int_{\gamma_i^F} P(\boldsymbol{\lambda}_i,\gamma_i^T,\gamma_i^F|\mathbf{b}_i)d\gamma_i^T d\gamma_i^F \tag{18}$$

$$= \int_{\gamma_i^T}\int_{\gamma_i^F} P(\boldsymbol{\lambda}_i|\gamma_i^T,\gamma_i^F,\mathbf{b}_i)P(\gamma_i^T,\gamma_i^F|\mathbf{b}_i)d\gamma_i^T d\gamma_i^F \tag{19}$$

$$= \int_{\gamma_i^T}\int_{\gamma_i^F} \left(\prod_{u=1}^{N} P(\lambda_i^{(u)}|\gamma_i^T,\gamma_i^F,b_i^{(u)})\right) P(\gamma_i^T)P(\gamma_i^F)d\gamma_i^T d\gamma_i^F \tag{20}$$

$$= \int_{\gamma_i^T}\int_{\gamma_i^F} (\gamma_i^T)^{m_{i+}}(1-\gamma_i^T)^{m_{i-}}(\gamma_i^F)^{\overline{m}_{i+}}(1-\gamma_i^F)^{\overline{m}_{i-}} P(\gamma_i^T)P(\gamma_i^F)d\gamma_i^T d\gamma_i^F \tag{21}$$

$$\propto \int_{\gamma_i^T}\int_{\gamma_i^F} (\gamma_i^T)^{m_{i+}+\alpha_T-1}(1-\gamma_i^T)^{m_{i-}+\beta_T-1}(\gamma_i^F)^{\overline{m}_{i+}+\alpha_F-1}(1-\gamma_i^F)^{\overline{m}_{i-}+\beta_F-1}d\gamma_i^T d\gamma_i^F \tag{22}$$

$$= B(m_{i+}+\alpha_T, m_{i-}+\beta_T)B(\overline{m}_{i+}+\alpha_F, \overline{m}_{i-}+\beta_F) \tag{23}$$

$$= \frac{\Gamma(m_{i+}+\alpha_T)\Gamma(m_{i-}+\beta_T)}{\Gamma(\alpha_T+\beta_T+m_{i+}+m_{i-})}\frac{\Gamma(\overline{m}_{i+}+\alpha_F)\Gamma(\overline{m}_{i-}+\beta_F)}{\Gamma(\alpha_F+\beta_F+\overline{m}_{i+}+\overline{m}_{i-})} \tag{24}$$

$$= \frac{\Gamma(m_{i+}+\alpha_T)\Gamma(m_{i-}+\beta_T)\Gamma(\overline{m}_{i+}+\alpha_F)\Gamma(\overline{m}_{i-}+\beta_F)}{\Gamma(\alpha_T+\beta_T+m_i)\Gamma(\alpha_F+\beta_F+N-m_i)}, \tag{25}$$

where:

- m_{i+} is the number of users **with** badge i and **with** rule i,
- m_{i-} is the number of users **with** badge i but **not** rule i,
- $\overline{m}_{i+}$ is the number of users **without** badge i but **with** rule i,
- $\overline{m}_{i-}$ is the number of users **without** badge i and **without** rule i, and,
- m_i is the total number of users with badge i.

Thus, if we are sampling $b_i^{(u)}$ where $\lambda_i^{(u)} = 1$, the expression in (25) takes the following value for $b_i^{(u)} = 0$:

$$\frac{\Gamma(m_{i+}^{(-u)} + \alpha_T)\Gamma(\overline{m}_{i+}^{(-u)} + \alpha_F + 1)}{\Gamma(\alpha_T + \beta_T + m_i^{(-u)})\Gamma(\alpha_F + \beta_F + N - m_i^{(-u)})}, \tag{26}$$

and this value for $b_i^{(u)} = 1$:

$$\frac{\Gamma(m_{i+}^{(-u)} + \alpha_T + 1)\Gamma(\overline{m}_{i+}^{(-u)} + \alpha_F)}{\Gamma(\alpha_T + \beta_T + m_i^{(-u)} + 1)\Gamma(\alpha_F + \beta_F + N - m_i^{(-u)} - 1)}, \tag{27}$$

where the $(-u)$ superscript indicates a value computed ignoring user u. We can simplify out the gamma functions by looking at the ratio of (26) to (27), which gives us,

$$\frac{\overline{m}_{i+}^{(-u)} + \alpha_F}{\alpha_F + \beta_F + N - m_i^{(-u)} - 1} \frac{\alpha_T + \beta_T + m_i^{(-u)}}{m_{i+}^{(-u)} + \alpha_T}. \tag{28}$$

