MLP SoME 8 big fixes Results

--- Starting Experiment: v3_baseline_orthogonal ---

--- Part 2: Data Preparation & Configuration ---
Training custom tokenizer...

README.md:
 1.06k/? [00:00<00:00, 97.9kB/s]
data/train-00000-of-00004-2d5a1467fff108(…): 100%
 249M/249M [00:02<00:00, 155MB/s]
data/train-00001-of-00004-5852b56a2bd28f(…): 100%
 248M/248M [00:01<00:00, 247MB/s]
data/train-00002-of-00004-a26307300439e9(…): 100%
 246M/246M [00:01<00:00, 120MB/s]
data/train-00003-of-00004-d243063613e5a0(…): 100%
 248M/248M [00:01<00:00, 186MB/s]
data/validation-00000-of-00001-869c898b5(…): 100%
 9.99M/9.99M [00:00<00:00, 13.6MB/s]
Generating train split: 100%
 2119719/2119719 [00:06<00:00, 332858.38 examples/s]
Generating validation split: 100%
 21990/21990 [00:00<00:00, 298206.05 examples/s]
Custom tokenizer loaded with vocab size: 8192

Tokenizing dataset...

Map (num_proc=12): 100%
 10000/10000 [00:03<00:00, 5454.01 examples/s]
Map (num_proc=12): 100%
 1000/1000 [00:00<00:00, 1614.70 examples/s]
Train dataset size (subset): 10000
Using 6 workers for DataLoader.

--- Part 3: Model Definition ---
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}

SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}

Compiling the model for faster training...

/tmp/ipython-input-310555793.py:454: FutureWarning: `torch.cuda.amp.GradScaler(args...)` is deprecated. Please use `torch.amp.GradScaler('cuda', args...)` instead.
  scaler = torch.cuda.amp.GradScaler()

--- Part 4: Training, Evaluation, and Metrics ---

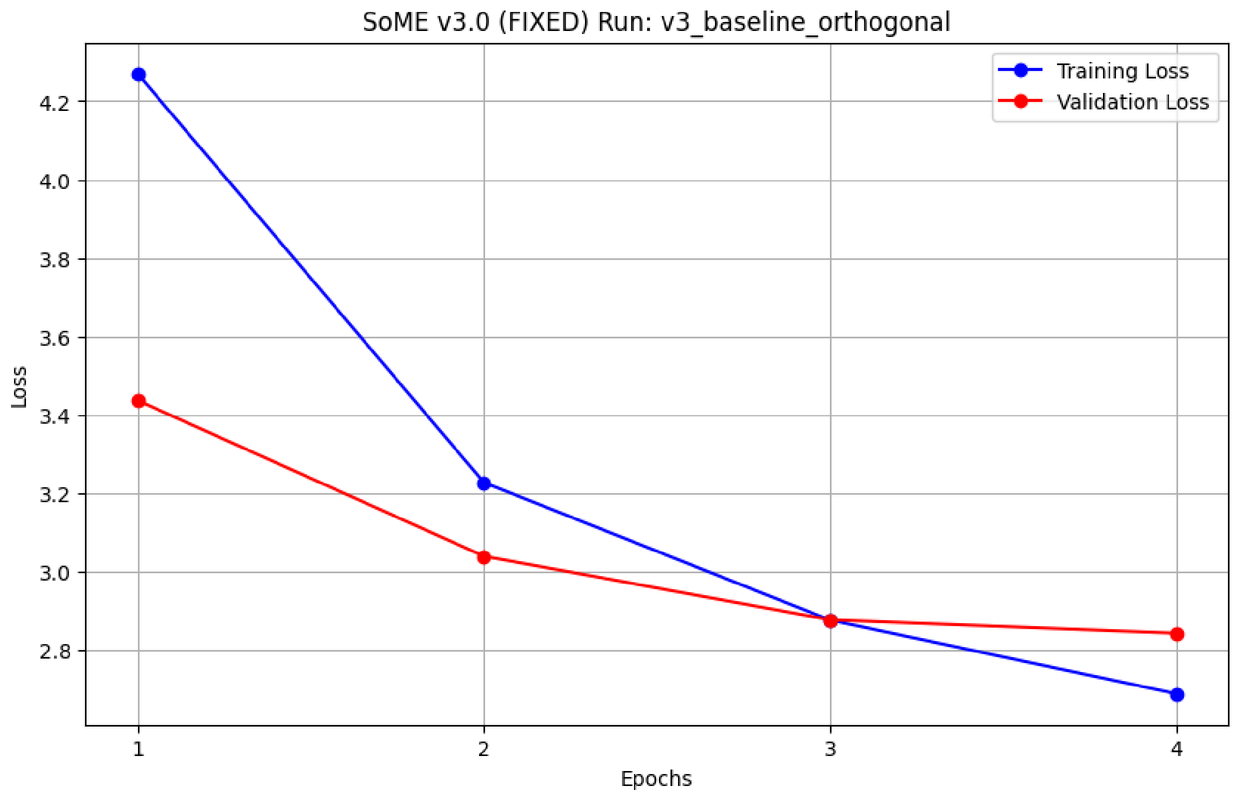
Total parameters: 2711.81M
Trainable parameters: 24.18M (0.89%)
Total training steps: 1664
Using expert initialization method: sparse

--- Epoch 1/4 ---

Training: 0%| | 0/416 [00:00<?, ?it/s]/tmp/ipython-input-310555793.py:462: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Training: 0%| | 1/416 [00:38<4:24:12, 38.20s/it, loss=9.1895, lr=6.0e-04]/tmp/ipython-input-310555793.py:462: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Training: 0%| | 2/416 [00:41<2:02:13, 17.71s/it, loss=8.5270, lr=6.0e-04]/tmp/ipython-input-310555793.py:462: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Training: 1%| | 5/416 [00:44<31:34, 4.61s/it, loss=7.2839, lr=6.0e-04]/tmp/ipython-input-310555793.py:462: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Evaluating: 0%| | 0/42 [00:00<?, ?it/s]/tmp/ipython-input-310555793.py:495: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Evaluating: 2%|▏ | 1/42 [00:07<04:59, 7.30s/it, loss=3.2004]/tmp/ipython-input-310555793.py:495: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.

