



The scientist
named the
population, after their
distinctive horn,
Ovid's Unicorn.

GPT2: Zero Shot的崛起

Language Model

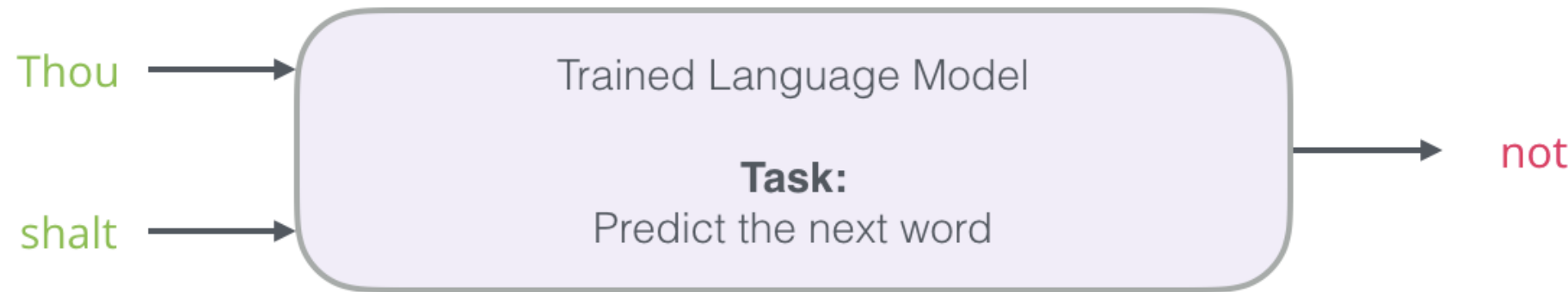
input/feature #1 input/feature #2 output/label