Similarly, this same ratio for the case where $\lambda_i^{(u)} = 0$ ends up being:

$$\frac{\overline{m}_{i-}^{(-u)} + \beta_F}{\alpha_F + \beta_F + N - m_i^{(-u)} - 1} \frac{\alpha_T + \beta_T + m_i^{(-u)}}{m_{i-}^{(-u)} + \beta_T}. \tag{29}$$

C:

$$P(b_i^{(u)}|\mathbf{b}_i^{(-u)}) = \int_{\omega_i} P(b_i^{(u)}, \omega_i|\mathbf{b}_i^{(-u)}) d\omega_i \tag{30}$$

$$= \int_{\omega_i} P(b_i^{(u)}|\omega_i, \mathbf{b}_i^{(-u)}) P(\omega_i|\mathbf{b}_i^{(-u)}) d\omega_i \tag{31}$$

$$= \int_{\omega_i} P(b_i^{(u)}|\omega_i) P(\omega_i|\mathbf{b}_i^{(-u)}) d\omega_i \tag{32}$$

$$\propto \int_{\omega_i} P(b_i^{(u)}|\omega_i) P(\mathbf{b}_i^{(-u)}|\omega_i) P(\omega_i) d\omega_i \tag{33}$$

$$= \int_{\omega_i} P(\omega_i) \prod_{v=1}^{N} P(b_i^{(v)}|\omega_i) d\omega_i \tag{34}$$

$$= \int_{\omega_i} Beta(\omega_i; \alpha_\omega, \beta_\omega) \prod_{v=1}^{N} Bernoulli(b_i^{(v)}; \omega_i) d\omega_i \tag{35}$$

$$= \frac{1}{B(\alpha_\omega, \beta_\omega)} \int_{\omega_i} \omega_i^{\alpha_\omega + m_i - 1} (1 - \omega_i)^{\beta_\omega + N - m_i - 1} d\omega_i \tag{36}$$

$$\propto B(\alpha_\omega + m_i, \beta_\omega + N - m_i), \tag{37}$$

where m_i is the number of total users with badge i (i.e., with $b_i^{(u)} = 1$).

1.2 Sampling $\phi_{bg,j}$

$$P(\phi_{bg,j}|\mathbf{a}_j, \boldsymbol{\phi}_j, \mathbf{s}_j, \mathbf{b}) \tag{38}$$

$$\propto P(\phi_{bg,j}) P(\mathbf{a}_j|\phi_{bg,j}, \boldsymbol{\phi}_j, \mathbf{s}_j, \mathbf{b}) \tag{39}$$

$$= Beta(\phi_{bg,j}; \alpha_\phi, \beta_\phi) \prod_{u=1}^{N} Bernoulli\left(a_j^{(u)}; 1 - (1 - \phi_{bg,j}) \prod_{i: b_i^{(u)}=1} (1 - \phi_{ij} s_{ij})\right) \tag{40}$$

Let's do a Metropolis Hastings step to sample from this conditional, with the following proposal:

$$\phi'_{bg,j}|\phi_{bg,j} = \phi \sim Beta(\phi\nu, (1-\phi)\nu).$$

This is a Beta distribution parameterized by a mean $\mu = \phi$ and a "sample size" ν.

Our acceptance probability is thus:

$$
\begin{aligned}
\rho =& \frac{P(\phi'_{bg,j}|\mathbf{a}_j, \boldsymbol{\phi}_j, \mathbf{s}_j, \mathbf{b})}{P(\phi_{bg,j}|\mathbf{a}_j, \boldsymbol{\phi}_j, \mathbf{s}_j, \mathbf{b})} \frac{Q(\phi_{bg,j}|\phi'_{bg,j})}{Q(\phi'_{bg,j}|\phi_{bg,j})} && (41)\\
=& \frac{Beta(\phi'_{bg,j}; \alpha_\phi, \beta_\phi)}{Beta(\phi_{bg,j}; \alpha_\phi, \beta_\phi)} \\
& \frac{\prod_{u=1}^{N} Bernoulli\left(a_j^{(u)}; 1 - \left(1 - \phi'_{bg,j}\right) \prod_{i:b_i^{(u)}=1} (1 - \phi_{ij} s_{ij})\right)}{\prod_{u=1}^{N} Bernoulli\left(a_j^{(u)}; 1 - (1 - \phi_{bg,j}) \prod_{i:b_i^{(u)}=1} (1 - \phi_{ij} s_{ij})\right)} \\
& \frac{Beta(\phi_{bg,j}; \phi'_{bg,j}\nu, (1 - \phi'_{bg,j})\nu)}{Beta(\phi'_{bg,j}; \phi_{bg,j}\nu, (1 - \phi_{bg,j})\nu)} && (42)\\
=& \frac{(\phi'_{bg,j})^{\alpha_\phi - 1}(1 - \phi'_{bg,j})^{\beta_\phi - 1}}{(\phi_{bg,j})^{\alpha_\phi - 1}(1 - \phi_{bg,j})^{\beta_\phi - 1}} \\
& \left(\frac{1 - \phi'_{bg,j}}{1 - \phi_{bg,j}}\right)^{n_{-j}} \prod_{u:a_j^{(u)}=1} \frac{1 - \left(1 - \phi'_{bg,j}\right) \prod_{i:b_i^{(u)}=1} (1 - \phi_{ij} s_{ij})}{1 - (1 - \phi_{bg,j}) \prod_{i:b_i^{(u)}=1} (1 - \phi_{ij} s_{ij})} \\
& \frac{\Gamma(\phi_{bg,j}\nu)\Gamma((1 - \phi_{bg,j})\nu)}{\Gamma(\phi'_{bg,j}\nu)\Gamma((1 - \phi'_{bg,j})\nu)} \frac{(\phi_{bg,j})^{\phi'_{bg,j}\nu - 1}(1 - \phi_{bg,j})^{(1 - \phi'_{bg,j})\nu - 1}}{(\phi'_{bg,j})^{\phi_{bg,j}\nu - 1}(1 - \phi'_{bg,j})^{(1 - \phi_{bg,j})\nu - 1}} && (43)\\
=& \frac{\Gamma(\phi_{bg,j}\nu)\Gamma((1 - \phi_{bg,j})\nu)}{\Gamma(\phi'_{bg,j}\nu)\Gamma((1 - \phi'_{bg,j})\nu)} \prod_{u:a_j^{(u)}=1} \frac{1 - \left(1 - \phi'_{bg,j}\right) \prod_{i:b_i^{(u)}=1} (1 - \phi_{ij} s_{ij})}{1 - (1 - \phi_{bg,j}) \prod_{i:b_i^{(u)}=1} (1 - \phi_{ij} s_{ij})} \\
& \frac{(\phi_{bg,j})^{\phi'_{bg,j}\nu - \alpha_\phi}(1 - \phi_{bg,j})^{(1 - \phi'_{bg,j})\nu - \beta_\phi - n_{-j}}}{(\phi'_{bg,j})^{\phi_{bg,j}\nu - \alpha_\phi}(1 - \phi'_{bg,j})^{(1 - \phi_{bg,j})\nu - \beta_\phi - n_{-j}}}, && (44)
\end{aligned}
$$

where n_{-j} is the number of users who do not perform action j (i.e., those with $a_j^{(u)} = 0$).

1.3 Sampling s_{ij}

$$
P(s_{ij}|\mathbf{a}_j, s_{-(ij)}, \boldsymbol{\phi}, \mathbf{b}) \propto \underbrace{P(s_{ij}|\mathbf{s}_{i(-j)})}_{A} \underbrace{P(\mathbf{a}_j|\boldsymbol{\phi}, \mathbf{s}, \mathbf{b})}_{B}, \tag{45}
$$

where we expand the two factors as follows:

A:

$$P(s_{ij}|\mathbf{s}_{i(-j)}) = \int_{\eta_i} P(s_{ij}, \eta_i|\mathbf{s}_{i(-j)}) d\eta_i \tag{46}$$

$$= \int_{\eta_i} P(s_{ij}|\eta_i) P(\eta_i|\mathbf{s}_{i(-j)}) d\eta_i \tag{47}$$

$$\propto \int_{\eta_i} P(\eta_i) \prod_{k=1}^{F} P(s_{ik}|\eta_i) d\eta_i \tag{48}$$

$$= \int_{\eta_i} Beta(\eta_i; \alpha_\eta, \beta_\eta) \prod_{k=1}^{F} Bernoulli(s_{ik}; \eta_i) d\eta_i \tag{49}$$

$$\propto \int_{\eta_i} \eta_i^{\alpha_\eta + v_i - 1} (1-\eta_i)^{\beta_\eta + F - v_i - 1} d\eta_i \tag{50}$$

$$= B(\alpha_\eta + v_i, \beta_\eta + F - v_i), \tag{51}$$

where v_i is the number of actions that are active for badge i (i.e., number of $s_{ik} = 1$ for some i).