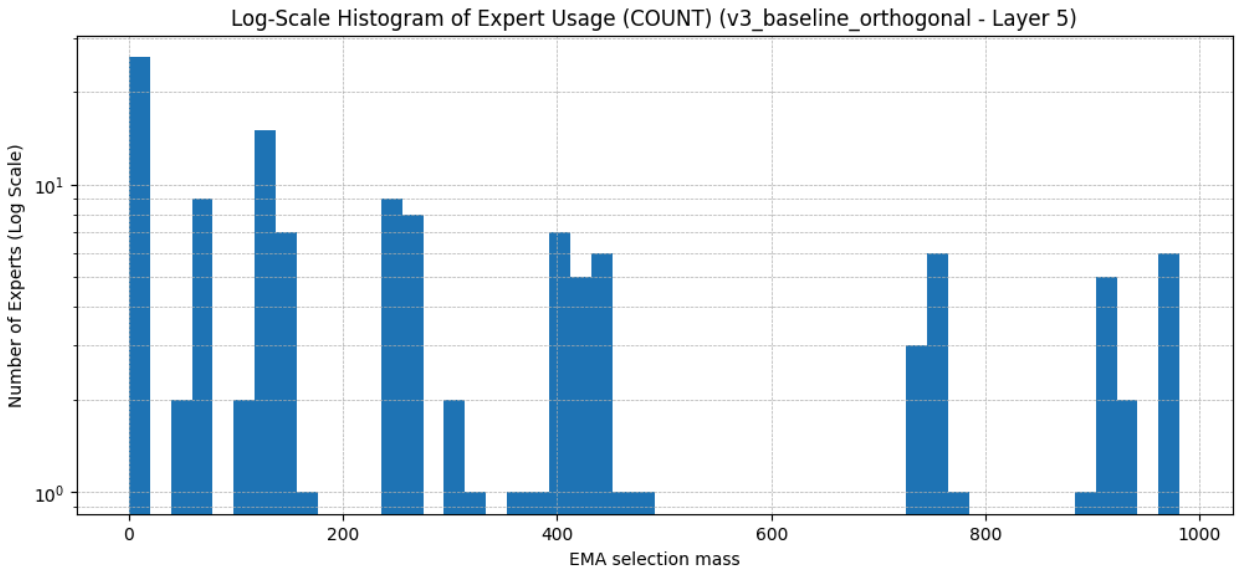
```
  with torch.cuda.amp.autocast():

Epoch 1: Train Loss = 4.2707, Val Loss = 3.4367, Val Perplexity = 31.09
  Middle Layer (count): Gini = 0.492, Entropy = 6.355
  Middle Layer (mass):  Gini = 0.491, Entropy = 6.357

/tmp/ipython-input-310555793.py:454: FutureWarning: `torch.cuda.amp.GradScaler(args...)` is
deprecated. Please use `torch.amp.GradScaler('cuda', args...)` instead.
  scaler = torch.cuda.amp.GradScaler()

 Model saved as best_model_v3_baseline_orthogonal_fixed.pth

--- Epoch 2/4 ---

Training:   0%|          | 0/416 [00:00<?, ?it/s]/tmp/ipython-input-310555793.py:462:
FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use
`torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Evaluating:   0%|          | 0/42 [00:00<?, ?it/s]/tmp/ipython-input-310555793.py:495:
FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use
`torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():

Epoch 2: Train Loss = 3.2273, Val Loss = 3.0392, Val Perplexity = 20.89
  Middle Layer (count): Gini = 0.569, Entropy = 6.175
  Middle Layer (mass):  Gini = 0.568, Entropy = 6.176
  Model saved as best_model_v3_baseline_orthogonal_fixed.pth

--- Epoch 3/4 ---

Epoch 3: Train Loss = 2.8759, Val Loss = 2.8767, Val Perplexity = 17.76
  Middle Layer (count): Gini = 0.545, Entropy = 6.243
  Middle Layer (mass):  Gini = 0.544, Entropy = 6.245
  Model saved as best_model_v3_baseline_orthogonal_fixed.pth

--- Epoch 4/4 ---

Epoch 4: Train Loss = 2.6881, Val Loss = 2.8424, Val Perplexity = 17.16
  Middle Layer (count): Gini = 0.531, Entropy = 6.277
  Middle Layer (mass):  Gini = 0.530, Entropy = 6.279
  Model saved as best_model_v3_baseline_orthogonal_fixed.pth

--- Training Complete for v3_baseline_orthogonal ---
```

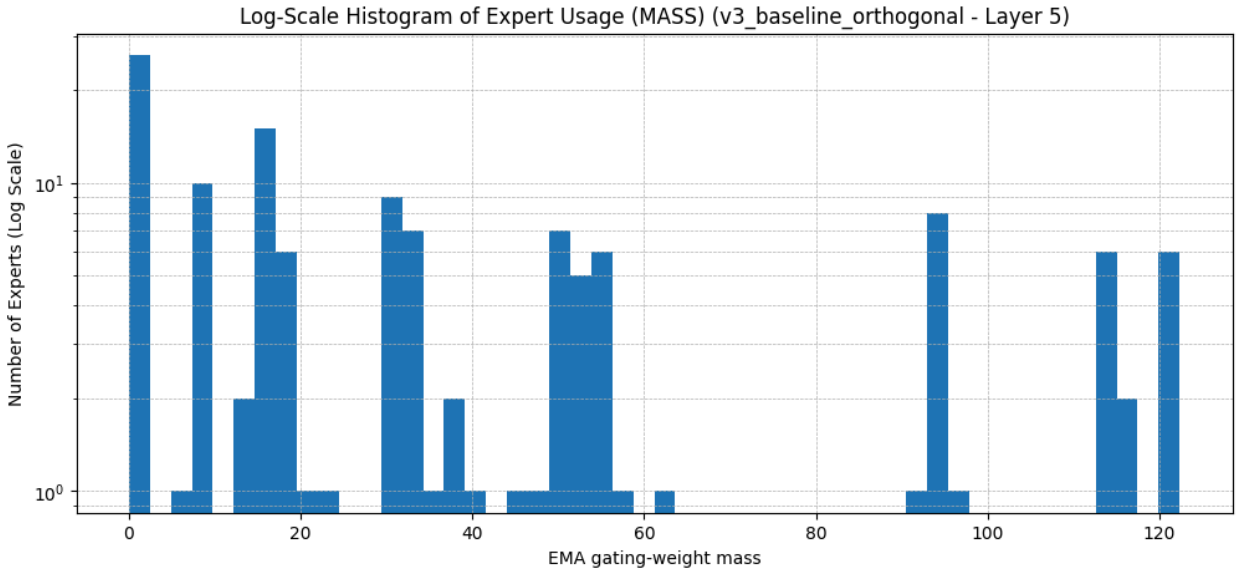
Loss curve plot saved to loss_curve_v3_baseline_orthogonal_fixed.png



--- Part 1: Dashboard Setup ---
Loading best model from: best_model_v3_baseline_orthogonal_fixed.pth
Loading tokenizer from: tinystories-tokenizer-v2.json
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}

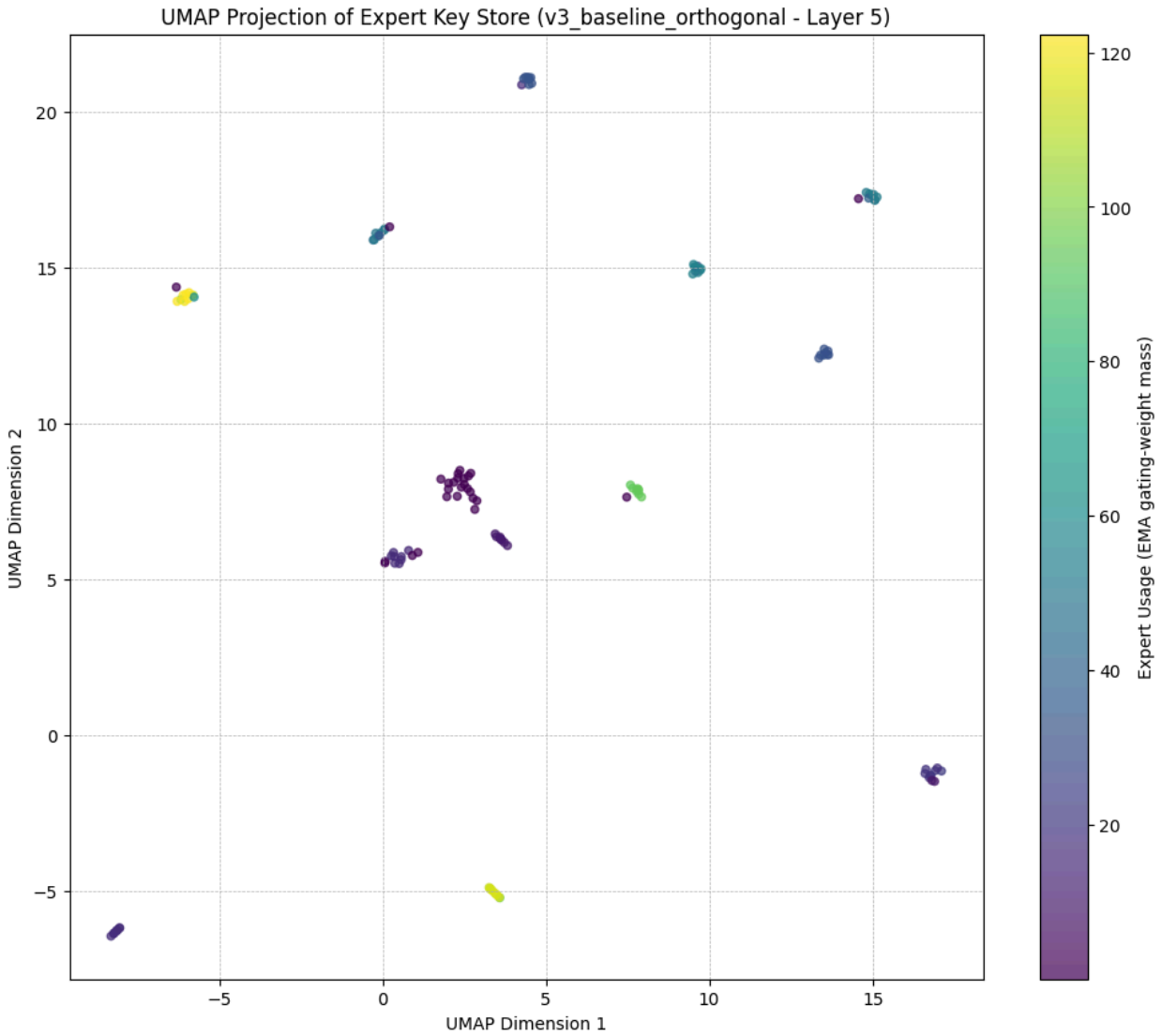
--- Part 2: Aggregate Utilization Analysis (from Middle Layer) ---
Expert Usage (Layer 5): 128/128 (100.00%)
Final Metrics (count) Layer 5: Gini=0.5306, Entropy=6.2774 (Max=7.0000)
Final Metrics (mass) Layer 5: Gini=0.5302, Entropy=6.2786 (Max=7.0000)





