

# Consciousness Turing Machine Handbook

Haofei Yu

Apr 23, 2024

# CTM-AI Input and Output

## System Input

Text-based query, e.g. How is the weather today?

Text-based information, e.g. The weather is 3 degrees to 15 degrees.

Image-based information, e.g. an outdoor image

Video-based information, e.g. a sky video

## System Output

Text-based answer to the query, e.g. Sunny

# CTM-AI Key Concept

## Terminology about architecture

**CTM:** the main system taking multimodal input and generating final output

For each CTM, it includes 2 subcomponents:

**Processor:** take modality-specific input and generate text-based gist

**Supervisor:** take one gist and output the final answer to the query

For each processor, it includes three subcomponents:

**Executor:** responsible for generating the output, e.g. LLM / VLM

**Scorer:** responsible for generating the relatedness, confidence, and surprise of the gist conditioned on the query

**Messenger:** responsible for updating the information and formatting

# CTM-AI Concept

## Terminology about input and output

**Query:** The question that user wants to get answer from.

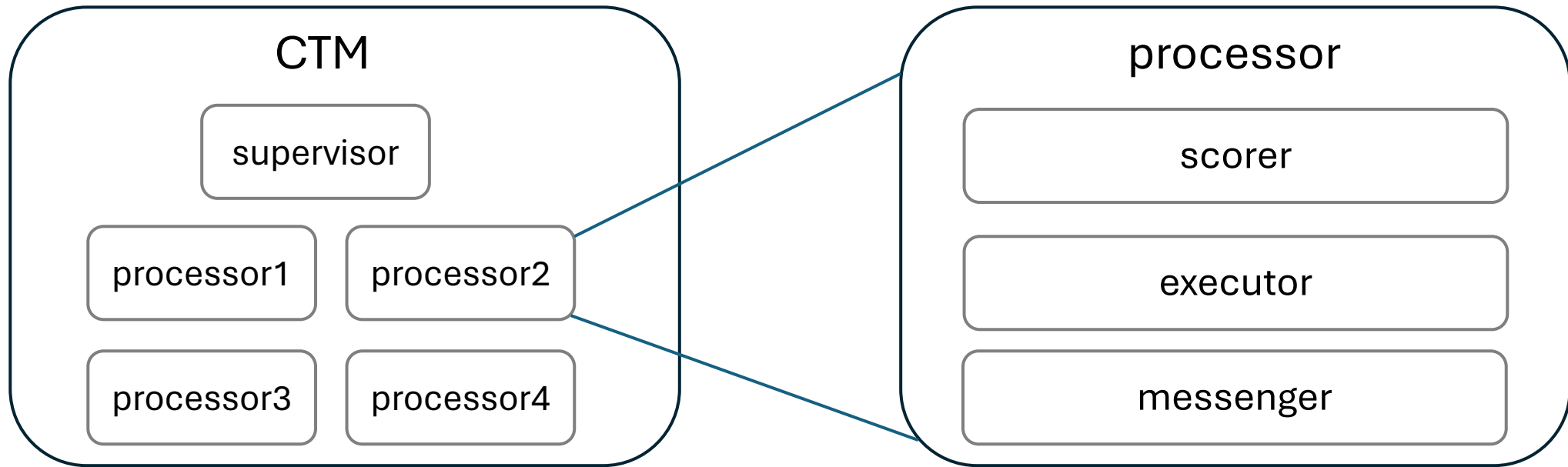
**Text Input:** Context information that the CTM system relies on to answer the query.

