To the Graduate Council:

I am submitting herewith a thesis written by Myles Brandon Bogner entitled "Creating a 'Conscious' Agent". I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Masters of Science, with a major in Mathematical Sciences.

Major Professor, Stan Franklin, Ph.D.

We have read this thesis and recommend its acceptance:

Edward Ordman, Ph.D.

Art Graesser, Ph.D.

Accepted for the Council:

Linda L. Brinkley, Ph.D.
Vice Provost for Research
& Dean of the Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Memphis, I agree that the Library shall make it available to the borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature ______________________________

Date ______________________________

Creating A "Conscious" Agent

A Thesis

Presented for the

Master of Science

Degree

The University of Memphis

Myles Brandon Bogner

May, 1998

ABSTRACT

Bernard Baars' Global Workspace Theory is a leading cognitive science model for human consciousness. *Creating A "Conscious" Agent* is a design specification for part of the first software implementation of Global Workspace Theory. Conscious Mattie, being developed by the University of Memphis' Conscious Software Research Group, is the name of the software agent where this implementation is occurring at the time of this thesis' submission.

There are several motivators for this research. First, it is hypothesized that a Global Workspace gives a multi-agent system several advantages. For example, it offers individual agents in the system a means of recruiting other agents in the system to help solve novel and ambiguous problems. Also, it gives a method for attentional focus for the overall system. This provides a means for associative learning and metacognitive functions to take place. Chapter five gives an in depth discussion of the specific functions of the Global Workspace in Conscious Mattie. Chapter five attempts to lend insight into the software implementation advantages gained when using a Global Workspace.

As a second motivation, the Conscious Software Research Group hopes to provide new hypotheses about human consciousness for cognitive scientists and neuroscientists. As a cognitive science theory, Baars' theory does not contain many of the low-level design specifications necessary for a computer scientist's implementation of the theory. Many of the design decisions made through this research can be considered hypotheses on human consciousness. It is hoped that the members of the above disciplines will view these implementation decisions as springboards for the further study of consciousness.

This thesis first gives an overview of Baars' Global Workspace Theory. Virtual Mattie, Conscious Mattie's predecessor agent, is then discussed. A general discussion of

Conscious Mattie's numerous components such as learning mechanisms, memory, and perception follows. The author has been actively involved in both Virtual Mattie's and Conscious Mattie's overall design. The author's primary research contribution is found in chapter five. Chapter five describes the design of a Global Workspace mechanism for Conscious Mattie. The Conscious Software Research Group's implementation decisions for Conscious Mattie are related back to Baars' theory in chapter six as justification for their use. Finally, a future direction of the Conscious Software Research Group is presented.

This thesis describes the design for a software implementation of Global Workspace Theory. It is beyond the scope of this thesis to address whether or not Conscious Mattie's "conscious experience" will be similar to the human conscious experience. This author plans to implement and experiment with a Global Workspace for Conscious Mattie. It is hoped that these results will generate several hypotheses about human consciousness.

PREFACE

Why do we pursue goals? Why is it that we actually feel pain in our hearts over emotional loss? Why does it seem that the majority of us always have a thought in our minds? I'm sure these are questions all of us have had at some point. I did for a long time. I still do. In college, however, through both industrial psychology and neuroscience oriented courses, I began to get answers to these questions. As a computer science major, however, my courses in these areas were limited. My desire to study how to integrate computer science with biological phenomena was a primary motivation in my attending graduate school.

Global Workspace Theory is a significant contribution to cognitive science. It is also the Conscious Software Research Group's belief, myself included, that it can make a significant impact on computer science. I try to illustrate this through *Creating A "Conscious" Agent*.

I would like to thank my advisor, Dr. Stan Franklin, for initiating a project which has allowed me to enjoy pursuing my graduate degree. Dr. Franklin has met me many times on late Friday afternoons when all others have deserted the university. I would like to thank all those who have and continue to participate in this project. I could not write this thesis without them. I always seem to leave our weekly meetings inspired.

I would like to thank my sister for listening to her big brother. Finally, I would like to thank my parents, who willingly continue to be wonderful guides.

Contents

List of Figures

Chapter 1

Thesis Introduction

1.1 Thesis Overview

The Conscious Software Research Group at the University of Memphis is investigating the possibilities of Conscious Software Agents (Franklin and the Conscious Software Research Group, 1997). A Conscious Software Agent is an intelligent agent which is an implementation of Bernard Baars' Global Workspace Theory (Baars, 1988, p.43). Global Workspace Theory provides a model for human consciousness. It does so within the framework of other mental components, such as emotion, learning, and memory modules. A Conscious Software agent may have many of such components.

This thesis presents the design for Conscious Mattie, the Conscious Software Research Group's first agent designed within the framework of Global Workspace Theory. The author has contributed to the design of each of the subjects presented by this thesis. The author's primary research focus and contribution is to the design of the agent's Global Workspace, described in chapter five.

The author's research is motivated by two overall themes. First, it is hypothesized that Global Workspace Theory can be a useful software engineering methodology for multi-agent systems. Global Workspace Theory provides such a system a method for attentional focus. It also offers a means for one collection of agents, working together, to communicate with all other agents in the system. This communication facilitates many of the system's functions such as learning and metacognition. Second, Baars' theory does not address many of the specific details necessary for a computer scientist to implement it. For example, it does not give detailed specifications on what each agent communicates to the system's other agents. The Conscious Software Research Group, in their design of Conscious Mattie, has made hypotheses about human consciousness. It is hoped that these hypotheses are studied by cognitive scientists and neuroscientists and lead to further knowledge in these fields.

Chapter two presents Global Workspace Theory. This includes the background for the theory, why it is needed, and what the theory states. Virtual Mattie is then discussed (Franklin et al., 1996). Virtual Mattie, an intelligent clerical agent, is the Conscious Software Research Group's first step in the creation of a Conscious Software Agent. This agent contains many of the underlying modules of Conscious Mattie (Franklin and the Conscious Software Research Group, 1997). Chapter four gives an overview of Conscious Mattie, the first agent to be implemented within the framework of Global Workspace Theory.

Chapter five describes the design for a Global Workspace in Conscious Mattie. There is discussion on the design decisions made for the implementation-level issues not addressed by Baars. This includes how coalitions of processes are created and managed, how the spotlight of attention is controlled, and the broadcasting technique used. This chapter also addresses how a Global Workspace is expected to enhance the agent. It is beyond this thesis' scope to address whether Conscious Mattie's "consciousness" will be similar to human consciousness.

Chapter six relates Conscious Mattie's components back to Global Workspace Theory's contexts. The chapter contains conclusions relating to the previous chapters. Possibilities for future conscious agents are also addressed.

Chapter 2

An Introduction To Global Workspace Theory

2.1 Introduction

This chapter describes Bernard Baars' Global Workspace Theory. Global Workspace Theory is a cognitive science model of the human conscious experience. This chapter focuses on the portion of Global Workspace Theory which most closely deals with consciousness. Global Workspace Theory encompasses additional subjects such as learning and metacognitive functions. This chapter first gives the operational definition and describes the functions of consciousness. The components of Global Workspace Theory are discussed. How Global Workspace Theory works is presented. Finally, a detailed example is used to trace through the theory.

It is hoped that by the end of this chapter the importance of Global Workspace Theory as a concrete description of the conscious experience is conveyed. Its computer science implications will be examined throughout the remainder of this work.

2.2 Consciousness

Baars states that throughout human history, consciousness has been extremely difficult to define. He states, "Even today, more and more nonsense is spoken of consciousness, probably, than of any other aspect of human functioning" (Baars, 1988, p.4). According to Baars, consciousness, while it can be inferred from reliable evidence, is a theoretical construct. Consciousness, therefore, is defined in terms of what makes up the human conscious experience. A two-part definition is necessary to define what it means to be conscious of an event. The first part states that a person is defined to be conscious of an event if the person states that they were conscious of it immediately after the event occurs. Events may be conscious for only hundreds of milliseconds. The second portion says that the experiencer's report must be able to be independently verified.

