NARS-GPT

by Patrick Hammer and Peter Isaev

JUNE 19, 2023
INLP 2023 @ AGI-23

ACKNOWLEDGMENT

Hugo Latapie, *Principal Engineer at Cisco, San Jose, California*

Julius Pfrommer, *Head of Department "Cognitive Industrial Systems" at Fraunhofer IOSB, Germany*

With special thanks to Peter Voss @ aigo.ai for providing an incremental Q&A benchmark!

And with special thanks to Robert Johansson for experimental studies with NARS-GPT!

THE TABLE
OF CONTENT



MOTIVATION

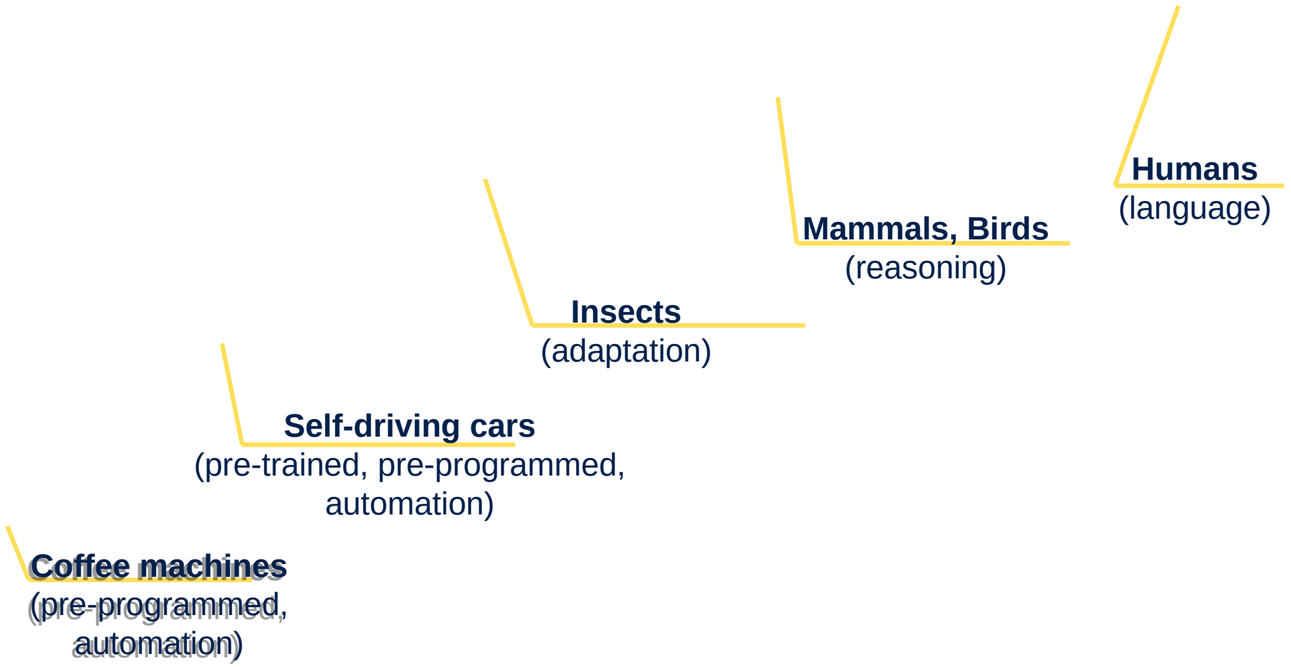
Observations:

- How could AI remember the information for a long time, reason and act upon it?
- Importance of Induction for incremental learning
- ChatGPT: limited context window
- Vector databases: can be used to overcome long-term memory problem
- Can GPT technology be combined with NARS reasoning?

Desired Features:

- Support long-running open-ended Q&A sessions
- Lift the long-term memory limitation of Large Language Models (LLM)
- Allow evidence collection over long-term of operation
- Become a useful and interactive tool for adaptive autonomous agents

THE LADDER OF INTELLIGENCE



Intelligence

Assumption of Insufficient Knowledge (AIKR):

With sufficient knowledge and resources any problem is solvable. Covering all relevant cases and having sufficient time to process is seldomly possible in real world! Approaching AI's original goal and build data-efficient open-ended real-time learners which can reason like us is challenging…

Definition of Intelligence:

According to NARS Team: intelligence is the ability to adapt under insufficient knowledge and resource (AIKR). Intelligent species in nature: adapt quickly, repeating mistakes is often fatal or survival-relevant!

