

# Object-oriented Neural Programming (OONP) for Document Understanding



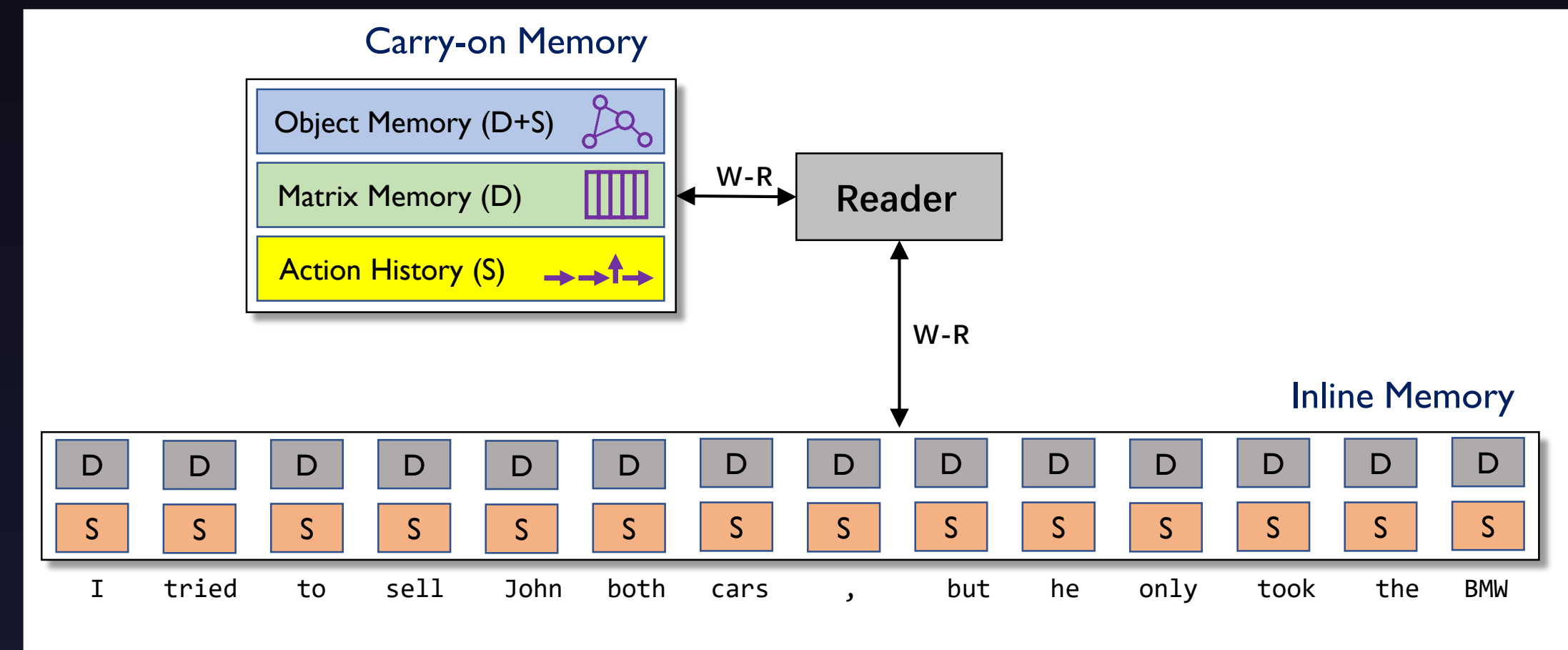
Zhengdong Lu<sup>1</sup>, Xianggen Liu<sup>1,2</sup>, Haofan Cui<sup>1,2</sup>, Yukun Yan<sup>1,2</sup>, Daqi Zheng<sup>1</sup>  
<sup>1</sup>DeeplyCurious.ai, <sup>2</sup>Tsinghua University