The reader may be wondering how this psychological definition ties into computer science. As seen further in chapter five, it contains two main implications for a software implementation of the theory. First, consciousness will contain elements which relate to events. Second, elements may be conscious in the agent for an extremely short period of time.

2.2.1 Functions of Consciousness

Baars states that consciousness has nine major functions. The reader will encounter these functions again in chapter five when it is shown how Conscious Mattie's implementation of consciousness fulfills these functions. The first function is Definition and Context-setting. An example of this occurs when one focuses on a distant tree in a forest. While multiple visual stimuli are present, a coherent image is able to be retrieved. The second function is Adaptation and Learning. For example, extremely difficult material is often pondered for a great deal of time when attempting to learn it. The third function is Editing, Flagging, and Debugging. This is evident in biofeedback training, where persons use flagging in order to gain conscious voluntary control over usually unconscious systems. Fourth is the Recruiting and Control function. An example of this function's use occurs when attempting to answer a question. While one is conscious of a question, the candidate answers to that question are recruited unconsciously and brought to consciousness. The fifth function is Prioritizing and Access-control. This occurs when learning a foreign language. One may wish to prioritize words which are difficult to pronounce.

The Decision-making or Executive function is useful in controlling thought and action. A decision-making question is “Should I go to the mall or to the park?” The Analogy-forming function occurs when people make analogies to compare a novel experience to known ones. A example of this is “Hate is the wrong road to travel.” The eighth function is the Metacognitive or Self-monitoring function. One example of this is humans’ ability to pinpoint and express their current feelings. The final function is Autoprogramming and Self-maintenance. This can be seen in the desire to exercise and eat properly in order to keep the body healthy.

2.3 The Components of Global Workspace Theory

Global Workspace Theory is an attempt to integrate the large amount of information about consciousness into one model. To comprehend the theory, its underlying components must first be understood.

2.3.1 Processors and Coalitions

Global Workspace Theory states as a premise that the nervous system is composed of processors. Each processor is autonomous and has a narrow focus. It is very efficient, works at high speeds, and makes very few errors. Each processor can act in parallel to others. This allows for the creation of a high capacity system such as the central nervous system.

A coalition is a set of processors which work together to perform a specific task. For example, it takes numerous processors for breathing to take place. Coalitions are recursive in nature. For example, a coalition may be composed of several coalitions. A processor may be a member of more than one coalition. Coalitions normally perform routine actions. However, coalitions may also perform duties relating to ambiguous, conflicting, or novel events for the system. These coalitions have the potential for entering consciousness. Baars states that coalitions may have an activation level. When performing routine actions, a coalition's activation level is low. Coalitions performing more uncommon tasks have a higher activation level. A higher activation level gives a coalition a greater chance to enter consciousness. Baars is careful to point out that a high activation level may be a necessary but possibly not a sufficient condition for a coalition to enter consciousness.

2.3.2 Contexts

Contexts are relatively stable coalitions of processors which affect consciousness. Contexts are normally unconscious. Therefore, contexts are not usually experienced directly. Contexts interact rapidly with what is occurring in consciousness. One example of a context is found in a collegiate classroom. Both students and faculty know that a certain classroom behavior is expected in an engineering class, while not always being conscious of it. This example illustrates cultural context, a main context according to Baars. Perceptual, conceptual, and goal contexts are the three other main contexts. Contexts provide an underlying level of stability. In the above classroom example, if the professor began singing during class, the students would quickly become conscious of the fact that the professor was exhibiting abnormal behavior. The professor's behavior went against the prevailing cultural context. In this sense, contexts allow for novel events to become conscious. New contexts can be learned, allowing for reality to be perceived in a further enhanced way.

2.3.3 The Global Workspace

The crux of the theory's architecture is the Global Workspace. The Global Workspace is intended to implement consciousness. The Global Workspace is a working memory which gives a central location for one coalition to interact with the system's other processors. Therefore, the Global Workspace can be considered analogous to a classroom blackboard.

The next section discuses how this information exchange works and what information this exchange provides.

2.4 How Global Workspace Theory Works

Global Workspace Theory can be compared to a theater's stage. In a theater, a spotlight is often used to focus the audience's attention. The spotlight roams around the stage, and there is normally only one person in the spotlight at a time. A person is usually in the spotlight when they are performing new actions. More is usually occurring in the production than what is currently in the spotlight.

This Global Workspace Theory theater is an interactive one. When members of the audience see something in the spotlight they can relate to, they begin acting. These members may join the actors in the spotlight, or they may begin acting outside the audience's main focus.

In Global Workspace Theory, consciousness is the spotlight which roams over the active unconscious processors. This spotlight shines on coalitions attempting to solve difficult problems. It shines on coalitions performing tasks relating to ambiguous and conflicting items. It also focuses on coalitions dealing with novel situations. Many unconscious coalitions and processes are working even while a particular coalition is in consciousness.

The audience is the unconscious processors not on the stage (Baars, 1997, p.42). When coalitions enter consciousness, they broadcast information to all processors. Some audience members which understand this information become active and perform their specific functions. These processors, therefore, potentially contribute to the work of the conscious coalition.

The spotlight can shine on only one coalition at a time. For example, humans must think of two alternatives one after the other; they cannot be addressed at the same time.

Due to its serial nature, consciousness is a much smaller capacity system compared to the large capacity system created by the numerous unconscious processors acting in parallel.

2.4.1 Overall Steps in Global Workspace Theory

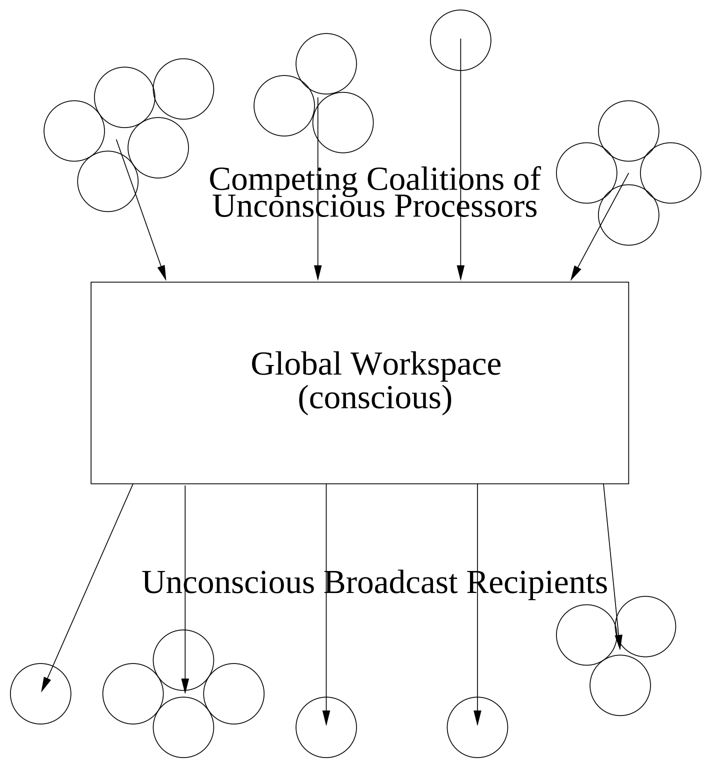


Figure 2.1: Global Workspace Theory

Figure 2.1 shows an illustration of Global Workspace Theory (Baars, 1988, p.88). It is important for the computer scientist to remember that Global Workspace Theory is a high level model. Chapters four and five propose a method for implementing this theory.

To understand Global Workspace Theory, it is helpful to think of processors going through five stages. These are:

1. Unconscious processes, each working towards achieving a portion of an overall goal, form a coalition.

2. Coalitions compete for access into the Global Workspace.

3. The coalition which enters consciousness broadcasts information to all unconscious processes. This broadcast allows the conscious coalition to recruit other processes which can contribute to the conscious coalition's tasks.

4. All unconscious processors will receive the broadcast message. However, only certain ones will be able to understand its contents.