B:

$$P(\mathbf{a}_j|\boldsymbol{\phi}, \mathbf{s}, \mathbf{b}) = \prod_{u=1}^{N} P(a_j^{(u)}|\boldsymbol{\phi}, \mathbf{s}, \mathbf{b}^{(u)}) \tag{52}$$

$$\propto \prod_{u:b_i^{(u)}=1} P(a_j^{(u)}|\boldsymbol{\phi}, \mathbf{s}, \mathbf{b}^{(u)}) \tag{53}$$

$$= \prod_{u:b_i^{(u)}=1} Bernoulli\left(a_j^{(u)}; 1 - (1-\phi_{bg,j}) \prod_{k:b_k^{(u)}=1} (1-\phi_{kj} s_{kj})\right). \tag{54}$$

We use a Metropolis step to sample this variable, with a deterministic proposal of flipping the value of s_{ij} from s to $\bar{s}$. This gives us the following acceptance probability:

$$
\begin{aligned}
\rho =& \frac{P(s'_{ij}|\mathbf{a}_j, s_{-(ij)}, \boldsymbol{\phi}, \mathbf{b})}{P(s_{ij}|\mathbf{a}_j, s_{-(ij)}, \boldsymbol{\phi}, \mathbf{b})} && (55)\\
=& \frac{B(\alpha_\eta + v'_i, \beta_\eta + F - v'_i)}{B(\alpha_\eta + v_i, \beta_\eta + F - v_i)} \\
& \frac{\prod_{u:b_i^{(u)}=1} Bernoulli\left(a_j^{(u)}; 1 - (1-\phi_{bg,j}) \prod_{k:b_k^{(u)}=1}\left(1 - \phi_{kj} s'_{kj}\right)\right)}{\prod_{u:b_i^{(u)}=1} Bernoulli\left(a_j^{(u)}; 1 - (1-\phi_{bg,j}) \prod_{k:b_k^{(u)}=1}\left(1 - \phi_{kj} s_{kj}\right)\right)} && (56)\\
=& \frac{B(\alpha_\eta + v'_i, \beta_\eta + F - v'_i)}{B(\alpha_\eta + v_i, \beta_\eta + F - v_i)} \\
& \left(\frac{1 - \phi_{ij} s'_{ij}}{1 - \phi_{ij} s_{ij}}\right)^{n_{i,(-j)}} \prod_{u:b_i^{(u)}=1 \wedge a_j^{(u)}=1} \frac{1 - (1-\phi_{bg,j}) \prod_{k:b_k^{(u)}=1}\left(1 - \phi_{kj} s'_{kj}\right)}{1 - (1-\phi_{bg,j}) \prod_{k:b_k^{(u)}=1}\left(1 - \phi_{kj} s_{kj}\right)}, && (57)
\end{aligned}
$$

where $n_{i,(-j)}$ is the number of users that have badge i but do not perform action j. Note that the proposal and reverse proposal terms don't appear here because it's a deterministic proposal.

We can simplify the first fraction further if we look at the two flip cases. When we flip $0 \to 1$, this fraction is:

$$
\frac{\alpha_\eta + v_i^{(-j)}}{\beta_\eta + F - v_i^{(-j)} - 1},
$$

where we write $v_i^{(-j)}$ to represent $v_i - s_{ij}$. Likewise, when we flip $1 \to 0$, we have:

$$
\frac{\beta_\eta + F - v_i^{(-j)} - 1}{\alpha_\eta + v_i^{(-j)}}.
$$

Note that in the case where we flip s_{ij} from 0 to 1, we sample the corresponding ϕ_{ij} from the prior, $Beta(\alpha_\phi, \beta_\phi)$.

1.4 Sampling ϕ_{ij}

$$
\begin{aligned}
P(\phi_{ij}|\mathbf{a}_j, \boldsymbol{\phi}_{-(ij)}, \mathbf{s}, \mathbf{b}) \propto & P(\phi_{ij})P(\mathbf{a}_j|\boldsymbol{\phi}, \mathbf{s}, \mathbf{b}) && (58)\\
\propto & Beta(\phi_{ij}; \alpha_\phi, \beta_\phi) \\
& \prod_{u:b_i^{(u)}=1} Bernoulli\left(a_j^{(u)}; 1-(1-\phi_{bg,j}) \prod_{k:b_k^{(u)}=1} (1-\phi_{kj}s_{kj})\right). && (59)
\end{aligned}
$$

Let's do a Metropolis Hastings step to sample from this conditional, with the same proposal we used for sampling $\phi_{bg,j}$.

$$\phi'_{ij}|\phi_{ij} = \phi \sim Beta(\phi\nu, (1-\phi)\nu).$$

Again, this is a Beta distribution parameterized by a mean $\mu = \phi$ and a "sample siz" ν.