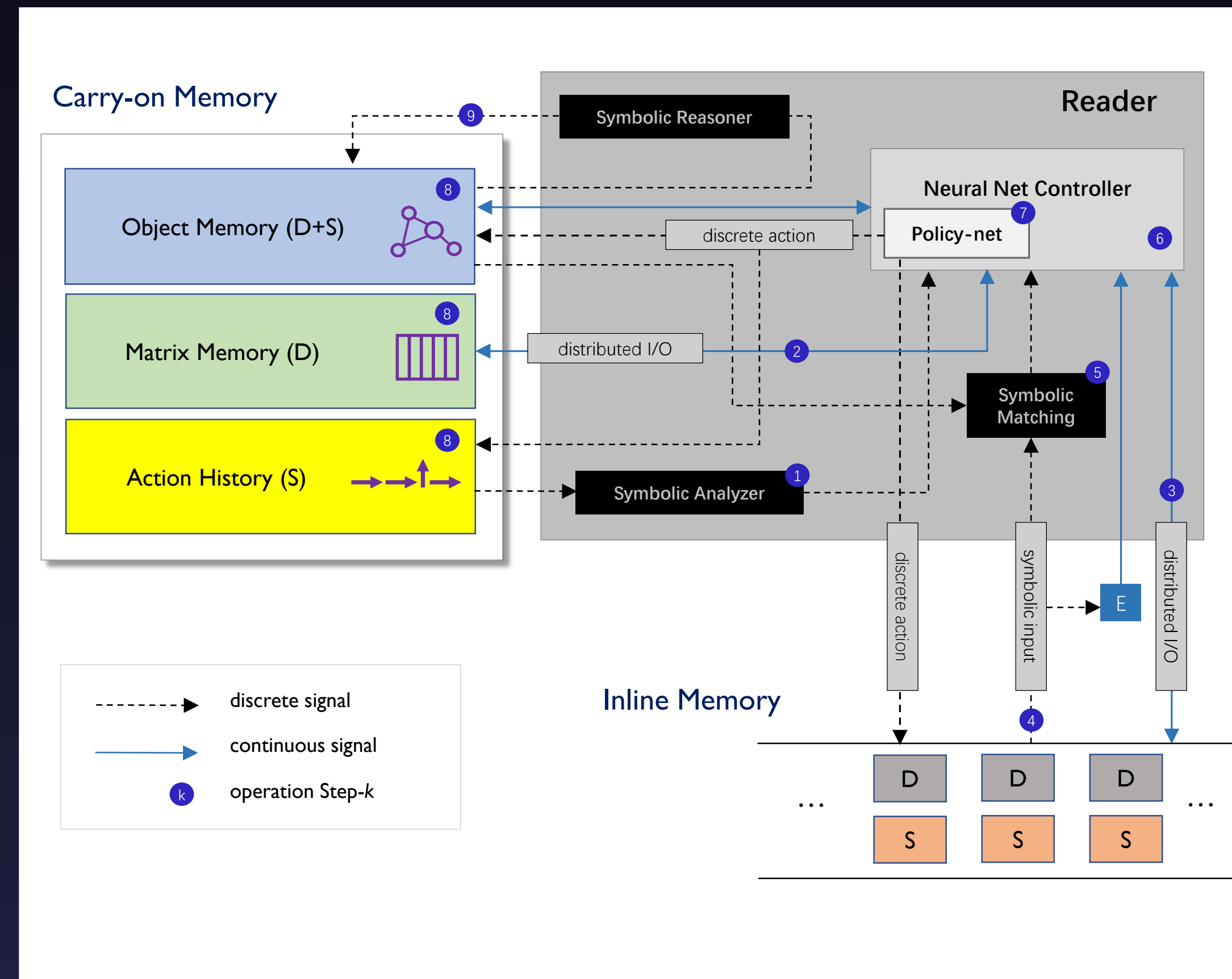
## OONP: Overview

An OONP parser consists of a Reader equipped with read/write heads, Inline Memory that represents the document, and Carry-on Memory that summarizes the current understanding of the document at each time step. For each document to parse, OONP first preprocesses it and puts it into the Inline Memory, and then Reader controls the read-heads to sequentially go through the Inline Memory and at the same time update the Carry-on Memory.