5. The processors which understand the message and which need to take action do so.

2.5 An Example to Illustrate Global Workspace Theory

Tracing through an example helps in understanding Global Workspace Theory. Patricia is a fourth year piano major who began playing the instrument her first collegiate year. At this point Patricia is well accomplished, but does not yet have a maestro's skill. She is currently performing the fifth piece of her memorized ten piece recital, and has so far faced no problems in her performance.

By this point in the recital, Patricia is relatively relaxed. She is playing habitually. While she is conscious of what she is hearing, she basically is letting her hands move themselves through this piece's long runs. Suddenly, however, in the quiet auditorium, Patricia hears the loud, sharp bark of a dog. This causes her to slightly lose concentration and hit two wrong notes. Patricia, with no recollection of where her fingers are to go next, jumps to a later portion of the piece where she continues to play. Only the keenest audience members know there was something in her playing out of place.

This example illustrates numerous aspects of Global Workspace Theory.

1. With her long hours of practicing, Patricia's ability to play her current piece has become habitual. Unconscious processors are controlling her finger and pedal

movement. Patricia is mainly focusing on each note she hears. In this case, the components involved in listening to the piano notes form one or more conscious coalitions.

2. The dog's loud bark introduces a new, unexpected, and potentially dangerous event. A coalition of processors involved in determining the presence of danger forms. This coalition may have a processor to determine a sound's volume, another to determine a sound's pitch, another to determine the direction a sound is coming from, another which detects novel sounds, etc. This coalition, having a very high activation level relative to the piano notes, gains access to consciousness. It then broadcasts information. This information is received by the unconscious processors. Visual and auditory processors which can help determine if there is any immediate danger respond.

3. Patricia, however, rapidly realizes that there is no immediate risk and she is in front of an audience. In this case, a new coalition arises into consciousness. This one's goal is to get her back on track in her piece. To do this, this coalition broadcasts asking for help. Processors in auditory, memory, and visual retrieval respond. Those processors containing the solution gain a high enough activation to reach consciousness.

4. While in the short term thoughts of the dog and the crowd's response come in and out of consciousness due to their remaining high activation level, as their activation falls Patricia relaxes into her practiced habitual playing mode.

2.6 Conclusions

Humans utilize consciousness extensively. It provides a means of cooperation for coalitions. This fosters conflict resolution, learning, and perceptual clarification. Global

Workspace Theory provides a high-level model describing human consciousness. Conscious Mattie is the first attempt to implement this experience as described by this theory. Before this Conscious Software Agent can be discussed, however, her predecessor, Virtual Mattie, must be examined.

Chapter 3

Virtual Mattie

3.1 Chapter Overview

Virtual Mattie is an intelligent autonomous agent. She functions in a clerical role. Specifically, she coordinates departmental seminar information, carrying out a role originally performed by the department's former secretary, Mattie. While Virtual Mattie is not a conscious software agent, she contains many of the underlying components of conscious software. The agent was developed by Stan Franklin and the Conscious Software Research Group and programmed by Honjung Song, Zhaohua Zhang, and Aregahegn Negatu. She is implemented in Java. Virtual Mattie will now be described in detail.

This chapter first discusses Virtual Mattie's job description. Next, the reasons why Virtual Mattie can be considered an autonomous agent are addressed. The agent's architecture is then presented. Finally, conclusions are drawn with mention of how this agent can be improved. The majority of the material contained in this chapter comes from *Artificial Minds* (Franklin, 1995), "Virtual Mattie — an Intelligent Clerical Agent" (Franklin et al., 1996), conference proceedings on Virtual Mattie (Franklin and the Conscious Software Research Group, 1997), and Conscious Software Research Group discussions.

3.2 Virtual Mattie's Job Description

Virtual Mattie is an unsupervised agent. She functions to announce the University of Memphis' Department of Mathematical Sciences' weekly seminars. Virtual Mattie communicates completely via email. Below is a discussion of the different components utilized by the agent.

Virtual Mattie gathers seminar information from seminar organizers. She accepts e-mail from organizers about their upcoming seminars. Since there is no predetermined format which the organizer's email messages must take, Virtual Mattie has natural language processing ability. The agent generates and sends acknowledgements to emailers for every incoming message.

Virtual Mattie composes the upcoming week's seminar announcement. She composes messages stating she has incomplete information. She also writes messages saying a received message was not understood.

Virtual Mattie emails the composed seminar announcements to a recipient's list at a specified time. To do this, the agent maintains a list of people who receive the weekly seminar announcements. Therefore, Virtual Mattie accepts incoming email for requests to join and leave the seminar list.

3.3 Virtual Mattie As An Autonomous Agent

Franklin and Graesser define an autonomous agent as "a system situated within and a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to effect what it senses in the future" (Franklin and Graesser, 1997, p.6). Based on this definition, Virtual Mattie has many properties which enable her to be an autonomous agent.

- Her environment is the UNIX operating system.

- The agent's niche is the maintenance of seminar announcements.
- Virtual Mattie can sense incoming email. The degree to which she actively understands these messages corresponds to different perceptual levels. She also is aware of dates.
- The agent's multiple drives are diverse and explicitly represented.
- Virtual Mattie has a distinctive action selection mechanism, known as the Behavior Network, which is not controlled by a central executive.

3.4 Virtual Mattie's Architecture

Virtual Mattie's architecture is an original high-level agent architecture. The architecture is largely based on the Behavior Networks developed by Pattie Maes (Maes, 1990) and the model of perception found in Hofstadter and Mitchell's Copycat project (Hofstadter and Mitchell, 1994). Both architectures have been modified and significantly extended for Virtual Mattie. Below is a diagram and discussion of Virtual Mattie's architectural components.

3.4.1 Drives

The drives portion of the architecture is based on Maes' goals. The agent's drives correspond to her tasks found in the above Job Description section. All of Virtual Mattie's drives are built into the agent. These drives can operate in parallel. Some drives vary in urgency. For example, the urgency level for sending out a seminar announcement may be higher as it gets closer to the time to send the announcement. This varying in the level of urgency is an addition to Maes' original work. Each drive activates behaviors which work to fulfill the drive.