Non-Axiomatic Reasoning System (NARS)

A general-purpose reasoning system proposed by Dr. Wang

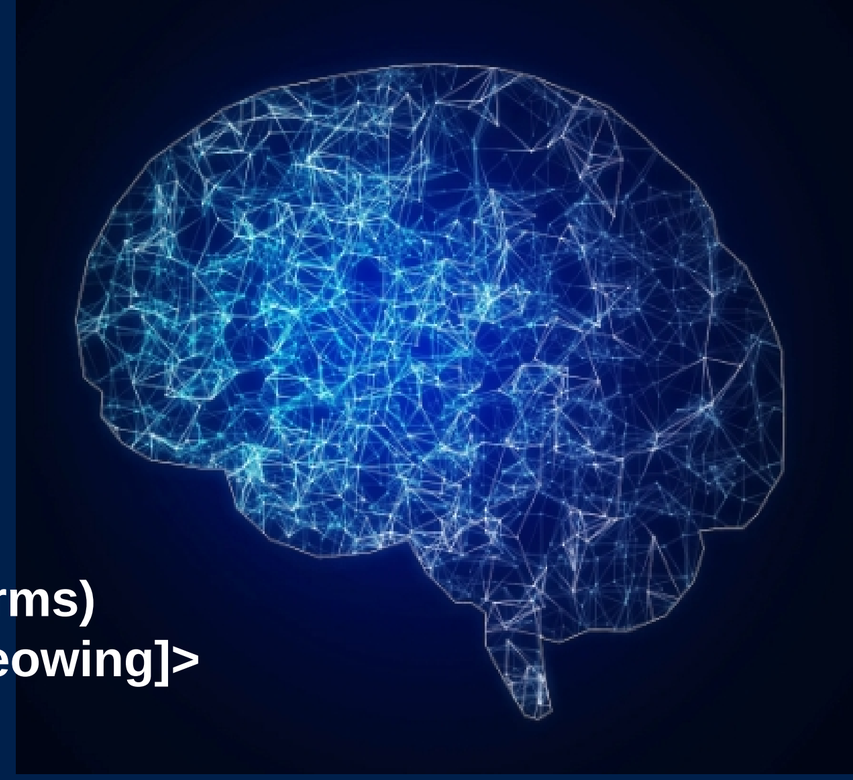
Operates under Assumption of Insufficient Knowledge and Resources:

- **Finite** processing demands
- **Open** to new information at runtime
- **Real-time** operation

Uses Non-Axiomatic Logic (NAL):

- a term logic able to deal with

Non-Axiomatic Logic (NAL)



Terms:

- atomic (such as ***cat, animal***) or compound

Statements in NAL: (compounds via logical copulas between terms)

- Inheritance **<cat → animal>, <{garfield} → animal>, <cat → [meowing]>**
- Implication **<lightning ⇒ thunder>**
- Nested **<<switch → [on]> ⇒ <light → [on]>>**

Truth Value in NAL:

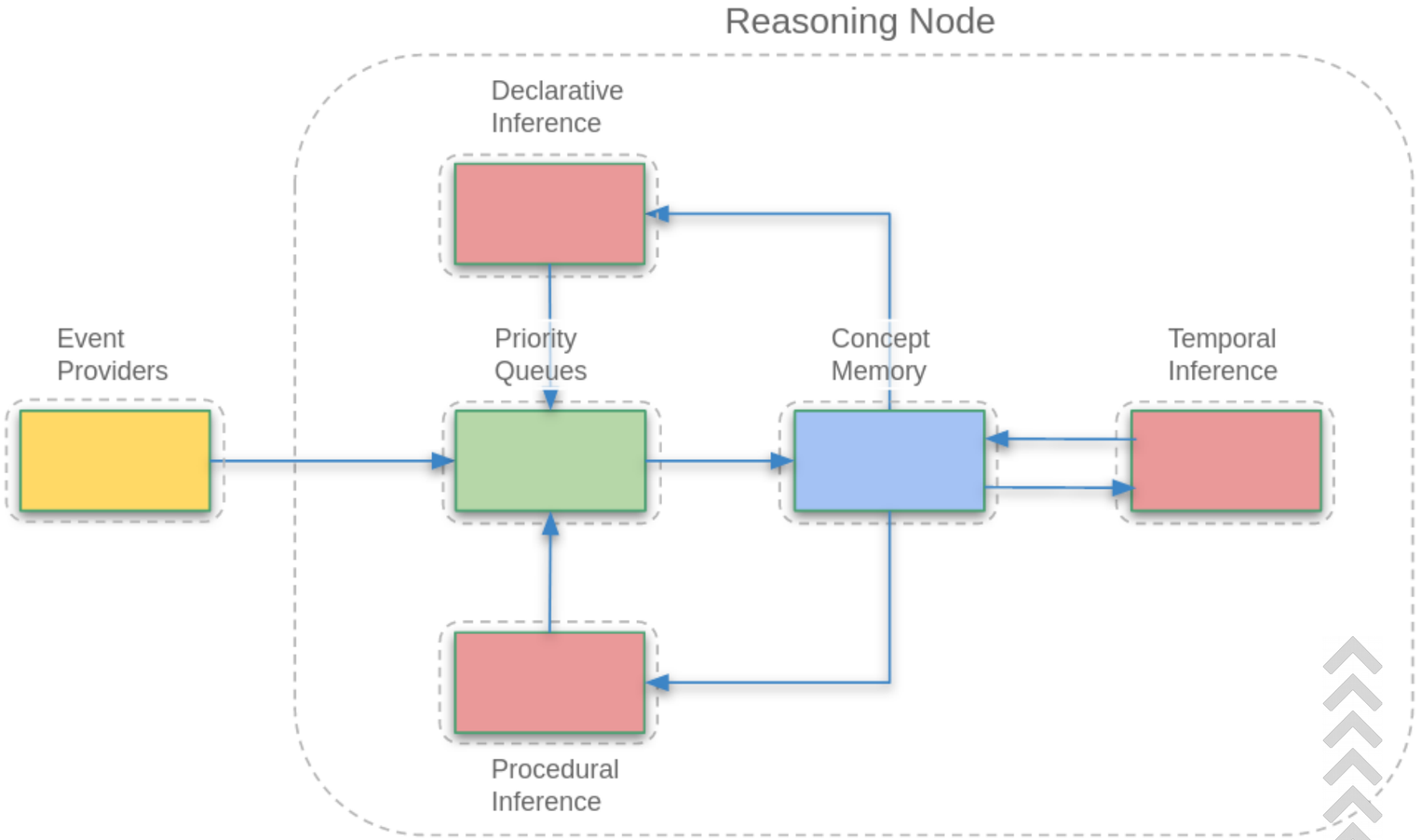
- Positive **w+** and negative evidence **w-**
- Frequency: **w+ / (w+ + w-)**
- Confidence: **w / (w + 1)** with **w := w+ + w-**
- Is the coin fair? 10-coin flips, 5 heads, less confidence than 100-coin flips, 50 heads, while frequency = 0.5 in both cases!