**Audio/Image/Video Input:** Multimodal information that the CTM system relies on to answer the query.

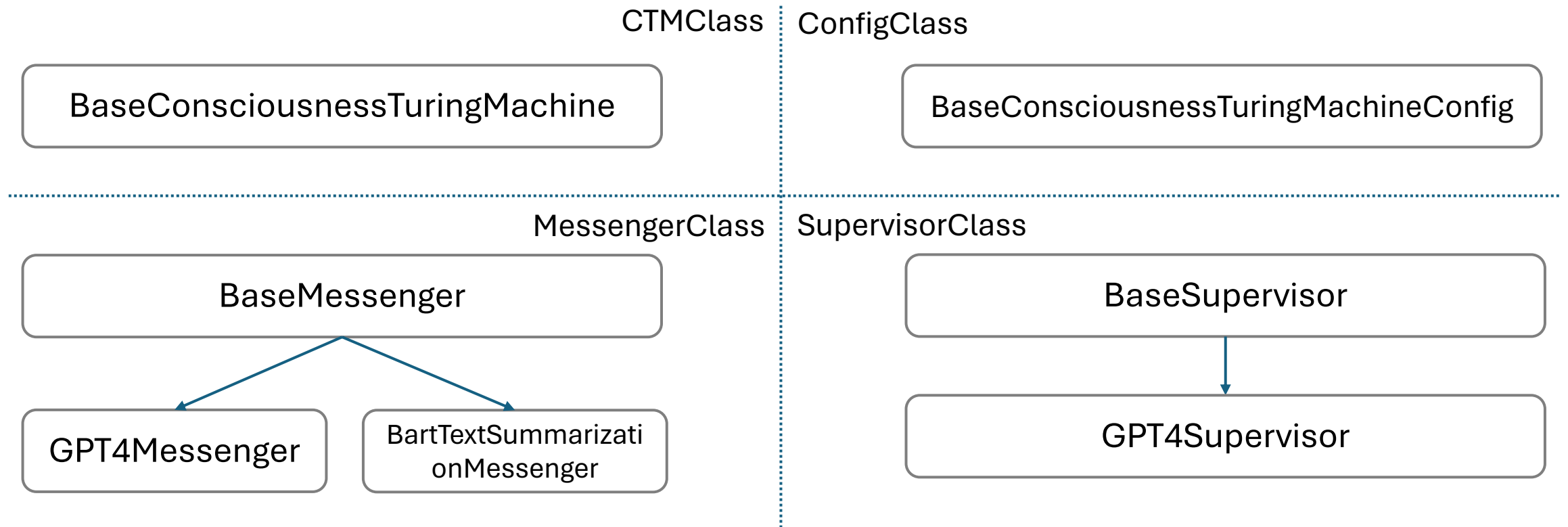
**Gist:** The output information from one processor. Normally just sentence-level key information about the input.