--- Part 3: Key Store Structure Visualization (from Middle Layer) ---
Running UMAP projection on the key store... (this may take a moment)

/usr/local/lib/python3.12/dist-packages/umap/umap_.py:1952: UserWarning: n_jobs value 1 overridden to 1 by setting random_state. Use no seed for parallelism.
  warn(



--- Part 4: Multi-Layer Generative Analysis with Expert Tracing ---

--- Prompt ---
Once upon a time, there was a little fox wholived
/tmp/ipython-input-310555793.py:121: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast(enabled=False):

inabighouse.Hewasveryhappyandlovedtoplaywithhisfriends.Oneday,hesawabig,shinyrockinthefo
rest.Hewassoexcitedtoseewhatitwas.Hewantedtoseewhatitwas.
--- End of Generation ---

--- Multi-Layer Expert Activation Trace ---

Token 'lived':

Layer 1: Used Experts -> [3, 121, 80, 41, 55, 117, 87, 47]
Layer 5: Used Experts -> [34, 93, 56, 119, 32, 124, 38, 35]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'in':
Layer 1: Used Experts -> [27, 90, 110, 72, 59, 98, 42, 56]
Layer 5: Used Experts -> [12, 50, 77, 99, 91, 57, 39, 7]
Layer 9: Used Experts -> [89, 49, 70, 56, 66, 78, 68, 106]

Token 'a':
Layer 1: Used Experts -> [57, 119, 122, 4, 16, 85, 49, 30]
Layer 5: Used Experts -> [73, 7, 39, 57, 91, 77, 99, 50]
Layer 9: Used Experts -> [70, 89, 56, 66, 68, 78, 49, 106]

Token 'big':
Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 92, 8]
Layer 5: Used Experts -> [67, 68, 125, 123, 101, 92, 94, 98]
Layer 9: Used Experts -> [16, 82, 52, 122, 38, 39, 23, 7]

Token 'house':
Layer 1: Used Experts -> [72, 90, 87, 98, 113, 42, 56, 111]
Layer 5: Used Experts -> [81, 96, 95, 71, 52, 102, 10, 8]
Layer 9: Used Experts -> [7, 39, 23, 38, 122, 52, 82, 16]

Token '.':
Layer 1: Used Experts -> [112, 1, 34, 66, 115, 96, 63, 39]
Layer 5: Used Experts -> [26, 82, 16, 43, 45, 28, 64, 87]
Layer 9: Used Experts -> [15, 110, 4, 27, 109, 8, 98, 65]

Token 'He':
Layer 1: Used Experts -> [114, 101, 11, 21, 68, 51, 100, 20]
Layer 5: Used Experts -> [10, 102, 52, 71, 96, 95, 8, 81]
Layer 9: Used Experts -> [81, 30, 107, 50, 22, 3, 36, 83]

Token 'was':
Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 8, 92]
Layer 5: Used Experts -> [34, 93, 119, 35, 38, 32, 124, 56]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'very':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [31, 42, 3, 61, 20, 37, 104, 48]
Layer 9: Used Experts -> [106, 78, 49, 68, 66, 56, 70, 61]

Token 'happy':
  Layer 1: Used Experts -> [48, 5, 83, 69, 126, 28, 78, 65]
  Layer 5: Used Experts -> [67, 125, 123, 94, 92, 101, 68, 98]
  Layer 9: Used Experts -> [9, 121, 87, 127, 112, 111, 104, 53]

Token 'and':
  Layer 1: Used Experts -> [41, 47, 3, 121, 55, 80, 117, 87]
  Layer 5: Used Experts -> [68, 101, 98, 92, 123, 125, 94, 67]
  Layer 9: Used Experts -> [98, 15, 110, 27, 65, 4, 109, 8]

Token 'loved':
  Layer 1: Used Experts -> [109, 64, 102, 45, 70, 97, 93, 94]
  Layer 5: Used Experts -> [106, 70, 34, 24, 55, 1, 46, 72]
  Layer 9: Used Experts -> [99, 49, 78, 68, 106, 66, 56, 70]

Token 'to':
  Layer 1: Used Experts -> [1, 34, 66, 115, 96, 63, 39, 91]
  Layer 5: Used Experts -> [73, 82, 26, 7, 39, 91, 57, 99]
  Layer 9: Used Experts -> [49, 78, 68, 106, 66, 56, 70, 89]

Token 'play':
  Layer 1: Used Experts -> [52, 24, 107, 22, 38, 124, 7, 104]
  Layer 5: Used Experts -> [19, 9, 40, 97, 116, 79, 25, 118]
  Layer 9: Used Experts -> [79, 100, 0, 77, 42, 11, 120, 118]

Token 'with':
  Layer 1: Used Experts -> [29, 71, 125, 44, 14, 43, 23, 120]
  Layer 5: Used Experts -> [67, 94, 125, 123, 92, 98, 101, 68]
  Layer 9: Used Experts -> [59, 93, 67, 91, 26, 119, 89, 5]

Token 'his':
  Layer 1: Used Experts -> [49, 85, 16, 4, 122, 119, 30, 57]
  Layer 5: Used Experts -> [67, 94, 125, 123, 92, 98, 101, 68]
  Layer 9: Used Experts -> [61, 106, 78, 68, 49, 66, 56, 70]

Token 'friends':
  Layer 1: Used Experts -> [120, 14, 71, 125, 43, 44, 23, 29]
  Layer 5: Used Experts -> [67, 94, 125, 123, 92, 98, 101, 68]
  Layer 9: Used Experts -> [16, 82, 52, 122, 38, 23, 7, 39]