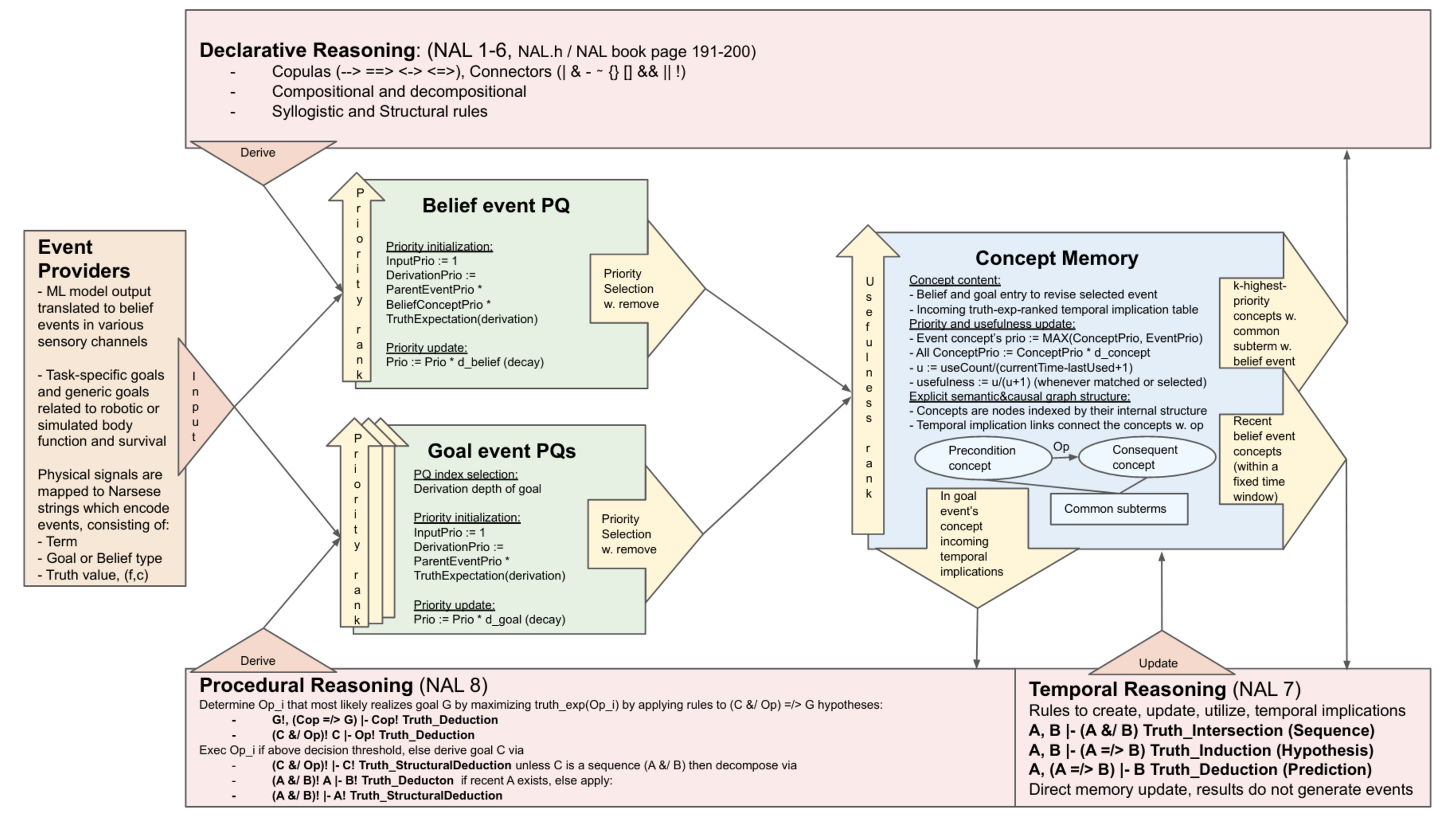
OpenNARS for Applications (ONA)

- A pragmatic design following NARS theory
- General-purpose reasoning system utilizing NAL
- Provides cumulative real-time learning
- Designed for reaching goals, establishing sub-goals and proceed with planning
- Capable of controlling intelligent agents including robots for enhanced autonomy
- Supports interfacing with state-of-the-art Deep Learning models

ONA architecture

- Learn from event streams in real time, without interruption
- Extract sensorimotor contingencies
- Use them to plan ahead to reach goals
- Enhance autonomy of intelligent agents

