

## Exploring Andrej Karpathy's llama2.c aka Baby Llama

Credit: <https://github.com/karpathy/llama2.c>

### Downloading the smaller model

```
! wget https://karpathy.ai/llama2c/model.bin -P out

--2023-07-24 18:08:14-- https://karpathy.ai/llama2c/model.bin
Resolving karpathy.ai (karpathy.ai)... 151.101.1.195, 151.101.65.195
Connecting to karpathy.ai (karpathy.ai)|151.101.1.195|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 60816028 (58M) [application/octet-stream]
Saving to: 'out/model.bin'

model.bin          100%[=====>]  58.00M  191MB/s   in 0.3s

2023-07-24 18:08:15 (191 MB/s) - 'out/model.bin' saved [60816028/60816028]
```

### Cloning the llama2.c repo to get the required files

```
!git clone https://github.com/karpathy/llama2.c

Cloning into 'llama2.c'...
remote: Enumerating objects: 184, done.
remote: Counting objects: 100% (96/96), done.
remote: Compressing objects: 100% (35/35), done.
remote: Total 184 (delta 76), reused 73 (delta 61), pack-reused 88
Receiving objects: 100% (184/184), 604.08 KiB | 2.25 MiB/s, done.
Resolving deltas: 100% (106/106), done.
```

```
%cd llama2.c

/content/llama2.c
```

### Compiling the C code with O3 (Refer - <https://github.com/karpathy/llama2.c/issues/20>)

```
!gcc -O3 -o run run.c -lm
```

```
! ./run ../out/model.bin
```

Once upon a time, a boy named Tim went to the museum with his mom. They saw big animal:  
In one room, Tim saw a big dinosaur. He wanted to touch it. His mom said, "No, Tim. We r

The man said, "I lost my toy. Can you help find it?" Tim's mom said, "No, we have to help him find it."  
<s>

Once upon a time, there was a little girl named Lily. She loved to have tea parties with her friends.  
As she was pouring her tea set into the cups, she saw a big crab walking in the forest.  
achieved tok/s: 38.197553

Compiling the C code with Ofast (Refer - <https://github.com/karpathy/llama2.c/issues/20>)

```
!gcc -Ofast -o run run.c -lm
```

```
! ./run ../out/model.bin
```

Once upon a time, there was a little girl named Lily. She loved to play with her toys and her friends.  
Max came over and saw the car. "Wow, that's a nice car!" he exclaimed.  
Lily was so happy to see her friend's toy car. She mentioned the car to Max and he played with it.  
Lily's mom smiled and said, "That's great, Lily. You should be persistent about finding your toys."  
<s>

Once upon a time, there was a boy named Timmy. Timmy loved to play in the rain, but his mom told him not to.  
achieved tok/s: 103.018109

```
!wget https://karpathy.ai/llama2c/model44m.bin -P out44m
```

```
!./run out44m/model44m.bin
```

```
--2023-07-24 18:08:28-- https://karpathy.ai/llama2c/model44m.bin
Resolving karpathy.ai (karpathy.ai)... 151.101.1.195, 151.101.65.195
Connecting to karpathy.ai (karpathy.ai)|151.101.1.195|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 167020572 (159M) [application/octet-stream]
Saving to: 'out44m/model44m.bin'
```

```
model44m.bin          100%[=====>] 159.28M   199MB/s   in 0.8s
```

```
2023-07-24 18:08:29 (199 MB/s) - 'out44m/model44m.bin' saved [167020572/167020572]
```

Once upon a time, there was a little boy named Timmy. Timmy had a little puppy named Max. Max opened his mouth wide and gave Timmy a little lick on the face. Timmy laughed and said, "Max is so cute!"  
Timmy decided to go outside and play with his friends. He saw a little bird on the ground. Timmy's friend asked, "Why is your dog so little?"  
Timmy said, "Max is very small and like to cry when he's scared. I don't want him to be big."  
Timmy's friend smiled and said, "You're very brave, Timmy!"  
<s>

Once upon a time, there was a big box. The box had a cover. The cover was to keep something safe.  
One day, the box opened. Inside the box, there was a perfect, small box. The small box was made of wood.  
The big, soft box and the perfect box with the perfect hole were happy. They were happy to be together.  
<s>

One day, Mark wanted to explore. He decided to go on an adventure. He put on his shoes and a backpack.  
Mark noticed that the door was a bit gloomy. He opened the door and saw a big tree. Mark was scared.  
Mark could see a lot of things outside. There were old chairs, a cute dog and a dusty picture.  
Mark was so excited that he opened the door again. He ran out of the door and into the forest.

Mark ran around the garden and looked at all of the things. The sun was shining and the Mark was happy. He closed the door and ran off into the sunset. It was a

<s>

Once upon a time, there was a boy named Timmy. Timmy loved to play outside in his back. His mom smiled and said, "That's a very large leaf, Timmy. Do you want to keep it?" Timmy nodded his head and said, "Yes, please! Can I tie it to my shoe?" His mom helped him tie the leaf to his shoe. Timmy was so happy to have his leaf as his. His dad smiled and said, "That's great, Timmy. You have a good eye for finding treasure!"

