| $$\begin{bmatrix}The\\brown\\fox\\jumps\end{bmatrix}\xrightarrow[tokenizer]{}\begin{bmatrix}\begin{bmatrix} 0.21 & -1.03 & 0.87 & 0.58 & -0.76 \end{bmatrix} \\ \begin{bmatrix} -0.45 & 0.92 & -0.23 & 1.12 & 0.09 \end{bmatrix} \\ \begin{bmatrix} 0.98 & -0.14 & 0.45 & -0.88 & 0.63 \end{bmatrix} \\ \begin{bmatrix} -0.67 & 0.73 & 0.36 & 0.21 & -0.47 \end{bmatrix}\end{bmatrix}$$ |
| --- |
|  |
| $$\begin{bmatrix}pos=0\\pos=1\\pos=2\\pos=3\end{bmatrix}\xrightarrow[encoder]{positional}\begin{bmatrix} \begin{bmatrix} PE{(0,0)} & PE{(0,1)} & PE{(0,2)} & PE{(0,3)} & PE{(0,4)} \end{bmatrix} \\ \begin{bmatrix} PE{(1,0)} & PE{(1,1)} & PE{(1,2)} & PE{(1,3)} & PE{(1,4)} \end{bmatrix} \\ \begin{bmatrix} PE{(2,0)} & PE{(2,1)} & PE{(2,2)} & PE{(2,3)} & PE{(2,4)} \end{bmatrix} \\ \begin{bmatrix} PE{(3,0)} & PE{(3,1)} & PE{(3,2)} & PE{(3,3)} & PE{(3,4)} \end{bmatrix} \end{bmatrix}$$ |
| $$=\begin{bmatrix} \begin{bmatrix} \sin\left(\frac{0}{10000^{0/5}}\right) & \cos\left(\frac{0}{10000^{0/5}}\right) & \sin\left(\frac{0}{10000^{2/5}}\right) & \cos\left(\frac{0}{10000^{2/5}}\right) & \sin\left(\frac{0}{10000^{4/5}}\right) \end{bmatrix} \\ \begin{bmatrix} \sin\left(\frac{1}{10000^{0/5}}\right) & \cos\left(\frac{1}{10000^{0/5}}\right) & \sin\left(\frac{1}{10000^{2/5}}\right) & \cos\left(\frac{1}{10000^{2/5}}\right) & \sin\left(\frac{1}{10000^{4/5}}\right) \end{bmatrix} \\ \begin{bmatrix} \sin\left(\frac{2}{10000^{0/5}}\right) & \cos\left(\frac{2}{10000^{0/5}}\right) & \sin\left(\frac{2}{10000^{2/5}}\right) & \cos\left(\frac{2}{10000^{2/5}}\right) & \sin\left(\frac{2}{10000^{4/5}}\right) \end{bmatrix} \\ \begin{bmatrix} \sin\left(\frac{3}{10000^{0/5}}\right) & \cos\left(\frac{3}{10000^{0/5}}\right) & \sin\left(\frac{3}{10000^{2/5}}\right) & \cos\left(\frac{3}{10000^{2/5}}\right) & \sin\left(\frac{3}{10000^{4/5}}\right) \end{bmatrix} \end{bmatrix}$$ |
|  |
| \[=\begin{bmatrix}\begin{bmatrix}0.000&1.000&0.000&1.000&0.000\end{bmatrix}\\\begin{bmatrix}0.841&1.000&0.001&1.000&0.000\end{bmatrix}\\\begin{bmatrix}0.909&0.999&0.001&1.000&0.000\end{bmatrix}\\\begin{bmatrix}0.141&0.997&0.002&1.000&0.000\end{bmatrix}\end{bmatrix}(3\:d.p.)\]  $$=\begin{bmatrix}\begin{bmatrix}0.000&1.000&0.000&1.000&0.000\end{bmatrix}\\\begin{bmatrix}0.841&1.000&0.001&1.000&0.000\end{bmatrix}\\\begin{bmatrix}0.909&0.999&0.001&1.000&0.000\end{bmatrix}\\\begin{bmatrix}0.141&0.997&0.002&1.000&0.000\end{bmatrix}\end{bmatrix}(3\:d.p.)$$ |
| $$\begin{bmatrix}0.000&1.000&0.000&1.000&0.000\\0.841&1.000&0.001&1.000&0.000\\0.909&0.999&0.001&1.000&0.000\\0.141&0.997&0.002&1.000&0.000\\\end{bmatrix}+\begin{bmatrix}0.21&-1.03&0.87&0.58&-0.76\\-0.45&0.92&-0.23&1.12&0.09\\0.98&-0.14&0.45&-0.88&0.63\\-0.67&0.73&0.36&0.21&-0.47\\\end{bmatrix}$$  $$=\begin{bmatrix}0.21&-0.03&0.87&1.58&-0.76\\0.39&1.92&-0.23&2.12&0.09\\1.89&0.86&0.45&0.12&0.63\\-0.53&1.73&0.36&1.21&-0.47\\\end{bmatrix}$$ |
| $$\begin{bmatrix}0.21&-0.03&0.87&1.58&-0.76\\0.39&1.92&-0.23&2.12&0.09\\1.89&0.86&0.45&0.12&0.63\\-0.53&1.73&0.36&1.21&-0.47\\\end{bmatrix}\xrightarrow[attention]{multi-head}\begin{bmatrix}0.18&-0.02&0.68&1.37&-0.66\\0.07&0.18&-0.02&0.32&0.01\\0.19&0.05&0.03&0.12&0.04\\-0.09&0.31&0.02&0.23&-0.06\end{bmatrix}$$ |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |
|  |