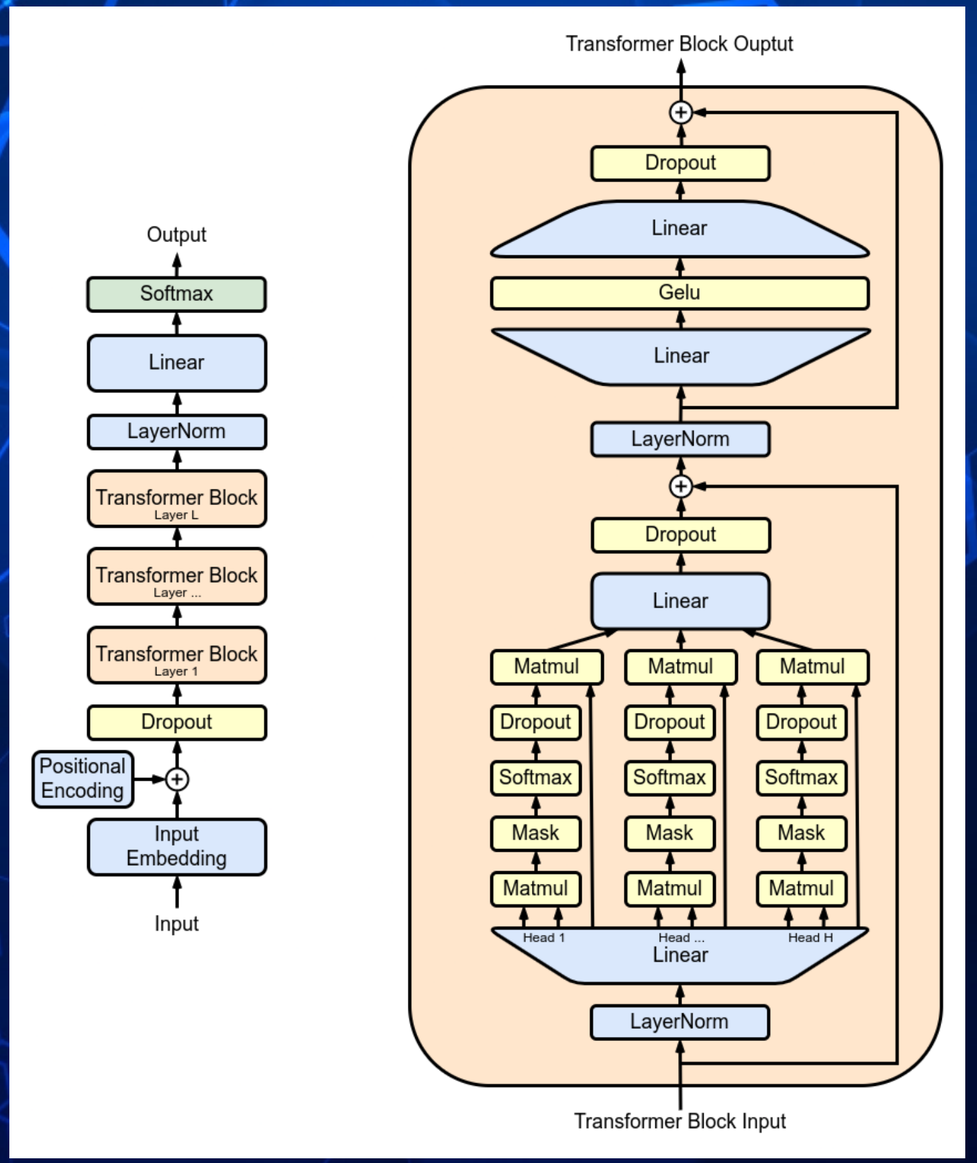



Declarative Reasoning: (NAL 1-6, NAL.h / NAL book page 191-200)
- Copulas (--> ==> <-> <=>), Connectors (| & - ~ {} [] && || !)
- Compositional and decompositional
- Syllogistic and Structural rules
Derive
Event Providers
- ML model output translated to belief events in various sensory channels
- Task-specific goals and generic goals related to robotic or simulated body function and survival
Physical signals are mapped to Narsese strings which encode events, consisting of:
- Term
- Goal or Belief type
- Truth value, (f,c)
Input
Priority rank
Belief event PQ
Priority initialization:
InputPrio := 1
DerivationPrio := ParentEventPrio * BeliefConceptPrio * TruthExpectation(derivation)
Priority update:
Prio := Prio * d_belief (decay)
Priority Selection w. remove
Priority rank
Goal event PQs
PQ index selection:
Derivation depth of goal
Priority initialization:
InputPrio := 1
DerivationPrio := ParentEventPrio * TruthExpectation(derivation)
Priority update:
Prio := Prio * d_goal (decay)
Priority Selection w. remove
Usefulness rank
Concept Memory
Concept content:
- Belief and goal entry to revise selected event
- Incoming truth-exp-ranked temporal implication table
Priority and usefulness update:
- Event concept's prio := MAX(ConceptPrio, EventPrio)
- All ConceptPrio := ConceptPrio * d_concept
- u := useCount/(currentTime-lastUsed+1)
- usefulness := u/(u+1) (whenever matched or selected)
Explicit semantic&causal graph structure:
- Concepts are nodes indexed by their internal structure
- Temporal implication links connect the concepts w. op
Precondition concept
Op
Consequent concept
Common subterms
In goal event's concept incoming temporal implications
k-highest-priority concepts w. common subterm w. belief event
Recent belief event concepts (within a fixed time window)
Derive
Update
Procedural Reasoning (NAL 8)
Determine Op_i that most likely realizes goal G by maximizing truth_exp(Op_i) by applying rules to (C &/ Op) =/> G hypotheses:
- G!, (Cop =/> G) |- Cop! Truth_Deduction
- (C &/ Op)! C |- Op! Truth_Deduction
Exec Op_i if above decision threshold, else derive goal C via
- (C &/ Op)! |- C! Truth_StructuralDeduction unless C is a sequence (A &/ B) then decompose via
- (A &/ B)! A |- B! Truth_Deducton if recent A exists, else apply:
- (A &/ B)! |- A! Truth_StructuralDeduction
Temporal Reasoning (NAL 7)
Rules to create, update, utilize, temporal implications
A, B |- (A &/ B) Truth_Intersection (Sequence)
A, B |- (A =/> B) Truth_Induction (Hypothesis)
A, (A =/> B) |- B Truth_Deduction (Prediction)
Direct memory update, results do not generate events