<s>

Anna liked to play in the garden with her dog, Rex. She liked to run and jump and dig. One day, Anna saw a big pear on the ground. It was black and shape and had a long stem. She ran to her house to show her mom. She wanted to share her peach with her. She opened. But then she felt a sharp pain in her mouth. She spit out the peach and saw blood. She achieved tok/s: 29.339293

!./run\_out44m/model44m.bin

Lily and Tom were friends who liked to play in the garden. One day, they saw a big, juicy pear. Lily was angry and said, "You are not nice, Tom. You are mean. I will get the pear and eat it!" Tom shouted, "Stop, Lily! Don't do that! The pear belongs to me!" He tried to pull the pear away from her. Lily saw the dog and screamed. She said, "Tom, watch out! The dog is coming! He will hurt you!" Tom was sad and hurt. He had no pear and he had no friend. He said, "Lily, I am sorry." Lily said, "Yes, Tom, I forgive you. But next time, you should listen to me. And don't do that again." They hugged and went to find something else to play with. They learned a lesson that day.

<s>

Once upon a time, there was a little seed named Sammy. Sammy was very small and always worried. Wally said, "Hello little seed! What are you worried about?" "I am worried about what I will be like. I am always seedy," said Sammy. Wally smiled and said, "Don't worry little seedie. You will be just as special as a big tree." Sammy was surprised by Wally's wisdom. He didn't know that he could grow from his worries. "Wow, my little seed became so big and strong now!" exclaimed Sammy. Wally chuckled and said, "You sure did! You are a clever little seed."

<s>

Lily and Sam were best friends. They liked to play in the park, draw pictures, and share secrets. "Maybe it is a surprise!" Sam said. "Let's open it and see!" Lily said. They pulled the tape and lifted the lid. Inside, they saw a big snake. It was green and yellow. "Ah! A snake!" Lily screamed. "Run!" Sam said. They dropped the lid and ran away. The snake slithered back to the box and hid. It was there for days. The next day, Lily felt sick. She had a fever and a sore throat. She stayed in bed and was worried. That night, Sam had a plan. He made a hole in the box and put some cream on Lily's fever. Lily did not think much of it. She stayed in bed and worried. The next morning, Sam came to visit her. He took off the bed and the pillow and the medicine. "Look what I found, Lily! A snake!" he said. Lily was shocked and scared. She did not want to be hurt. She did not want to see the snake. "Who are you, Sam?" she asked. "I am Sam. I am your friend. Don't worry, I can reverse the snake back into the box," he said. Lily sighed. She knew that Sam was right. She was angry and sad at the same time. She wished they could be like the snake and hide away from the world. achieved tok/s: 28.755967

## Learning the code from ChatGPT Code Interpreter

### 1. Configuration and Initialization

- Define structures: Config, TransformerWeights, RunState
- Allocate memory for RunState and TransformerWeights

### 2. Read Checkpoint

- Initialize the transformer weights from a checkpoint file

### 3. Main Function

- Read model configuration and weights from a checkpoint file
- Read vocabulary from a tokenizer file
- Initialize the RunState

### 4. Start Loop for Sequence Generation

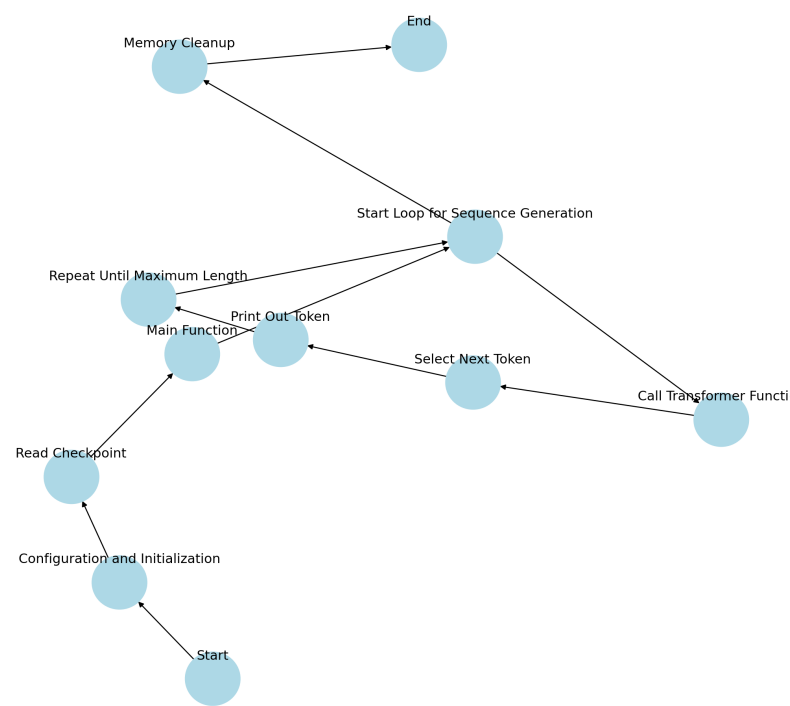
- Call the transformer function to get the output logits for the next token
  - Apply attention mechanism, softmax, RMS normalization, etc.
- Select the next token using sampling or argmax
- Print out the token
- Repeat until a sequence of the maximum length is generated

### 5. Memory Cleanup

- Deallocate memory for RunState and TransformerWeights

## Flow Chart from Code Interpreter

Flowchart of Transformer Model Execution



Start coding or generate with AI.