AFRL-RH-WP-TR-2024-0025

SENSEMAKING IN NOVEL ENVIRONMENTS: HOW HUMAN COGNITION CAN INFORM ARTIFICIAL AGENTS

Robert E. Patterson, Ph.D. / Regina Buccello-Stout / Justin Nelson
711th HPW/RHWI
2255 H. Street
Wright-Patterson AFB, OH 45433

Mary E. Frame / Anna M. Maresca /
Barbara Acker-Mills / Erica Curtis
Parallax Advanced Research
4035 Colonel Glenn Hwy
Beavercreek, OH 45431

Jared Culburtson / Kevin Schmidt
Autonomy Capability Team
2255 H. Street
Wright-Patterson AFB, OH 45433

MARCH 2024

FINAL REPORT

DISTRIBUTION A. Approved for public release; distribution unlimited.

AIR FORCE RESEARCH LABORATORY
711 HUMAN PERFORMANCE WING
HUMAN EFFECTIVENESS DIRECTORATE
WRIGHT-PATTERSON AFB, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2024-0025 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

PATTERSON.ROBERT.E.1282458973 Digitally signed by PATTERSON.ROBERT.E.1282458973 Date: 2024.05.30 09:47:06 -04'00'

ROBERT E. PATTERSON, DR-III, Ph.D.
Work Unit Manager
Mission Analytics Branch
Human Effectiveness Directorate
711 Human Performance Wing
Air Force Research Laboratory

MURDOCK.WILLIAM.P.1048742161 Digitally signed by MURDOCK.WILLIAM.P.1048742161 Date: 2024.05.30 10:16:18 -04'00'

WILLIAM P. MURDOCK, DR-IV, Ph.D.
Chief, Mission Analytics Branch
Human Effectiveness Directorate
711 Human Performance Wing
Air Force Research Laboratory

CARTER.LOUISE.ANN.1230249128 Digitally signed by CARTER.LOUISE.ANN.1230249128 Date: 2024.06.17 10:11:15 -04'00'

LOUISE A. CARTER, DR-IV, Ph.D.
Chief, Warfighter Interactions and Readiness Division
Human Effectiveness Directorate
711 Human Performance Wing
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED – START DATE	END DATE
March 2024	Final	01 September 2016	03 March 2024

3. TITLE AND SUBTITLE
Sensemaking In Novel Environments: How Human Cognition Can Inform Artificial Agents

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
In House		
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
		H0SJ

6. AUTHOR(S)
[1]Robert Patterson / [1]Regina Buccello-Stout / [1]Justin Nelson / [2]Mary E. Frame / [2]Barbara Acker-Mills / [2]Erica Curtis / [3]Jared Culbertson / [3]Kevin Schmidt

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
[1]711 HPW/RHWI, 2255 H. Street, Wright-Patterson AFB, OH 45433
[2]Parallax Advanced Research, 4035 Col Glenn Hwy, Beavercreek, OH 45431
[3] Autonomy Capability Team 3, 2255 H. Street, Wright-Patterson AFB, OH 45433

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
711 Human Performance Wing
Human Effectiveness Directorate
Warfighter Interactions and Readiness Division
Wright-Patterson Air Force Base, OH 45433

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'SREPORT NUMBER(S)
AFRL-RH-WP-TR-2024-0025

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A. Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
AFRL-2024-2815; Cleared 23 May 2024 (Previously cleared as a journal article: AFRL-2023-3021; Cleared 21 Jun 2023)

14. ABSTRACT
One of the most vital cognitive skills to possess is the ability to make sense of objects, events, and situations in the world. In the current paper, we offer an approach for creating artificially-intelligent agents with the capacity for sensemaking in novel environments. First, sensemaking is represented as sign relations embedded within and across frames (i.e., schemata). Such sign relations are represented as probabilities in a Bayesian Network to reflect uncertainty. Moreover, synthesized patterns can arise from an interplay among different sources of knowledge—a feature of distributed representations—thus, synthesized patterns could be recognized as signs during the sensemaking process in a novel environment. Finally, those aspects of memories that get synthesized can be determined via an unconscious, embodied simulation process that aligns different combinations of attributes from different memories to find a solution to the sensemaking. In sum, we offer a novel approach by suggesting that attributes across memories can be shared and recombined in novel ways to create synthesized signs, which can denote certain outcomes in novel environments (i.e., sensemaking).
.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
U	U	U	SAR	26

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
Dr. Robert A. Patterson	

PREVIOUS EDITION IS OBSOLETE.

STANDARD FORM 298 (REV. 5/2020)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

One of the most vital cognitive skills to possess is the ability to make sense of objects, events, and situations in the world. In the current paper, we offer an approach for creating artificially-intelligent agents with the capacity for sensemaking in novel environments. First, sensemaking is represented as sign relations embedded within and across frames (i.e., schemata). Such sign relations are represented as probabilities in a Bayesian Network to reflect uncertainty. Moreover, synthesized patterns can arise from an interplay among different sources of knowledge—a feature of distributed representations—thus, synthesized patterns could be recognized as signs during the sensemaking process in a novel environment. Finally, those aspects of memories that get synthesized can be determined via an unconscious, embodied simulation process that aligns different combinations of attributes from different memories to find a solution to the sensemaking. In sum, we offer a novel approach by suggesting that attributes across memories can be shared and recombined in novel ways to create synthesized signs, which can denote certain outcomes in novel environments (i.e., sensemaking).

1.0 INTRODUCTION

One of the most important cognitive skills to possess is the ability to make sense of objects, events, and situations in the world. The capability for sensemaking is critical when comprehending everyday situations, such as when individuals engage in social interaction with coworkers, make critical decisions at work, navigate through traffic during rush hour, or avoid injury in dangerous situations. It is likely that all other cognitive processes, such as perception, attention, and memory, act in the service of sensemaking ([1], [2]). Sensemaking is important for biology as a whole ([3]).

Despite sensemaking's importance for biology, and general agreement that it entails the cognitive process of utilizing past experiences to interpret new situations**,** the concept has not been consistently defined in the literature. For instance, Russell, Stefik, Pirolli and Card ([4]) argued that sensemaking is the process of encoding data in a representation to answer task-specific questions. In doing so, representations are chosen and changed to reduce the cost of information processing. On the other hand, Dervin ([5]) suggested that sensemaking is a label for the process of how people construct sense of their worlds, and how they use information during that process. In this case, sensemaking is defined as both internal (i.e. cognitive) and external (i.e. procedural) behavior. Weick, Sutcliffe and Obstfeld ([6]) posited that sensemaking is a process in which circumstances and situations are understood in explicit language that enables rationalizing what one is doing. Thus, sensemaking helps inform and constrain identity and action. Finally, Lebiere, Pirolli, Thomson, Paik, Rutledge-Taylor, Staszewski and Anderson ([7]) defined sensemaking as the mental process of constructing a 'meaningful' representation of some complex aspect of the world. These authors also assumed that sensemaking involved information-foraging and hypothesis-updating processes.