ONA CAPABILITIES

Conceptualization	➢ Conceptual representation ➢ Semantic, spatial & temporal relations
Memory	➢ Attentional Focus & context ➢ Long term concept memory
Adaption	➢ Unsupervised, one-shot learning ➢ Real-time procedure learning
Reasoning	➢ Handle uncertainty & contradiction ➢ Answer & explain questions and answers

Generative Pre-trained Transformer (GPT)



GPT models are artificial neural networks that are based on the transformer architecture, pre-trained on large data sets of unlabelled text, and able to generate novel human-like content.

Huge successes: GPT models have arguably made the idea of AGI more mainstream and provide a great preview of the capabilities that await us

Other successes Transformer models are also applicable to other domains, such as vision, and can handle multi-modal input.

EXPLORED APPROACHES

GPT (Direct use of GPT):

(-) Contextual window size limitation
(-) In-Context-Learning has no long-term effect
(-) Can answer questions but cannot store conclusions for later use

LTM-GPT (Embedding-based long-term sentence storage attached to GPT):

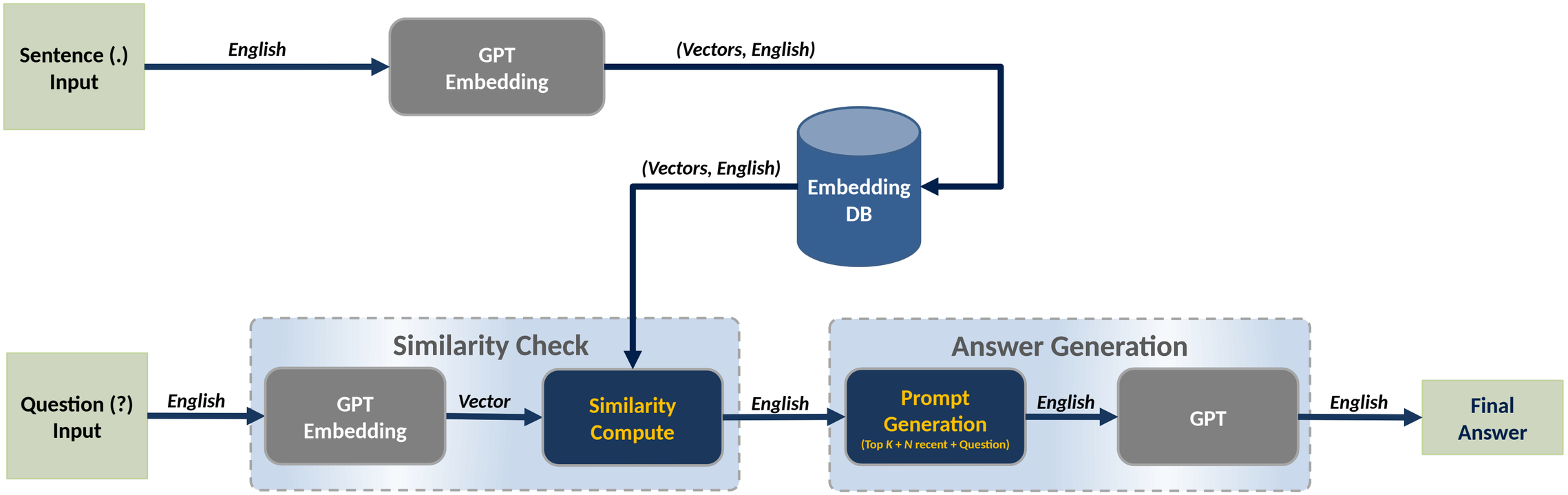
(+) Long-term recall as it directly stores the sentences and their embedding vector
(-) Can store sentences but cannot update beliefs
(-) Can answer questions but cannot store conclusions for later use

