Linear 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, 95.1kB/s]
data/train-00000-of-00004-2d5a1467fff108(…): 100%
249M/249M [00:02<00:00, 5.23MB/s]
data/train-00001-of-00004-5852b56a2bd28f(…): 100%
248M/248M [00:01<00:00, 200MB/s]
data/train-00002-of-00004-a26307300439e9(…): 100%
246M/246M [00:01<00:00, 283MB/s]
data/train-00003-of-00004-d243063613e5a0(…): 100%
248M/248M [00:01<00:00, 266MB/s]
data/validation-00000-of-00001-869c898b5(…): 100%
9.99M/9.99M [00:00<00:00, 14.8MB/s]
Generating train split: 100%
2119719/2119719 [00:06<00:00, 341241.28 examples/s]
Generating validation split: 100%
21990/21990 [00:00<00:00, 293373.28 examples/s]
Custom tokenizer loaded with vocab size: 8192

Tokenizing dataset...

Map (num_proc=12): 100%
10000/10000 [00:02<00:00, 3719.04 examples/s]
Map (num_proc=12): 100%
1000/1000 [00:00<00:00, 174.07 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-538142859.py:434: 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: 2709.18M
Trainable parameters: 21.55M (0.80%)
Total training steps: 1664
Using expert initialization method: sparse

--- Epoch 1/4 ---

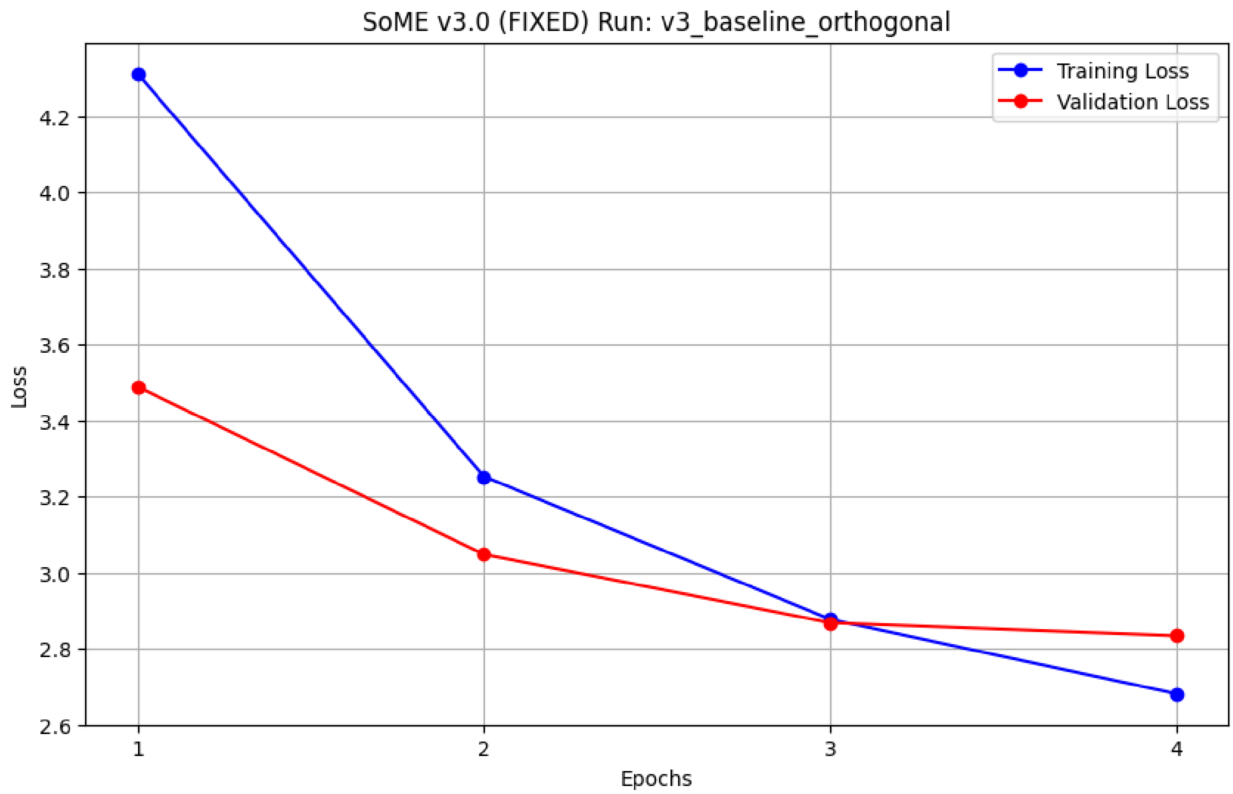
Training: 0%| | 0/416 [00:00<?, ?it/s]/tmp/ipython-input-538142859.py:442: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Training: 0%| | 1/416 [00:36<4:13:19, 36.63s/it, loss=9.2097, lr=6.0e-04]/tmp/ipython-input-538142859.py:442: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Training: 1%| | 3/416 [00:41<1:09:50, 10.15s/it, loss=7.6983, lr=6.0e-04]/tmp/ipython-input-538142859.py:442: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():
Training: 1%| | 4/416 [00:42<46:00, 6.70s/it, loss=7.3452, lr=6.0e-04]
/tmp/ipython-input-538142859.py:442: 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-538142859.py:475: 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<05:06, 7.47s/it, loss=3.2327]/tmp/ipython-input-538142859.py:475: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast():

```
Evaluating:  12%|█          | 5/42 [00:10<00:48,  1.31s/it,
loss=3.1660]/tmp/ipython-input-538142859.py:475: 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 = 4.3130, Val Loss = 3.4889, Val Perplexity = 32.75
  Middle Layer (count): Gini = 0.383, Entropy = 6.651
  Middle Layer (mass):  Gini = 0.381, Entropy = 6.654

/tmp/ipython-input-538142859.py:434: 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-538142859.py:442:
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-538142859.py:475:
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 = 3.2530, Val Loss = 3.0480, Val Perplexity = 21.07
  Middle Layer (count): Gini = 0.366, Entropy = 6.684
  Middle Layer (mass):  Gini = 0.365, Entropy = 6.687
  Model saved as best_model_v3_baseline_orthogonal_fixed.pth

--- Epoch 3/4 ---

Epoch 3: Train Loss = 2.8771, Val Loss = 2.8681, Val Perplexity = 17.60
  Middle Layer (count): Gini = 0.370, Entropy = 6.679
  Middle Layer (mass):  Gini = 0.369, Entropy = 6.680
  Model saved as best_model_v3_baseline_orthogonal_fixed.pth

--- Epoch 4/4 ---

Epoch 4: Train Loss = 2.6812, Val Loss = 2.8337, Val Perplexity = 17.01
  Middle Layer (count): Gini = 0.373, Entropy = 6.673
  Middle Layer (mass):  Gini = 0.372, Entropy = 6.675
```