**System Output:** the supervisor's answer to the query

# CTM Architecture



# CTM Class Diagram



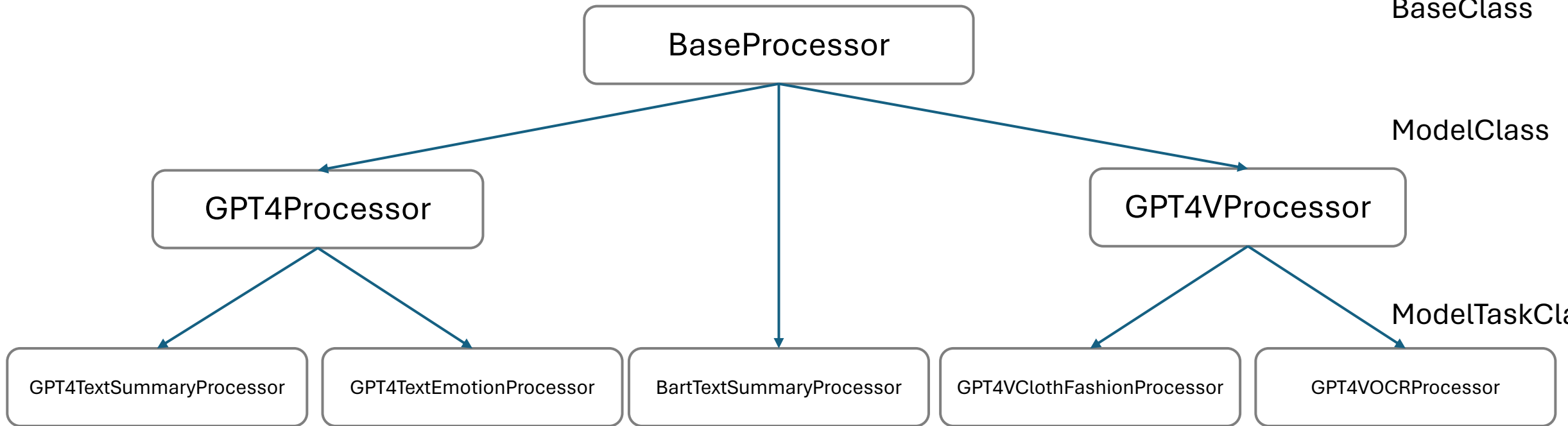
# CTM Class Diagram

ProcessorClass

BaseClass

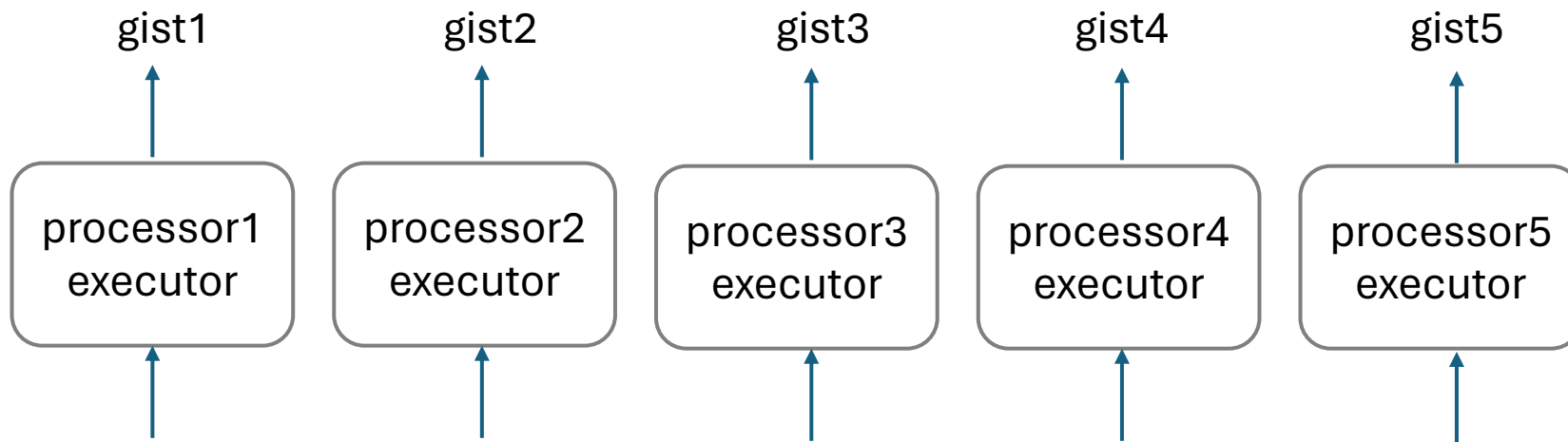
ModelClass

ModelTaskClass



