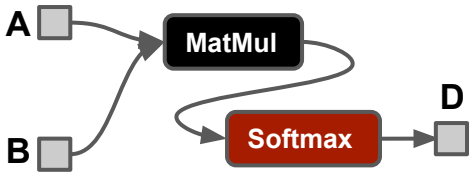
The mathmetical expressions:

$$M[i, j] = \sum_{K} A[i, k] \times B[k, j]$$

$$D[i, j] = \text{softmax}(M[i, j])$$

The computation graph:



Ansor

Generated Sketch 1:

```
[ Placeholder: A, B
for i.0 (None)
 for j.0 (None)
  ...
   for i_c.2 (None)
    for j_c.2 (None)
     for k.0 (None)
      for k.1 (None)
       for k.2 (None)
        C.local = ...
     for i.3 (None)
      for j.3 (None)
       C = ...

for i0 (0,1050)
 for k.0 (0,33)
  threadIdx.x k.1 (0,32)
   T_softmax_maxelem = ...
for i0 (0,1050)
 for k (0,1050)
  T_softmax_expsum = ...
for i0 (0,1050)
 for i1 (0,1050)
  T_softmax_norm = ...]
```

Generated Sketch 2:

```
[ Placeholder: A, B
for i.0 (None)
 for j.0 (None)
  ...
   for i_c.2 (None)
    for j_c.2 (None)
     for k.0 (None)
      for k.1 (None)
       for k.2 (None)
        C.local = ...
     for i0 (0,1050)
      for k.0 (0,33)
       threadIdx.x k.1 (0,32)
        T_softmax_maxelem = ...
     for i0 (0,1050)
      for k (0,1050)
       T_softmax_expsum = ...
     for i0 (0,1050)
      for i1 (0,1050)
       T_softmax_norm = ...
     for i.3 (None)
      for j.3 (None)
       C = ... ]
```

Our Sketch Customization and Policy