Token '.':
  Layer 1: Used Experts -> [78, 65, 83, 5, 126, 28, 69, 48]
  Layer 5: Used Experts -> [108, 98, 94, 92, 123, 125, 101, 68]
  Layer 9: Used Experts -> [65, 27, 110, 15, 4, 109, 8, 86]

Token 'One':
  Layer 1: Used Experts -> [114, 101, 11, 21, 68, 51, 100, 20]
  Layer 5: Used Experts -> [10, 102, 52, 71, 96, 95, 8, 81]
  Layer 9: Used Experts -> [83, 36, 3, 22, 50, 107, 30, 81]

Token 'day':
  Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 92, 8]
  Layer 5: Used Experts -> [37, 42, 61, 20, 3, 48, 104, 31]
  Layer 9: Used Experts -> [16, 7, 38, 23, 52, 39, 82, 122]

Token ',':
  Layer 1: Used Experts -> [23, 120, 43, 44, 14, 125, 71, 29]
  Layer 5: Used Experts -> [68, 101, 67, 98, 92, 125, 123, 94]
  Layer 9: Used Experts -> [59, 67, 91, 93, 26, 119, 5, 58]

Token 'he':
  Layer 1: Used Experts -> [17, 92, 8, 116, 82, 6, 36, 89]
  Layer 5: Used Experts -> [25, 116, 97, 40, 9, 79, 19, 118]
  Layer 9: Used Experts -> [106, 66, 56, 68, 70, 78, 49, 89]

Token 'saw':
  Layer 1: Used Experts -> [29, 71, 125, 44, 14, 43, 23, 120]
  Layer 5: Used Experts -> [34, 93, 119, 35, 38, 124, 32, 56]
  Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'a':
  Layer 1: Used Experts -> [19, 25, 0, 12, 73, 108, 46, 33]
  Layer 5: Used Experts -> [99, 50, 91, 77, 57, 39, 7, 73]
  Layer 9: Used Experts -> [106, 78, 68, 49, 66, 56, 70, 89]

Token 'big':
  Layer 1: Used Experts -> [53, 92, 89, 8, 36, 116, 6, 82]
  Layer 5: Used Experts -> [67, 68, 101, 125, 92, 123, 98, 94]
  Layer 9: Used Experts -> [16, 82, 52, 38, 7, 122, 23, 39]

Token ',':
  Layer 1: Used Experts -> [87, 117, 80, 121, 55, 3, 47, 41]
  Layer 5: Used Experts -> [67, 68, 101, 125, 92, 98, 123, 94]
  Layer 9: Used Experts -> [7, 23, 39, 38, 52, 122, 82, 16]

Token 'shiny':
  Layer 1: Used Experts -> [17, 92, 8, 116, 82, 6, 36, 89]
  Layer 5: Used Experts -> [25, 116, 9, 40, 97, 79, 19, 118]

Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'rock':
Layer 1: Used Experts -> [109, 70, 64, 102, 45, 93, 94, 97]
Layer 5: Used Experts -> [94, 123, 92, 125, 98, 101, 68, 67]
Layer 9: Used Experts -> [62, 7, 57, 23, 37, 39, 38, 97]

Token 'in':
Layer 1: Used Experts -> [83, 5, 48, 69, 126, 28, 78, 65]
Layer 5: Used Experts -> [65, 6, 100, 107, 63, 27, 51, 69]
Layer 9: Used Experts -> [10, 86, 8, 4, 15, 41, 109, 110]

Token 'the':
Layer 1: Used Experts -> [57, 119, 122, 4, 16, 85, 49, 30]
Layer 5: Used Experts -> [7, 99, 91, 39, 57, 50, 77, 73]
Layer 9: Used Experts -> [89, 70, 56, 66, 68, 49, 78, 106]

Token 'forest':
Layer 1: Used Experts -> [12, 46, 33, 73, 108, 19, 0, 25]
Layer 5: Used Experts -> [10, 102, 81, 52, 71, 96, 95, 8]
Layer 9: Used Experts -> [17, 25, 1, 43, 103, 75, 116, 124]

Token '.':
Layer 1: Used Experts -> [112, 83, 5, 48, 126, 69, 78, 28]
Layer 5: Used Experts -> [51, 100, 107, 27, 63, 69, 6, 65]
Layer 9: Used Experts -> [86, 15, 8, 4, 109, 110, 27, 65]

Token 'He':
Layer 1: Used Experts -> [114, 11, 68, 101, 21, 51, 20, 100]
Layer 5: Used Experts -> [10, 102, 52, 8, 71, 96, 95, 81]
Layer 9: Used Experts -> [83, 36, 3, 22, 50, 107, 30, 81]

Token 'was':
Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 8, 92]
Layer 5: Used Experts -> [34, 93, 119, 35, 38, 124, 32, 56]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'so':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [31, 104, 3, 61, 42, 20, 48, 37]
Layer 9: Used Experts -> [89, 70, 49, 56, 66, 68, 78, 106]

Token 'excited':
Layer 1: Used Experts -> [48, 65, 126, 69, 28, 5, 78, 83]

Layer 5: Used Experts -> [67, 125, 94, 123, 92, 98, 101, 68]
Layer 9: Used Experts -> [99, 49, 78, 68, 106, 66, 56, 70]

Token 'to':
Layer 1: Used Experts -> [112, 76, 113, 111, 59, 56, 42, 98]
Layer 5: Used Experts -> [51, 100, 107, 69, 63, 27, 6, 12]
Layer 9: Used Experts -> [119, 91, 67, 26, 93, 5, 59, 58]

Token 'see':
Layer 1: Used Experts -> [52, 24, 107, 22, 124, 38, 7, 104]
Layer 5: Used Experts -> [86, 19, 9, 40, 97, 116, 79, 25]
Layer 9: Used Experts -> [79, 0, 42, 100, 77, 11, 120, 118]

Token 'what':
Layer 1: Used Experts -> [62, 88, 119, 122, 85, 16, 4, 49]
Layer 5: Used Experts -> [12, 6, 69, 63, 27, 107, 100, 51]
Layer 9: Used Experts -> [61, 106, 78, 68, 49, 66, 56, 70]

Token 'it':
Layer 1: Used Experts -> [65, 28, 69, 5, 48, 126, 83, 78]
Layer 5: Used Experts -> [108, 118, 19, 9, 79, 40, 97, 116]
Layer 9: Used Experts -> [89, 70, 56, 66, 49, 68, 78, 106]

Token 'was':
Layer 1: Used Experts -> [78, 65, 83, 5, 28, 69, 126, 48]
Layer 5: Used Experts -> [81, 96, 95, 71, 52, 102, 10, 8]
Layer 9: Used Experts -> [41, 10, 18, 62, 97, 51, 57, 37]

Token '.':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [31, 104, 3, 61, 42, 20, 48, 37]
Layer 9: Used Experts -> [93, 59, 67, 119, 91, 26, 5, 58]

Token 'He':
Layer 1: Used Experts -> [20, 100, 51, 68, 11, 21, 101, 114]
Layer 5: Used Experts -> [10, 102, 52, 71, 96, 8, 95, 81]
Layer 9: Used Experts -> [83, 36, 3, 22, 50, 107, 30, 81]

Token 'wanted':
Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 8, 92]
Layer 5: Used Experts -> [34, 93, 119, 35, 38, 124, 32, 56]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'to':

  Layer 1: Used Experts -> [23, 120, 43, 44, 14, 125, 71, 29]
  Layer 5: Used Experts -> [73, 7, 39, 57, 91, 77, 99, 50]
  Layer 9: Used Experts -> [49, 56, 70, 66, 68, 78, 106, 89]