# CTM Key Function

Forward Function1: ask processor in parallel



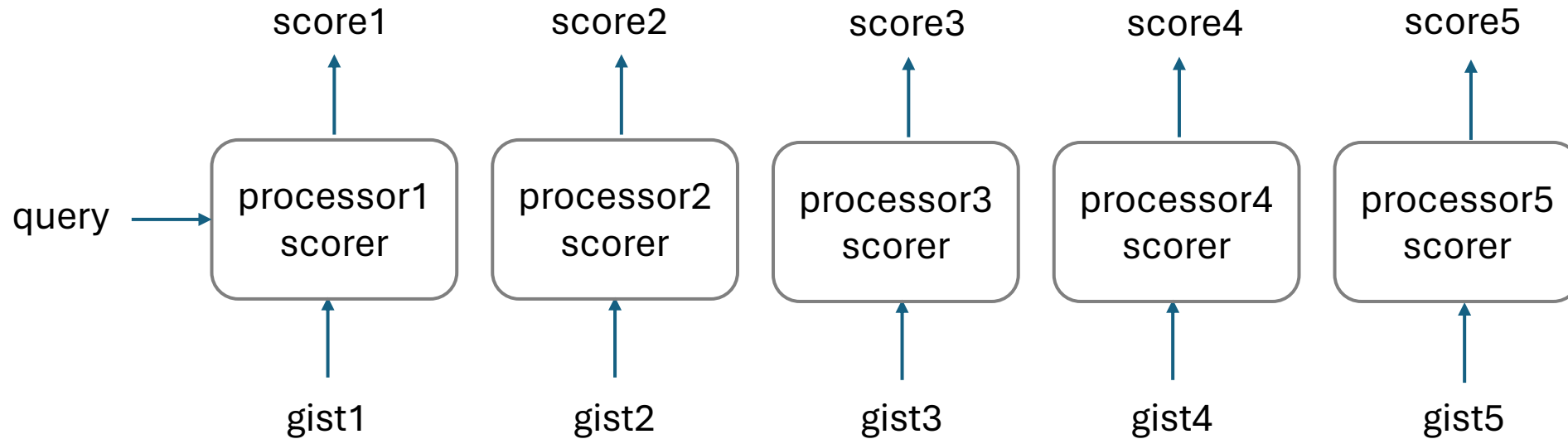
Multimodal Input



# CTM Key Function

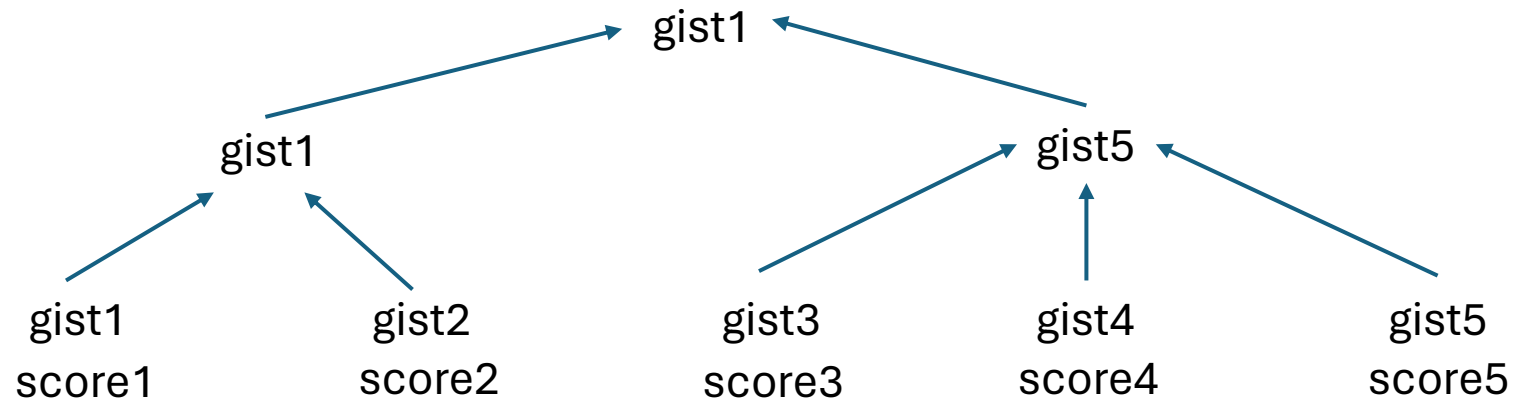
Forward Function2: ask score

score = (relatedness, confidence, surprise)