Model saved as best_model_v3_baseline_orthogonal_fixed.pth

--- Training Complete for v3_baseline_orthogonal ---

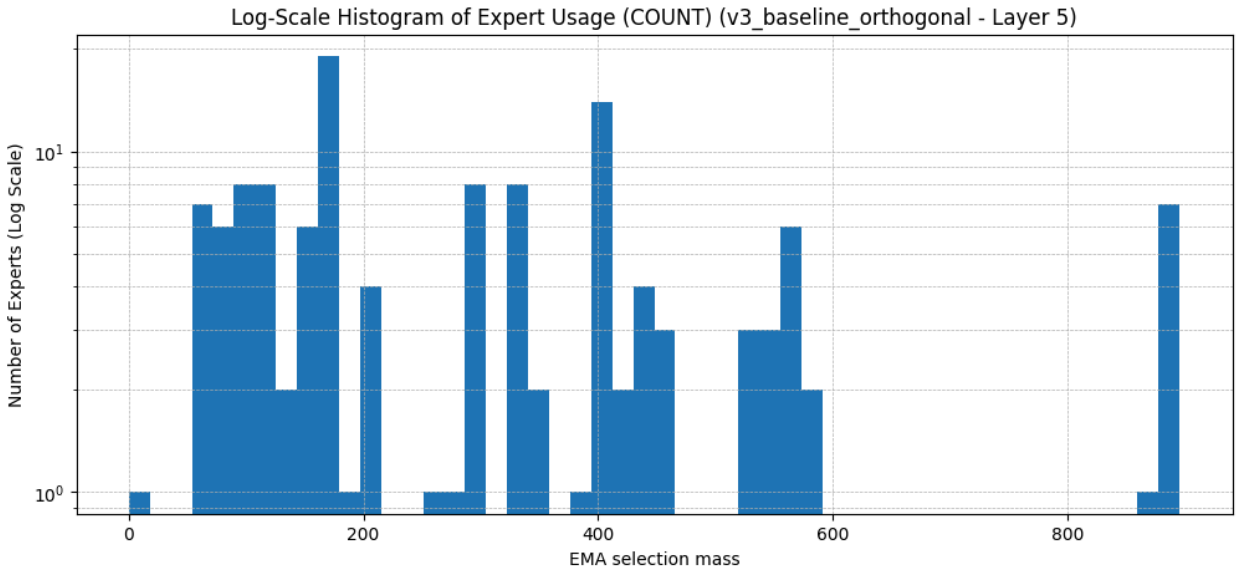
Loss curve plot saved to loss_curve_v3_baseline_orthogonal_fixed.png

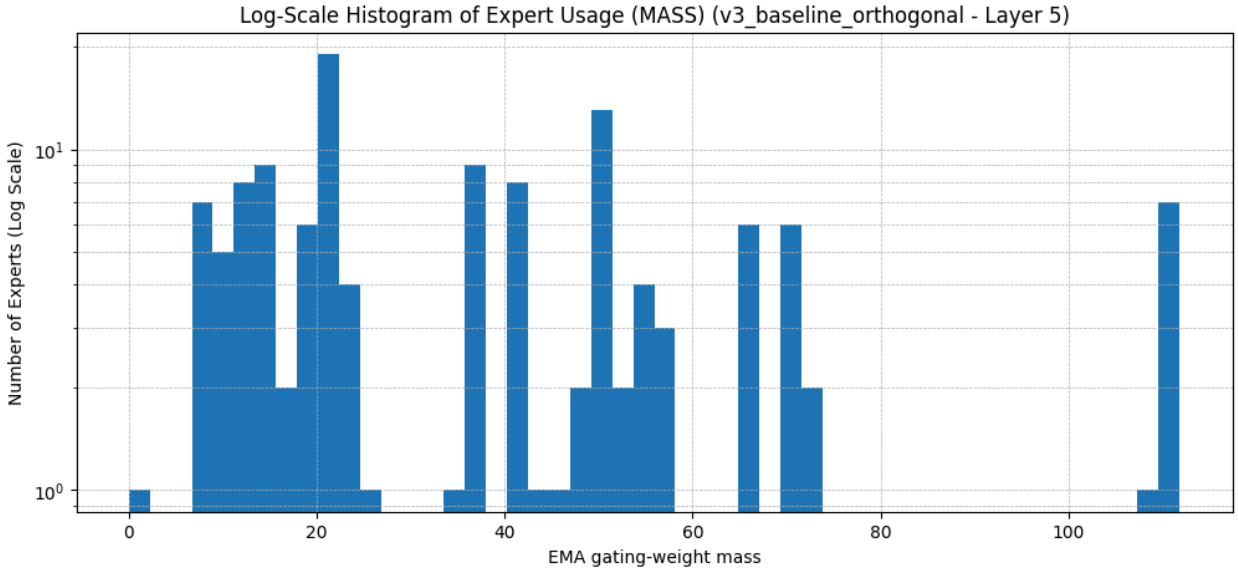


--- Part 1: Dashboard Setup ---
Loading best model from: best_model_v3_baseline_orthogonal_fixed.pth
Loading tokenizer from: tinystories-tokenizer-v2.json
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True, 'use_delta': True}
SoME Layer Ablation Flags: {'use_alpha': True, 'use_beta': True,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

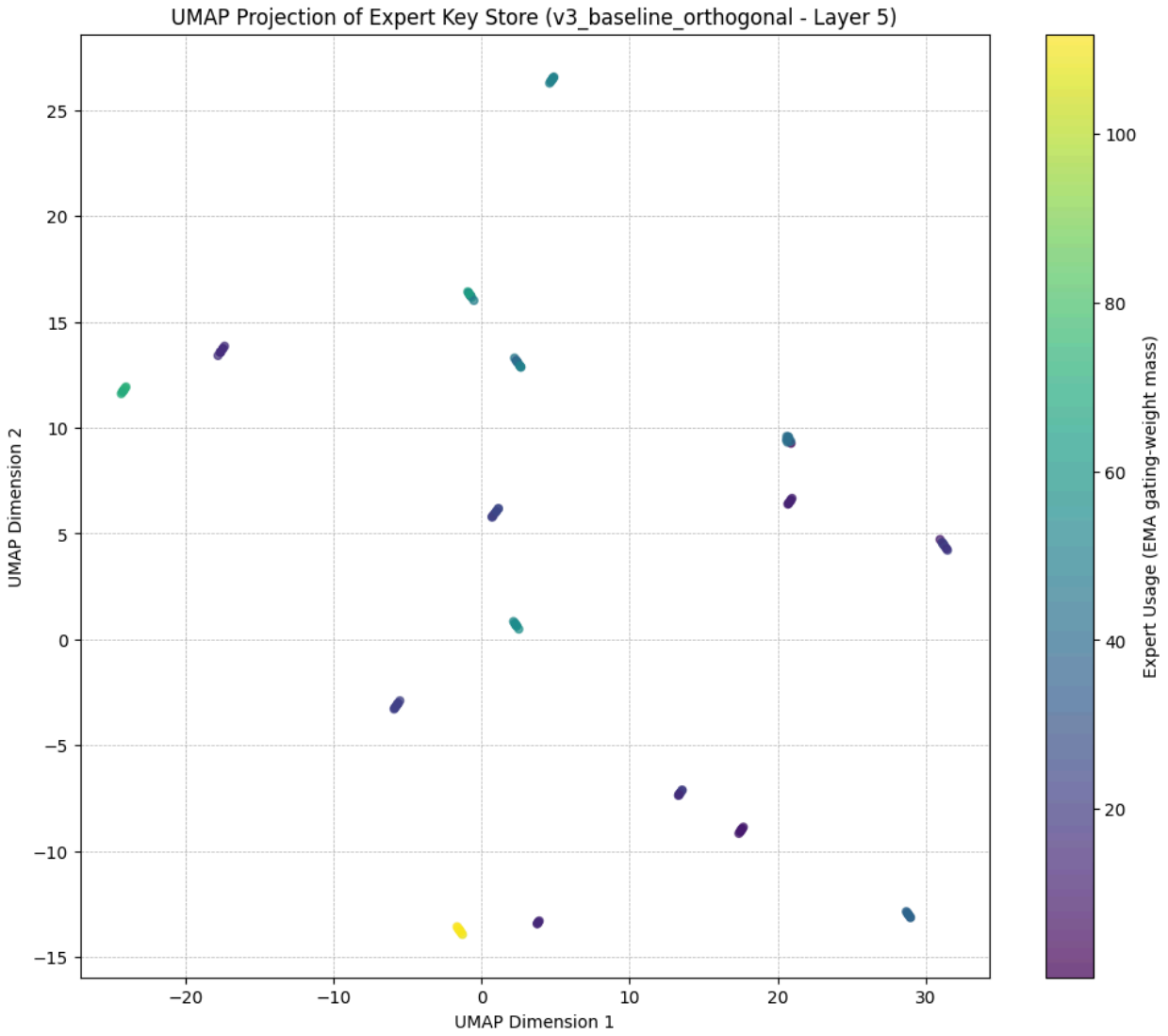
Expert Usage (Layer 5): 128/128 (100.00%)
Final Metrics (count) Layer 5: Gini=0.3731, Entropy=6.6731 (Max=7.0000)
Final Metrics (mass) Layer 5: Gini=0.3718, Entropy=6.6752 (Max=7.0000)