Many of these approaches, discussed above, may contain useful elements for defining and conceptualizing sensemaking in certain situations, such as the potential need for information-foraging or hypothesis updating. Yet many of these approaches contains elements that are too narrow (e.g., answering task-specific questions) or too vague (e.g., constructing meaningful representations) to be useful for our purposes. We wish to employ an approach toward sensemaking that is precise enough to enable predictions yet flexible enough to apply to many different types of context and situation. Therefore, in the present case, we present a precise, yet flexible, unified framework for sensemaking ([1],[8]) and consider novel environments in which sensemaking can be particularly challenging. In novel environments, context is confusing—which elements in a scene should be attended to and which elements can be ignored? In novel environments, cause and effect may be unknown—which variables are related and which variables are independent?

Given the importance of sensemaking in novel environments, the present paper offers the following innovative ideas: (a) a unified framework for sensemaking in which sensemaking is represented as sign relations embedded within and across frames (i.e., schemata), the former of which are represented as probabilities in a Bayesian Network to reflect uncertainty; (b) the interplay among different sources of knowledge—a feature of distributed representations—can create synthesized patterns, each of which could be recognized as a sign during the sensemaking process in a novel environment; (c) distributed representations; and (d) embodied simulations that attempt to align different combinations of attributes from different memories to find a solution to the sensemaking. We also offer an actual example of sensemaking and end with discussion. We begin with a unified framework for sensemaking.

2.0 UNIFIED FRAMEWORK FOR SENSEMAKING

The capability to make sense of our environment would seem to require sophisticated cognitive processing. As Fuster [2] notes, in humans, all cognitive functions are interdependent: language depends on perception, attention, and memory, and intelligence depends on perception, attention, memory, language, and reasoning. The neural foundation of these cognitive functions is amazingly complex. The human brain typically contains slightly less than 100 billion neurons and has approximately 150 trillion synapses [9]. This complexity entails systems containing thousands of non-linear feedback loops, both positive (self- reinforcing) and negative (self-correcting) feedback, coupled to one another with multiple time delays, non-linearities, and accumulations, which can generate very complex endogenous behavior in a system ([10]; [11]). Such complex behavior implies that it will be very challenging to model human sensemaking.

2.1 Sensemaking

2.1.1 Sign Relations

Sensemaking can be conceptualized as sign interpretation, which is called 'meaning making' in the field of semiotics, a field originating from the writings of Peirce [12] and de Saussure ([13]; see also [14], [15], [16], [17]). Peirce [12] treated meaning making as sign interpretation—the meaning of a given thought occurs due to a triadic relation among the thought, the interpretation of the thought as a sign (its meaning), and a determining thought that the sign denotes. In simpler terms, the meaning or sense of an object or event is found in its *interpretation as a sign denoting some other (determining) object or event*, which can be causal or correlational. Bains [17] discussed how sensemaking entails relations that function as signs—sign relations. The interpretation of such relations as signs is based on knowledge about situated patterns, which includes the grounding of objects or events to the real world ([18], [19]). For example, making sense of a traffic jam during a morning commute—due to an accident—would be found in its interpretation as a sign denoting that the person will be late for work.

2.1.2 Frames

Sensemaking can also be conceptualized as involving 'frames,' as suggested by Minsky ([20]; see also [21], [22]). Minsky argued that when a new situation is encountered, a *mental data-structure or mental model for representing stereotypical situations, called a frame*, is retrieved from memory. The context of a given situation is represented by higher, fixed levels of a frame whereas the specifics of the situation are represented by arguments that fit within the frame (i.e., details of a frame are adapted to fit within current environmental context). According to Minsky [20], the relations among the lower-level terminals and their frame denote meaning. The concepts of scripts and plans, which are designed for representing sequences of events, are similar to frames [23]; [24]). Another concept closely allied to that of a frame is a schema (plural: schemata), which refers to a flexible mental data structure representing context-dependent relations among concepts in memory ([25]; [26]; [27]; [28], [29]; [30]).

2.1.3 Signs Embedded Within/Across Frames

Sign relations can be bound together in a frame. That is, within any given frame, a given element of the frame may serve as a sign denoting another element of the same frame (within-frame sign relation). For example, recall the commuter who is stuck in traffic during a morning commute due to an accident. Being stuck in traffic could activate a traffic-accident frame composed of elements like injuries, ambulance, debris, and stalled or stationary traffic. The presence of injuries would serve as a sign denoting the arrival of an ambulance, while the significant debris on the road would serve as a sign denoting stationary traffic. The ambulance itself could also serve as a sign denoting the possibility of stationary traffic. These are within-frame sign relations.

Moreover, an element of a given frame may serve as a sign denoting an element of a different frame (across-frame sign relation). The stationary traffic of this traffic-accident frame would serve as a sign denoting that the commuter will be late for work, which would activate a new late-for-work frame. Thus, the stationary traffic (of the traffic-accident frame) and the late for work event (of the late-for-work frame) would comprise an across-frame sign relation. The late-for-work frame could be composed of elements such as missed meeting, work late, unfinished projects, and boss unhappy. The missed meeting could serve as a sign (its sense) denoting the need to work late, while the unfinished projects could serve as a sign (its sense) denoting an unhappy boss. The working late itself could also serve as a sign denoting the possibility of an unhappy boss. These latter relationships would be within-frame sign relations. See Figure 1.

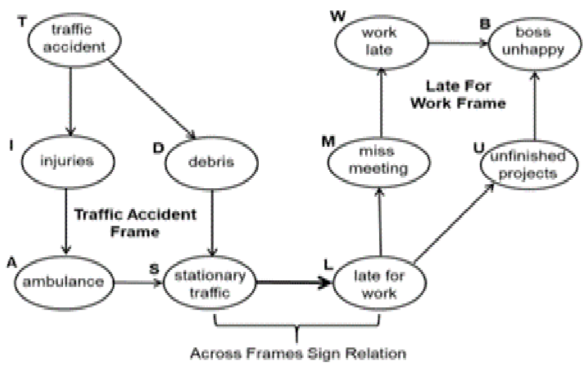

Figure 1. Diagram depicting Two Frames, Traffic-Accident frame and Late-for-Work Frame, and their Within-Frame Sign Relations and an Across-Frames Sign Relation.
The links shown in the traffic-accident frame T→I, D; I→A; D→S; A→S; and the links shown in the late-for-work frame L→M, U; M→W; U→B; W→B; represent within-frame sign relations. The link S→L represents an across-frame sign relation.

Sensemaking also affords humans the opportunity to look back and interpret the potential causes of events. For example, the traffic-accident frame could have been preceded by a speeding- cars frame, which would have served as a sign denoting the traffic-accident frame; and the late- for-work frame could be followed by a possible-demotion frame, which would be denoted by the late-for-work frame. The sense derived from these within- and between-frame sign relations would allow humans to generate predictions about potential future events, which could influence subsequent decision making ([1],[8]).

