$$\mathcal{L}_{\text{NCE}}(\{x_i, y_i\}_{i=1}^K; f, h) = \sum_{i=1}^K \log\left(\frac{e^{f(y_i)^T h(x_i)}}{\sum_{j=1}^K e^{f(y_i)^T h(x_j)}}\right) + \sum_{j=1}^K \log\left(\frac{e^{f(y_i)^T h(x_i)}}{\sum_{i=1}^K e^{f(y_i)^T h(x_j)}}\right)$$

$$\mathcal{L}_{\text{BC}}\left(\{s_i, a_i, s_i^+, \ell_i\}_{i=1}^K; \pi\right) = \sum_{i=1}^K \log \pi\left(a_i \mid s_i, \xi(\ell_i)\right) + \log \pi\left(a_i \mid s_i, \psi(s_i^+)\right)$$

$$\mathcal{L}_{\text{TRA}}\left(\{s_i, a_i, s_i^+, g_i, \ell_i\}_{i=1}^K; \pi, \phi, \psi, \xi\right) = \underbrace{\mathcal{L}_{\text{BC}}\left(\{s_i, a_i, s_i^+, \ell_i\}_{i=1}^K; \pi, \psi, \xi\right)}_{\text{behavioral cloning}} + \underbrace{\mathcal{L}_{\text{NCE}}\left(\{s_i, s_i^+\}_{i=1}^K; \phi, \psi\right)}_{\text{temporal alignment}} + \underbrace{\mathcal{L}_{\text{NCE}}\left(\{g_i, \ell_i\}_{i=1}^K; \psi, \xi\right)}_{\text{task alignment}}$$