--- 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-538142859.py:101: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.
  with torch.cuda.amp.autocast(enabled=False):

inabigforest.Thefoxlovedtoplaywithhisfriends.Oneday,thefoxsawabig,shinyrock.Thefoxwasscaredandwantedtoplaywiththefox.Thefoxsaid,"Iwanttoplaywithyou
--- End of Generation ---

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

Token 'lived':

Layer 1: Used Experts -> [86, 26, 69, 81, 95, 72, 75, 70]
Layer 5: Used Experts -> [63, 116, 11, 113, 0, 8, 80, 95]
Layer 9: Used Experts -> [63, 70, 19, 96, 64, 54, 92, 9]

Token 'in':
Layer 1: Used Experts -> [95, 75, 81, 72, 69, 26, 86, 70]
Layer 5: Used Experts -> [3, 26, 29, 111, 60, 65, 76, 52]
Layer 9: Used Experts -> [62, 34, 100, 21, 99, 102, 26, 53]

Token 'a':
Layer 1: Used Experts -> [124, 19, 38, 93, 42, 22, 30, 66]
Layer 5: Used Experts -> [126, 3, 26, 60, 76, 65, 111, 29]
Layer 9: Used Experts -> [20, 13, 6, 118, 123, 67, 23, 68]

Token 'big':
Layer 1: Used Experts -> [124, 19, 38, 42, 93, 22, 66, 30]
Layer 5: Used Experts -> [52, 78, 69, 37, 6, 96, 107, 70]
Layer 9: Used Experts -> [103, 106, 4, 97, 94, 83, 52, 48]

Token 'forest':
Layer 1: Used Experts -> [60, 122, 49, 57, 82, 94, 65, 9]
Layer 5: Used Experts -> [98, 42, 101, 92, 72, 74, 46, 122]
Layer 9: Used Experts -> [48, 78, 8, 44, 16, 27, 120, 76]

Token '.':
Layer 1: Used Experts -> [8, 15, 121, 46, 100, 39, 99, 73]
Layer 5: Used Experts -> [86, 104, 120, 40, 14, 9, 45, 123]
Layer 9: Used Experts -> [102, 99, 21, 34, 26, 100, 23, 62]

Token 'The':
Layer 1: Used Experts -> [12, 101, 113, 106, 14, 16, 7, 96]
Layer 5: Used Experts -> [36, 44, 71, 110, 24, 90, 82, 30]
Layer 9: Used Experts -> [17, 31, 85, 127, 1, 50, 35, 126]

Token 'fox':
Layer 1: Used Experts -> [70, 69, 75, 72, 81, 26, 95, 86]
Layer 5: Used Experts -> [98, 42, 101, 72, 92, 74, 46, 122]
Layer 9: Used Experts -> [76, 27, 120, 16, 44, 8, 78, 48]

Token 'loved':
Layer 1: Used Experts -> [20, 40, 108, 31, 0, 88, 29, 71]
Layer 5: Used Experts -> [11, 63, 116, 0, 113, 80, 8, 95]
Layer 9: Used Experts -> [70, 63, 92, 19, 96, 54, 64, 9]

Token 'to':
  Layer 1: Used Experts -> [109, 59, 120, 55, 118, 110, 1, 58]
  Layer 5: Used Experts -> [104, 45, 9, 40, 14, 120, 123, 86]
  Layer 9: Used Experts -> [13, 6, 118, 20, 123, 67, 68, 36]

Token 'play':
  Layer 1: Used Experts -> [11, 76, 68, 77, 114, 107, 47, 115]
  Layer 5: Used Experts -> [68, 84, 38, 54, 43, 102, 15, 35]
  Layer 9: Used Experts -> [117, 55, 15, 47, 14, 73, 41, 59]

Token 'with':
  Layer 1: Used Experts -> [8, 86, 95, 26, 81, 72, 69, 75]
  Layer 5: Used Experts -> [75, 58, 50, 121, 48, 28, 20, 91]
  Layer 9: Used Experts -> [26, 23, 102, 99, 21, 34, 100, 62]

Token 'his':
  Layer 1: Used Experts -> [87, 23, 90, 33, 21, 126, 104, 98]
  Layer 5: Used Experts -> [126, 26, 3, 76, 60, 65, 111, 29]
  Layer 9: Used Experts -> [20, 13, 6, 118, 123, 67, 68, 36]

Token 'friends':
  Layer 1: Used Experts -> [127, 51, 27, 25, 53, 28, 123, 54]
  Layer 5: Used Experts -> [18, 127, 88, 125, 21, 23, 7, 112]
  Layer 9: Used Experts -> [78, 48, 27, 120, 16, 44, 76, 8]

Token '.':
  Layer 1: Used Experts -> [8, 121, 100, 15, 46, 39, 99, 73]
  Layer 5: Used Experts -> [11, 0, 116, 63, 95, 80, 113, 8]
  Layer 9: Used Experts -> [34, 21, 102, 99, 100, 62, 26, 23]

Token 'One':
  Layer 1: Used Experts -> [12, 101, 113, 106, 14, 16, 7, 96]
  Layer 5: Used Experts -> [71, 36, 44, 110, 30, 24, 90, 82]
  Layer 9: Used Experts -> [17, 31, 85, 127, 35, 50, 126, 1]

Token 'day':
  Layer 1: Used Experts -> [62, 117, 4, 78, 92, 13, 56, 71]
  Layer 5: Used Experts -> [103, 65, 76, 111, 60, 29, 26, 3]
  Layer 9: Used Experts -> [78, 48, 27, 120, 16, 8, 44, 76]

Token ',':
  Layer 1: Used Experts -> [69, 75, 70, 72, 81, 26, 95, 86]
  Layer 5: Used Experts -> [115, 2, 81, 117, 66, 12, 114, 47]
  Layer 9: Used Experts -> [57, 65, 88, 36, 68, 67, 123, 118]

Token 'the':
  Layer 1: Used Experts -> [57, 49, 82, 60, 65, 122, 94, 9]
  Layer 5: Used Experts -> [64, 108, 119, 109, 77, 106, 22, 87]
  Layer 9: Used Experts -> [112, 109, 110, 77, 69, 116, 111, 65]

Token 'fox':
  Layer 1: Used Experts -> [70, 69, 75, 81, 72, 26, 95, 86]
  Layer 5: Used Experts -> [98, 42, 101, 72, 92, 74, 46, 122]
  Layer 9: Used Experts -> [27, 76, 120, 16, 44, 8, 78, 48]

Token 'saw':
  Layer 1: Used Experts -> [20, 108, 40, 88, 31, 0, 29, 71]
  Layer 5: Used Experts -> [11, 0, 116, 63, 80, 95, 113, 8]
  Layer 9: Used Experts -> [89, 116, 111, 69, 70, 77, 110, 63]

Token 'a':
  Layer 1: Used Experts -> [50, 48, 3, 2, 64, 61, 105, 111]
  Layer 5: Used Experts -> [62, 57, 4, 27, 85, 94, 89, 49]
  Layer 9: Used Experts -> [36, 68, 67, 123, 118, 6, 13, 20]

Token 'big':
  Layer 1: Used Experts -> [66, 30, 22, 42, 38, 93, 124, 19]
  Layer 5: Used Experts -> [52, 78, 69, 37, 6, 96, 107, 70]
  Layer 9: Used Experts -> [27, 120, 16, 44, 78, 76, 48, 8]