2.1.4 Bayesian Network (BN) Modeling

A BN can be used to represent our framework of sensemaking. BNs are a form of probabilistic graphical model that employ a directed acyclic graph for encoding multi-dimensional probability distributions [31]. Bayesian networks represent information about a given domain using variables to symbolize propositions and directed edges to convey dependencies [32]. Probabilities are employed to quantify variables. Representing sensemaking and its attendant sign relations as probabilities in a BN reflects uncertainty [33], a natural feature of sensemaking in any realistic domain due to a lack of absolute knowledge. Such sign relations reflect conditional dependency.

Returning to the traffic-accident frame and late-for-work frame (Figure 1), but now represented as a BN with associated hypothetical conditional probability tables (CPT) as in Figure 2. We simulated the BN model shown in Figure 2 with a software tool called *Hugin* (version 8.9; [34]).

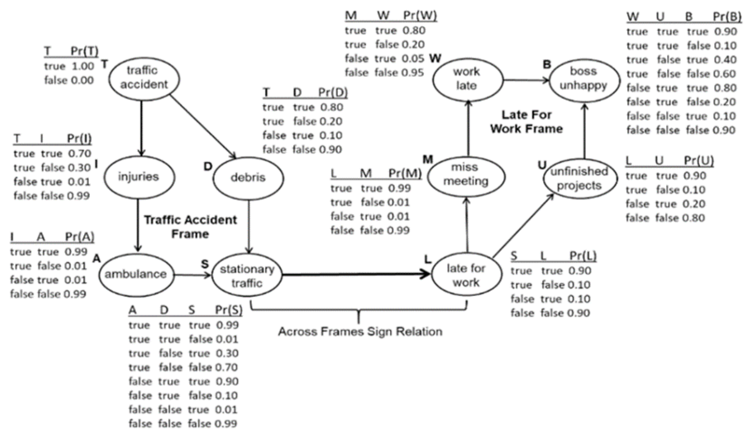

Figure 2. Diagram Depicting Two Frames, Traffic-Accident Frame and Late-for-Work Frame, and their Within-Frame Sign Relations and An Across-Frames Sign Relation, Previously Depicted in Figure 1.

In the present figure, each of these two frames are shown as a Bayesian network with its conditional probability table.

Prior probabilities of the model shown in Figure 2 are calculated as the following: T (traffic accident): 1.0; I (injuries): 0.7; A (ambulance): 0.7; D (debris): 0.8; S (stationary traffic): 0.81; L (late for work): 0.75; M (miss meeting): 0.75; W (work late): 0.61; U (unfinished projects): 0.73; and B (boss unhappy): 0.68. Thus, when the probability of debris is 80%, probability of being late for work is 75% and probability of boss being unhappy is 68%. The 80% probability of debris serves as a sign denoting that the commuter is more likely to be late than on time, the latter of which serves as a sign denoting that the boss is more likely to be unhappy than happy. When the probability of debris is 0%, the posterior probability of being late for work is 27% and the posterior probability of the boss being unhappy is 41%. Here, the absence of debris serves as a sign denoting that the commuter is more likely to be on time than late, the latter of which serves as a sign denoting that the boss is more likely to be happy than unhappy.

2.1.5 Sensemaking in Artificial Intelligence (AI) Systems

Concepts similar to sign relations and frames have been implemented in current systems in a variety of ways, often under different names. For example, Garcia et al. [35] used a reinforcement learning paradigm, along with encoded and decoder networks, to learn an emergent language between two agents given a categorization task. This two-component system may implement a version of probabilistic sign relations as described above. More broadly, representations such as large language models have been shown to have predictable implicit semantics ([36]). Nonetheless, the functional limitations of artificial cognition suggest that such 'sensemaking' falls well short of the successes seen with human sensemaking in naturalistic environments (see also [37]).

3.0 PATTERN SYNTHESIS AND MEMORY RECOMBINATION

Sensemaking in novel situations may entail a form of pattern synthesis. Differences between a current novel situation and previous situations can be significant and sensemaking may require the synthesis of different attributes from different memories rather than retrieval of a single memory (Patterson & Eggleston, [1], [8]). Such pattern synthesis would entail interactions among knowledge structures.

3.1 Interactions Among Knowledge Structures

The generation of synthesized patterns as signs would require the interplay among various sources of knowledge. Based on this idea, knowledge structures would interact with each other to capture the generative capacity of human understanding in novel situations [38]. In humans, several types of memory have been identified [39]; [40]; [41]; [42], [43]): declarative memory, the conscious recollection of facts and events; and nondeclarative memory, which includes procedural memory (unconscious memory of invariant, relational knowledge supporting skill and behavioral dispositions), priming, and classical conditioning. The generation of synthesized patterns as signs may entail the interaction among two or more of these memory systems.

3.2 Insight and Sensemaking

Pattern synthesis and memory recombination during sensemaking is revealed by the insight literature. Insight refers to the sudden conscious realization of a problem solution (i.e., its sense) following a period of impasse, which typically occurs with nonroutine problem solving (e.g., [44]; [45]; [46]). Wallas [47] gave the original description of the phenomenon and suggested four stages of insight, which he called 'illumination': (1) preparation—conscious investigation of a problem; (2) incubation—unconscious processing; (3) illumination—"Aha" experience or sudden conscious insight due to previous unconscious processing during incubation; and (4) verification—conscious assessment of the insight. The first and fourth stages entail conscious processing whereas the second and third stages involve unconscious processing. Specifically, incubation is where the person recognizes the sense of a problem solution intuitively ([1], [8]).

Insight problem solving is unconscious and intuitive as shown by studies revealing that (1) cognitive processing leading to the problem solution during incubation was largely unconscious (e.g., [48]); (2) performance on insight problems was not linked with executive functions associated with working memory ([49]); and (3) participants with impaired neurology in a region of cortex associated with working memory solved 50% more insight problems than healthy participants ([50]). Disconnection from consciousness and working memory is a feature of intuitive cognition.

3.3 Maier's (1931) Study

In a classic study of insight by Maier [51], which is discussed by Patterson and Eggleston ([1], [8]), participants had to tie together ends of two long cords hung from the ceiling of a room and separated by a large distance. A solution was to make one cord a pendulum by tying an object (pliers) to its end and swinging it closer to the other cord so the person could grab both cords simultaneously. Some participants discovered the pendulum solution only after seeing the experimenter casually bump into and sway one of the cords while walking across the room ("help 1"). For 85% of the successful participants (with or without the help), the pendulum solution was discovered suddenly and intuitively, from unconscious processing, and with insight ("Aha" experience). For the successful participants, sensemaking of the pendulum would entail its interpretation as a sign denoting a solution to the two-cord problem.