Token 'see':
  Layer 1: Used Experts -> [52, 24, 107, 22, 124, 38, 7, 104]
  Layer 5: Used Experts -> [19, 9, 40, 97, 116, 79, 25, 118]
  Layer 9: Used Experts -> [42, 0, 79, 100, 77, 120, 118, 11]

Token 'what':
  Layer 1: Used Experts -> [62, 88, 119, 122, 85, 16, 4, 49]
  Layer 5: Used Experts -> [12, 6, 69, 63, 27, 107, 100, 51]
  Layer 9: Used Experts -> [61, 118, 120, 11, 77, 42, 100, 0]

Token 'it':
  Layer 1: Used Experts -> [65, 28, 69, 5, 48, 126, 83, 78]
  Layer 5: Used Experts -> [108, 19, 118, 9, 79, 40, 97, 116]
  Layer 9: Used Experts -> [89, 70, 56, 66, 49, 68, 78, 106]

Token 'was':
  Layer 1: Used Experts -> [65, 78, 83, 5, 28, 69, 126, 48]
  Layer 5: Used Experts -> [81, 95, 96, 71, 52, 102, 8, 10]
  Layer 9: Used Experts -> [10, 41, 51, 18, 97, 37, 57, 62]

Token '.':
  Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
  Layer 5: Used Experts -> [31, 104, 3, 61, 42, 20, 48, 37]
  Layer 9: Used Experts -> [93, 59, 67, 119, 91, 26, 5, 65]

--- Prompt ---
The recipe for the perfect cake is to
firsttheotherchildrentolook.Hewasveryexcitedtoseethebigandhewasveryexcited.Hewantedtosee
whatitwas.Heaskedtheotherchildren,"Whatareyoudoing?"Thechildrenreplied,"I'mgoingto
--- End of Generation ---

--- Multi-Layer Expert Activation Trace ---

Token 'the':
  Layer 1: Used Experts -> [87, 117, 80, 121, 3, 55, 47, 41]
  Layer 5: Used Experts -> [12, 65, 6, 63, 69, 27, 107, 100]
  Layer 9: Used Experts -> [89, 70, 56, 66, 49, 68, 78, 106]

Token 'other':
  Layer 1: Used Experts -> [12, 46, 33, 73, 108, 19, 0, 25]

Layer 5: Used Experts -> [10, 102, 52, 71, 96, 81, 95, 8]
Layer 9: Used Experts -> [124, 116, 75, 103, 43, 25, 1, 17]

Token 'children':
Layer 1: Used Experts -> [99, 10, 67, 13, 40, 50, 74, 9]
Layer 5: Used Experts -> [31, 104, 3, 61, 20, 42, 48, 37]
Layer 9: Used Experts -> [15, 98, 110, 4, 8, 109, 27, 86]

Token 'to':
Layer 1: Used Experts -> [87, 117, 80, 121, 55, 3, 47, 41]
Layer 5: Used Experts -> [16, 45, 28, 64, 43, 87, 82, 26]
Layer 9: Used Experts -> [86, 8, 109, 4, 27, 110, 65, 15]

Token 'look':
Layer 1: Used Experts -> [52, 24, 107, 22, 124, 7, 38, 104]
Layer 5: Used Experts -> [19, 118, 9, 79, 97, 40, 116, 25]
Layer 9: Used Experts -> [61, 118, 120, 11, 42, 77, 0, 100]

Token '.':
Layer 1: Used Experts -> [112, 83, 87, 48, 5, 78, 126, 69]
Layer 5: Used Experts -> [108, 94, 98, 92, 123, 101, 125, 68]
Layer 9: Used Experts -> [15, 110, 98, 27, 4, 65, 109, 8]

Token 'He':
Layer 1: Used Experts -> [20, 100, 101, 114, 51, 68, 11, 21]
Layer 5: Used Experts -> [10, 102, 52, 71, 96, 8, 95, 81]
Layer 9: Used Experts -> [3, 36, 22, 30, 107, 50, 83, 81]

Token 'was':
Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 8, 92]
Layer 5: Used Experts -> [34, 93, 119, 35, 38, 32, 124, 56]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'very':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [37, 42, 31, 61, 20, 3, 48, 104]
Layer 9: Used Experts -> [89, 49, 70, 56, 66, 78, 68, 99]

Token 'excited':
Layer 1: Used Experts -> [48, 5, 83, 126, 69, 28, 78, 65]
Layer 5: Used Experts -> [67, 68, 125, 101, 123, 92, 94, 98]
Layer 9: Used Experts -> [99, 9, 121, 87, 127, 112, 111, 104]

Token 'to':

Layer 1: Used Experts -> [112, 76, 113, 111, 59, 56, 42, 98]
Layer 5: Used Experts -> [69, 51, 100, 27, 107, 63, 6, 12]
Layer 9: Used Experts -> [65, 98, 27, 93, 110, 15, 4, 109]

Token 'see':
Layer 1: Used Experts -> [52, 24, 107, 22, 124, 38, 7, 104]
Layer 5: Used Experts -> [25, 19, 9, 116, 40, 97, 79, 118]
Layer 9: Used Experts -> [79, 42, 0, 100, 77, 11, 120, 118]

Token 'the':
Layer 1: Used Experts -> [62, 88, 119, 122, 85, 16, 4, 49]
Layer 5: Used Experts -> [99, 91, 7, 39, 57, 12, 50, 77]
Layer 9: Used Experts -> [89, 70, 56, 66, 68, 49, 78, 106]

Token 'big':
Layer 1: Used Experts -> [12, 46, 33, 19, 73, 108, 0, 25]
Layer 5: Used Experts -> [10, 102, 52, 71, 96, 81, 95, 8]
Layer 9: Used Experts -> [124, 116, 103, 75, 43, 25, 1, 17]

Token 'and':
Layer 1: Used Experts -> [87, 117, 80, 121, 55, 3, 47, 72]
Layer 5: Used Experts -> [68, 67, 101, 98, 92, 125, 123, 94]
Layer 9: Used Experts -> [10, 41, 51, 18, 97, 62, 57, 37]

Token 'he':
Layer 1: Used Experts -> [97, 64, 102, 45, 94, 109, 70, 93]
Layer 5: Used Experts -> [106, 70, 24, 46, 55, 1, 72, 5]
Layer 9: Used Experts -> [89, 49, 70, 56, 66, 78, 68, 106]

Token 'was':
Layer 1: Used Experts -> [29, 71, 125, 44, 14, 43, 120, 23]
Layer 5: Used Experts -> [34, 93, 119, 35, 124, 38, 32, 56]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'very':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [37, 31, 42, 61, 3, 20, 48, 104]
Layer 9: Used Experts -> [86, 8, 10, 109, 41, 65, 51, 18]

Token 'excited':
Layer 1: Used Experts -> [48, 5, 126, 69, 83, 28, 78, 65]
Layer 5: Used Experts -> [34, 67, 93, 119, 38, 35, 32, 56]
Layer 9: Used Experts -> [99, 62, 57, 37, 97, 18, 51, 41]

Token '.':
  Layer 1: Used Experts -> [112, 76, 113, 111, 56, 59, 42, 98]
  Layer 5: Used Experts -> [51, 100, 69, 6, 63, 107, 27, 12]
  Layer 9: Used Experts -> [98, 65, 27, 110, 15, 4, 109, 93]

Token 'He':
  Layer 1: Used Experts -> [20, 100, 51, 68, 21, 11, 101, 114]
  Layer 5: Used Experts -> [10, 102, 52, 71, 96, 95, 8, 81]
  Layer 9: Used Experts -> [83, 36, 3, 22, 50, 107, 30, 81]