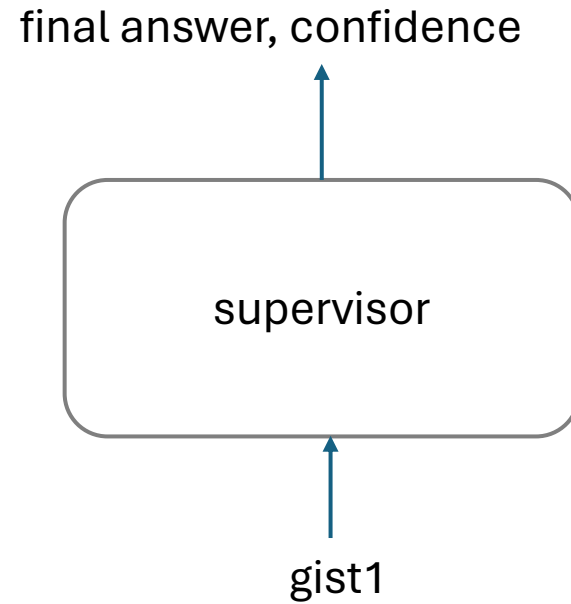
# CTM Key Function

Forward Function3: up-tree competition



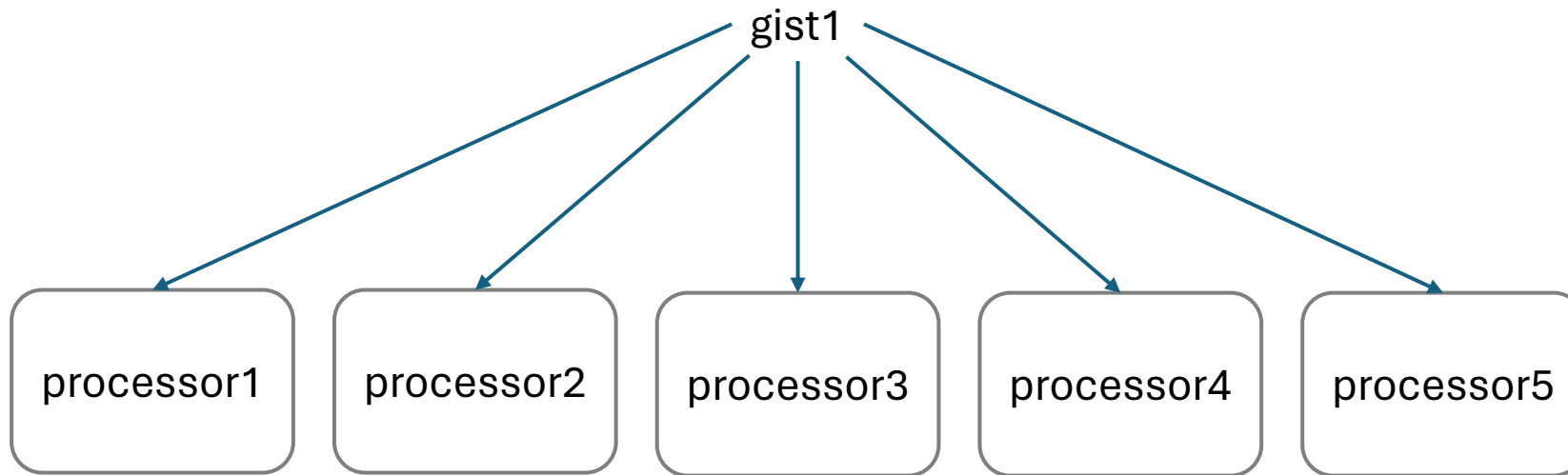
# CTM Key Function

Forward Function4: ask supervisor



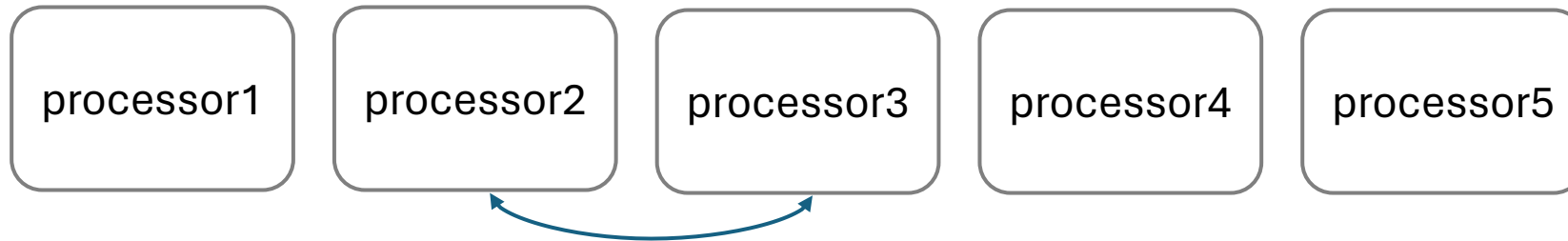
# CTM Key Function

Backward Function1: down-tree broadcasting



# CTM Key Function

Backward Function2: link form



# CTM Key Function

Backward Function3: processor fusion