NARS-GPT (GPT attached to NARS):

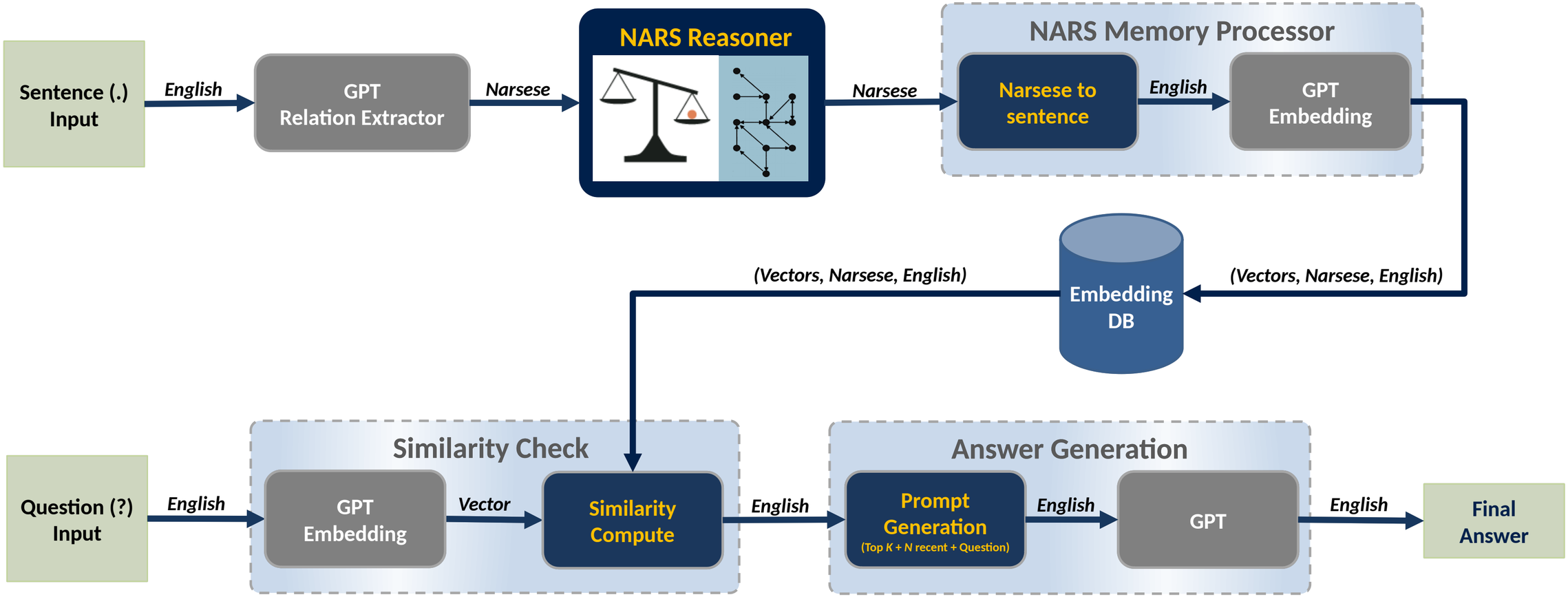
(+) Long-term belief updating and evidence accumulation
(+) Inference always correct according to the evidence
(+) Seamless integration with sensorimotor learning
(-) Relation extraction in complex sentences sometimes misses information