Token 'wanted':
  Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 8, 92]
  Layer 5: Used Experts -> [34, 93, 119, 35, 38, 32, 124, 56]
  Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'to':
  Layer 1: Used Experts -> [120, 23, 43, 14, 44, 125, 71, 29]
  Layer 5: Used Experts -> [73, 7, 39, 57, 77, 91, 50, 99]
  Layer 9: Used Experts -> [89, 70, 56, 66, 49, 68, 78, 106]

Token 'see':
  Layer 1: Used Experts -> [52, 24, 38, 107, 22, 124, 7, 104]
  Layer 5: Used Experts -> [19, 9, 40, 97, 79, 116, 118, 25]
  Layer 9: Used Experts -> [42, 0, 79, 100, 77, 120, 118, 11]

Token 'what':
  Layer 1: Used Experts -> [62, 88, 119, 122, 85, 16, 4, 49]
  Layer 5: Used Experts -> [12, 99, 91, 50, 57, 39, 7, 77]
  Layer 9: Used Experts -> [61, 106, 78, 68, 49, 66, 56, 70]

Token 'it':
  Layer 1: Used Experts -> [65, 28, 69, 5, 126, 48, 78, 83]
  Layer 5: Used Experts -> [108, 118, 19, 9, 79, 97, 40, 116]
  Layer 9: Used Experts -> [89, 70, 56, 49, 66, 68, 78, 106]

Token 'was':
  Layer 1: Used Experts -> [65, 78, 28, 69, 126, 5, 83, 48]
  Layer 5: Used Experts -> [81, 10, 52, 71, 102, 96, 95, 8]
  Layer 9: Used Experts -> [10, 41, 51, 18, 97, 37, 57, 62]

Token '.':
  Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
  Layer 5: Used Experts -> [31, 42, 61, 3, 20, 37, 104, 48]
  Layer 9: Used Experts -> [93, 65, 59, 27, 109, 86, 8, 67]

Token 'He':
  Layer 1: Used Experts -> [20, 100, 51, 68, 21, 11, 101, 114]
  Layer 5: Used Experts -> [10, 102, 52, 71, 96, 95, 8, 81]
  Layer 9: Used Experts -> [83, 36, 3, 22, 50, 107, 30, 81]

Token 'asked':
  Layer 1: Used Experts -> [53, 89, 36, 6, 82, 116, 8, 92]
  Layer 5: Used Experts -> [34, 93, 119, 35, 38, 124, 32, 56]
  Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'the':
  Layer 1: Used Experts -> [30, 57, 49, 4, 16, 85, 122, 119]
  Layer 5: Used Experts -> [87, 64, 43, 28, 45, 16, 82, 26]
  Layer 9: Used Experts -> [89, 70, 56, 66, 68, 78, 106, 49]

Token 'other':
  Layer 1: Used Experts -> [46, 12, 33, 73, 108, 19, 0, 25]
  Layer 5: Used Experts -> [10, 102, 52, 71, 81, 96, 95, 8]
  Layer 9: Used Experts -> [124, 1, 116, 43, 75, 25, 17, 103]

Token 'children':
  Layer 1: Used Experts -> [99, 13, 10, 67, 40, 50, 9, 74]
  Layer 5: Used Experts -> [67, 94, 125, 123, 92, 101, 98, 68]
  Layer 9: Used Experts -> [98, 15, 110, 4, 27, 8, 109, 86]

Token ',':
  Layer 1: Used Experts -> [23, 120, 87, 43, 44, 14, 125, 71]
  Layer 5: Used Experts -> [12, 6, 65, 63, 27, 69, 107, 100]
  Layer 9: Used Experts -> [93, 65, 27, 110, 109, 59, 15, 4]

Token '"':
  Layer 1: Used Experts -> [17, 92, 8, 116, 82, 6, 36, 89]
  Layer 5: Used Experts -> [25, 116, 97, 40, 79, 9, 118, 19]
  Layer 9: Used Experts -> [86, 8, 109, 4, 27, 65, 110, 15]

Token 'What':
  Layer 1: Used Experts -> [9, 74, 50, 40, 67, 10, 13, 99]
  Layer 5: Used Experts -> [54, 15, 59, 76, 60, 120, 78, 86]
  Layer 9: Used Experts -> [3, 36, 22, 50, 107, 30, 83, 81]

Token 'are':
  Layer 1: Used Experts -> [2, 127, 60, 106, 118, 105, 84, 103]
  Layer 5: Used Experts -> [25, 118, 116, 79, 97, 40, 9, 19]

Layer 9: Used Experts -> [20, 58, 60, 117, 105, 102, 5, 47]

Token 'you':
Layer 1: Used Experts -> [48, 126, 69, 28, 5, 83, 78, 65]
Layer 5: Used Experts -> [3, 61, 20, 42, 31, 37, 48, 104]
Layer 9: Used Experts -> [61, 106, 68, 66, 78, 56, 70, 49]

Token 'doing':
Layer 1: Used Experts -> [63, 66, 115, 96, 39, 34, 1, 91]
Layer 5: Used Experts -> [25, 9, 19, 40, 116, 97, 79, 118]
Layer 9: Used Experts -> [79, 100, 0, 77, 11, 42, 120, 118]

Token '?"':
Layer 1: Used Experts -> [83, 5, 28, 69, 48, 126, 78, 65]
Layer 5: Used Experts -> [50, 77, 99, 91, 57, 39, 7, 73]
Layer 9: Used Experts -> [20, 58, 60, 105, 92, 47, 5, 34]

Token 'The':
Layer 1: Used Experts -> [83, 78, 5, 126, 69, 48, 28, 65]
Layer 5: Used Experts -> [25, 9, 116, 19, 40, 97, 79, 118]
Layer 9: Used Experts -> [89, 70, 56, 66, 68, 5, 106, 26]

Token 'children':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [81, 95, 96, 71, 52, 102, 10, 8]
Layer 9: Used Experts -> [1, 116, 43, 25, 75, 17, 103, 124]

Token 'replied':
Layer 1: Used Experts -> [87, 117, 80, 121, 55, 3, 47, 41]
Layer 5: Used Experts -> [6, 65, 63, 27, 107, 100, 69, 51]
Layer 9: Used Experts -> [86, 8, 109, 4, 10, 15, 110, 41]

Token ',':
Layer 1: Used Experts -> [99, 10, 67, 40, 50, 13, 74, 9]
Layer 5: Used Experts -> [94, 125, 123, 92, 98, 101, 67, 68]
Layer 9: Used Experts -> [58, 5, 26, 91, 119, 67, 59, 20]

Token '"':
Layer 1: Used Experts -> [17, 92, 8, 116, 82, 6, 36, 89]
Layer 5: Used Experts -> [25, 116, 97, 40, 79, 9, 118, 19]
Layer 9: Used Experts -> [65, 93, 27, 109, 86, 8, 4, 110]