Thou shalt _____



Input
Features

Output
Prediction



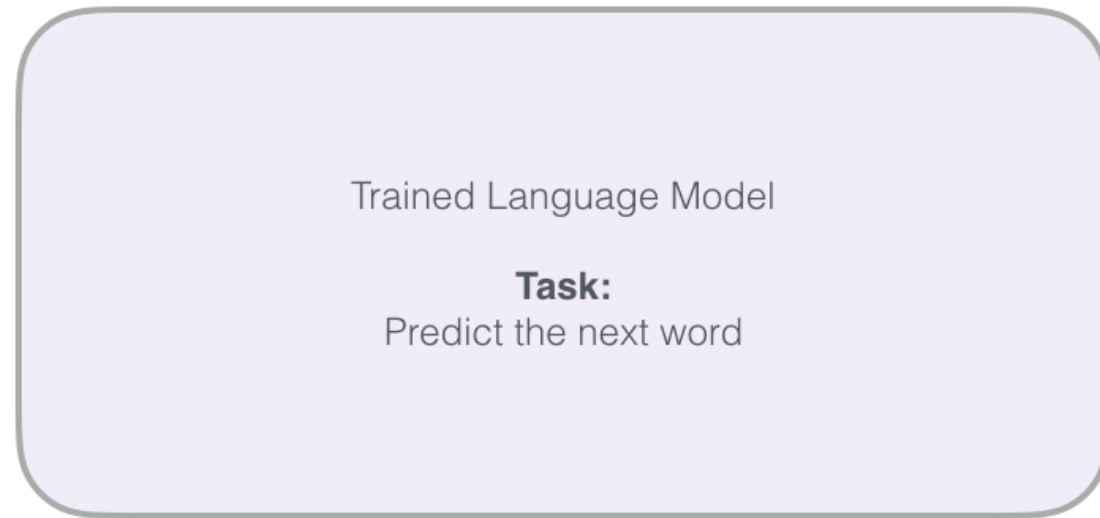
Language Model

Input
Features

Thou



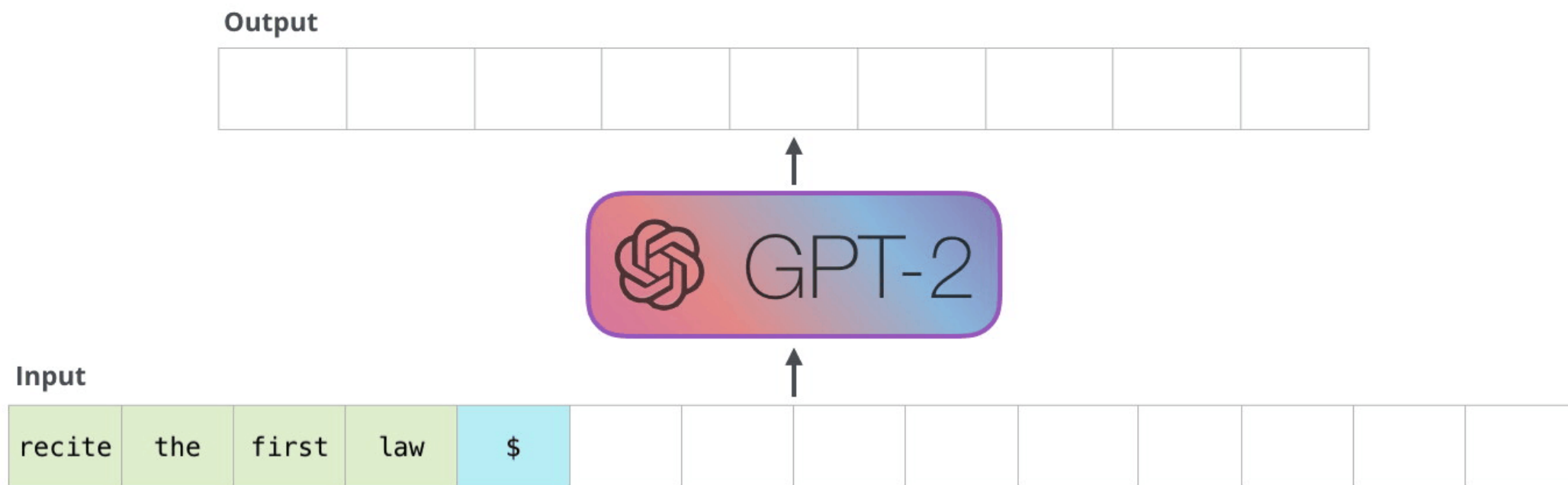
shalt



Output
Prediction

0%	aardvark
0%	aarhus
0.1%	aaron
...	
40%	not
...	
0.01	zyzzyva

Auto Regressive



Learning Objectives and Concepts

- Task Conditioning
- Zero Shot Learning and Zero Short Task Transfer

Task Conditioning

Training objective of language model is formulated as:

$$P(\text{output}|\text{input})$$

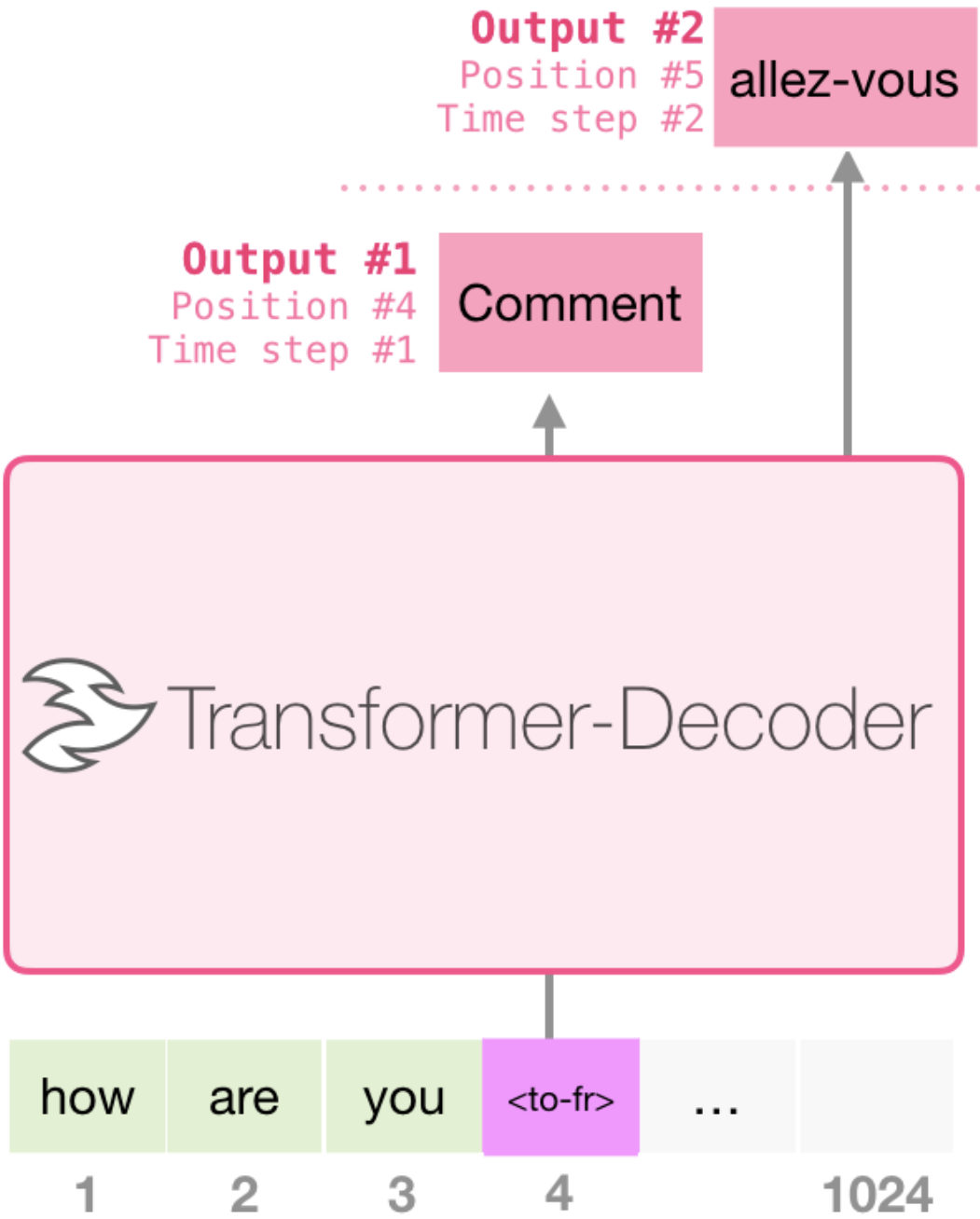
However, GPT-2 aimed at learning multiple tasks using the same unsupervised model. To achieve that, the learning objective should be modified to:

$$P(\text{output}|\text{input}, \text{task})$$

This modification is known as task conditioning, where the model is expected to produce different output for same input for different tasks. Some models implement task conditioning at an architectural level where the model is fed both, the input and the task. For language models, the output, input and task, all are sequences of natural language.

Training Dataset

I	am	a	student	<to-fr>	je	suis	étudiant
let	them	eat	cake	<to-fr>	Qu'ils	mangent	de
good	morning	<to-fr>	Bonjour				





Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

Languages

Esperanto
Français
Italiano
日本語
Português
Svenska
Türkçe

30 more

Edit links

Article Talk

Read Edit View history Search Wikipedia

Positronic brain

From Wikipedia, the free encyclopedia
(Redirected from *Positronic robot*)

*This article is about a fictional technological device. For the manufacturing company based in Springfield, Missouri, see *Positronic* (company).*



This article **needs additional citations for verification**. Please help improve this article by adding citations to reliable sources. Unourced material may be challenged and removed.
Find sources: "Positronic brain" – news – newspapers – books – scholar – JSTOR (July 2008) (Learn how and when to remove this template message)

A **positronic brain** is a fictional technological device, originally conceived by science fiction writer Isaac Asimov.^{[1][2]} It functions as a central processing unit (CPU) for robots, and, in some unspecified way, provides them with a form of consciousness recognizable to humans. When Asimov wrote his first robot stories in 1939 and 1940, the positron was a newly discovered particle, and so the buzz word positronic added a contemporary gloss of popular science to the concept. The short story "Runaround", by Asimov, elaborates on the concept, in the context of his fictional Three Laws of Robotics.