## #1 Global View

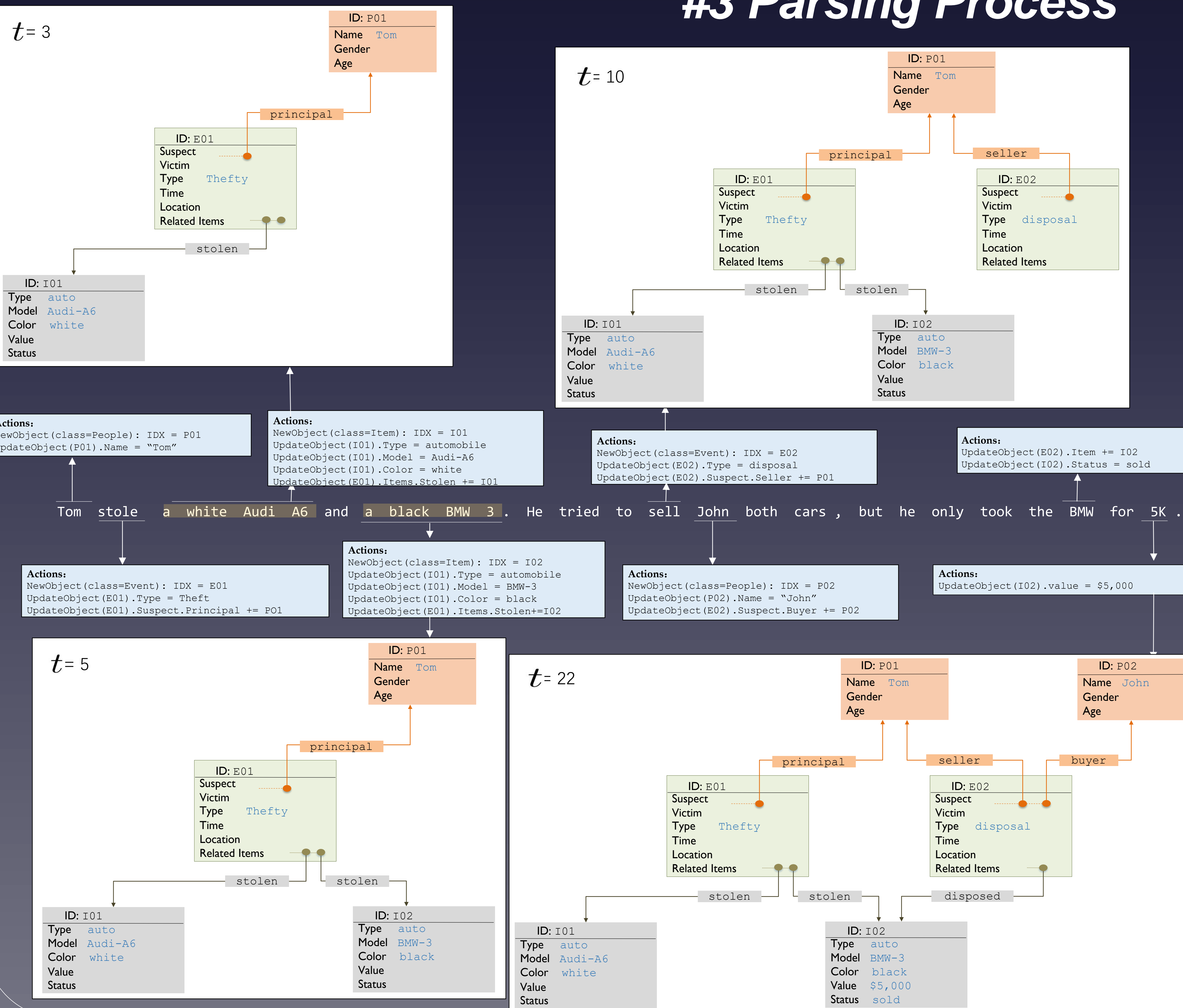


## #2 Zoom in Controller

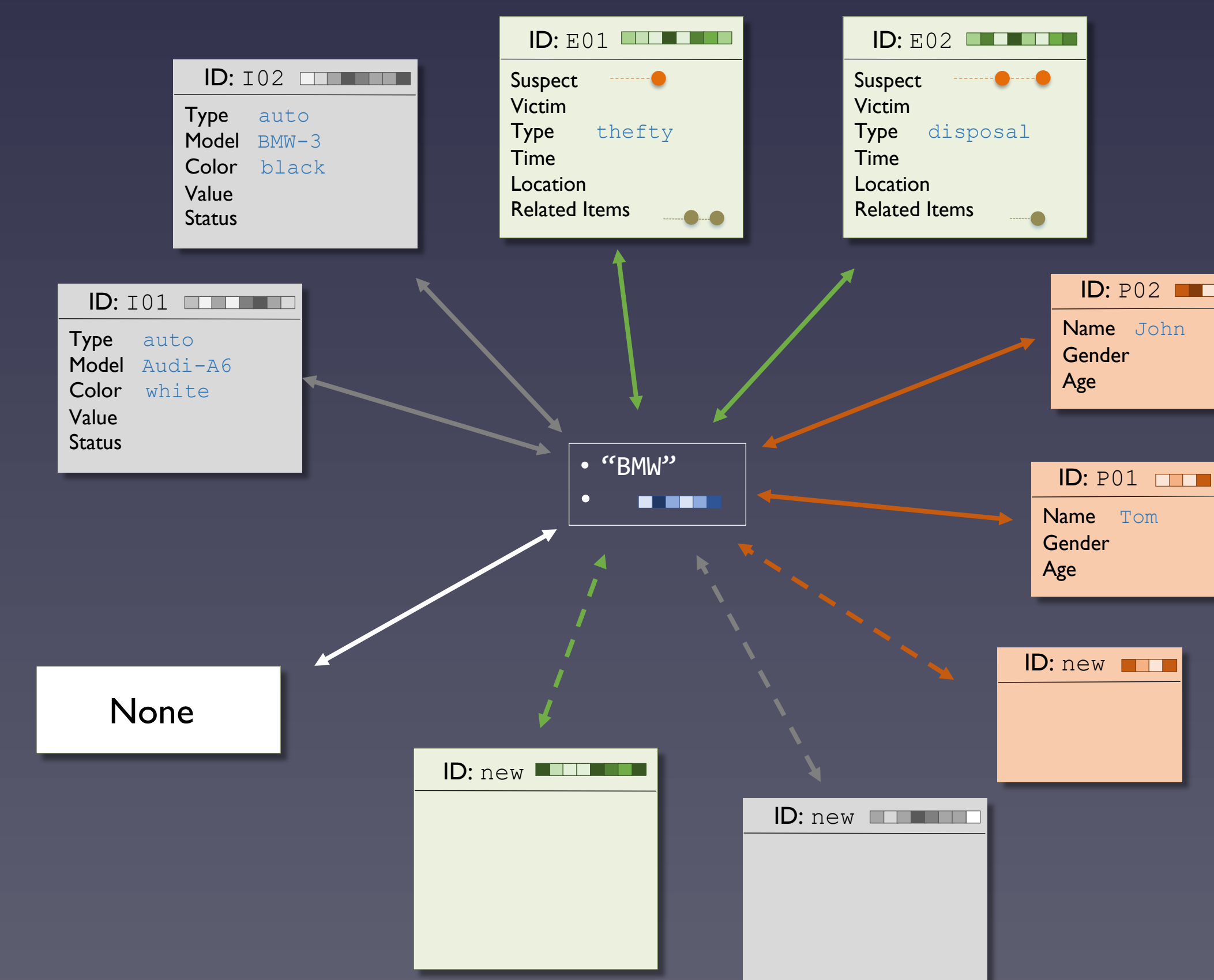


Reader is the control center of OONP, coordinating and managing all the operations of OONP. More specifically, it takes the input of different forms (reading), processes it (thinking), and updates the memory (writing). As shown in Figure 4, Reader contains Neural Net Controller (NNC) and multiple symbolic processors, and NNC also has Policy-net as its sub-component. Similar to the controller in Neural Turing, NNC is equipped with multiple read-heads and write-heads for differentiable read/write over Matrix Memory and (the distributed part of) Inline Memory, with a variety of addressing strategies. Policy-net however issues discrete outputs (i.e., actions), which gradually builds and updates the Object Memory in time. The symbolic processors are designed to handle information in symbolic form from Object Memory, Inline Memory, Action History, and Policy-net, while information from Inline Memory and Action History is eventually generated by Policy-net.

## #3 Parsing Process



## #4 New or Assign



With any information at hand (denoted as  $S_t$ ) at time  $t$ , the choices of New-Assign include the following three categories of actions:

- 1) creating (New) an object of a certain type,
- 2) assigning  $S_t$  to an existed object,
- 3) doing nothing for Stand moving on.

## Key Properties

**OONP in a nutshell** The key properties of OONP can be summarized as follows

- 1 OONP models parsing as a decision process: as the “reading and comprehension” agent goes through the text, it gradually forms the ontology as the representation of the text through its action;
- 2 OONP uses a symbolic memory with graph structure as part of the state of the parsing process. This memory will be created and updated through the sequential actions of the decision process, and will be used as the semantic representation of the text at the end;
- 3 OONP can blend supervised learning (SL) and reinforcement learning (RL) in tuning its parameters to suit the supervision signal in different forms and strengths.