Token 'I':
Layer 1: Used Experts -> [9, 74, 50, 40, 67, 10, 13, 99]

```
  Layer 5:       Used Experts -> [54, 15, 59, 76, 60, 120, 78, 86]
  Layer 9:       Used Experts -> [3, 36, 22, 30, 107, 50, 81, 83]

Token '"':
  Layer 1:       Used Experts -> [30, 57, 103, 84, 105, 118, 106, 60]
  Layer 5:       Used Experts -> [53, 17, 80, 111, 127, 49, 30, 11]
  Layer 9:       Used Experts -> [79, 81, 30, 107, 3, 50, 22, 36]

Token 'm':
  Layer 1:       Used Experts -> [1, 34, 66, 115, 63, 96, 39, 91]
  Layer 5:       Used Experts -> [117, 4, 84, 2, 122, 44, 23, 74]
  Layer 9:       Used Experts -> [98, 102, 92, 105, 60, 47, 34, 20]

Token 'going':
  Layer 1:       Used Experts -> [103, 84, 105, 118, 106, 60, 127, 2]
  Layer 5:       Used Experts -> [48, 37, 104, 42, 61, 20, 3, 31]
  Layer 9:       Used Experts -> [61, 106, 78, 118, 49, 68, 42, 120]

Token 'to':
  Layer 1:       Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
  Layer 5:       Used Experts -> [31, 104, 48, 3, 61, 20, 42, 37]
  Layer 9:       Used Experts -> [65, 27, 109, 110, 4, 8, 15, 86]

--- Prompt ---
The robot opened its eyes and
sawalittlegirl.Shewantedtogotothepark.Sheaskedhermom,"Whatisthelittlegirl?"Hermomsaid,"It'sa
big,littlegirl.It'saverypretty."Thelittlegirlwasexcited
--- End of Generation ---

--- Multi-Layer Expert Activation Trace ---

Token 'a':
  Layer 1:       Used Experts -> [108, 73, 33, 46, 12, 19, 25, 0]
  Layer 5:       Used Experts -> [37, 42, 104, 61, 20, 3, 48, 31]
  Layer 9:       Used Experts -> [98, 15, 109, 4, 8, 110, 27, 86]

Token 'little':
  Layer 1:       Used Experts -> [53, 89, 36, 6, 82, 116, 8, 92]
  Layer 5:       Used Experts -> [67, 68, 101, 125, 92, 98, 123, 94]
  Layer 9:       Used Experts -> [7, 23, 38, 39, 52, 16, 82, 122]

Token 'girl':
  Layer 1:       Used Experts -> [30, 57, 49, 4, 16, 85, 122, 119]
  Layer 5:       Used Experts -> [8, 52, 71, 102, 95, 96, 10, 81]
```

Layer 9: Used Experts -> [62, 10, 57, 41, 18, 37, 97, 51]

Token '.':
Layer 1: Used Experts -> [76, 59, 98, 56, 42, 90, 111, 113]
Layer 5: Used Experts -> [65, 6, 63, 27, 107, 100, 69, 51]
Layer 9: Used Experts -> [10, 41, 51, 18, 97, 37, 57, 62]

Token 'She':
Layer 1: Used Experts -> [114, 101, 11, 21, 68, 51, 100, 20]
Layer 5: Used Experts -> [10, 102, 52, 71, 96, 95, 81, 8]
Layer 9: Used Experts -> [81, 30, 107, 50, 22, 83, 3, 36]

Token 'wanted':
Layer 1: Used Experts -> [44, 71, 125, 14, 43, 29, 23, 120]
Layer 5: Used Experts -> [34, 93, 119, 35, 38, 124, 32, 56]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'to':
Layer 1: Used Experts -> [23, 120, 43, 44, 14, 125, 71, 29]
Layer 5: Used Experts -> [73, 7, 39, 57, 77, 91, 50, 99]
Layer 9: Used Experts -> [49, 78, 106, 68, 66, 56, 70, 89]

Token 'go':
Layer 1: Used Experts -> [52, 24, 38, 107, 22, 124, 7, 104]
Layer 5: Used Experts -> [19, 9, 25, 40, 116, 97, 79, 118]
Layer 9: Used Experts -> [79, 0, 100, 42, 77, 11, 120, 118]

Token 'to':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [31, 104, 3, 61, 20, 42, 48, 37]
Layer 9: Used Experts -> [93, 65, 27, 110, 109, 59, 4, 8]

Token 'the':
Layer 1: Used Experts -> [52, 24, 38, 107, 22, 124, 7, 104]
Layer 5: Used Experts -> [19, 9, 25, 40, 97, 116, 79, 118]
Layer 9: Used Experts -> [42, 0, 118, 100, 77, 120, 11, 79]

Token 'park':
Layer 1: Used Experts -> [46, 33, 73, 12, 108, 19, 0, 25]
Layer 5: Used Experts -> [10, 102, 52, 71, 96, 81, 95, 8]
Layer 9: Used Experts -> [124, 1, 116, 43, 75, 25, 17, 103]

Token '.':
Layer 1: Used Experts -> [91, 29, 71, 125, 14, 44, 43, 120]

Layer 5: Used Experts -> [65, 6, 63, 27, 107, 100, 69, 51]
Layer 9: Used Experts -> [93, 119, 98, 91, 67, 26, 5, 59]

Token 'She':
Layer 1: Used Experts -> [20, 100, 51, 101, 68, 21, 11, 114]
Layer 5: Used Experts -> [10, 102, 8, 52, 71, 96, 95, 81]
Layer 9: Used Experts -> [81, 30, 107, 50, 22, 3, 83, 36]

Token 'asked':
Layer 1: Used Experts -> [29, 71, 125, 44, 14, 43, 23, 120]
Layer 5: Used Experts -> [34, 93, 119, 35, 38, 124, 32, 56]
Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 'her':
Layer 1: Used Experts -> [30, 57, 49, 4, 16, 85, 122, 119]
Layer 5: Used Experts -> [105, 33, 87, 64, 28, 45, 43, 16]
Layer 9: Used Experts -> [61, 70, 106, 56, 66, 68, 78, 49]

Token 'mom':
Layer 1: Used Experts -> [72, 90, 98, 42, 59, 56, 111, 113]
Layer 5: Used Experts -> [67, 94, 125, 123, 92, 98, 101, 68]
Layer 9: Used Experts -> [16, 82, 52, 38, 122, 7, 23, 39]

Token ',':
Layer 1: Used Experts -> [112, 76, 113, 111, 56, 59, 42, 98]
Layer 5: Used Experts -> [34, 93, 119, 35, 38, 32, 124, 56]
Layer 9: Used Experts -> [58, 5, 119, 26, 91, 67, 59, 20]

Token '"':
Layer 1: Used Experts -> [17, 92, 8, 116, 82, 6, 36, 89]
Layer 5: Used Experts -> [25, 116, 9, 40, 97, 79, 19, 118]
Layer 9: Used Experts -> [59, 93, 67, 91, 119, 26, 5, 65]

Token 'What':
Layer 1: Used Experts -> [9, 74, 50, 40, 67, 10, 99, 13]
Layer 5: Used Experts -> [54, 15, 59, 76, 60, 120, 78, 86]
Layer 9: Used Experts -> [81, 30, 107, 50, 22, 3, 36, 83]

Token 'is':
Layer 1: Used Experts -> [2, 127, 60, 106, 118, 105, 84, 103]
Layer 5: Used Experts -> [19, 9, 40, 97, 79, 116, 118, 25]
Layer 9: Used Experts -> [48, 117, 102, 34, 47, 92, 105, 60]

Token 'the':

Layer 1: Used Experts -> [65, 78, 28, 69, 5, 126, 83, 48]
Layer 5: Used Experts -> [37, 42, 61, 20, 3, 48, 104, 31]
Layer 9: Used Experts -> [61, 106, 78, 68, 49, 66, 56, 70]

Token 'little':
Layer 1: Used Experts -> [108, 33, 73, 46, 12, 19, 25, 0]
Layer 5: Used Experts -> [10, 102, 81, 52, 71, 96, 95, 8]
Layer 9: Used Experts -> [39, 122, 23, 52, 82, 38, 7, 16]

Token 'girl':
Layer 1: Used Experts -> [30, 57, 49, 4, 16, 85, 122, 119]
Layer 5: Used Experts -> [81, 52, 96, 71, 95, 102, 10, 8]
Layer 9: Used Experts -> [39, 7, 23, 38, 122, 52, 82, 16]

Token '?"':
Layer 1: Used Experts -> [76, 59, 56, 111, 42, 98, 113, 90]
Layer 5: Used Experts -> [106, 70, 24, 55, 72, 1, 46, 5]
Layer 9: Used Experts -> [60, 105, 47, 34, 92, 102, 117, 20]