The pendulum solution, as noted by Maier [51], was derived from a conceptual reorganization (pattern synthesis) of the weight of the pliers, length and position of the cord, and (for those who needed the help) the cord-swaying aspect of “help 1”. Such *synthesis of pattern to derive sense* likely entailed memory recombination: knowledge about (1) how a length of cord behaves when weighted on one end, and (2) what a pair of pliers would weigh, would be retrieved and recombined from procedural memories (unconscious relational knowledge supporting skill development and behavioral tendencies tuned through experience; [42], [43]).

Memory recombination supporting intuitive situational pattern synthesis was fundamental to the insightful creation of the pendulum and its interpretation as a sign (its sense) denoting a solution to the two-cord problem. Other studies on insight can be interpreted analogously (e.g., [44],[48]).

3.4 Compositional AI

The ability to recombine learned concepts, behaviors, and skills has long been recognized as a critical ability for AI systems (e.g., see descriptions of necessary functionality in the original proposal on artificial intelligence by McCarthy, Minsky, Rochester, and Shannon [52]). Particularly in reinforcement learning, there is extensive literature on recombination of behaviors, including Sutton et al. [53] on learning policies over low level options as well as more recent work on multi-task learning by Andreas et al. [54]. In computer vision, many approaches have attempted to explicitly model the decomposition and recombination of conceptual primitives (e.g., deformable part-based object recognition; [55]). Yet, such necessary functionality is usually shown to be less effective in machine systems than more simplistic methods; [56]; see also [57]).

4.0 DISTRIBUTED NETWORK REPRESENTATIONS

Cognitive processing frequently entails an interplay between various sources of knowledge in memory. For example, the insightful creation of the pendulum and its interpretation as a sign denoting a solution to the two-cord problem, discussed above, was derived from a memory recombination process supporting intuitive pattern synthesis. Interactive processing, involving an interplay between bottom-up and top-down information, has been recognized in humans as being critical for behavior as well as for consciousness ([58]; [59]; [60]). This interplay among different sources of knowledge is a feature of distributed representations [61].

4.1 Distributed Representation and Content Addressable Memory (CAM)

The existence of a distributed network representation assumes numerous highly interconnected units and no central processing center [62]. Theoretically, in one kind of distributed representation involving CAM, each memory involves network units that have mutually excitatory interactions with units signifying each of its properties. Thus, activating an attribute of the memory would tend to activate the whole memory; and activating the whole memory would tend to activate all of its attributes. There would also be mutually inhibitory interactions between mutually incompatible attribute units. Thus, different memories would correspond to different patterns of activity over the same hardware units ([63], [64]). One key feature of human memory is that it is content addressable—we can access information in memory based on most, if not all, attributes of the representation we are trying to retrieve [65]. The presence of distributed representations has been directly observed in human brain [66].

The results of the study by Maier [51] can be viewed in a distributed network representation scheme as shown in Figure 3. In Panel A, we have distributed representations of activated memories of a cord (e.g., with attributes strength, color, length, hung, knot on end) and of a pair of pliers (e.g., with attributes shape, weight, grip, color). As indicated in the figure, we hypothesize that the memory of the cord came from rope climbing in a gym class; and memory of the pliers came from carpentry work. In Panel B, we have the same representation as in Panel A except that we have added a new (non-activated) memory of a pendulum (which hypothetically came from seeing clocks) which has the attribute of 'cord sways' and also shares the attributes of 'weight', 'hung', and 'length' with the other memories. We call this a 'shared attribute' principle of distributed representations. In Panel B, we assume that the memory of the pendulum is not yet activated.

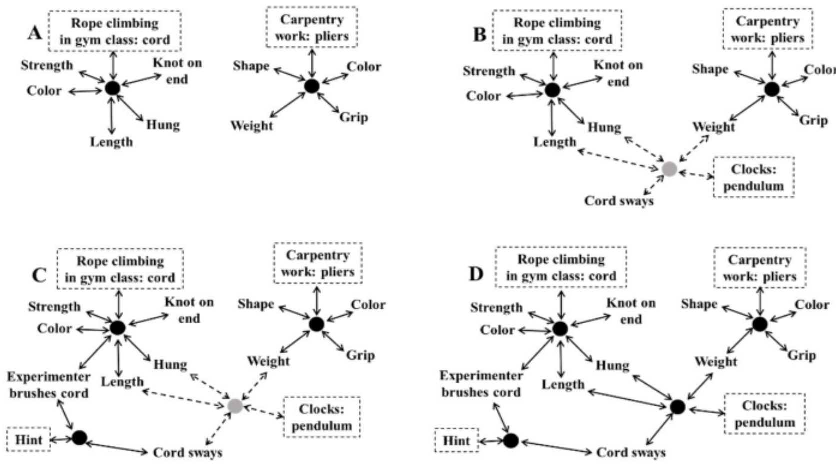

Figure 3. Diagram Depicting Memories as a Set of Distributed Network Representations of Items from a Study by Maier [51].

Panel A shows activated memories of a cord (attributes of strength, color, length, hung, knot on end) and of a pair of pliers (e.g., attributes of shape, weight, grip, color). In Panel B, a new (non- activated) memory of a pendulum has been added, which has the attribute of 'cord sways' and which also shares the attributes of 'weight', 'hung', and 'length' with the other memories. In Panel C, a new (activated) memory of the hint has been added, which shares the attribute of 'experimenter brushes cord' with the memory of the cord, and also shares the attribute of 'cord sways' with the memory of a pendulum. In Panel D, the pendulum memory has become activated, due to the presence of the hint and the activation of the attribute 'cord sways' being added to the coactivation of the attributes of 'length', 'hung', 'weight'.

In Panel C, a new (activated) memory of the hint (which came from seeing the experimenter brush against the cord) is added, which shares the attribute of 'experimenter brushes cord' with the memory of the cord, and also shares the attribute of 'cord sways' with the memory of a pendulum. In Panel D, the pendulum memory has now become activated due to the presence of the hint and the activation of the attribute 'cord sways' being added to the coactivation of the attributes of 'length', 'hung', and 'weight'. In short, the memory of the hint interacts with the memories of the cord and the pliers to cause the dormant representation of the problem solution (pendulum) to become activated, and the participant has the 'Aha' (insight) experience.

4.2 Hypothetical Example

Now consider a hypothetical example of sensemaking adopted from Patterson and Eggleston [1]. Imagine that an individual is on her way to work in the morning and stopped at a traffic light next to a gas station. Suddenly she sees a large dump truck speeding down a nearby steep hill. The truck is out of control and will likely strike the gas pumps. The site of the truck speeding toward the gas pumps would indicate that she could be injured by an explosion, and she would try to escape the situation. Yet how would she know about the danger of exploding gas pumps given that she has never experienced such an event in the past?

Because that exact event has never been previously experienced, the recognized danger could not be based on any single memory. Rather, the sight of the truck speeding toward the gas pumps would serve as a cue for the synthesis of a new frame that would entail a sign involving exploding gas pumps, derived from different memories from past experiences, and what the sign denoted, which would be the possibility of being injured. See Figure 4.