Token ',':
  Layer 1: Used Experts -> [122, 60, 82, 57, 65, 94, 49, 9]
  Layer 5: Used Experts -> [98, 74, 46, 92, 42, 101, 72, 122]
  Layer 9: Used Experts -> [27, 76, 120, 16, 44, 8, 78, 48]

Token 'shiny':
  Layer 1: Used Experts -> [57, 49, 60, 82, 122, 65, 94, 9]
  Layer 5: Used Experts -> [64, 87, 108, 119, 109, 77, 22, 106]
  Layer 9: Used Experts -> [30, 66, 60, 81, 24, 79, 10, 2]

Token 'rock':
  Layer 1: Used Experts -> [35, 73, 15, 99, 100, 46, 39, 121]
  Layer 5: Used Experts -> [42, 92, 74, 101, 98, 72, 46, 122]
  Layer 9: Used Experts -> [48, 78, 44, 8, 16, 27, 120, 76]

Token '.':
  Layer 1: Used Experts -> [105, 2, 61, 3, 50, 64, 48, 111]
  Layer 5: Used Experts -> [20, 50, 48, 28, 58, 121, 75, 91]

Layer 9: Used Experts -> [34, 21, 100, 99, 62, 102, 26, 23]

Token 'The':
Layer 1: Used Experts -> [12, 101, 113, 106, 14, 16, 7, 96]
Layer 5: Used Experts -> [71, 30, 36, 44, 110, 24, 90, 82]
Layer 9: Used Experts -> [17, 31, 85, 127, 1, 50, 35, 126]

Token 'fox':
Layer 1: Used Experts -> [70, 69, 75, 81, 26, 72, 95, 86]
Layer 5: Used Experts -> [42, 98, 101, 92, 72, 74, 46, 122]
Layer 9: Used Experts -> [76, 27, 120, 16, 44, 8, 78, 48]

Token 'was':
Layer 1: Used Experts -> [20, 29, 108, 40, 88, 31, 0, 111]
Layer 5: Used Experts -> [0, 11, 80, 116, 95, 63, 8, 113]
Layer 9: Used Experts -> [70, 63, 92, 19, 96, 54, 64, 9]

Token 'scared':
Layer 1: Used Experts -> [98, 21, 23, 90, 87, 33, 126, 104]
Layer 5: Used Experts -> [52, 78, 69, 37, 6, 96, 107, 70]
Layer 9: Used Experts -> [32, 66, 60, 24, 79, 81, 53, 30]

Token 'and':
Layer 1: Used Experts -> [30, 22, 42, 93, 38, 66, 19, 124]
Layer 5: Used Experts -> [1, 124, 123, 14, 45, 40, 9, 120]
Layer 9: Used Experts -> [102, 26, 99, 21, 34, 23, 100, 62]

Token 'wanted':
Layer 1: Used Experts -> [85, 89, 74, 116, 52, 63, 67, 10]
Layer 5: Used Experts -> [55, 118, 79, 39, 25, 31, 73, 53]
Layer 9: Used Experts -> [89, 116, 69, 77, 110, 109, 111, 112]

Token 'to':
Layer 1: Used Experts -> [39, 15, 8, 99, 46, 100, 73, 121]
Layer 5: Used Experts -> [104, 120, 9, 45, 40, 14, 123, 86]
Layer 9: Used Experts -> [13, 6, 118, 20, 123, 67, 68, 36]

Token 'play':
Layer 1: Used Experts -> [76, 11, 68, 77, 114, 107, 47, 115]
Layer 5: Used Experts -> [68, 84, 38, 54, 43, 102, 15, 35]
Layer 9: Used Experts -> [117, 47, 73, 15, 59, 14, 41, 55]

Token 'with':
Layer 1: Used Experts -> [86, 26, 95, 81, 72, 69, 75, 70]

```
  Layer 5:        Used Experts -> [75, 58, 121, 50, 48, 28, 20, 91]
  Layer 9:        Used Experts -> [20, 13, 6, 118, 123, 67, 23, 68]

Token 'the':
  Layer 1:        Used Experts -> [87, 23, 33, 90, 21, 126, 104, 98]
  Layer 5:        Used Experts -> [126, 84, 68, 38, 102, 15, 54, 43]
  Layer 9:        Used Experts -> [67, 123, 68, 118, 6, 13, 36, 20]

Token 'fox':
  Layer 1:        Used Experts -> [70, 75, 69, 81, 72, 95, 26, 86]
  Layer 5:        Used Experts -> [98, 42, 101, 72, 92, 74, 46, 122]
  Layer 9:        Used Experts -> [48, 78, 8, 44, 16, 120, 27, 76]

Token '.':
  Layer 1:        Used Experts -> [20, 29, 108, 40, 88, 31, 0, 71]
  Layer 5:        Used Experts -> [95, 80, 0, 8, 113, 11, 116, 63]
  Layer 9:        Used Experts -> [23, 20, 26, 13, 6, 118, 123, 102]

Token 'The':
  Layer 1:        Used Experts -> [12, 113, 106, 101, 14, 16, 7, 96]
  Layer 5:        Used Experts -> [30, 71, 36, 110, 44, 90, 24, 82]
  Layer 9:        Used Experts -> [17, 31, 85, 127, 1, 50, 35, 126]

Token 'fox':
  Layer 1:        Used Experts -> [70, 69, 75, 81, 72, 26, 95, 86]
  Layer 5:        Used Experts -> [42, 98, 101, 92, 72, 74, 46, 122]
  Layer 9:        Used Experts -> [76, 27, 120, 16, 8, 44, 78, 48]

Token 'said':
  Layer 1:        Used Experts -> [20, 29, 108, 40, 88, 31, 0, 71]
  Layer 5:        Used Experts -> [95, 80, 8, 0, 113, 11, 116, 63]
  Layer 9:        Used Experts -> [89, 70, 63, 116, 92, 69, 77, 110]

Token ',':
  Layer 1:        Used Experts -> [31, 0, 40, 20, 108, 88, 71, 29]
  Layer 5:        Used Experts -> [1, 124, 105, 32, 19, 51, 67, 97]
  Layer 9:        Used Experts -> [36, 68, 67, 123, 118, 6, 13, 20]

Token '"':
  Layer 1:        Used Experts -> [60, 122, 57, 82, 49, 65, 94, 9]
  Layer 5:        Used Experts -> [87, 108, 119, 64, 109, 106, 77, 22]
  Layer 9:        Used Experts -> [112, 109, 110, 77, 69, 116, 111, 65]

Token 'I':
```

  Layer 1: Used Experts -> [86, 26, 81, 72, 95, 75, 69, 70]
  Layer 5: Used Experts -> [13, 5, 97, 51, 19, 32, 67, 105]
  Layer 9: Used Experts -> [17, 31, 85, 127, 50, 1, 35, 126]

Token 'want':
  Layer 1: Used Experts -> [86, 95, 26, 81, 72, 69, 75, 70]
  Layer 5: Used Experts -> [10, 93, 83, 34, 47, 49, 89, 94]
  Layer 9: Used Experts -> [42, 18, 45, 76, 27, 120, 16, 44]

Token 'to':
  Layer 1: Used Experts -> [41, 37, 5, 119, 6, 43, 32, 45]
  Layer 5: Used Experts -> [123, 45, 14, 40, 9, 120, 1, 104]
  Layer 9: Used Experts -> [20, 13, 6, 43, 118, 123, 67, 68]

Token 'play':
  Layer 1: Used Experts -> [76, 11, 68, 77, 107, 114, 47, 115]
  Layer 5: Used Experts -> [68, 84, 38, 54, 15, 102, 43, 35]
  Layer 9: Used Experts -> [117, 15, 73, 47, 14, 41, 59, 55]

Token 'with':
  Layer 1: Used Experts -> [86, 26, 95, 81, 72, 69, 75, 70]
  Layer 5: Used Experts -> [75, 121, 58, 91, 50, 48, 28, 20]
  Layer 9: Used Experts -> [20, 13, 6, 118, 123, 67, 23, 68]