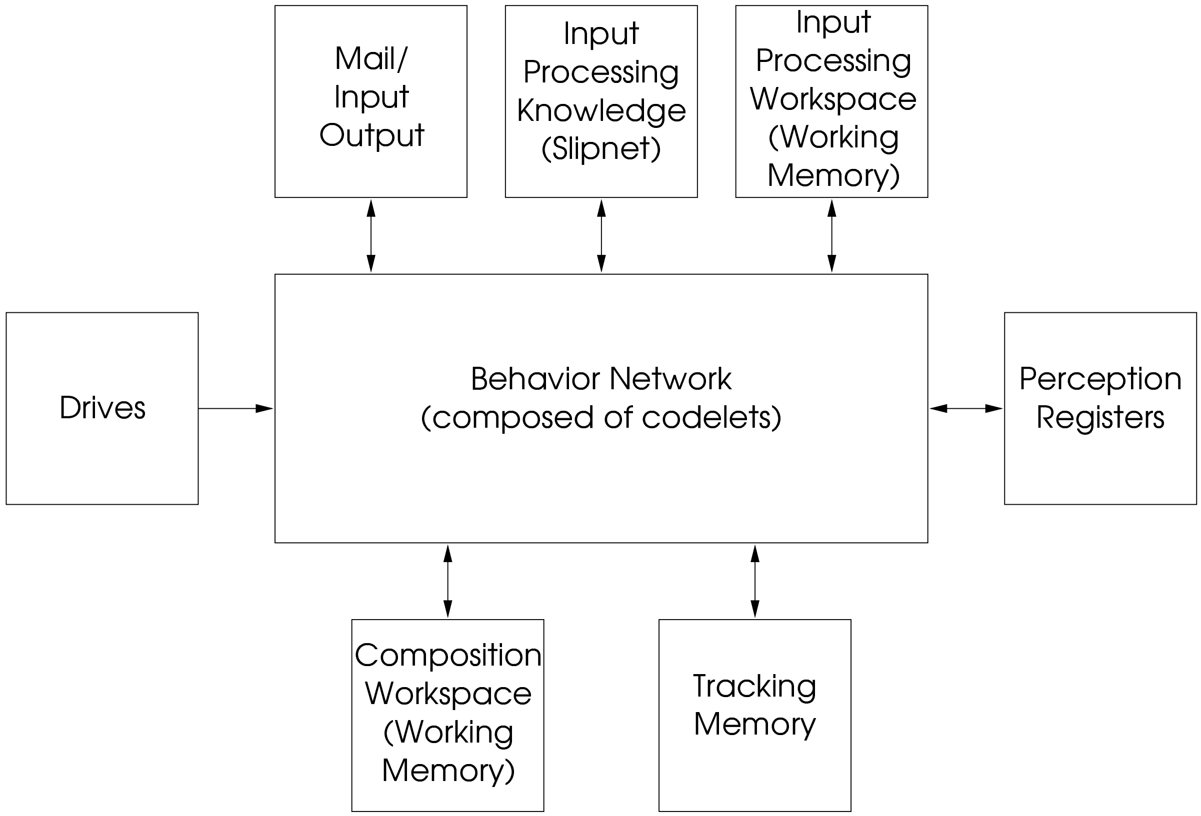


Figure 3.1: Virtual Mattie's Architecture

3.4.2 Behavior Network

The Behavior Network is composed of behaviors whose role is to fulfill the drives. Behaviors in Virtual Mattie's architecture correspond to Maes' competences. Behaviors have an activation level. In general, this activation level is affected by the Drives, the agent's internal conditions, and the Perception Registers. The Perception Registers serve as the Behavior Network's environmental inputs. Behaviors have preconditions that must be met. For example, a behavior's preconditions might be fulfilled if an organizer's email message contains specific items of information. A behavior's activation level increases as more of its preconditions are met.

A behavior's activation level is spread to other behaviors. Broadly speaking, its activation spreads to three locations. A behavior's activation is spread to those behaviors which can fulfill this behavior's unmet preconditions. Also, the behavior's activation is spread to the behaviors whose preconditions can be filled by this behavior. Third, the specific behavior sends inhibition, causing a reduction in activation level, to all behaviors which can remove one of its currently met preconditions. Due to the way behaviors spread forward and backward activation, each behavior can be thought to be a part of a behavior stream. If a behavior has a high enough activation level and all of its preconditions are met, it has the potential to become active. Only one behavior in a behavior stream can be active at a time. The active behavior is determined by choosing the executable behavior in the behavior stream with the highest activation level above a threshold level. Each behavior stream can also be thought of as a plan, created and executed without the building of a search tree.

As a behavior's activation spreads, it diminishes in strength. Also, activation level continually decays at a slow rate. Once a behavior performs its function, its activation level returns to zero. The behavior network is tunable through global parameters.

3.4.3 Perception

Virtual Mattie's sensory data are the incoming email messages she receives. Perception for the agent occurs by her comprehending these email messages. Virtual Mattie contains three percpetual components: the Input Processing Knowledge, the Input Processing Workspace, and the Perception Registers.

Input Processing Knowledge (Slipnet)

Virtual Mattie's Input Processing Knowledge, also known as the Slipnet as it is based on Copycat's Slipnet (Hofstadter and Mitchell, 1994), contains the knowledge needed to understand incoming email messages. Two years worth of email messages to the department secretary were studied in order to generate the knowledge utilized by Virtual Mattie's Slipnet. Items in the Input Processing Knowledge include the message type, ways to identify the different portions of email messages such as the name of the seminar and speaker, and the abbreviation of words. A common case of abbreviation found in departmental email is seen in the writing of building names, such as Dunn Hall being abbreviated DH, D.H., or D. Hall. Another common example is in the days of the week, where Thursday can be abbreviated Thu, Thurs, Th, etc.

Virtual Mattie's Input Processing Knowledge contains knowledge of nine message types. Examples of message types include messages declaring the establishment of a new seminar and messages stating the upcoming speaker and topic for a seminar. Virtual Mattie uses surface level natural language processing in conjunction with a feed-forward neural network to determine an email's message type. The highest output value from the neural network is taken to be the candidate message type. If the output of the neural network is inconclusive, Virtual Mattie sends back an acknowledgement saying the message was not understood.

Input Processing Workspace (A Working Memory)

Sometime after a candidate message type is determined, a message template of this type is placed in the Input Processing Workspace. Codelets, described in the next section, work to fill the template's fields. Similar to Copycat, as mandatory fields in the template are filled, the temperature, representing the proximity to completeness, falls. If the temperature falls low enough a message is considered understood. However, if a certain number of mandatory fields remain empty after the codelets have completed their tasks, the next highest output of the neural network is tried as the appropriate message type. For safety in the event a message is classified incorrectly, Virtual Mattie acknowledges every message. If the acknowledgement conveys an incorrect understanding of the message, the seminar organizer can send a reworded message.

Perception Registers

Once a message template has been filled, its contents are moved to the Perception Registers, and the perception module can begin working to understand another message. The Perception Registers are similar to a blackboard. Information which is placed in the Perception Registers is available for utilization by Virtual Mattie's other modules such as the message composition component described below.

3.4.4 Codelets

Each codelet can be thought of as a small distinct agent designed to perform one task. The term codelet originated with Copycat. Virtual Mattie's codelets correspond to Baars' processors described in chapter two. The above two sections described Virtual Mattie's Behavior Network and perceptual module. These two portions are largely implemented via codelets. For example, one codelet's task in the Slipnet is to fill a message template's speaker name field. Codelets perform the vast majority of Virtual Mattie's actions.

Most codelets serve to implement a behavior or a portion of the Slipnet. However, Primitive Codelets also exist. A Primitive Codelet does not serve the Behavior Network or Slipnet. Instead, it functions independently to perform housekeeping functions. For example, a Primitive Codelet might poll for an incoming email message addressed to Virtual Mattie.

3.4.5 Tracking Memory

Tracking Memory contains the information utilized in composing outgoing email messages. Tracking memory contains the default information on seminars, such as the day of the week each one occurs. It saves the current seminar announcement mailing list. Both the seminar and mailing list information are updated via codelets attached to behaviors. Tracking Memory also stores the templates for different types of outgoing messages. The corpus of email messages collected for two years contributed to the building of Tracking Memory.

3.4.6 Composition Workspace (A Working Memory)

All outgoing messages are composed in the Composition Workspace. Message composition consists of filling the fields of an outgoing message template. The information used to fill the template fields comes from the Tracking Memory and the Perception Registers. There is always a copy of the current seminar announcement being generated in the Composition Workspace. As new information arrives in the Perception Registers and Tracking Memory, the template fields are filled. When a seminar is announcement is mailed, a new default announcement template is placed in the Composition Workspace.

3.4.7 Mail Input and Output

The Mail Input and Output portion of Virtual Mattie deals with the actual receipt of incoming email messages and the sending of outgoing ones. Incoming email messages are first received by the Mail Input portion. Messages are moved from here to the perceptual module. Once an outgoing message is fully composed, it can be moved to the Mail Output portion. Mail Output hands off the outgoing message to the operating system.

3.5 Conclusions

At the time of this writing, the initial development on Virtual Mattie has just been completed. The agent is currently in the experimentation phase. It appears that Virtual Mattie is able to accurately perform the vast majority of her duties.

To effectively coordinate departmental seminars, Virtual Mattie's job description needs expansion. Currently the agent cannot accurately handle an incoming message containing two message types. She cannot deal with one time events such as a colloquium. Virtual Mattie does not perform any learning. Learning would be very useful in several areas, such as learning new message types and new behaviors.

Virtual Mattie's modules provide an agent implemented by combining and extending several recent artificial intelligent mechanisms. The agent implements several portions of Baars' Global Workspace Theory. This correspondence will be discussed in chapter six. It is hoped that by implementing the remainder of the theory many of the areas needing improvement can be enhanced. The next chapter presents a design overview for Conscious Mattie, the successor to Virtual Mattie which is designed under the framework of Global Workspace Theory.

