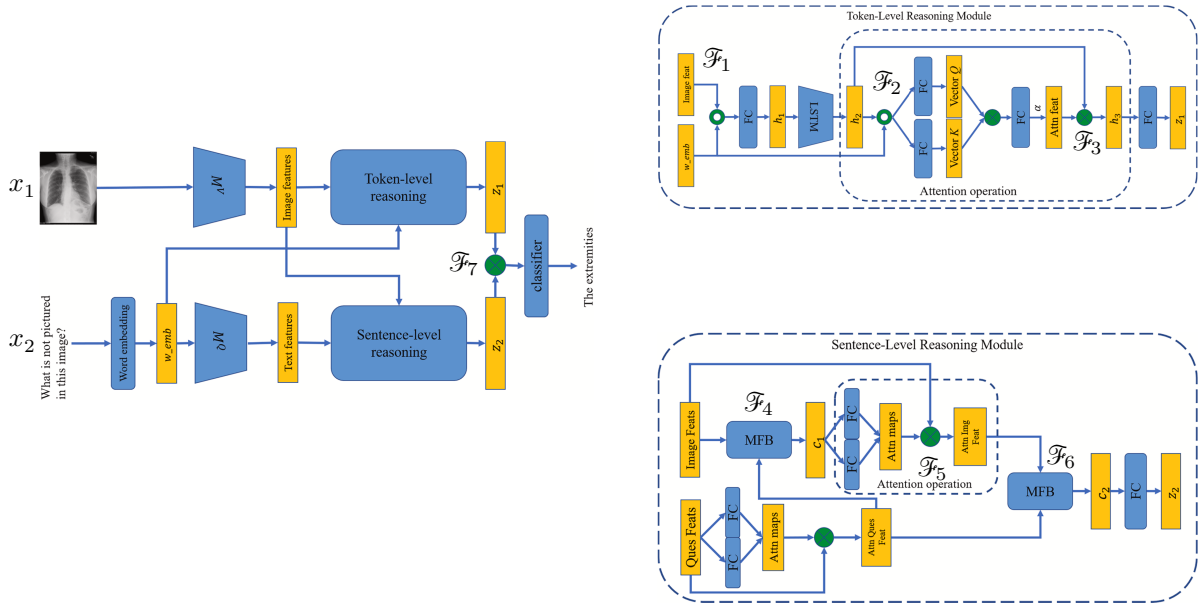
1 A Bi-level representation learning model for medical visual question answering [1]

Model Architecture

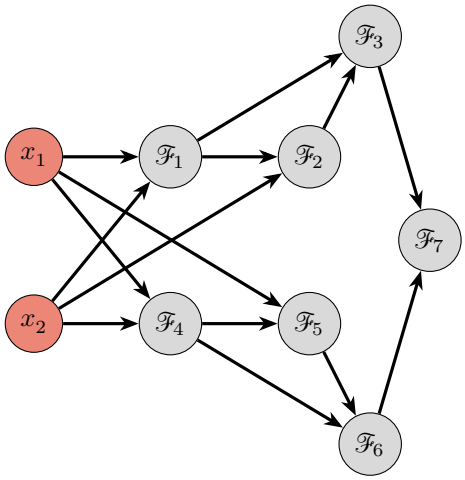


Fusion Formulas

$\mathscr{F}_1 = \oplus(x_1^x, x_2^{y_1})$ $\quad \mathscr{F}_2 = \boxplus(x_2^{y_1}, \mathscr{F}_1^2)$ $\quad \mathscr{F}_3 = \mathbb{A}(\mathscr{F}_1^2, \mathscr{F}_2^1)$ $\quad \mathscr{F}_4 = \boxplus(x_1^x, x_2^{y_2+1})$

$\mathscr{F}_5 = \mathbb{A}(x_1^x, \mathscr{F}_4^1)$ $\quad \mathscr{F}_6 = \boxplus(\mathscr{F}_5^0, x_2^{y_2+1})$ $\quad \mathscr{F}_7 = \oplus(\mathscr{F}_3^1, \mathscr{F}_6^1)_{\rightarrow}$

Fusion Graph Representation



Fusion Analysis

How Many: Single or Multiple?	Multiple, 7
Number of Fusion Flows	2
Multiple Type Sudden, Gradual or Multi-Flow?	Multi-Flow
|Sudden| Synchro?	No