Token 'you':
  Layer 1: Used Experts -> [87, 33, 23, 90, 21, 104, 126, 98]
  Layer 5: Used Experts -> [68, 84, 38, 102, 15, 54, 43, 35]
  Layer 9: Used Experts -> [43, 68, 67, 123, 118, 6, 13, 36]

--- Prompt ---
The recipe for the perfect cake is to
firstthestore.Thelittlegirlwasveryexcited.Sheaskedhermomiftheycouldbuysomefood.Hermomsaid,
"Yes,wecan'thavetobuysomefood."Thelittlegirlwasexcitedandexcited.Shewantedtoshow
--- End of Generation ---

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

Token 'the':
  Layer 1: Used Experts -> [48, 64, 3, 111, 50, 2, 61, 79]
  Layer 5: Used Experts -> [56, 17, 59, 99, 33, 100, 41, 16]
  Layer 9: Used Experts -> [32, 53, 66, 60, 24, 79, 81, 62]

Token 'store':
  Layer 1: Used Experts -> [70, 69, 75, 26, 81, 72, 95, 86]

Layer 5: Used Experts -> [98, 122, 72, 101, 46, 42, 74, 92]
Layer 9: Used Experts -> [76, 27, 8, 120, 16, 44, 78, 48]

Token '.':
Layer 1: Used Experts -> [104, 126, 33, 90, 87, 23, 21, 98]
Layer 5: Used Experts -> [103, 29, 111, 60, 65, 76, 3, 26]
Layer 9: Used Experts -> [53, 32, 62, 100, 34, 21, 99, 102]

Token 'The':
Layer 1: Used Experts -> [12, 106, 16, 113, 7, 14, 101, 96]
Layer 5: Used Experts -> [71, 24, 90, 82, 44, 110, 36, 30]
Layer 9: Used Experts -> [101, 1, 31, 85, 127, 50, 126, 35]

Token 'little':
Layer 1: Used Experts -> [70, 75, 69, 81, 72, 26, 95, 86]
Layer 5: Used Experts -> [122, 98, 46, 72, 101, 74, 42, 92]
Layer 9: Used Experts -> [76, 27, 120, 16, 44, 8, 78, 48]

Token 'girl':
Layer 1: Used Experts -> [51, 54, 27, 28, 123, 127, 25, 53]
Layer 5: Used Experts -> [62, 57, 4, 27, 85, 94, 89, 49]
Layer 9: Used Experts -> [76, 27, 120, 16, 44, 8, 78, 48]

Token 'was':
Layer 1: Used Experts -> [1, 110, 118, 120, 55, 59, 58, 109]
Layer 5: Used Experts -> [63, 11, 116, 0, 113, 8, 80, 95]
Layer 9: Used Experts -> [70, 63, 92, 19, 96, 54, 64, 9]

Token 'very':
Layer 1: Used Experts -> [98, 21, 23, 90, 87, 33, 126, 104]
Layer 5: Used Experts -> [52, 78, 69, 37, 6, 107, 96, 70]
Layer 9: Used Experts -> [32, 66, 60, 24, 53, 79, 81, 30]

Token 'excited':
Layer 1: Used Experts -> [53, 51, 127, 25, 54, 28, 123, 27]
Layer 5: Used Experts -> [52, 69, 78, 107, 37, 6, 70, 96]
Layer 9: Used Experts -> [66, 60, 24, 79, 81, 30, 63, 70]

Token '.':
Layer 1: Used Experts -> [43, 32, 5, 6, 119, 37, 79, 41]
Layer 5: Used Experts -> [121, 58, 75, 91, 20, 50, 48, 28]
Layer 9: Used Experts -> [53, 32, 62, 100, 34, 21, 99, 102]

Token 'She':

Layer 1: Used Experts -> [12, 113, 101, 106, 14, 16, 7, 96]
Layer 5: Used Experts -> [71, 82, 90, 30, 110, 36, 44, 24]
Layer 9: Used Experts -> [17, 31, 1, 85, 126, 127, 50, 35]

Token 'asked':
Layer 1: Used Experts -> [47, 77, 11, 114, 68, 107, 76, 115]
Layer 5: Used Experts -> [63, 11, 116, 0, 113, 8, 80, 95]
Layer 9: Used Experts -> [63, 92, 19, 70, 96, 54, 64, 9]

Token 'her':
Layer 1: Used Experts -> [35, 73, 99, 39, 46, 100, 15, 121]
Layer 5: Used Experts -> [123, 120, 14, 40, 9, 45, 104, 1]
Layer 9: Used Experts -> [36, 68, 67, 123, 118, 6, 13, 57]

Token 'mom':
Layer 1: Used Experts -> [98, 21, 23, 87, 90, 33, 126, 104]
Layer 5: Used Experts -> [112, 23, 125, 21, 7, 88, 18, 127]
Layer 9: Used Experts -> [8, 76, 78, 44, 16, 48, 120, 27]

Token 'if':
Layer 1: Used Experts -> [105, 50, 2, 61, 3, 48, 64, 111]
Layer 5: Used Experts -> [113, 8, 63, 80, 95, 116, 11, 0]
Layer 9: Used Experts -> [3, 57, 11, 107, 121, 7, 124, 46]

Token 'they':
Layer 1: Used Experts -> [32, 5, 6, 37, 43, 119, 41, 45]
Layer 5: Used Experts -> [28, 48, 91, 50, 20, 75, 121, 58]
Layer 9: Used Experts -> [65, 112, 109, 110, 77, 69, 116, 111]

Token 'could':
Layer 1: Used Experts -> [104, 126, 87, 90, 23, 33, 21, 98]
Layer 5: Used Experts -> [8, 113, 80, 95, 63, 116, 11, 0]
Layer 9: Used Experts -> [28, 95, 88, 65, 71, 80, 113, 72]

Token 'buy':
Layer 1: Used Experts -> [35, 53, 24, 54, 25, 28, 123, 51]
Layer 5: Used Experts -> [127, 18, 7, 88, 21, 23, 125, 112]
Layer 9: Used Experts -> [59, 41, 117, 47, 15, 73, 14, 55]

Token 'some':
Layer 1: Used Experts -> [35, 8, 121, 100, 46, 99, 39, 73]
Layer 5: Used Experts -> [54, 35, 43, 102, 38, 15, 124, 84]
Layer 9: Used Experts -> [13, 6, 118, 20, 123, 67, 68, 36]

Token 'food':
  Layer 1: Used Experts -> [10, 89, 116, 67, 74, 63, 85, 52]
  Layer 5: Used Experts -> [104, 86, 120, 40, 9, 14, 45, 123]
  Layer 9: Used Experts -> [119, 46, 121, 11, 124, 7, 107, 3]

Token '.':
  Layer 1: Used Experts -> [1, 110, 120, 118, 59, 55, 109, 58]
  Layer 5: Used Experts -> [123, 14, 120, 40, 9, 45, 104, 1]
  Layer 9: Used Experts -> [23, 20, 26, 13, 6, 118, 102, 123]

Token 'Her':
  Layer 1: Used Experts -> [101, 96, 113, 106, 14, 16, 7, 12]
  Layer 5: Used Experts -> [82, 90, 24, 110, 44, 71, 36, 30]
  Layer 9: Used Experts -> [17, 31, 85, 127, 1, 50, 126, 35]

Token 'mom':
  Layer 1: Used Experts -> [35, 73, 15, 99, 39, 46, 100, 121]
  Layer 5: Used Experts -> [127, 18, 88, 7, 21, 125, 23, 112]
  Layer 9: Used Experts -> [76, 27, 120, 16, 8, 44, 78, 48]

Token 'said':
  Layer 1: Used Experts -> [105, 50, 2, 61, 3, 48, 64, 1]
  Layer 5: Used Experts -> [8, 113, 95, 80, 63, 116, 11, 0]
  Layer 9: Used Experts -> [9, 64, 54, 96, 19, 92, 63, 70]