Chapter 4

Conscious Mattie

4.1 Introduction

Conscious Mattie, an extension to Virtual Mattie, is the first software agent designed to be an implementation of Baars' Global Workspace Theory. Therefore, Conscious Mattie can be considered the first "conscious" software agent. This chapter gives a two-part overview of Conscious Mattie. First, the additions to Virtual Mattie's job description are addressed. Second, a broad overview of Conscious Mattie's new architectural components is given. At the time of this chapter's writing, Conscious Mattie is in the design phase.

4.2 Conscious Mattie's Job Description

Conscious Mattie performs all the tasks of Virtual Mattie. In addition, Conscious Mattie can:

- Understand messages which contain multiple message types. For example, an email message might state, "I am establishing a new seminar entitled the 'Artificial Life Seminar'. It will meet on Wednesdays from 1:30 to 3:30 in room 222 in the Math Building. Mr. Lee McCauley will be the first speaker, talking on 'Defining Artificial Life.' " In this case, the two message types are the new seminar establishment and speaker-topic message types.

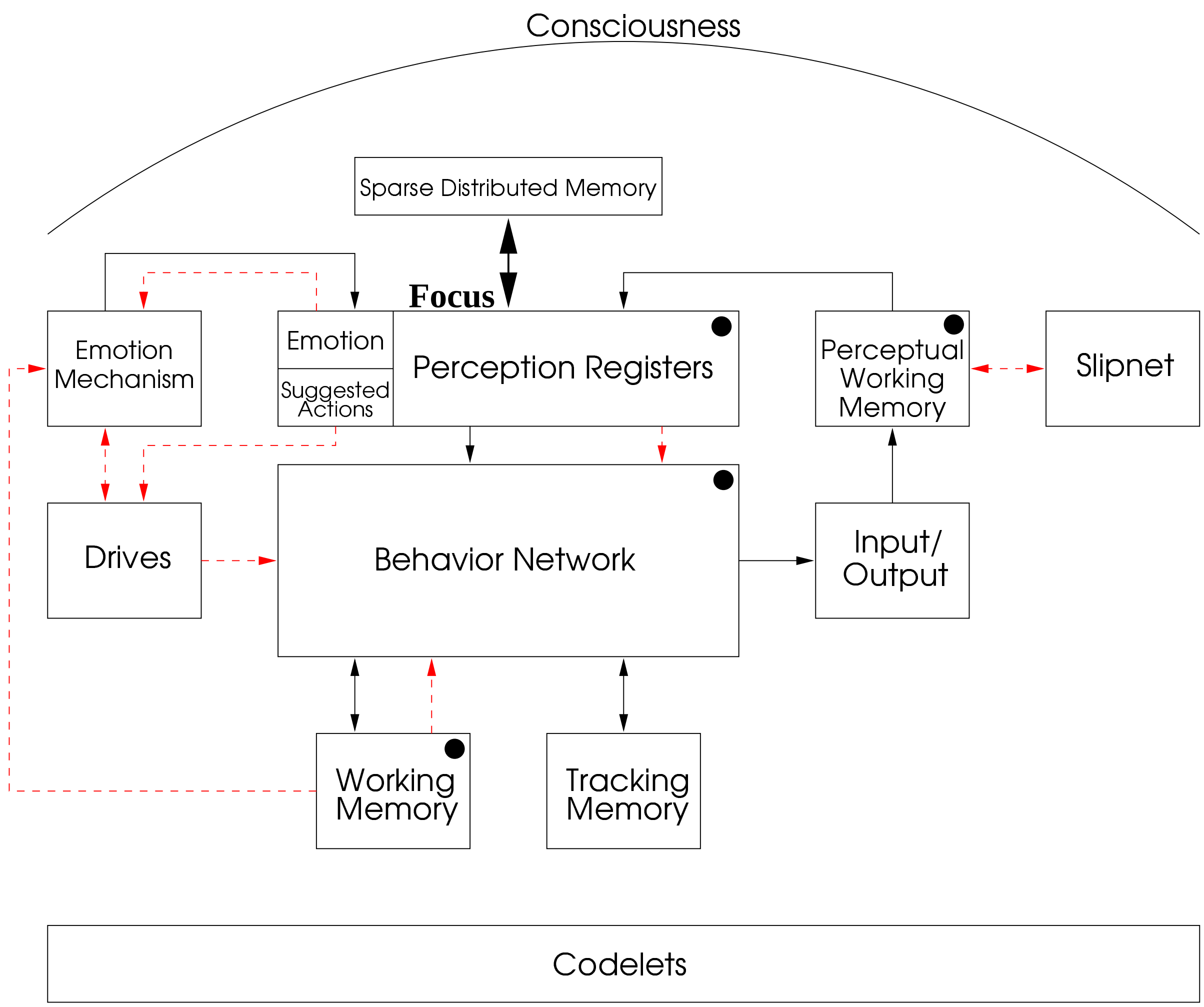
- Coordinate one-time events such as a colloquium.
- Detect conflicts such as in the time and place of seminars.
- Detect problems in the operating system and properly prepare herself for a system shutdown.

4.3 Conscious Mattie's Architectural Additions To Virtual Mattie

Conscious Mattie contains several architectural additions to Virtual Mattie, enabling the agent to be an implementation of Global Workspace Theory. Figure 4.1 shows an overall architectural diagram of Conscious Mattie. The additions to the architecture include consciousness, the emotion module, and Sparse Distributed Memory. Also added, although not illustrated in figure 4.1, are several learning modules and a metacognition module.

4.3.1 Emotion Module

The emotion portion of Conscious Mattie is still in the early design phase. Emotions most likely play a role in two major areas. First, emotions directly affect a behavior stream's activation level. For example, emotions allow Conscious Mattie to be pleased about sending out a seminar announcement on time and to have anxiety about an impending system shutdown. In these cases, emotion might increase a behavior stream's activation level since it is pleasing for Conscious Mattie to complete these streams promptly. Second, emotions influence the suggested actions which are the output of Sparse Distributed Memory, which is described next. Therefore, Conscious Mattie may be more or less inclined to actively pursue a suggested action based on the action's associated emotional level. Lee McCauley is responsible for the agent's emotion mechanisms.



Key:

Solid arrow signifies regular data transfer.

Dotted arrow signifies potential activation of target can occur with data transfer.

Filled circle indicates modules where spotlight can shine.

Figure 4.1: Conscious Mattie's Architecture

4.3.2 Sparse Distributed Memory

Artificial Minds (Franklin, 1995) contains a detailed summary of Pentti Kanerva's Sparse Distributed Random Access Memory (Kanerva, 1988). This content addressable memory serves as a long term episodic memory for Conscious Mattie. Sparse Distributed Memory helps Conscious Mattie in the understanding of incoming email messages. This memory stores the contents of the Perceptual Registers, and it can help fill them with the appropriate information when needed. Sparse Distributed Memory also allows the agent to remember sequences of actions and events. Suggested actions for the agent to pursue, therefore, can be queried from Sparse Distributed Memory. Sparse Distributed Memory can also store the emotional state of Conscious Mattie during a particular action sequence. Since this memory allows for events along with their associated emotions to be stored, their retrieval gives Conscious Mattie capability for a degree of prediction. Figure 4.1 illustrates where Sparse Distributed Memory fits into Conscious Mattie's framework. The Focus serves as the location where Sparse Distributed Memory interacts with Conscious Mattie's other components. Ashraf Anwar is responsible for Conscious Mattie's Sparse Distributed Memory.

4.3.3 Metacognition

Metacognition plays the role of a manager making sure everything is in order. Metacognition monitors Conscious Mattie's internal conditions. Specifically, Metacognition monitors what is in consciousness. In addition, this module monitors the activation of drives, emotional states, parameters in the Behavior Network, and the perceptual module. Using a fuzzy production system, Metacognition then makes inferences on the state of Conscious Mattie. If necessary, Metacognition can influence the Spotlight Controller, perception, learning, and the Behavior Network. For example, Metacognition can change the Behavior Network's activation level threshold to make the agent more goal-oriented or more opportunistic. It can cause voluntary attention by directing the Spotlight Controller to pay attention to certain coalitions of codelets. Zhaohua Zhang is working on the Metacognition portion.

