Algorithm 1. Cross-MPM	
1	Input: Patient and doctor embeddings learned from consultation dialogue, online medical knowledge and offline hospital visit trace. $\mathrm{E} = \{\mathrm{e}_{\mathrm{mod,\,ent}} \mid \mathrm{mod} \in \{\mathrm{dialogue, online, offline}\}, \mathrm{ent} \in \{\mathrm{patient, doctor, hospital, disease}\}\}$
2	Output: Fused representation $v_{mix} \in R^d$.
3	**procedure** Cross-MPM(E)
4	Divide embeddings E into a group of paired embeddings(x_i, x_j), i ≠ j
5	**for** each pair (x_i, x_j) **do**
6	$v_{i,j} = \mathrm{MPM}(x_i, x_j, c, d)$
7	Compute the attentive factors by formula $f(v_{i,j}) = leaky\_ReLU(v_{i,j} \cdot W_{pair})$ where $W_{pair} \in R^{d\times 1}$ is a projection matrix.
8	Normalize the attentive factors by formula $\alpha_{i,j} = \frac{exp(f(v_{i,j}))}{\sum_{(x_i,x_j)\in G} exp(f(v_{i,j}))}$
9	Weighted fusion $v_{mix} = \sum_{(x_i,x_j)\in G} \alpha_{i,j} v_{i,j}$
10	**Return** v_{mix}
11	**procedure** MPM(x_i, x_j, c, d)
12	**for** $k \leftarrow i, j$ **do**
13	**if** h_k, s_k not initialized **then**
14	Generate random but fixed $h_k \in R^c$ and $s_k \in \{+1. -1\}^c$ where $h_k(i)$ is uniformly drawn from $\{1 \ldots d\}$, s_k is uniformly drawn from $\{-1, +1\}$.
15	$x_k^{'} = \psi(x_k, h_k, s_k, c)$
16	Generate random but fixed $W_k \in R^{d\times c}$, where each entry is -1 or +1 with equal probability.
17	$x_{RM} = \frac{1}{\sqrt{d}}(W_i x_i) \odot (W_j x_j)$
18	$\phi_{RM} = sign(x_{RM})\sqrt{\lvert x_{RM}\rvert}$
19	$x_{TS} = FFT^{-1}\left(FFT(x_i') \odot FFT(x_j')\right)$
20	$\phi_{TS} = sign(x_{TS})\sqrt{\lvert x_{TS}\rvert}$
21	$v_{i,j} = \phi_{RM} + \phi_{TS}$
22	**return** $v_{i,j}$
23	**procedure** $\psi(v, h, s, c)$
24	$y = [0, \ldots, 0]$
25	**for** $i \leftarrow 1, \ldots, c$ **do**
26	$y[h[i]] = y[h[i]] + s[i] \cdot v[i]$
27	**return** y

Note: the embedding size of input $\mathrm{c} = 176$, the output embedding size $\mathrm{d} = 12000$