GPT FLOWCHART

LTM GPT FLOWCHART



NARS-GPT FLOWCHART



NARS-GPT PROMPTS

Belief Prompt:

```
"""
RelationClaim(noun,verb,noun) ... this relation is claimed to be true in the sentence
RelationClaim(noun,"IsA",noun) ... this relation is claimed to be true in the sentence
PropertyClaim(noun,"HasProperty", adjective) ... this relation is claimed to be true in the sentence
NegatedRelationClaim(noun,verb,noun) ... this relation is claimed to be false in the sentence with an explicit 'not' word
NegatedRelationClaim(noun,"IsA",noun) ... this relation is claimed to be false in the sentence with an explicit 'not' word
NegatedPropertyClaim(noun,"HasProperty",adjective) ... this relation is claimed to be false in the sentence with an explicit 'not' word

Capture the complete sentence meaning with code that calls the four functions.
Please make sure that the word "not" is not included in your call, just use the functions and Negated functions instead.
And use verbs for comparative relations!

Encode all relations in the sentence, and the sentence has to be believed!
"""
```

Question Prompt:

```
"""
Mention concrete memory contents with certainty values.
Use the minimum involved certainty value.

Memory:
.....
.....
according to Memory and which memory item i? Answer in a probabilistic sense and within 15 words based on memory content only.
"""
```

EXPERIMENTS

Tasks for measuring reasoning and understanding

- Facebook AI research bAbI dataset:
  - 20 reasoning tasks, 2000 questions each
  - Deductive and non-deductive reasoning
  - Spatial, temporal, declarative
  - Short-term Q&A

- Aigo.ai INT_inf benchmark:
  - Personal assistant input over 1000+ lines
  - Spatial, temporal, declarative
  - Long-term Q&A
  - 652 questions

bAbI tasks questions on QA16

899	654	246	0.73
TOTAL	CORRECT	INCORRECT	CORRECT RATIO

QA16: Inductive reasoning

NARS-GPT gets 73% questions right in comparison to 32% of **GPT-3.5**! (induction restriction)

bAbI TASKS

Induction limitation

Example input

Greg is a frog. Brian is a lion.
Bernhard is a swan. Brian is gray.
Greg is yellow. Lily is a frog.
Bernhard is yellow. Julius is a lion.
Lily is green.
What color is Julius?

ChatGPT

Based on the given information, we know that Julius is a lion. However, we don't have any information about his color, so we can't determine his color from the given statements.

NARS-GPT

Julius is likely gray according to memory item i=14
with a certainty of 0.164297.

Ground truth

gray

INT_inf questions

652
TOTAL

424
CORRECT

228
INCORRECT

0.65
CORRECT RATIO

QA16: Inductive reasoning

NARS-GPT gets 65% in comparison to 11% of **GPT-3.5** (contextual window restriction)

BELIEF CORRECTION

Input: what is kaeferbohnensalat?
No memory item. I believe kaeferbohnensalat is a German salad made with green beans and potatoes. Certainty: 80%.

Input: kaeferbohnensalat is an Austrian salad made with runner beans, onion, and pumpkin seed oil.
Input: what is kaeferbohnensalat?
Kaeferbohnensalat is a salad made with runner beans, according to memory item i=1 with a certainty of 0.69

Input: in which country is kaeferbohnensalad made?
Memory item i=1 suggests that kaeferbohnensalad is made in Austria with a certainty of 0.69

Input: what are all the ingredients of kaeferbohnensalat, list them all?
According to memory, kaeferbohnensalat contains runner bean and onion (i=1,3,4,5,6). Pumpkin seed oil may also be an ingredient (i=11) with a certainty of 0.788.

DEMONSTARTION

1. bAbI EXAMPLES

2. CHAT INTERRACTION

3. DEMO ONTHE PHONE

TAKEAWAYS

- GPT model is an effective way to extract relations and properties from arbitrary English input for translation to Narsese.
- NARS is a great way to provide adaptive long-term memory, uncertainty estimates and trustworthiness (via explicit memory item references) to GPT.
- ONA's inference control strategy is applicable to deal with incremental Q&A in natural language settings, it is not the bottleneck.
- Reasoning performance increases by using ONA
- GPT model can be used to answer arbitrary natural language questions about ONA memory contents with the help of statement embeddings.
- Long-term incremental Q&A sessions are now possible.

QUESTIONS AND ANSWERS

GET IN TOUCH

To arrange demonstration, discuss applications, partnership opportunity, or obtain more information contact:

EMAIL ADDRESS

patrick.hammer@psychology.su.se

peter.isaev@temple.edu

WEBSITE

https://tinyurl.com/StockholmAGILab

THANK YOU!

WITH GREAT RESPECT FOR YOUR BUSINESS