Token ',':
  Layer 1: Used Experts -> [0, 31, 40, 108, 20, 88, 71, 29]
  Layer 5: Used Experts -> [105, 32, 19, 67, 51, 97, 5, 13]
  Layer 9: Used Experts -> [57, 36, 65, 88, 68, 67, 123, 118]

Token '"':
  Layer 1: Used Experts -> [49, 57, 82, 65, 94, 9, 122, 60]
  Layer 5: Used Experts -> [87, 106, 108, 119, 77, 109, 64, 22]
  Layer 9: Used Experts -> [112, 109, 110, 77, 65, 69, 116, 111]

Token 'Yes':
  Layer 1: Used Experts -> [75, 70, 69, 72, 81, 26, 95, 86]
  Layer 5: Used Experts -> [13, 5, 97, 51, 19, 67, 32, 105]
  Layer 9: Used Experts -> [101, 50, 35, 127, 126, 1, 85, 31]

Token ',':
  Layer 1: Used Experts -> [18, 84, 102, 112, 44, 34, 24, 45]
  Layer 5: Used Experts -> [91, 20, 121, 58, 50, 48, 28, 75]
  Layer 9: Used Experts -> [3, 46, 121, 124, 119, 11, 7, 107]

Token 'we':
  Layer 1: Used Experts -> [49, 57, 82, 65, 94, 60, 9, 122]
  Layer 5: Used Experts -> [87, 108, 106, 119, 64, 109, 77, 22]
  Layer 9: Used Experts -> [112, 109, 110, 77, 69, 116, 111, 65]

Token 'can':
  Layer 1: Used Experts -> [0, 31, 88, 40, 108, 29, 20, 71]
  Layer 5: Used Experts -> [83, 93, 10, 34, 47, 49, 89, 94]
  Layer 9: Used Experts -> [42, 18, 45, 39, 56, 9, 72, 64]

Token '':
  Layer 1: Used Experts -> [105, 50, 61, 64, 2, 3, 48, 79]
  Layer 5: Used Experts -> [10, 93, 83, 34, 49, 89, 94, 47]
  Layer 9: Used Experts -> [117, 15, 73, 14, 47, 41, 59, 55]

Token 't':
  Layer 1: Used Experts -> [8, 15, 73, 39, 99, 46, 100, 121]
  Layer 5: Used Experts -> [114, 66, 12, 81, 117, 2, 115, 47]
  Layer 9: Used Experts -> [86, 33, 25, 75, 0, 122, 61, 115]

Token 'have':
  Layer 1: Used Experts -> [32, 43, 5, 6, 119, 37, 41, 79]
  Layer 5: Used Experts -> [96, 70, 6, 37, 107, 78, 69, 52]
  Layer 9: Used Experts -> [55, 41, 59, 14, 47, 15, 117, 73]

Token 'to':
  Layer 1: Used Experts -> [1, 120, 110, 118, 55, 59, 109, 58]
  Layer 5: Used Experts -> [1, 124, 123, 14, 45, 40, 120, 9]
  Layer 9: Used Experts -> [20, 43, 23, 13, 26, 6, 118, 123]

Token 'buy':
  Layer 1: Used Experts -> [76, 68, 11, 107, 77, 47, 114, 115]
  Layer 5: Used Experts -> [68, 84, 38, 15, 102, 54, 43, 35]
  Layer 9: Used Experts -> [55, 41, 59, 47, 14, 15, 117, 73]

Token 'some':
  Layer 1: Used Experts -> [35, 8, 121, 100, 46, 99, 39, 73]
  Layer 5: Used Experts -> [124, 35, 54, 43, 102, 15, 38, 84]
  Layer 9: Used Experts -> [43, 119, 107, 46, 7, 124, 121, 11]

Token 'food':
  Layer 1: Used Experts -> [43, 5, 6, 119, 32, 79, 37, 10]
  Layer 5: Used Experts -> [104, 120, 86, 14, 40, 9, 45, 123]

Layer 9: Used Experts -> [119, 46, 124, 121, 7, 11, 107, 43]

Token '."':
Layer 1: Used Experts -> [1, 110, 118, 120, 55, 59, 109, 58]
Layer 5: Used Experts -> [123, 14, 120, 40, 45, 9, 1, 104]
Layer 9: Used Experts -> [43, 20, 13, 6, 119, 23, 118, 107]

Token 'The':
Layer 1: Used Experts -> [32, 43, 5, 6, 119, 37, 41, 45]
Layer 5: Used Experts -> [25, 79, 53, 39, 31, 73, 118, 55]
Layer 9: Used Experts -> [17, 31, 85, 127, 1, 50, 35, 126]

Token 'little':
Layer 1: Used Experts -> [70, 75, 69, 72, 81, 26, 95, 86]
Layer 5: Used Experts -> [42, 74, 101, 92, 46, 72, 98, 122]
Layer 9: Used Experts -> [76, 27, 120, 16, 44, 8, 78, 48]

Token 'girl':
Layer 1: Used Experts -> [51, 54, 27, 123, 127, 28, 25, 53]
Layer 5: Used Experts -> [85, 62, 27, 4, 57, 94, 89, 49]
Layer 9: Used Experts -> [76, 27, 120, 16, 44, 8, 78, 48]

Token 'was':
Layer 1: Used Experts -> [1, 110, 118, 120, 55, 59, 58, 109]
Layer 5: Used Experts -> [63, 11, 113, 116, 8, 0, 80, 95]
Layer 9: Used Experts -> [96, 54, 19, 92, 64, 63, 9, 70]

Token 'excited':
Layer 1: Used Experts -> [98, 21, 23, 90, 33, 87, 126, 104]
Layer 5: Used Experts -> [52, 78, 69, 37, 6, 96, 107, 70]
Layer 9: Used Experts -> [32, 53, 66, 60, 24, 79, 89, 81]

Token 'and':
Layer 1: Used Experts -> [43, 79, 119, 5, 6, 32, 37, 41]
Layer 5: Used Experts -> [20, 50, 48, 28, 58, 121, 75, 91]
Layer 9: Used Experts -> [62, 100, 34, 21, 99, 102, 53, 26]

Token 'excited':
Layer 1: Used Experts -> [89, 74, 116, 63, 67, 10, 85, 52]
Layer 5: Used Experts -> [55, 118, 39, 25, 79, 31, 73, 53]
Layer 9: Used Experts -> [92, 70, 19, 63, 96, 54, 64, 9]

Token '.':
Layer 1: Used Experts -> [79, 43, 119, 6, 5, 32, 37, 41]

  Layer 5: Used Experts -> [20, 50, 48, 28, 58, 121, 75, 91]
  Layer 9: Used Experts -> [26, 62, 102, 100, 99, 23, 21, 34]

Token 'She':
  Layer 1: Used Experts -> [101, 113, 106, 14, 16, 96, 7, 12]
  Layer 5: Used Experts -> [82, 30, 90, 71, 110, 36, 24, 44]
  Layer 9: Used Experts -> [1, 126, 31, 35, 85, 127, 50, 17]

Token 'wanted':
  Layer 1: Used Experts -> [114, 11, 77, 68, 47, 107, 76, 115]
  Layer 5: Used Experts -> [63, 116, 11, 0, 113, 8, 80, 95]
  Layer 9: Used Experts -> [9, 64, 54, 96, 19, 92, 63, 70]

Token 'to':
  Layer 1: Used Experts -> [15, 39, 99, 73, 46, 100, 121, 8]
  Layer 5: Used Experts -> [120, 45, 123, 9, 40, 14, 104, 1]
  Layer 9: Used Experts -> [36, 68, 67, 123, 118, 6, 13, 20]