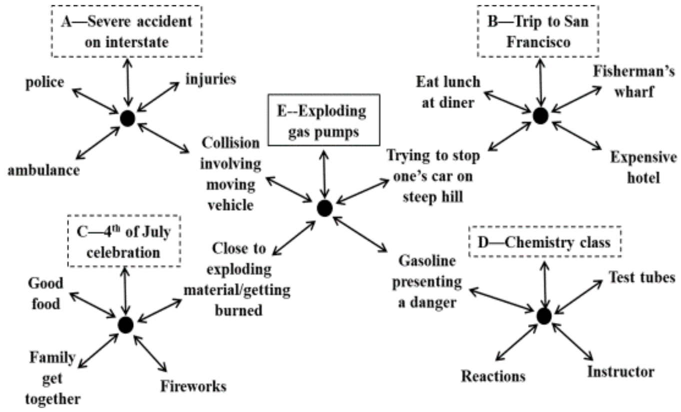

Figure 4. Diagram Depicting Memories as a Set of Distributed Network Representations During Sensemaking (Sign Interpretation) Involving a Hypothetical Accidental Explosion (adapted from Patterson & Eggleston [1]).

Assume that an individual is stopped at a traffic light next to a gas station and suddenly sees an out-of-control speeding dump truck that will likely strike the gas pumps. At the moment the out-of-control dump truck is sighted, different attributes from various memories—such as a collision involving a moving vehicle (panel A: from memory of severe accident on an interstate), trying to stop one's own car on a steep hill (panel B: from memory of trip to San Francisco), being close to exploding material (panel C: from memory of getting burned at 4th of July celebration), and gasoline presenting a danger (panel D: memory from chemistry class)—would be synthesized into a meaningful pattern, a sign, of exploding gas pumps (panel E). This synthesis of sign would be based on the 'shared attribute' principle of distributed representations (see also Rumelhart & McClelland, 1986, vol. 1, p. 28-30). The sign of potential exploding gas pumps would denote possible injury (its sense) and the individual would immediately leave the situation.

In synthesizing this new frame, there could be an episodic memory of not stopping one's car on a steep hill during a previous trip to San Francisco; another episodic memory of a collision involving a moving vehicle on an interstate (truck can't stop → collision, a within-frame sign relation); declarative memory of gasoline being a dangerous liquid during chemistry class; and an episodic memory of a friend being too close to exploding material and getting burned during a 4th of July celebration (collision + gasoline danger → explosion, a within-frame sign relation). This synthesized element 'explosion' would serve as a new sign and its sense, or what that sign denoted, would be the possibility of injury (an across-frame sign relation). Thus, what we have is *sign relations produced by interaction among different representations of knowledge involving different kinds of memory*. See Figure 4. The idea of being injured would set up a new injury frame (not shown). Recognition of potential injury would be posted to consciousness as a feeling of fear.

We can represent this synthesized exploding-gas-pumps frame and attendant sign relations in a BN. The associated CPTs are given in Figure 5 (with the values hypothesized to come from previous memories), and this BN model was simulated with the software tool *Hugin* (version 8.9 [34]). In our simulation, the prior probability of an explosion is 0.73 and the prior probability of an injury is 0.58. However, if the probability of an explosion becomes 100% (i.e., the explosion variable is instantiated), then the posterior probability of an injury becomes 0.80.

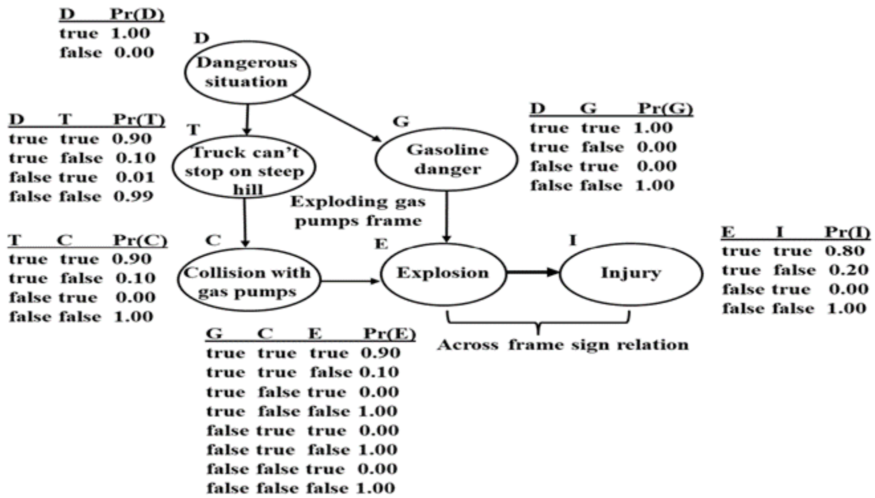

Figure 5. Diagram Depicting an Exploding Gas-Pumps Frame, Synthesized from Different Memories.

. Hypothetically, an individual sees that a large dump truck is out of control as it speeds down a steep hill and will strike nearby gas pumps. The sight of the truck speeding toward the gas pumps would serve as a cue for the synthesis of a new frame with a sign involving an explosion. This synthesized frame and sign would be derived from aspects of the individual's various memories (e.g., unable to stop one's car on a steep hill; collision involving a moving vehicle; gasoline presenting a danger; getting burned from exploding material). The sense of this sign called 'explosion' would be the possibility of being injured. The links D→T, G; T→C; G→E; C→E represent within-frame sign relations. The link E→I represents an across-frame sign relation.

In summary, the frame and accompanying sign relations used to inform sense making in humans can arise from a mental pattern-synthesis and memory-recombination operation carried out in novel environments. The benefit of this distributed network of knowledge for artificial agents is that it would allow for improved dynamic assessment of novel situations and better AI decision-making. This structure would afford AI more flexibility and ameliorate the problem of brittleness from supervised learning approaches, such as the problem of subtle superficial changes in image appearance leading to AI classification errors [67].

We now address the following issue: when sensemaking in novel situations, how does an individual determine which aspects of which memories are to be synthesized?

5.0 EMBODIED SIMULATION

How does an individual determine which aspects of which memories are to be synthesized? We propose that each individual engages in a set of unconscious, embodied, intuitive simulations that attempt to align different combinations of various attributes from different memories until a solution to the sense of a given object, event, or situation is found; that sense is posted to consciousness as a certain gut feeling. With embodied cognition, the same neural substrate that underlies perception and action also underlies central cognitive processing ([68],[69],[70]). Embodied cognitive processes can be conceptualized as high-level "simulators" operating within modal systems of representation ([71], [72]).

This idea of unconscious embodied intuitive simulation stands in opposition to standard cognitive theories based on the concept that representations in modal systems (e.g., vision) are transduced into amodal symbols for representing knowledge. In an amodal cognitive system, perceptual and motor systems would not be playing any significant role in "central" cognitive processing ([68]; [69]). However, the problem with amodal cognitive systems is that their symbols are arbitrary and lack grounding in the natural world—and grounding may be a necessary precondition for sensemaking ([18], [19]). A solution to the symbol grounding problem is to make cognition embodied [70]. We now consider an actual example.