4.3.4 Learning

Learning allows Conscious Mattie to be more closely coupled to her environment. Conscious Mattie contains several types of learning mechanisms. Conscious Mattie can learn new behaviors. For example, the agent may learn a new step in preparing for a system shutdown. She might also learn a new strategy for sending out reminders to seminar organizers. Reinforcement learning techniques are currently being used for this area of learning. The agent can learn new Slipnet message types. This allows Conscious Mattie to learn to better understand incoming email messages. Case based memory is the current method used for perceptual learning. It is also hoped that Conscious Mattie can learn new codelets by basing them on existing codelets. This allows Conscious Mattie to be able to perform the new learned behaviors and perceptual techniques. Coalitions of codelets will also be learned via associative learning. This allows the agent's codelets greater ease in communicating and recruiting other processors to help in performing a task. This learning, based on John Jackson's Pandemonium Theory, will be discussed in greater depth in chapter five (Jackson, 1987). Associative learning also occurs in Sparse Distributed Memory as actions, events, and emotions are associated with one another when placed in this memory. Aregahegn Negatu and Uma Ramamurthy are working on Conscious Mattie's learning.

4.3.5 Consciousness

Conscious Mattie contains a Global Workspace. Based on Baars' Theory, consciousness allows the agent to focus attention on a specific situation. This shining of a spotlight on a coalition of codelets serves numerous purposes, described in chapter two. How Conscious Mattie's spotlight fulfills Baars' functions of consciousness is described in depth in the next chapter. The agent's conscious component consists of three parts. First, a Coalition Manager groups associated codelets into coalitions. Second, a Spotlight Controller shines the spotlight of attention on certain coalitions, placing them into consciousness. The spotlight shines on every incoming message, as these messages are novel to the agent. Third, a Broadcast Manager takes information from the coalition in the spotlight and broadcasts it to all of Conscious Mattie's other codelets. The details of these three components are discussed in chapter five. The author of this work is responsible for Conscious Mattie's Global Workspace design and implementation.

4.4 Additional Enhancements and Conclusions

As discussed in chapter three, in Virtual Mattie's Behavior Network there can be one active behavior in each behavior stream. In Conscious Mattie, there can be only one active behavior throughout the Behavior Network. Discussed further in chapter six, the active behavior corresponds to Baars' dominant goal context.

As seen in the above section, Conscious Mattie has several additions to Virtual Mattie, both in terms of job description and architectural components. The architectural components taken from Virtual Mattie remain mostly intact in Conscious Mattie. Both Conscious Mattie's Behavior Network and perceptual module are enhanced to allow for the agent's learning and additional performed tasks. In addition, the Slipnet is designed in a manner closer to Hofstadter and Mitchell's original Copycat architecture. Each of the architectural additions made for Conscious Mattie allow for the creation of an agent which implements the elements of Global Workspace Theory.

While Conscious Mattie is a significant extension to Virtual Mattie, her environment is still a very limited domain. Chapter six will discuss possibilities for other conscious software agents with more complex environmental and job description domains. The upcoming chapter discusses Conscious Mattie's Global Workspace implementation, including its functions in Conscious Mattie and its architectural components.

Chapter 5

A Global Workspace Implementation

5.1 Introduction

Conscious Mattie's design includes a Global Workspace. While this Global Workspace is based on Baars' theory, it must be remembered that the theory is a high-level cognitive science theory. Therefore, many implementation specific decisions are made by the author of this work in coordination with the Conscious Software Research Group. These decisions can be considered postulations on how human consciousness works. It is up to the cognitive scientists and neuroscientists to assess whether these hypotheses are correct (Franklin, 1997).

Conscious Mattie is the first software agent known to include a Global Workspace based on Baars' Theory in its design. This chapter describes the mid-level design of the agent's Global Workspace. It is written after nine months of research by this author. This chapter first describes the role of consciousness in Conscious Mattie. In conjunction with this, how this Global Workspace implementation fulfills Baars' nine major functions of consciousness is covered. Next, Conscious Mattie's Playing Field, based on the one described in Pandemonium Theory, is discussed (Jackson, 1987). Pandemonium Theory's playing field provides a basis for understanding Conscious Mattie's Global Workspace implementation. The Conscious Mattie architecture significantly extends Pandemonium Theory. Third, the agent's Coalition Manager is addressed. The Coalition Manager groups codelets into coalitions, making them available for attention. Fourth, the Spotlight Controller is covered. The Spotlight Controller shines conscious attention on a coalition of codelets. Finally, the Broadcast Manager is discussed. This module takes information from the coalition and sends it to all of Conscious Mattie's codelets.

5.2 The Global Workspace's Role in Conscious Mattie

Conscious Mattie's Global Workspace gives the agent several performance features. It allows for coalitions of codelets to gain attention. Information about these codelets is broadcast to all of the agent's other codelets. Recipients of this broadcast become active if enough of the information is understood and it is applicable. In this way, the broadcast recipients have the potential to contribute to the solving of the problem raised by the conscious coalition. This broadcast also allows Metacognition a view of the events taking place in the system. Learning also uses the information in consciousness to learn to associate codelets as a coalition.

Consciousness shines on every incoming email message as a new message contains novel information. In this way, consciousness' attention is directly focused on perception. Specifically, consciousness can be said to shine on the Perception Registers as it focuses on the coalitions associated with filling this module's fields. In addition, if there is difficulty understanding an email message, consciousness can shine on the Slipnet coalitions attempting to fill the trial message template in the Perceptual Working Memory. By consciousness shining on the perception portion, it allows information from Sparse Distributed Memory and Tracking Memory to be retrieved if needed to fill in the missing information. It also is used to help foster the sending of an acknowledgment to a seminar organizer stating that the message is not fully understood.

Consciousness also shines on the Working Memory associated with the Behavior Network. At times it may focus on the Behavior Network as well. More specifically, in these cases consciousness shines on the codelets associated with each behavior. The codelets associated with each instantiated behavior often form a coalition. This coalition's information, if it comes to consciousness, can be broadcast to recruit other codelets to help fulfill the coalition's tasks. Consciousness' importance to the Behavior Network can be seen in the following example. Consciousness has just shined on a perceived message from the System's Administrator stating that the system is going to be shut down "soon", but the message does not give a specific time. The information in the Perception Registers and the conscious broadcast of what is being perceived activates the self-preservation behavior stream. A behavior in this stream might be to keep track of how much time is left before the system goes down. This behavior may itself come into consciousness when it cannot easily determine what "soon" is. Its broadcast might entice a codelet associated with Sparse Distributed Memory to look up the time remaining in the last several messages the system administrator sent. While this does not give a definitive answer, it may help Conscious Mattie prioritize the steps she takes in preparing for the shutdown.

Consciousness shines on coalitions of codelets associated with the Behavior Network. Unlike Virtual Mattie's Behavior Network, Conscious Mattie's coalition activation works both in a top-down and bottom-up manner. As in Virtual Mattie, a behavior can activate the lower level codelets associated with it. In Conscious Mattie, codelets associated with each behavior can also activate the higher level behavior. This becomes extremely important with the addition of consciousness. The Behavior Network serves as Conscious Mattie's high-level action selection mechanism. One way for this mechanism to begin occurs with the activation of new codelets associated with the ones currently in consciousness. Since the newly activated codelets send activation to their higher level behaviors, the conscious experience may result in Conscious Mattie pursuing a new behavior stream.

Consciousness shines on Metacognition's codelets. This is one method used by Metacognition to recruit codelets to help gather the information which Metacognition needs. It also enables the changes Metacognition desires for the system to be broadcast to all of the agent's other modules.

5.2.1 Relating Baars' Functions of Consciousness to Conscious Mattie

Chapter two addresses Baars' nine functions of consciousness. Each of these roles is present in Conscious Mattie.