Token 'Her':
Layer 1: Used Experts -> [47, 55, 41, 121, 80, 117, 3, 83]
Layer 5: Used Experts -> [9, 19, 25, 116, 40, 97, 79, 118]
Layer 9: Used Experts -> [36, 83, 3, 22, 50, 107, 30, 81]

Token 'mom':
Layer 1: Used Experts -> [119, 122, 85, 16, 4, 49, 57, 30]
Layer 5: Used Experts -> [67, 94, 125, 123, 92, 98, 101, 68]
Layer 9: Used Experts -> [53, 104, 111, 112, 127, 87, 121, 9]

Token 'said':
Layer 1: Used Experts -> [112, 76, 113, 111, 56, 59, 42, 98]
Layer 5: Used Experts -> [124, 35, 119, 38, 32, 34, 93, 56]
Layer 9: Used Experts -> [15, 4, 8, 110, 86, 109, 27, 65]

Token ',':
Layer 1: Used Experts -> [41, 47, 55, 121, 80, 3, 117, 87]
Layer 5: Used Experts -> [67, 94, 125, 123, 92, 98, 101, 68]
Layer 9: Used Experts -> [5, 26, 58, 91, 119, 67, 59, 93]

Token '"':
Layer 1: Used Experts -> [17, 92, 8, 116, 82, 6, 36, 89]
Layer 5: Used Experts -> [25, 116, 97, 40, 9, 79, 19, 118]
Layer 9: Used Experts -> [86, 8, 109, 4, 27, 110, 65, 15]

Token 'It':
  Layer 1: Used Experts -> [9, 74, 50, 40, 67, 10, 13, 99]
  Layer 5: Used Experts -> [54, 15, 59, 76, 60, 78, 120, 86]
  Layer 9: Used Experts -> [81, 30, 107, 50, 22, 3, 36, 83]

Token '"':
  Layer 1: Used Experts -> [120, 23, 43, 44, 14, 125, 71, 29]
  Layer 5: Used Experts -> [93, 56, 32, 119, 124, 38, 35, 34]
  Layer 9: Used Experts -> [62, 57, 48, 37, 97, 18, 51, 41]

Token 's':
  Layer 1: Used Experts -> [1, 34, 66, 115, 96, 63, 39, 91]
  Layer 5: Used Experts -> [4, 117, 84, 122, 44, 74, 23, 2]
  Layer 9: Used Experts -> [15, 110, 4, 98, 27, 8, 109, 86]

Token 'a':
  Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
  Layer 5: Used Experts -> [94, 125, 123, 92, 98, 101, 67, 68]
  Layer 9: Used Experts -> [99, 49, 78, 106, 68, 66, 56, 70]

Token 'big':
  Layer 1: Used Experts -> [53, 89, 92, 36, 6, 8, 82, 116]
  Layer 5: Used Experts -> [67, 68, 101, 125, 92, 98, 123, 94]
  Layer 9: Used Experts -> [16, 82, 52, 38, 122, 7, 23, 39]

Token ',':
  Layer 1: Used Experts -> [87, 117, 80, 121, 55, 3, 47, 41]
  Layer 5: Used Experts -> [88, 68, 101, 98, 92, 123, 125, 94]
  Layer 9: Used Experts -> [7, 23, 39, 38, 52, 122, 82, 16]

Token 'little':
  Layer 1: Used Experts -> [17, 92, 8, 116, 82, 6, 36, 89]
  Layer 5: Used Experts -> [25, 116, 97, 40, 9, 79, 19, 118]
  Layer 9: Used Experts -> [104, 53, 111, 127, 112, 87, 121, 9]

Token 'girl':
  Layer 1: Used Experts -> [30, 57, 49, 4, 16, 85, 122, 119]
  Layer 5: Used Experts -> [81, 95, 96, 71, 52, 102, 8, 10]
  Layer 9: Used Experts -> [7, 23, 39, 38, 122, 52, 82, 16]

Token '.':
  Layer 1: Used Experts -> [76, 59, 56, 42, 111, 98, 113, 90]
  Layer 5: Used Experts -> [70, 106, 24, 55, 72, 1, 5, 46]
  Layer 9: Used Experts -> [102, 92, 34, 47, 105, 60, 48, 117]

Token 'It':
  Layer 1: Used Experts -> [20, 100, 51, 101, 68, 114, 21, 11]
  Layer 5: Used Experts -> [8, 95, 96, 71, 52, 81, 102, 10]
  Layer 9: Used Experts -> [81, 30, 107, 50, 22, 3, 36, 83]

Token '"':
  Layer 1: Used Experts -> [120, 23, 43, 44, 14, 125, 71, 29]
  Layer 5: Used Experts -> [93, 56, 32, 38, 119, 124, 35, 34]
  Layer 9: Used Experts -> [62, 57, 37, 97, 18, 51, 41, 10]

Token 's':
  Layer 1: Used Experts -> [1, 34, 66, 115, 96, 63, 39, 91]
  Layer 5: Used Experts -> [4, 117, 84, 122, 74, 44, 23, 2]
  Layer 9: Used Experts -> [15, 86, 4, 8, 110, 109, 27, 65]

Token 'a':
  Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
  Layer 5: Used Experts -> [94, 125, 123, 92, 98, 101, 67, 68]
  Layer 9: Used Experts -> [99, 9, 121, 87, 127, 112, 111, 104]

Token 'very':
  Layer 1: Used Experts -> [53, 89, 92, 36, 8, 6, 116, 82]
  Layer 5: Used Experts -> [67, 68, 101, 125, 92, 98, 123, 94]
  Layer 9: Used Experts -> [16, 82, 52, 38, 7, 122, 23, 39]

Token 'pretty':
  Layer 1: Used Experts -> [48, 5, 83, 126, 69, 28, 78, 65]
  Layer 5: Used Experts -> [67, 68, 101, 92, 98, 125, 123, 94]
  Layer 9: Used Experts -> [16, 82, 52, 38, 7, 23, 122, 39]

Token '."':
  Layer 1: Used Experts -> [65, 48, 126, 69, 28, 5, 78, 83]
  Layer 5: Used Experts -> [98, 101, 92, 123, 94, 125, 68, 67]
  Layer 9: Used Experts -> [39, 7, 23, 38, 122, 52, 82, 16]

Token 'The':
  Layer 1: Used Experts -> [2, 127, 60, 106, 118, 105, 84, 103]
  Layer 5: Used Experts -> [8, 95, 81, 96, 71, 52, 102, 10]
  Layer 9: Used Experts -> [81, 30, 107, 50, 22, 3, 36, 83]

Token 'little':
  Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
  Layer 5: Used Experts -> [88, 56, 32, 38, 35, 124, 119, 93]

Layer 9: Used Experts -> [122, 82, 52, 39, 23, 38, 16, 7]

Token 'girl':
Layer 1: Used Experts -> [30, 57, 49, 4, 16, 85, 122, 119]
Layer 5: Used Experts -> [81, 95, 96, 71, 52, 8, 102, 10]
Layer 9: Used Experts -> [7, 23, 39, 38, 52, 122, 82, 16]

Token 'was':
Layer 1: Used Experts -> [76, 59, 56, 42, 111, 98, 113, 90]
Layer 5: Used Experts -> [70, 24, 72, 5, 55, 1, 46, 106]
Layer 9: Used Experts -> [62, 57, 97, 18, 37, 41, 51, 10]

Token 'excited':
Layer 1: Used Experts -> [91, 39, 63, 96, 115, 66, 34, 1]
Layer 5: Used Experts -> [31, 104, 48, 61, 42, 3, 20, 37]
Layer 9: Used Experts -> [89, 49, 70, 56, 66, 78, 68, 106]