Token 'show':
  Layer 1: Used Experts -> [76, 11, 68, 107, 77, 114, 47, 115]
  Layer 5: Used Experts -> [68, 84, 38, 15, 102, 54, 43, 35]
  Layer 9: Used Experts -> [117, 15, 73, 47, 14, 41, 59, 55]

--- Prompt ---
The robot opened its eyes and
sawwereplayinginthepark.Thesunwasshiningbrightlyandthesunwasshiningbrightly.Suddenly,abig
dogcamerunningandthedogwasscared.Thedogranawayandthedogwasscared.Thedogranawayan
dthedogranaway
--- End of Generation ---

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

Token 'were':
  Layer 1: Used Experts -> [64, 48, 3, 50, 2, 61, 111, 105]
  Layer 5: Used Experts -> [99, 17, 56, 33, 59, 16, 100, 41]
  Layer 9: Used Experts -> [39, 56, 72, 113, 80, 71, 95, 2]

Token 'playing':
  Layer 1: Used Experts -> [83, 80, 125, 103, 97, 17, 36, 91]
  Layer 5: Used Experts -> [29, 3, 111, 26, 60, 65, 76, 126]
  Layer 9: Used Experts -> [66, 60, 24, 79, 81, 30, 10, 2]

Token 'in':
  Layer 1: Used Experts -> [29, 0, 88, 31, 40, 108, 20, 71]

Layer 5: Used Experts -> [126, 76, 65, 111, 29, 60, 26, 3]
Layer 9: Used Experts -> [62, 100, 34, 21, 99, 102, 32, 53]

Token 'the':
Layer 1: Used Experts -> [124, 19, 38, 93, 42, 22, 66, 30]
Layer 5: Used Experts -> [68, 84, 38, 102, 15, 54, 43, 35]
Layer 9: Used Experts -> [20, 13, 6, 23, 118, 123, 67, 26]

Token 'park':
Layer 1: Used Experts -> [70, 69, 75, 81, 72, 26, 95, 86]
Layer 5: Used Experts -> [126, 26, 3, 60, 76, 29, 111, 65]
Layer 9: Used Experts -> [78, 8, 48, 44, 16, 120, 27, 76]

Token '.':
Layer 1: Used Experts -> [105, 50, 2, 3, 48, 61, 64, 111]
Layer 5: Used Experts -> [59, 56, 100, 41, 99, 16, 17, 33]
Layer 9: Used Experts -> [62, 100, 53, 34, 21, 99, 102, 32]

Token 'The':
Layer 1: Used Experts -> [12, 101, 106, 113, 96, 14, 16, 7]
Layer 5: Used Experts -> [82, 90, 24, 71, 110, 44, 36, 30]
Layer 9: Used Experts -> [101, 1, 31, 85, 127, 50, 126, 35]

Token 'sun':
Layer 1: Used Experts -> [70, 75, 69, 72, 81, 26, 95, 86]
Layer 5: Used Experts -> [98, 42, 101, 72, 92, 74, 46, 122]
Layer 9: Used Experts -> [27, 120, 76, 16, 78, 44, 8, 48]

Token 'was':
Layer 1: Used Experts -> [73, 121, 100, 99, 46, 39, 15, 8]
Layer 5: Used Experts -> [0, 11, 116, 95, 63, 80, 113, 8]
Layer 9: Used Experts -> [63, 70, 92, 19, 96, 54, 64, 9]

Token 'shining':
Layer 1: Used Experts -> [98, 21, 23, 90, 87, 33, 126, 104]
Layer 5: Used Experts -> [29, 111, 65, 60, 76, 3, 26, 126]
Layer 9: Used Experts -> [66, 60, 24, 79, 81, 30, 32, 10]

Token 'brightly':
Layer 1: Used Experts -> [64, 61, 3, 2, 48, 50, 111, 105]
Layer 5: Used Experts -> [86, 104, 126, 120, 9, 40, 45, 14]
Layer 9: Used Experts -> [32, 53, 66, 60, 24, 79, 81, 30]

Token 'and':

Layer 1: Used Experts -> [35, 24, 8, 121, 46, 100, 99, 73]
Layer 5: Used Experts -> [86, 104, 120, 40, 14, 9, 45, 123]
Layer 9: Used Experts -> [53, 32, 62, 100, 34, 21, 99, 102]

Token 'the':
Layer 1: Used Experts -> [116, 74, 89, 63, 67, 85, 10, 52]
Layer 5: Used Experts -> [25, 55, 39, 118, 79, 31, 53, 73]
Layer 9: Used Experts -> [89, 116, 111, 69, 77, 110, 109, 112]

Token 'sun':
Layer 1: Used Experts -> [70, 75, 69, 72, 81, 26, 95, 86]
Layer 5: Used Experts -> [98, 72, 101, 42, 122, 92, 74, 46]
Layer 9: Used Experts -> [78, 48, 27, 120, 16, 44, 76, 8]

Token 'was':
Layer 1: Used Experts -> [73, 121, 100, 99, 46, 15, 39, 8]
Layer 5: Used Experts -> [28, 20, 48, 50, 58, 75, 121, 91]
Layer 9: Used Experts -> [89, 70, 63, 92, 116, 111, 19, 69]

Token 'shining':
Layer 1: Used Experts -> [98, 21, 23, 90, 87, 33, 126, 104]
Layer 5: Used Experts -> [29, 3, 111, 65, 60, 26, 76, 126]
Layer 9: Used Experts -> [32, 53, 66, 60, 24, 62, 79, 81]

Token 'brightly':
Layer 1: Used Experts -> [64, 61, 3, 2, 48, 50, 111, 105]
Layer 5: Used Experts -> [86, 104, 120, 9, 40, 45, 14, 123]
Layer 9: Used Experts -> [32, 53, 62, 100, 34, 21, 99, 102]

Token '.':
Layer 1: Used Experts -> [24, 35, 18, 44, 84, 112, 102, 34]
Layer 5: Used Experts -> [86, 104, 20, 50, 48, 28, 58, 121]
Layer 9: Used Experts -> [62, 100, 34, 21, 99, 102, 53, 26]

Token 'Suddenly':
Layer 1: Used Experts -> [12, 106, 101, 113, 14, 16, 7, 96]
Layer 5: Used Experts -> [71, 90, 44, 24, 110, 36, 82, 30]
Layer 9: Used Experts -> [17, 31, 85, 127, 1, 50, 126, 35]

Token ',':
Layer 1: Used Experts -> [35, 73, 99, 15, 39, 46, 100, 121]
Layer 5: Used Experts -> [53, 73, 31, 25, 39, 79, 118, 55]
Layer 9: Used Experts -> [111, 110, 112, 77, 109, 69, 53, 116]

Token 'a':
  Layer 1: Used Experts -> [60, 122, 82, 57, 65, 94, 49, 9]
  Layer 5: Used Experts -> [77, 106, 119, 22, 64, 109, 108, 87]
  Layer 9: Used Experts -> [111, 110, 77, 109, 112, 69, 116, 53]

Token 'big':
  Layer 1: Used Experts -> [124, 19, 38, 42, 93, 22, 66, 30]
  Layer 5: Used Experts -> [52, 69, 78, 37, 6, 107, 96, 70]
  Layer 9: Used Experts -> [48, 27, 78, 120, 44, 16, 76, 8]

Token 'dog':
  Layer 1: Used Experts -> [122, 82, 57, 49, 94, 65, 9, 60]
  Layer 5: Used Experts -> [98, 46, 122, 74, 72, 101, 42, 92]
  Layer 9: Used Experts -> [78, 48, 27, 120, 44, 16, 8, 76]

Token 'came':
  Layer 1: Used Experts -> [15, 73, 121, 100, 99, 46, 39, 8]
  Layer 5: Used Experts -> [0, 11, 116, 63, 80, 95, 113, 8]
  Layer 9: Used Experts -> [32, 53, 89, 111, 116, 69, 77, 62]