1. *Definition and Context-Setting.* Consciousness shines on the perception module. This allows additional codelets in the Slipnet, Sparse Distributed Memory, and Tracking Memory to potentially help define the input data.

2. *Adaptation and Learning.* Described further below, codelets can learn their associations with other codelets. This leads to adaptation as codelets learn they are part of a new or different behavior. Codelets can then send activation to their new higher level behavior if necessary.

3. *Editing, Flagging, and Debugging.* The Behavior Network, Metacognition, and perception module have associated codelets designed for checks and balances purposes. Their input can come from consciousness. While the design for these codelets is in the early stages, it appears they are Primitive Codelets which can send activation to themselves when detecting a problem. In this way, they give themselves the potential for attention.

4. *Recruiting and Control Function.* By viewing the information broadcast from consciousness, Conscious Mattie's other codelets can come into action if they understand the message and it is applicable.

5. *Prioritizing and Access-Control.* New codelets become active due to what is broadcast in consciousness. These new codelets may join an existing coalition. This

new greater coalition may become learned. Therefore, these new codelets now have greater access to consciousness if the coalition they have joined is one that frequently is conscious.

6. *Decision-making or Executive Function.* Uncertainty is a prime reason why coalitions enter consciousness. Other codelets in the Behavior Network, when viewing what is in consciousness, may activate their higher level behavior streams. An action is then taken if a behavior in this stream executes.

7. *Analogy-forming Function.* Codelets, when receiving a conscious broadcast, search for parameters which enable them to become active. After a certain number of parameters, possibly not all, the codelet becomes active. The codelet may use its built in knowledge, possibly learned, to make analogies about the information not found in the broadcast parameters.

8. *Metacognitive or Self-Monitoring Function.* Described in chapter four, Conscious Mattie contains a separate metacognitive module. This module gets a significant portion of its input from consciousness. It also executes voluntary control of attention by controlling consciousness' Spotlight.

9. *Autoprogramming and Self-maintenance Function.* The Coalition Manager, described below, forms coalitions based on a threshold level of association between codelets. The Spotlight Controller, in determining what enters consciousness, uses a threshold level as well. Both learning and metacognitive modules can influence these threshold levels to control access to consciousness. This control affects the overall stability of the agent as more or less competing actions are successfully executed based on the ability to recruit helpful processors.

5.3 Conscious Mattie's Playing Field

5.3.1 Pandemonium Theory

Artificial Minds (Franklin, 1995) contains a detailed summary of Pandemonium Theory first described by Oliver Selfridge in 1959 for perceptual uses and extended by John Jackson in 1987 (Jackson, 1987). Pandemonium Theory's components interact like people in a sports arena. Both the fans and players are known as demons. Demons can cause external actions, they can act on other internal demons, and they are involved in perception. The vast majority of demons are the audience in the stands. There are a small number of demons on the playing field. These demons are attempting to excite the fans. Audience members respond in varying degrees to these attempts to excite them, with the more excited fans yelling louder. The loudest fan goes down to the playing field and joins the performers. This may cause a performer to return to the stands. The louder fans are those which are most closely linked to the performers. There are initial links in the system. Links are created and strengthed by the time demons spend together on the playing field and by the system's overall motivational level at this time.

5.3.2 Conscious Mattie's Utilization of Pandemonium Theory

Conscious Mattie utilizes Pandemonium Theory's notion of a playing field. In Conscious Mattie, a group of codelets which act as demons are instantiated upon the program's initialization. Each of these demon codelets represent a specific codelet type. If one of these codelet types is to become active, the demon codelet instantiates a copy of itself with the appropriate information. This allows for multiple copies of the same codelet type, each working with different information, to run in parallel. These demon codelets, which receive the broadcast information from coalitions in consciousness, can be considered

Pandemonium's Theory's fans in the arena's seats.

Excluding these demon codelets, all codelets which are performing some function are considered active. These active codelets are the performers on the Playing Field. The Playing Field is a shared space in memory. All active codelets exist in this shared memory space.

Codelets on the Playing Field may be associated with one another. Some of these links are initially determined. Codelets derived from the same higher level module such as a Behavior are more likely associated with one another. Codelets associations also develop by codelets being in the spotlight at the same time.

Codelets contain one or more markers. A marker signifies from where a codelet on the Playing Field is derived, such as a particular behavior. A codelet's marker has two parts. The first part is one of the higher level constructs with which the codelet is associated. The second part is the strength of this association. Since codelets can contain more than one marker, they can be simultaneous members of multiple coalitions. Primitive Codelets, while not directly associated with higher level constructs, do have information in the first part of their markers. This information helps to distinguish the roles of these codelets.

Markers are both static and dynamic in terms of their assignment. Static markers are built into the codelet's initial design or learned. Codelets can be activated from their higher level behaviors or perceptual nodes. These assignments may be dynamic, especially if new behaviors are learned. Also, some codelets may have to activate other codelets on the fly, such as where there is a conflict needing rapid resolution. A codelet's job may be to notice a particular conflict and recruit other codelets to start the process of resolving this conflict.

Codelets on the Playing Field contain an activation level, derived from where they are instantiated: the Behavior Network, Metacognition, or the perception module. Also, in some cases Primitive Codelets activate themselves. Normally, activation decays over time. This activation level becomes an important determinate in deciding which coalitions

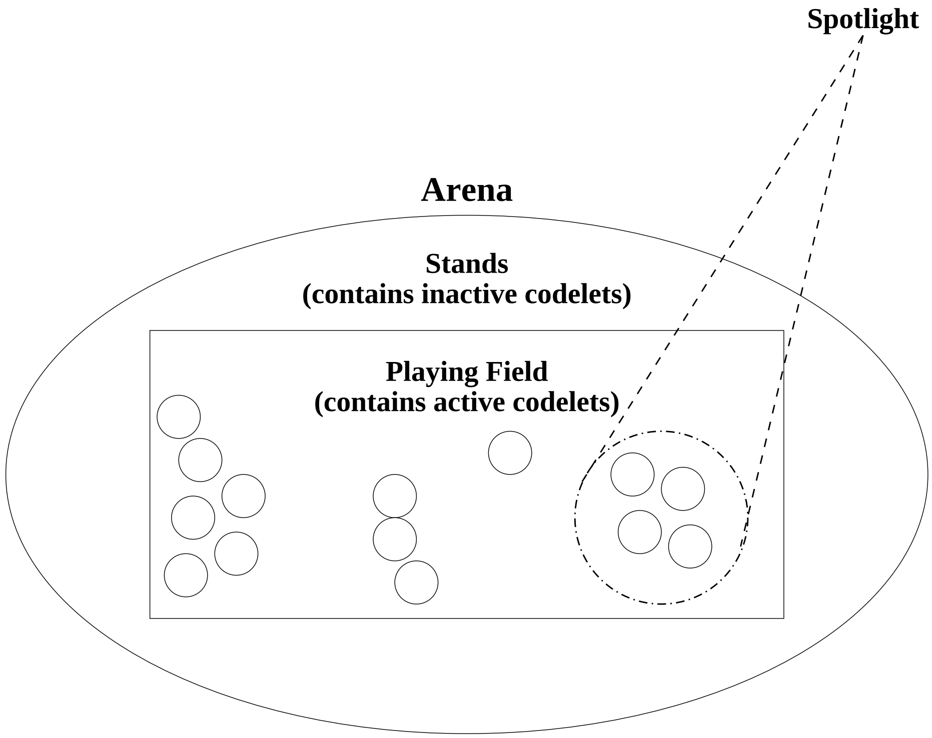


Figure 5.1: Conscious Mattie's Playing Field

gain conscious attention.

Figure 5.1 illustrates Conscious Mattie's Playing Field. Two components of Conscious Mattie's Global Workspace implementation are actors on the Playing Field: the Coalition Manager and Spotlight Controller. Each of these will now be discussed.