Contents

- Conceptual overview
- In Allen's trilogy
- References in other fiction and films
 - Abbott and Costello Go To Mars
 - The Avengers*
 - Doctor Who*
 - Star Trek*
 - Perry Rhodan*
 - I, Robot*, 2004 Film
 - Bicentennial Man*
 - Buck Rogers in the 25th Century*
 - Mystery Science Theater 3000*
 - Spectruman*
 - Stallone*
- References
- External links

Conceptual overview

Asimov remained vague about the technical details of positronic brains except to assert that their substructure was formed from an alloy of platinum and indium. They were said to be vulnerable to radiation and apparently involve a type of *volatile memory* (since robots in storage required a power source keeping their brains "alive"). The focus of Asimov's stories was directed more towards the *software* of robots—such as the Three Laws of Robotics—than the hardware in which it was implemented, although it is stated in his stories that to create a positronic brain without the Three Laws, it would have been necessary to spend years redesigning the fundamental approach towards the brain itself.

Within his stories of *robotics* on *Earth* and their development by *U.S. Robots*, Asimov's positronic brain is less of a *plot device* and more of a technological item worthy of study.

A positronic brain cannot ordinarily be built without incorporating the Three Laws; any modification thereof would drastically modify robot behavior. Behavioral dilemmas resulting from conflicting potentials set by inexperienced and/or malicious users of the robot for the Three Laws make up the bulk of Asimov's stories concerning robots. They are resolved by applying the *science* of logic and *psychology* together with *mathematics*, the supreme solution finder being Dr. Susan Calvin, Chief Robopsychologist of U.S. Robots.

The Three Laws are also a *bottleneck* in brain sophistication. Very complex brains designed to handle world economy interpret the First Law in expanded sense to include humanity as opposed to a single human; in Asimov's later works like *Robots and Empire* this is referred to as the "Zeroth Law". At least one brain constructed as a calculating *machine*, as opposed to being a robot control circuit, was designed to have a flexible, childlike personality so that it was able to pursue difficult problems without the Three Laws inhibiting it completely. Specialized brains created for overseeing world economics were stated to have no personality at all.

Under specific conditions, the Three Laws can be obviated, with the modification of the actual robotic design.

- Robots that are of low enough value can have the **Third Law** deleted; they do not have to protect themselves from harm, and the brain size can be reduced by half.
- Robots that do not require orders from a human being may have the **Second Law** deleted, and therefore require smaller brains again, providing they do not require the Third Law.
- Robots that are disposable, cannot receive orders from a human being and are not able to harm a human, will not require even the **First Law**. The sophistication of positronic circuitry renders a brain so small that it could comfortably fit within the skull of an insect.

Robots of the latter type directly parallel contemporary industrial robotics practice, though real-life robots do contain safety sensors and systems, in a concern for human safety (a weak form of the First Law; the robot is a safe tool to use, but has no "judgment", which is implicit in Asimov's own stories).

In Allen's trilogy

Several robot stories have been written by other authors following Asimov's death. For example, in Roger MacBride Allen's *Caliban* trilogy, a *Spacer* robotist called Gubber Anshaw invents the **gravitronic brain**. It offers speed and capacity improvements over traditional positronic designs, but the strong influence of tradition make robotics labs reject Anshaw's work. Only one robotist, Freda Leving, chooses to adopt gravitronics, because it offers her a blank slate on which she could explore alternatives to the Three Laws. Because they are not dependent upon centuries of earlier research, gravitronic brains can be programmed with the standard Laws, variations of the Laws, or even empty pathways which specify no Laws at all.



Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

Languages

Esperanto
Français
Italiano
日本語
Português
Svenska
Türkçe

30 more

Edit links

Article Talk

Read Edit View history Search Wikipedia

Positronic brain

From Wikipedia, the free encyclopedia
(Redirected from *Positronic robot*)

*This article is about a fictional technological device. For the manufacturing company based in Springfield, Missouri, see *Positronic* (company).*



This article **needs additional citations for verification**. Please help improve this article by adding citations to reliable sources. Unourced material may be challenged and removed.
Find sources: "Positronic brain" – news – newspapers – books – scholar – JSTOR (July 2008) (Learn how and when to remove this template message)

A **positronic brain** is a fictional technological device, originally conceived by science fiction writer Isaac Asimov.^{[1][2]} It functions as a central processing unit (CPU) for robots, and, in some unspecified way, provides them with a form of consciousness recognizable to humans. When Asimov wrote his first robot stories in 1939 and 1940, the positron was a newly discovered particle, and so the buzz word positronic added a contemporary gloss of popular science to the concept. The short story "Runaround", by Asimov, elaborates on the concept, in the context of his fictional Three Laws of Robotics.

Contents

- Conceptual overview
- In Allen's trilogy
- References in other fiction and films
 - Abbott and Costello Go To Mars
 - The Avengers*
 - Doctor Who*
 - Star Trek*
 - Perry Rhodan*
 - I, Robot*, 2004 Film
 - Bicentennial Man*
 - Buck Rogers in the 25th Century*
 - Mystery Science Theater 3000*
 - Spectruman*
 - Stallone*
- References
- External links

Conceptual overview

Asimov remained vague about the technical details of positronic brains except to assert that their substructure was formed from an alloy of platinum and indium. They were said to be vulnerable to radiation and apparently involve a type of *volatile memory* (since robots in storage required a power source keeping their brains "alive"). The focus of Asimov's stories was directed more towards the *software* of robots—such as the Three Laws of Robotics—than the hardware in which it was implemented, although it is stated in his stories that to create a positronic brain without the Three Laws, it would have been necessary to spend years redesigning the fundamental approach towards the brain itself.

Within his stories of *robotics* on *Earth* and their development by *U.S. Robots*, Asimov's positronic brain is less of a *plot device* and more of a technological item worthy of study.

A positronic brain cannot ordinarily be built without incorporating the Three Laws; any modification thereof would drastically modify robot behavior. Behavioral dilemmas resulting from conflicting potentials set by inexperienced and/or malicious users of the robot for the Three Laws make up the bulk of Asimov's stories concerning robots. They are resolved by applying the *science* of logic and *psychology* together with *mathematics*, the supreme solution finder being Dr. Susan Calvin, Chief Robopsychologist of U.S. Robots.

The Three Laws are also a *bottleneck* in brain sophistication. Very complex brains designed to handle world economy interpret the First Law in expanded sense to include humanity as opposed to a single human; in Asimov's later works like *Robots and Empire* this is referred to as the "Zeroth Law". At least one brain constructed as a calculating *machine*, as opposed to being a robot control circuit, was designed to have a flexible, childlike personality so that it was able to pursue difficult problems without the Three Laws inhibiting it completely. Specialized brains created for overseeing world economics were stated to have no personality at all.

Under specific conditions, the Three Laws can be obviated, with the modification of the actual robotic design.

- Robots that are of low enough value can have the **Third Law** deleted; they do not have to protect themselves from harm, and the brain size can be reduced by half.
- Robots that do not require orders from a human being may have the **Second Law** deleted, and therefore require smaller brains again, providing they do not require the Third Law.
- Robots that are disposable, cannot receive orders from a human being and are not able to harm a human, will not require even the **First Law**. The sophistication of positronic circuitry renders a brain so small that it could comfortably fit within the skull of an insect.