6.0 ACTUAL EXAMPLE

Now contemplate the following actual example that involved measuring human sensemaking using the approach that we are currently advocating (i.e., sign relations embedded within and across frames). To determine whether certain human-machine systems improve the sensemaking capability of intelligence analysts, Frame, Maresca, Christensen-Salem and Patterson [31] investigated *Project Maven*, a machine learning-based recognition aid developed for the US Department of Defense. Project Maven automates the process of analyzing drone surveillance video by classifying objects, such as vehicles or people, in a scene ([73],[74]). Frame et al. tested the idea that, by embracing detection and identification, Project Maven may alleviate the cognitive workload of analysts so that their sensemaking would be enhanced.

Frame et al. [31] measured sensemaking by creating a set of ten simulated intelligence, surveillance and reconnaissance (ISR) compound overwatch scenarios (3-minutes each) viewed as full-motion-video. Each scenario depicted a plot in a terrorist narrative that served as a sensemaking frame (e.g. hostage taking). Sensemaking was measured by event prediction and sign identification—that is, by determining how well the participants (n = 73) could: (1) predict the final ending of each of the ten scenarios ('prediction') after seeing its beginning; and (2) identify the initial sign or cue that denoted the final ending of each scenario ('identification').

Five overwatch scenarios contained signs that were people and/or vehicles and highlighted/tracked by simulated Maven: Improvised Explosive Device (IED) explosion (sign: digging by road); people attacking hospital (sign: someone stealing ambulance); school bus explosion (sign: someone tampering with school bus); attack on first responders (sign: slum lit on fire); and running into protesters (sign: someone stealing car). Five overwatch scenarios contained signs that were not people/vehicles and not highlighted/tracked by simulated Maven: hostage taking (sign: rope); quick change evasion (sign: laundry/clothing); blowing up building (sign: gas canisters); taking over ambulance (sign: injured bodies); shooting at market (sign: dead drop gun). The results showed that having Project Maven direct the participants' attention to people or vehicles serving as a sign did improve the participants' ability to predict the final ending of the scenarios and thus enhance sensemaking.

7.0 DISCUSSION

The ideas and observations expressed in the present paper offer unique information toward creating artificially-intelligent agents with the capacity for sensemaking in novel environments. With respect to sensemaking, our approach entails a unified framework for sensemaking that posits sign relations embedded within and across frames. That is, an element of a given frame may serve as a sign denoting an element of the same frame - i.e., within-frame sign relation. Or an element of a given frame may serve as a sign denoting an element of a different frame—i.e., across-frame sign relation. This unified framework of sensemaking is represented as a Bayesian network, which reflects uncertainty in the sensemaking process.

With regard to novel environments, our approach involves the idea of creating interaction among various distributed-knowledge structures. This type of memory representation is a form of CAM. Such interaction would be mediated via shared attributes among the various memories [75]. Thus, activating a memory would tend to activate all of its attributes, including any shared attributes. Activation of those shared attributes would tend to activate other memories that also share those attributes (there would also be mutually inhibitory interactions between mutually incompatible attribute units). For those activated memories agreeing on an attribute, the node for that attribute would be strongly activated, which would tend to activate other representations. And for those activated memories not agreeing on an attribute, the different attributes would tend to cancel each other out and be suppressed [75]. The net response of the ensemble of network units would be a *synthesized representation of some object, event, or situation that would serve as a unique sign for sensemaking in a novel environment*.

In summary, aspects and attributes across many memories can be shared and recombined in unique ways to create synthesized signs which then denote certain outcomes. In this way, the frame and accompanying sign relations used to define sensemaking can arise from a mental pattern- synthesis and memory-recombination operation carried out in novel environments. Such synthesized signs can be produced by distributed representations of interacting memories that share attributes. The interplay among different sources of knowledge in memory is a key feature of distributed representations underlying human cognitive processing.

7.1 Markov Decision Processes (MDP)

Currently, the ability to make sense in novel environments is outside the scope of contemporary computing science. For instance, consider the field of reinforcement learning (RL). RL refers to a computational approach for automating goal-directed learning and decision making that comes from the direct interaction with the environment. RL uses the formal framework of MDPs that entails the interaction between a learning agent and its environment in terms of states, actions, and rewards. This type of trial-and-error-learning involves learning how to map situations onto actions so as to maximize a reward signal. The concept of trial-and-error learning emerged early in the field of artificial intelligence ([76]; [77]; [78]).

With MDPs, the probability of each possible value for the state and reward depends only on the immediately preceding state and action and not at all on earlier states and actions. This is called the Markov property. With the Markov property, the conditional probability distribution of future states depends only upon the present state; given the present, the future does not depend on the past. This means that the information from past memories would not be available for any pattern-synthesis and memory-recombination operation for sensemaking in novel environments.

7.2 Constraint Satisfaction Network

A distributed representation of the kind discussed in this paper can be seen as a constraint satisfaction network. In a constraint-satisfaction network, we have the following: (1) each unit represents an attribute of a memory; (2) each connection represents restrictions or constraints among the attributes ([75]; [79]); and (3) values for the attributes that permit simultaneous satisfaction of most or all of the constraints. Constraints defined by the pattern of connections among the units determines the set of possible stable states of the system and therefore the set of possible interpretations of the input [80].

A constraint satisfaction network processes input by moving from state to state until it eventually settles into an optimal stable state of 'relaxation' for a given input [80]. In such a state, as many as possible of the constraints are satisfied, with priority given to the strongest constraints. As noted by Hopfield [81], with symmetric weights and asynchronous updates, such systems can be conceptualized as always moving from a state that satisfies fewer constraints to a state that satisfies more constraints—the system has "settled" on a solution to the problem or settled into an interpretation of the input. Such systems can be conceptualized as minimizing a global measure of energy of the system [82].

7.3 Summary

Our approach to sensemaking in novel environments entails sensemaking represented as sign relations embedded within and across frames. The interchange among different sources of knowledge—an aspect of distributed representations—can create synthesized patterns, each of which could be recognized as a sign during the sensemaking process in a novel environment. Aspects of memories that get synthesized can be determined via unconscious, embodied simulations. The benefit of this approach for artificial agents is that it would allow for improved dynamic assessment of novel situations and better AI decision-making.

8.0 REFERENCES

[1] Patterson, R.E., Eggleston R.G., "Human–Machine Synergism in High-Level Cognitive Functioning: The Human Component," in *IEEE Transactions on Emerging Topics in Computational Intelligence*, *2*, 249-257, Aug. 2018, doi: 10.1109/TETCI.2018.2816584.

[2] Fuster, J.M. (2000). The Module: Crisis of a Paradigm. *Neuron, 26*, 51–53.

[3] von Uexkull, J. (1982). The Theory of Meaning. *Semiotica, 42*, 25–82.