Token 'running':
  Layer 1: Used Experts -> [109, 59, 1, 120, 118, 55, 110, 58]
  Layer 5: Used Experts -> [86, 104, 120, 9, 40, 45, 14, 123]
  Layer 9: Used Experts -> [62, 100, 34, 21, 99, 102, 53, 26]

Token 'and':
  Layer 1: Used Experts -> [35, 105, 53, 54, 25, 123, 28, 27]
  Layer 5: Used Experts -> [86, 104, 112, 23, 125, 21, 7, 88]
  Layer 9: Used Experts -> [53, 32, 62, 100, 34, 21, 99, 102]

Token 'the':
  Layer 1: Used Experts -> [10, 89, 74, 116, 63, 67, 85, 52]
  Layer 5: Used Experts -> [55, 118, 39, 25, 79, 31, 53, 73]
  Layer 9: Used Experts -> [89, 111, 116, 69, 77, 110, 109, 112]

Token 'dog':
  Layer 1: Used Experts -> [70, 75, 69, 81, 72, 26, 95, 86]
  Layer 5: Used Experts -> [98, 72, 101, 42, 122, 92, 74, 46]
  Layer 9: Used Experts -> [78, 48, 27, 44, 120, 16, 8, 76]

Token 'was':
  Layer 1: Used Experts -> [15, 73, 121, 99, 100, 46, 39, 8]
  Layer 5: Used Experts -> [0, 11, 116, 63, 80, 95, 113, 8]
  Layer 9: Used Experts -> [53, 32, 89, 111, 116, 69, 77, 110]

Token 'scared':
  Layer 1: Used Experts -> [98, 21, 23, 90, 33, 87, 126, 104]
  Layer 5: Used Experts -> [29, 111, 65, 76, 60, 3, 26, 126]
  Layer 9: Used Experts -> [32, 53, 62, 100, 34, 21, 99, 102]

Token '.':
  Layer 1: Used Experts -> [22, 30, 42, 93, 38, 19, 124, 66]
  Layer 5: Used Experts -> [20, 1, 50, 48, 28, 58, 121, 91]
  Layer 9: Used Experts -> [34, 21, 100, 99, 102, 62, 26, 23]

Token 'The':
  Layer 1: Used Experts -> [12, 106, 101, 113, 14, 16, 7, 96]
  Layer 5: Used Experts -> [82, 90, 24, 110, 44, 36, 71, 30]
  Layer 9: Used Experts -> [17, 31, 1, 85, 127, 50, 126, 35]

Token 'dog':
  Layer 1: Used Experts -> [70, 75, 69, 72, 81, 26, 95, 86]
  Layer 5: Used Experts -> [98, 72, 42, 101, 92, 74, 122, 46]
  Layer 9: Used Experts -> [27, 78, 120, 16, 44, 8, 76, 48]

Token 'ran':
  Layer 1: Used Experts -> [15, 73, 99, 100, 121, 46, 39, 8]
  Layer 5: Used Experts -> [0, 11, 116, 63, 80, 95, 113, 8]
  Layer 9: Used Experts -> [89, 111, 116, 70, 69, 63, 32, 77]

Token 'away':
  Layer 1: Used Experts -> [105, 2, 61, 50, 3, 48, 64, 111]
  Layer 5: Used Experts -> [86, 104, 120, 9, 40, 14, 45, 123]
  Layer 9: Used Experts -> [62, 100, 34, 21, 99, 53, 102, 32]

Token 'and':
  Layer 1: Used Experts -> [105, 50, 48, 2, 3, 61, 64, 111]
  Layer 5: Used Experts -> [20, 50, 48, 28, 58, 75, 121, 91]
  Layer 9: Used Experts -> [62, 100, 34, 21, 99, 102, 26, 53]

Token 'the':
  Layer 1: Used Experts -> [89, 116, 10, 74, 63, 67, 85, 52]
  Layer 5: Used Experts -> [55, 118, 39, 25, 79, 31, 53, 73]
  Layer 9: Used Experts -> [89, 116, 111, 69, 77, 110, 109, 112]

Token 'dog':
  Layer 1: Used Experts -> [70, 75, 81, 69, 72, 95, 26, 86]
  Layer 5: Used Experts -> [98, 72, 101, 42, 92, 122, 74, 46]

Layer 9: Used Experts -> [78, 48, 8, 27, 44, 120, 16, 76]

Token 'was':
Layer 1: Used Experts -> [15, 73, 99, 100, 121, 46, 39, 8]
Layer 5: Used Experts -> [0, 11, 116, 63, 80, 95, 113, 8]
Layer 9: Used Experts -> [89, 111, 116, 69, 77, 110, 53, 109]

Token 'scared':
Layer 1: Used Experts -> [98, 21, 23, 90, 33, 87, 126, 104]
Layer 5: Used Experts -> [52, 78, 69, 37, 6, 107, 96, 70]
Layer 9: Used Experts -> [32, 53, 62, 100, 34, 21, 99, 102]

Token '.':
Layer 1: Used Experts -> [22, 30, 42, 93, 38, 19, 124, 66]
Layer 5: Used Experts -> [1, 20, 50, 48, 28, 58, 121, 91]
Layer 9: Used Experts -> [102, 99, 21, 34, 100, 62, 26, 23]

Token 'The':
Layer 1: Used Experts -> [12, 106, 113, 101, 14, 16, 7, 96]
Layer 5: Used Experts -> [90, 44, 110, 24, 36, 82, 71, 30]
Layer 9: Used Experts -> [17, 1, 126, 31, 85, 127, 50, 35]

Token 'dog':
Layer 1: Used Experts -> [70, 75, 69, 72, 81, 26, 95, 86]
Layer 5: Used Experts -> [98, 72, 42, 101, 122, 74, 92, 46]
Layer 9: Used Experts -> [78, 48, 27, 8, 44, 120, 16, 76]

Token 'ran':
Layer 1: Used Experts -> [15, 73, 99, 100, 121, 46, 39, 8]
Layer 5: Used Experts -> [0, 11, 116, 63, 80, 95, 113, 8]
Layer 9: Used Experts -> [89, 111, 116, 69, 77, 70, 32, 110]

Token 'away':
Layer 1: Used Experts -> [105, 2, 61, 50, 3, 48, 64, 111]
Layer 5: Used Experts -> [86, 104, 120, 9, 40, 14, 45, 123]
Layer 9: Used Experts -> [62, 100, 34, 21, 99, 102, 26, 23]

Token 'and':
Layer 1: Used Experts -> [105, 50, 48, 2, 3, 61, 64, 111]
Layer 5: Used Experts -> [50, 48, 28, 20, 58, 75, 121, 91]
Layer 9: Used Experts -> [62, 100, 34, 21, 99, 102, 26, 23]

Token 'the':
Layer 1: Used Experts -> [116, 89, 74, 10, 63, 67, 85, 52]

Layer 5: Used Experts -> [55, 118, 39, 25, 79, 31, 53, 73]
Layer 9: Used Experts -> [89, 116, 69, 111, 77, 110, 109, 112]

Token 'dog':
Layer 1: Used Experts -> [70, 75, 81, 72, 69, 95, 26, 86]
Layer 5: Used Experts -> [98, 72, 101, 42, 122, 92, 74, 46]
Layer 9: Used Experts -> [78, 48, 8, 44, 27, 16, 120, 76]

Token 'ran':
Layer 1: Used Experts -> [15, 73, 99, 100, 121, 46, 39, 8]
Layer 5: Used Experts -> [0, 11, 116, 63, 80, 95, 113, 8]
Layer 9: Used Experts -> [89, 111, 116, 69, 77, 110, 109, 53]

Token 'away':
Layer 1: Used Experts -> [105, 2, 61, 50, 3, 48, 64, 111]
Layer 5: Used Experts -> [86, 104, 120, 9, 40, 14, 45, 123]
Layer 9: Used Experts -> [100, 62, 34, 21, 99, 102, 26, 23]