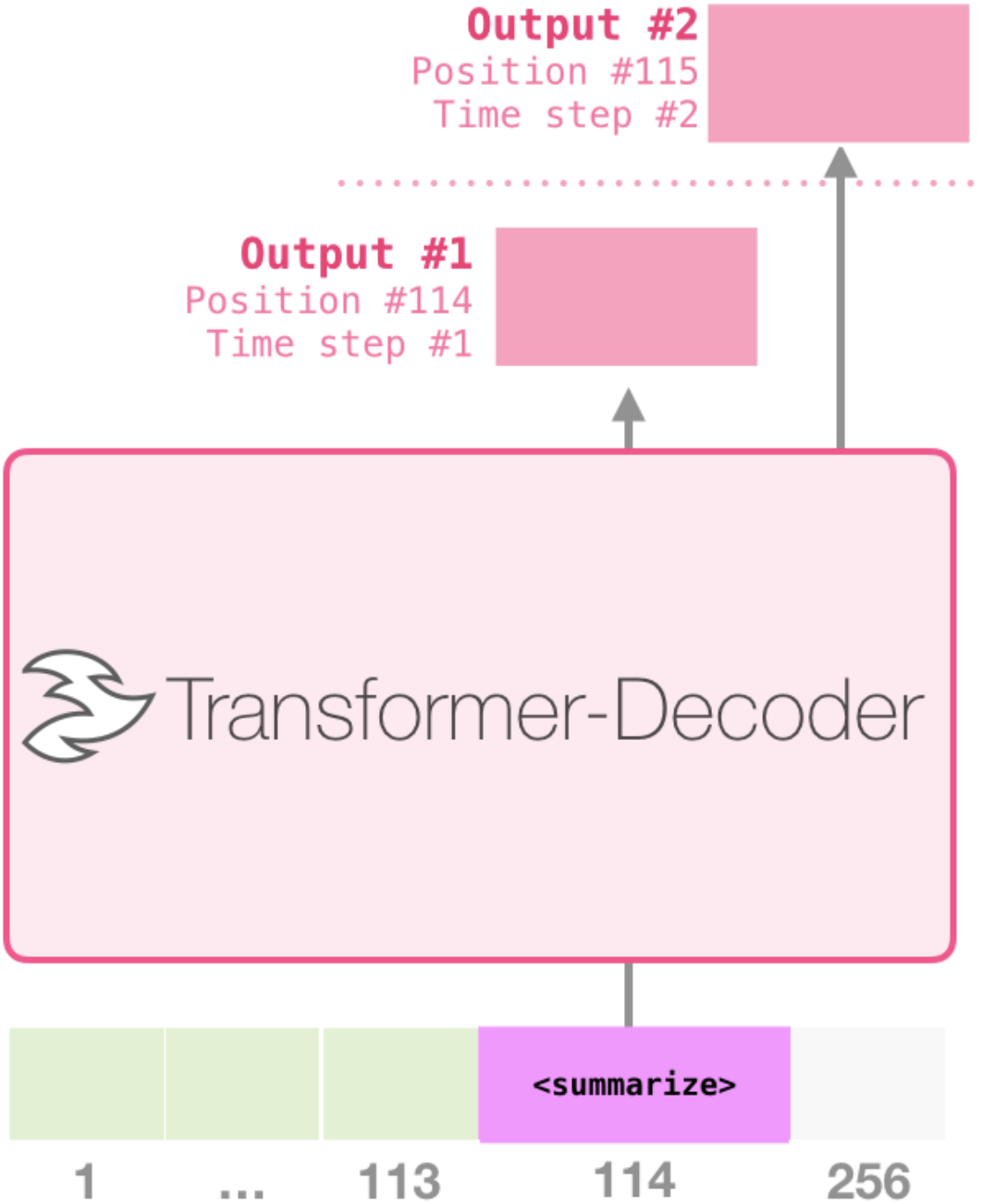
Robots of the latter type directly parallel contemporary industrial robotics practice, though real-life robots do contain safety sensors and systems, in a concern for human safety (a weak form of the First Law; the robot is a safe tool to use, but has no "judgment", which is implicit in Asimov's own stories).

In Allen's trilogy

Several robot stories have been written by other authors following Asimov's death. For example, in Roger MacBride Allen's *Caliban* trilogy, a *Spacer* robotist called Gubber Anshaw invents the **gravitronic brain**. It offers speed and capacity improvements over traditional positronic designs, but the strong influence of tradition make robotics labs reject Anshaw's work. Only one robotist, Freda Leving, chooses to adopt gravitronics, because it offers her a blank slate on which she could explore alternatives to the Three Laws. Because they are not dependent upon centuries of earlier research, gravitronic brains can be programmed with the standard Laws, variations of the Laws, or even empty pathways which specify no Laws at all.