[4] Russell, D.M., Stefik, M.J., Pirolli, P. & Card, S.K. (1993). The Cost Structure of Sensemaking. In Proceedings of the INTERACT 1993 and CHI 1993 Conference on Human Factors in Computing Systems, N.Y., 269–276.

[5] Dervin, B. (1983). An Overview of Sense-Making Research: Concepts, Methods and Results. Presented at the Annual Meeting of the International Communication Association, Dallas, TX.

[6] Weick, K.E., Sutcliffe, K.M. & Obstfeld, D. (2005). Organizing and the Process of Sensemaking. *Organization Science, 16*, 409-421.

[7] Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Staszewski, J., & Anderson, J.R. (2013). A Functional Model of Sensemaking in a Neurocognitive Architecture. Computational Intelligence and Neuroscience, Article 921695.

[8] Patterson, R.E., & Eggleston, R.G. (2017). Intuitive Cognition. *Journal of Cognitive Engineering and Decision Making*, *11*(1), 5-22.

[9] Pakkenberg B, Pelvig D, Marner L, Bundgaard MJ, Gundersen HJ, Nyengaard JR, Regeur L. Aging and the Human Neocortex. Exp Gerontol. 2003 Jan-Feb;38(1-2):95-9. doi: 10.1016/s0531-5565(02)00151-1. PMID: 12543266.

[10] Ford, A. (2009, 2nd edition). *Modeling the Environment*. Washington, D.C.: Island Press.

[11] Sterman, J.D. (2000). *Business Dynamics: Systems Thinking and Modeling for a Complex World*. NY: McGraw Hill/Irwin.

[12] Peirce, C. S. (1960). *Collected Papers of Charles Sanders Peirce*(Vol. 2, p. 228). Cambridge, MA: Harvard University Press.

[13] De Saussure, F. (1916/1972). Nature of the Linguistic Sign. Course in General Linguistics, 1, 65-70.

[14] Sebeok, T. A. (1994). Signs: An Introduction to Semiotics. Toronto, Canada: University of Toronto Press.

[15] Hoffmeyer, J. (1996). Signs of Meaning in the Universe. Bloomington: Indiana University Press.

[16] Hoopes, J. (1991). Peirce on Signs. Chapel Hill, NC: University of North Carolina Press.

[17] Bains, P. (2006). The Primacy of Semiosis: An Ontology of Relations. University of Toronto Press.

[18] Harnad, S. (1990). The Symbol Grounding Problem. *Physica D, Nonlinear Phenomena, 42*, 335– 346.

[19] Harnad, S. (2003). The Symbol Grounding Problem. In *Encyclopedia of Cognitive Science*. New York: Nature.

[20] Minsky, M. (1975). A Framework for Representing Knowledge. In Psychology of Computer Vision. New York: McGraw-Hill.

[21] Klein, G., Moon, B. & Hoffman, R.R. (2006a). Making Sense of Sensemaking I: Alternative Perspectives. *IEEE Intelligent Systems, 21*, 70–73.

[22] Klein, G., Moon, B. & Hoffman, R.R. (2006b). Making Sense of Sensemaking II: A Macrocognitive model. *IEEE Intelligent Systems*, *21*, 88–92.

[23] Schank, R. C. (1976). Research At Yale in Natural Language Processing. Yale Univ New Haven Conn Dept of Computer Science.

[24] Schank, R. C., & Abelson, R. P. Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures. Hillsdale , New Jersey : Lawrence Erlbaum Associates, 1977.

[25] Bartlett F.C. (1932). Remembering: A Study in Experimental and Social Psychology. Cambridge, England: Cambridge University Press.

[26] Bobrow, D. G., & Norman, D. A. (1975). Some Principles of Memory Schemata. In Representation and understanding (pp. 131-149). Morgan Kaufmann.

[27] Norman, D. A., & Bobrow, D. G. (1976). On the Analysis of Performance Operating Characteristics. Psychological Review, 83(6), 508.

[28] Rumelhart, D.E. (1975). Notes on a Schema for Stories. In D.G. Bobrow & A. Collins (Eds.), Representation and Understanding: Studies in Cognitive Science (pp. 211-236). New York: Academic Press.

[29] Rumelhart, D.E. (1980). Schemata: The Building Blocks of Cognition. In R. Spiro, B. Bruce & W. Brewer (Eds.), Theoretical Issues in Reading Comprehension (pp. 33-58). Hillsdale, N.J.: Erlbaum.

[30] Rumelhart, D. E., & Ortony, A. (1977). The Representation of Knowledge in Memory, Schooling and the Acquiation of Knowledge Eds: RC Anderson, RJ Spiro, & WE Montague.

[31] Frame, M.E., Maresca, A.M., Christensen-Salem, A., & Patterson, R.E. (Sept 3, 2022). Evaluation of Simulated Recognition Aids for Human Sensemaking in Applied Surveillance Scenarios. *Human Factors*, *0*(0). https://doi.org/10.1177/00187208221120461.

[32] Koller, Daphne, and Nir Friedman. Probabilistic Graphical Models: Principles and Techniques. MIT press, 2009.

[33] Darwiche, A. (2009). Modeling and Reasoning with Bayesian Networks. Cambridge University Press.

[34] Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann.

[35] Hugin New release – hugin v. 6.2. Hugin Expert. (n.d.). https://www.hugin.com/new-release-hugin-v-6-2/

[36] Garcia, W., Clouse, H., & Butler, K. (2022, July). Disentangling Categorization in Multi-agent Emergent Communication. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (pp. 4523-4540).

[37] Li, B. Z., Nye, M., & Andreas, J. (2021). Implicit Representations of Meaning in Neural Language Models. arXiv preprint arXiv:2106.00737.

[38] Pearl, J. (2009). Causality. Cambridge University Press.

[39] Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel Distributed Processing. Volume 1: Foundations* (Vol. 1). The MIT Press.

[40] Adams, S., Arel, I., Bach, J., Coop, R., Furlan, R., Goertzel, B., Hall, J.S., Samsonovich, A., Scheutz, M., Schlesinger, M., et al. (2012). Mapping the Landscape of Human-Level Artificial General Intelligence. AI Magazine 33, 25–42.

[41] Goertzel, B. (2014). Artificial General Intelligence: Concept, State of The Art, and Future Prospects. *Journal of Artificial General Intelligence, 5*(1) 1-46.

[42] Sherry, D. F., & Schacter, D. L. (1987). The Evolution of Multiple Memory Systems. Psychological Review, 94, 439–454.

[43] Squire, L. R. (2004). Memory Systems of the Brain: A Brief History and Current Perspective. *Neurobiology of Learning and Memory*, *82*, 171–177.

[44] Squire, L. R. (2009). Memory and Brain Systems: 1969 –2009. *Journal of Neuroscience*, *29*, 12711–12716.

[45] Sternberg, R. J., & Davidson, J. E. (1995). *The Nature of Insight*. Cambridge, MA: MIT Press.