Our acceptance probability is thus:

$$
\begin{aligned}
\rho = & \frac{P(\phi'_{ij}|\mathbf{a}_j, \boldsymbol{\phi}_{-(ij)}, \mathbf{s}, \mathbf{b})}{P(\phi_{ij}|\mathbf{a}_j, \boldsymbol{\phi}_{-(ij)}, \mathbf{s}, \mathbf{b})} \frac{Q(\phi_{ij}|\phi'_{ij})}{Q(\phi'_{ij}|\phi_{ij})} && (60)\\
= & \frac{Beta(\phi'_{ij}; \alpha_\phi, \beta_\phi)}{Beta(\phi_{ij}; \alpha_\phi, \beta_\phi)} \\
& \frac{\prod_{u:b_i^{(u)}=1} Bernoulli\left(a_j^{(u)}; 1-(1-\phi_{bg,j}) \prod_{i:b_i^{(u)}=1}\left(1-\phi'_{ij}s_{ij}\right)\right)}{\prod_{u:b_i^{(u)}=1} Bernoulli\left(a_j^{(u)}; 1-(1-\phi_{bg,j}) \prod_{i:b_i^{(u)}=1}(1-\phi_{ij}s_{ij})\right)} \\
& \frac{Beta(\phi_{ij}; \phi'_{ij}\nu, (1-\phi'_{ij})\nu)}{Beta(\phi'_{ij}; \phi_{ij}\nu, (1-\phi_{ij})\nu)} && (61)\\
= & \frac{\Gamma(\phi_{ij}\nu)\Gamma((1-\phi_{ij})\nu)}{\Gamma(\phi'_{ij}\nu)\Gamma((1-\phi'_{ij})\nu)} \prod_{u:b_i^{(u)}=1 \wedge a_j^{(u)}=1} \frac{1-(1-\phi_{bg,j}) \prod_{i:b_i^{(u)}=1}\left(1-\phi'_{ij}s_{ij}\right)}{1-(1-\phi_{bg,j}) \prod_{i:b_i^{(u)}=1}(1-\phi_{ij}s_{ij})} \\
& \frac{(\phi_{ij})^{\phi'_{ij}\nu-\alpha_\phi}(1-\phi_{ij})^{(1-\phi'_{ij})\nu-\beta_\phi-n_{i,(-j)}}}{(\phi'_{ij})^{\phi_{ij}\nu-\alpha_\phi}(1-\phi'_{ij})^{(1-\phi_{ij})\nu-\beta_\phi-n_{i,(-j)}}}. && (62)
\end{aligned}
$$

Figure 1: Initialization condition for the Apple fanboy badge.

2 Experimental Details

2.1 Hyperparameters

In the experiments described in this paper, we run our model with the following hyperparameter settings:

- $\alpha_\eta = 1, \beta_\eta = 999$: gives us an expected sparsity level per badge of 0.1 percent of all possible actions.
- $\alpha_\omega = 5, \beta_\omega = 25$: indicate that our intuition that a given badge will only be active in a small proportion of users.
- $\alpha_{\mathrm{T}} = 10, \beta_{\mathrm{T}} = 90, \alpha_{\mathrm{F}} = 1, \beta_{\mathrm{F}} = 1000$: encodes our assumption that rules are high precision and low recall, making it very unlikely for a user without a particular badge to activate its corresponding rule.
- $\alpha_\phi = 1, \beta_\phi = 99$: encodes the belief that, on average, we expect that when a badge explains an action (i.e., $s_{ij} > 0$), only 1 % of users with that badge will actually observe the action.

We note that we can also take the fully Bayesian approach and sample our hyperparameters, saving us having to set them in this way.

2.2 Initialization

We initialize the $\mathbf{s}$ and $\boldsymbol{\phi}$ variables for our sampler by assuming that active rules indicate positive examples, and looking at the actions performed by these users. For example, if we want to figure out which actions are explained by the "Apple fanboy" badge, we can look at the users with this rule active, and for each action, compute the proportion of these users that perform it. We then compare this proportion to the overall proportion for this action, across all users. If the within-badge proportion is greater, we assume this badge can explain this action, and initialize $s_{ij} = 1$ and ϕ_{ij} to the within-badge proportion. Figure 1 shows, for example, the initial state for the "Apple fanboy" badge as computed in this way (which can be compared to the posterior learned from our model in Figure 6 of our main paper).

The other variables are initialized from their prior distributions.