Training Dataset

Article #1 tokens		<summarize>	Article #1 Summary
Article #2 tokens	<summarize>	Article #2 Summary	padding
Article #3 tokens		<summarize>	Article #3 Summary



Zero Shot Learning and Task Transfer

- An interesting capability of GPT 2 is zero shot task transfer. Zero shot learning is a special case of zero shot task transfer where no examples are provided at all and the model understands the task based on the given instruction. Instead of rearranging the sequences, as was done for GPT-1 for fine-tuning, input to GPT-2 was given in a format which expected the model to understand the nature of task and provide answers. This was done to emulate zero-shot task transfer behavior.
- E.g. for English to French translation task, the model was given an English sentence followed by the word French and a prompt (:). The model was supposed to understand that it is a translation task and give French counterpart of English sentence.

Model architecture

- GPT-2 had 48 layers and used 1600 dimensional vectors for word embedding.
- Larger vocabulary of 50,257 tokens was used.
- Larger batch size of 512 and larger context window of 1024 tokens were used.

```
class GPT2Config(PreTrainedConfig):
    """
    Configuration for gpt2-base
    """

    pretrained_config_archive_map = CONFIG_ARCHIVE_MAP

    def __init__(
        self,
        vocab_size=50257,
        n_positions=1024,
        n_embd=768,
        n_layer=12,
        n_head=12,
        n_inner=None,
        activation_function="gelu_new",
        resid_pdrop=0.1,
        embd_pdrop=0.1,
        attn_pdrop=0.1,
        layer_norm_epsilon=1e-5,
        initializer_range=0.02,
        summary_type="cls_index",
        summary_use_proj=True,
        summary_activation=None,
        summary_proj_to_labels=True,
        summary_first_dropout=0.1,
        scale_attn_weights=True,
        use_cache=True,
        bos_token_id=50256,
        eos_token_id=50256,
        scale_attn_by_inverse_layer_idx=False,
        reorder_and_upcast_attn=False,
        **kwargs,
    ):
```

Model architecture

```
def construct(
    self,
    hidden_states: Tuple[Tensor],
    layer_past: Optional[Tuple[Tensor]] = None,
    attention_mask: Optional[Tensor] = None,
    head_mask: Optional[Tensor] = None,
    encoder_hidden_states: Optional[Tensor] = None,
    encoder_attention_mask: Optional[Tensor] = None,
    use_cache: Optional[bool] = False,
):
    residual = hidden_states
    hidden_states = self.ln_1(hidden_states)
    attn_outputs = self.attn(
        hidden_states,
        layer_past=layer_past,
        attention_mask=attention_mask,
        head_mask=head_mask,
        use_cache=use_cache,
    )
    attn_output = attn_outputs[0] # output_attn: a, present, (attentions)
    outputs = attn_outputs[1:]
    # residual connection
    hidden_states = attn_output + residual
```

- Layer normalization was moved to input of each sub-block

Model architecture

```
for i, block in enumerate(self.h):
    if self.output_hidden_states:
        all_hidden_states = all_hidden_states + (hidden_states,)

    outputs = block(
        hidden_states,
        layer_past=past_key_values[i],
        attention_mask=attention_mask,
        head_mask=head_mask[i],
        encoder_hidden_states=encoder_hidden_states,
        encoder_attention_mask=encoder_attention_mask,
        use_cache=self.use_cache,
    )

    hidden_states = outputs[0]
    if self.use_cache:
        presents = presents + (outputs[1],)

    if self.output_attentions:
        all_self_attentions = all_self_attentions + (outputs[2 if self.use_cache else 1],)
        if self.add_cross_attention:
            all_cross_attentions = all_cross_attentions + (outputs[3 if self.use_cache else 2],)

    hidden_states = self.ln_f(hidden_states)
```

- an additional layer normalization was added after final self-attention block.