[46] Broderbauer, S., Huemer, M. & Riffert, F. (2013). On the Effectiveness of Incidental Hints in Problem Solving Revisiting Norman Maier and Karl Duncker. *Gestalt Theory, 35*, 349–364.

[47] Duncker, K., & Lees, L. S. (1945). On Problem-Solving. *Psychological monographs*, *58*(5), i.

[48] Wallas, G. (1926). *The Art of Thought*. Kent, UK: Solis.

[49] Moss, J., Kotovsky, K., & Cagan, J. (2011). The Effect of Incidental Hints when Problems are Suspended Before, During, or After an impasse. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 140.

[50] Gilhooly, K.J. & Fioratou, E. (2009). Executive Functions in Insight versus Non-Insight Problem Solving: An Individual Differences Approach. *Think Reason, 15*, 355–376.

[51] Reverberi, C., Toraldo, A., D'Agostini, S. & Skrap, M. (2005). Better without (Lateral) Frontal Cortex? Insight Problems Solved by Frontal Patients. *Brain, 128*, 2882–2890.

[52] Maier, N.R.F. (1931). Reasoning in Humans: II. The Solution of a Problem and its Appearance in Consciousness. *Journal of Comparative Psychology, 12*, 181–194.

[53] McCarthy, J., Minsky, M.L., Rochester, N. & Shannon, C.E. (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31. Re-printed in *AI Magazine*, *27*(*4*), 2006, 12-14.

[54] Sutton, R. S., Precup, D., & Singh, S. (1999). Between MDPs and Semi-MDPs: A Framework for Temporal Abstraction in Reinforcement Learning. Artificial intelligence, 112(1-2), 181-211.

[55] Andreas, J., Klein, D., & Levine, S. (2017, July). Modular Multitask Reinforcement Learning with Policy Sketches. In International Conference on Machine Learning (pp. 166-175). PMLR.

[56] Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., .. & Zhang, Y. (2023). Sparks of Artificial General Intelligence: Early Experiments with GPT-4. arXiv Preprint arXiv:2303.12712.

[57] Sutton, R. (2019). The Bitter Lesson. Incomplete Ideas (blog), 13(1).

[58] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2009). Object Detection with Discriminatively Trained Part-Based Models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9), 1627-1645.

[59] Dehaene, S., Sergent, C., & Changeux, J. P. (2003). A Neuronal Network Model Linking Subjective Reports and Objective Physiological Data during Conscious perception. *Proceedings of the National Academy of Sciences*, *100*(14), 8520-8525.

[60] Gilbert, C. D., & Sigman, M. (2007). Brain States: Top-Down Influences in Sensory Processing. *Neuron*, *54*(5), 677-696.

[61] Lupyan, G., & Ward, E. (2013). *Language can Boost Otherwise Unseen Objects into Visual Awareness*. PNAS. https://www.pnas.org/doi/full/10.1073/pnas.1303312110

[62] Mayor, J., Gomez, P., Chang, F. & Lupyan, G. (2014). Connectionism Coming of Age: Legacy and Future Challenges. *Frontiers in Psychology, 5*, 187.

[63] Gibbons, M. (2019). Attaining landmark status: Rumelhart and McClelland's PDP Volumes and the Connectionist Paradigm. *Journal of History of Behavioral Science, 55*, 54–70.

[64] Feldman, J. A., & Ballard, D. H. (1982). Connectionist Models and Their Properties. *Cognitive science*, *6*(3), 205-254.

[65] Kriegeskorte N, Mur M and Bandettini P (2008). Representational Similarity Analysis – Connecting the Branches of Systems Neuroscience. *Front. Syst. Neurosci.* **2**:4. doi: 10.3389/neuro.06.004.2008

[66] Chang, E., Rieger, J., Johnson, K. *et al.* Categorical Speech Representation in Human Superior Temporal Gyrus. *Nat Neurosci* **13**, 1428–1432 (2010).

[67] Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erham, D., Goodfellow, I. & Fergus, R. (2014). Intriguing Properties of Neural Networks. arXiv: 1312.6199v4. [cs.CV] 19 Feb.

[68] Barsalou, L.W. (1999). Perceptual Symbol Systems. *Behavioral Brain Sciences, 22*, 577–660.

[69] Barsalou, L.W. (2005). Abstraction as Dynamic Interpretation in Perceptual Symbol Systems in *Building Object Categories*. Mahwah, NJ, USA: Erlbaum, 2005, pp. 389–431.

[70] Barsalou, L.W. (2008). Grounded Cognition. *Annual Review of Psychology, 59*, 617–645.

[71] Clark, A. (1998). Where Brain, Body, and World Collide. Daedalus: Journal of the American Academy of Arts and Sciences, 127, 257–280.

[72] Wilson, M. (2002). Six Views of Embodied Cognition. *Psychonomic Bulletin & Review*, *9*, 625–636.

[73] Clark, C. (2019). Air Combat Commander doesn't Trust Project Maven's Artificial Intelligence—Yet. Retrieved May, 24, 2020.

[74] McLeary, P. (2018). Pentagon's Big AI Program, Maven, Already Hunts Data in Middle East, Africa. *Breaking Defense*.

[75] McClelland, J.L. (1981). Retrieving General and Specific Knowledge from Stored Knowledge of Specifics. Proceedings of the Third Annual Conference of the Cognitive Science Society, Berkley, CA.

[76] Turing, A.M. (1948). Intelligent Machinery, A Heretical Theory. The Turing Test: Verbal Behavior as the Hallmark of Intelligence, 105.

[77] Minsky, M.L. (1961). Steps toward artificial intelligence. *Proceedings of the Institute of Radio Engineers, 49*, 8–30. Reprinted in E.A. Feigenbaum and J. Feldman (eds.), Computers and Thought, 406–450. McGraw-Hill, New York, 1963.

[78] Michie, D. (1974). On Machine Intelligence. Edinburgh: Edinburgh University Press.

[79] Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel Distributed Processing. Volume 1: Foundations* (Vol. 2). The MIT Press.

[80] Rumelhart, D.E., Smolensky, P., McClelland, J.L. & Hinton, G.E. (1986). Schemata and Sequential thought Processes in PDP Models (pp. 8-57). In J.L. McClelland, D.E. Rumelhart et al. (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Volume II.* Cambridge, MA: MIT Press.

[81] Hopfield, J.J. (1984). Neurons with Graded Response have Collective Computational Properties like those of Two-State Neurons. *Proceedings of the National Academy of Sciences, 81*, 3088-3092.

[82] Hinton, Geoffrey E., and Terrence J. Sejnowski. "Learning and Relearning in Boltzmann Machines." *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* 1.282-317 (1986): 2.

7.0 LIST OF ABBREVIATIONS, ACRONYMS AND SYMBOLS

AI	Artificial Intelligence
BN	Bayesian Network
CAM	Content Addressable Memory
CPT	Conditional Probability Tables
IED	Improvised Explosive Device
ISR	Intelligence, Surveillance and Reconnaissance
MDP	Markov Decision Processes
RL	Reinforcement Learning