Special Scaled Initialization

```
# Reinitialize selected weights subject to the OpenAI GPT-2 Paper Scheme:
#   > A modified initialization which accounts for the accumulation on the residual path with model depth. Scale
#   > the weights of residual layers at initialization by a factor of  $1/\sqrt{N}$  where N is the # of residual layers.
#   > -- GPT-2 :: https://openai.com/blog/better-language-models/
#
# Reference (Megatron-LM): https://github.com/NVIDIA/Megatron-LM/blob/main/megatron/model/gpt\_model.py
for name, p in module.named_parameters():
    if name == "c_proj.weight":
        # Special Scaled Initialization --> There are 2 Layer Norms per Transformer Block
        p.data.normal_(mean=0.0, std=(self.config.initializer_range / math.sqrt(2 * self.config.n_layer)))
```

[在线竞赛](#) > MindSpore开源项目贡献赛

MindSpore开源项目贡献赛

奖金	浏览	报名战队
暂无	213	7

竞赛时间：2023-04-22~2023-10-22

报名截止：2023-09-30 23:30:00

竞赛状态：进行中

[立即报名](#)[通知公告](#)[赛事介绍](#)[赛题说明](#)

大赛公告

第六届“中国软件开源创新大赛”在国家自然科学基金委信息科学部的指导下，由中国计算机学会（CCF）主办，西北工业大学、绿色计算产业联盟、CCF 开源发展委员会联合承办。大赛面向国家“十四五”开源生态发展战略布局，聚焦“卡脖子”软件领域以及人工智能、大数据、芯片设计、物联网等前沿技术领域的开源软件，旨在为国内开源社区提供展示、交流、合作的平台，激发开源创新活力，培养开源实践人才，助力开源生态建设的高质量发展。

[在线竞赛](#) > MindSpore代码评注赛

MindSpore代码评注赛

奖金	浏览	报名战队
暂无	139	8

竞赛时间: 2023-04-22~2023-10-22

报名截止: 2023-09-30 23:30:00

竞赛状态: 进行中

[立即报名](#)[通知公告](#)[赛事介绍](#)[赛题说明](#)

大赛公告

第六届“中国软件开源创新大赛”在国家自然科学基金委信息科学部的指导下，由中国计算机学会（CCF）主办，西北工业大学、绿色计算产业联盟、CCF 开源发展委员会联合承办。大赛面向国家“十四五”开源生态发展战略布局，聚焦“卡脖子”软件领域以及人工智能、大数据、芯片设计、物联网等前沿技术领域的开源软件，旨在为国内开源社区提供展示、交流、合作的平台，激发开源创新活力，培养开源实践人才，助力开源生态建设的高质量发展。

昇腾AI 创新大赛

数智未来 因你而来

指导单位：新一代人工智能产业技术创新战略联盟 (AITISA)、中国人工智能产业发展联盟 (AIIA)

主办单位：华为技术有限公司

协办单位：OpenI 启智社区

赞助单位：华鲲振宇、湘江鲲鹏

昇思赛道

[M]^s

01 赛道介绍



昇思赛道分为创新算法与数据打榜两类赛题，围绕高质量数据集，结合行业应用场景与前沿技术设置赛题，并提供免费的算力支撑。

02 赛程介绍

报名时间
2023.4.25 - 8.15

作品提交时间
2023.4.25 - 8.30

第一批任务发布
2023.5.12

第二批任务发布
2023.6月中旬

全国赛时间
2023年10月

全国赛颁奖时间
2023年10月

03 奖金设置

数据打榜赛题（每题）

奖项名称	奖项数量(个/题)	奖项金额(元)(税前金额)	其他权益
金奖 🏆	2	5万	Y
银奖 🥈	4	1万	Y
铜奖 🥉	6	5千	Y

算法创新赛题（每题）

奖项名称	奖项数量(个/题)	奖项金额(元)(税前金额)	其他权益
金奖 🏆	1	1.5万	Y