5.4 Coalition Manager

The Coalition Manager is the component of Conscious Mattie which groups active codelets into coalitions. It also keeps track of these coalitions. The Coalition Manager's domain is the Playing Field.

To make coalitions, the Coalition Manager groups codelets according to the strength of association between codelets. Only if codelets are associated above a certain threshold level are the codelets considered to be in a coalition. Therefore, all codelets associated with the same higher level construct are not necessarily part of the same coalition.

The Playing Field is an extremely dynamic environment as codelets decay in activation. Codelet associations are also dynamic as a codelet can serve more than one higher level construct, newly activated codelets can join existing coalitions, and codelets can leave one coalition and possibly join another. In addition, the strength of each codelet's association to its higher level construct may potentially change. Due to its dynamic environment, the Coalition Manager must be continually searching the Playing Field with great efficiency in order to keep its record of coalitions up to date. Multi-threading techniques may be necessary in dealing with this large, dynamic environment.

5.5 Spotlight Controller

The Spotlight Controller determines which coalition becomes conscious. The Spotlight Controller calculates the average activation level of each of the Coalition Manager's coalitions. The average activation level of a coalition is calculated by averaging the activation level of the coalition's codelets. The Spotlight shines on the coalition with the highest average activation level. Average activation among a coalition's codelets, not the total activation, is taken to prevent larger coalitions from having an advantage over smaller ones. For the same reasons as for the Coalition Manager, the Spotlight Controller's domain is extremely dynamic.

Specifications for the Spotlight Controller are still being designed, and several different approaches may actually be implemented to determine the method which yields the best results. While currently the highest average activation level is used to determine consciousness, a fuzzy-rule based system may yield better results. There are several factors which could go into this fuzzy-based system besides the activation level of each codelet. One factor is the size of a coalition. It is the Conscious Software Research Group's current hypothesis that the size of a coalition should not be taken into account. When this agent is implemented this hypothesis may prove to be incorrect. Also, factors such as what function the coalitions are serving may need to be taken into account. For example, a coalition serving for self-preservation may in certain circumstances by default take precedence over a coalition working to send an acknowledgment message.

5.6 Broadcast Manager

Once the Spotlight Controller has determined a conscious coalition, it notifies the Broadcast Manager. The Broadcast Manager is responsible for gathering information from this coalition, and sending it to all of Conscious Mattie's codelets. As in Baars' theory, messages are small and not understood by all of the agent's codelets.

Specifically, the Broadcast Manager looks at each codelet in the conscious coalition. It scans for objects labeled for it, such as "INFO_TO_BE_BROADCAST". This object contains the information which needs to be broadcast to the audience of codelets if this codelet becomes conscious. The Broadcast Manager gathers each object of this type from the coalition's codelets.

The Broadcast Manager then broadcasts this information to all of Conscious Mattie's codelets. Currently, this broadcast is to be done via a blackboard model. Information gathered from a coalition is placed in a shared memory space. When each demon codelet polls this blackboard, it searches for parameters it understands. If a demon codelet finds enough relevant information on the blackboard, it will instantiate an active codelet.

Baars' theory does not specify how the conscious information is broadcast. Using a blackboard type model with polling was selected as a first attempt to implement this broadcasting. Other methods were considered and may need to be attempted. Polling appears to be most applicable due to the dynamic nature of codelets. Conscious Mattie contains a large number of codelets. If the Broadcast Manager were to send the information to each codelet, it would have to keep track of the memory addresses of each of the

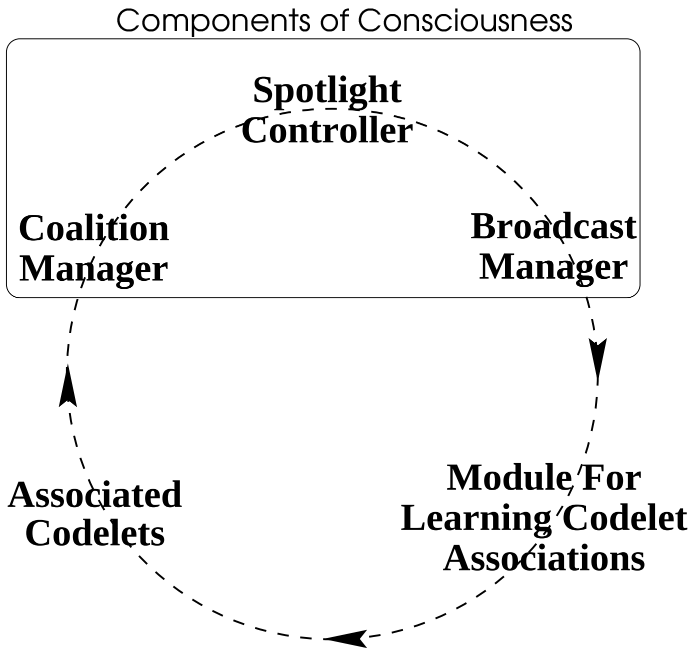


Figure 5.2: Associative Learning of Coalitions in Conscious Mattie

agent's codelets. Also, with the learning of new codelets, it appears to be more efficient for the Broadcast Manager not to have to communicate with the various learning modules and not to have to keep track of each new codelet learned. Having the Broadcast Manager send information to each codelet would most likely be a very time intensive task.

There are many considerations necessary in the implementation-level design of the blackboard. For example, it must be decided how long the information will remain on the blackboard. Another example is the determination of the maximum amount of information which can be present on the blackboard at any time.

The broadcast is critical to implementing Baars' Theory. Whether directly or indirectly, all of Conscious Mattie's modules are affected by the information broadcast. Figure 5.2 illustrates how this is the case in associative learning. A similar model applies to the Behavior Network, Metacognition, and perception.

5.7 Conclusions

This chapter has presented a design for an implementation of "consciousness" in a software agent. In addition to the Playing Field, where active codelets carry out their tasks, this implementation consists of three main components. First, the Coalition Manager forms coalitions from these active coalitions. Second, the Spotlight Controller shines its attentional spotlight on a coalition, bringing it into consciousness. Third, the Broadcast Manager sends the coalition's information to all of the agent's codelets.

The design of Conscious Mattie's Global Workspace fits within the framework of Global Workspace Theory. As the role of consciousness in the agent continues to become clearer, it is likely that the agent's Global Workspace design will be expanded. Creating a Global Workspace as modular as possible for ease in portability is of great concern. As this is the first attempt at a Global Workspace implementation, many of the design decisions will most likely be refined after initial experimentation.

Chapter 6

Contextual Synergy Between Conscious Software and Baars, Conclusions, and Future Directions

6.1 Introduction

This work presents the design for a Global Workspace in a Software Agent. This chapter first discusses a concordance between portions of Conscious Mattie's architecture and Baars' contexts. Next, overall conclusions are presented. Third, the Conscious Software Research Group's future directions are addressed.

6.2 Concordance Between Baars' Contexts and Conscious Mattie

Conscious Mattie is designed under the framework of Baars' Global Workspace Theory. There is a correlation between portions of the agent's modules and Baars' contexts described in *A Cognitive Theory of Consciousness*. Conscious Mattie's Perception Registers serve as Baars' perceptual images. The Slipnet's classification of messages into certain types relates to Baar's perceptual contexts. The Slipnet nodes which represent the specific message types for each of the commonly expected message types correspond to Baars' conceptual contexts. Baars' goal context hierarchy is also represented. This hierarchy moves from overall broad-based goals down to specific goals. Conscious Mattie's drives correspond to broad-based goals and the behaviors each drive activates correspond to the specific goals. The dominant goal context is the current active behavior.

6.3 Conclusions

This thesis gives an overview of Bernard Baars' Global Workspace Theory. It describes Virtual Mattie, the Conscious Software Research Group's first software agent. This paper discusses the overall design for Conscious Mattie, the research group's first "conscious" agent. It presents this author's design for Conscious Mattie's Global Workspace implementation.

Conscious Mattie has the properties of a cognitive agent as described by the Cognitive Agent Architecture and Theory (CAAT) framework. (Franklin, 1997) While structurally tied to her environment, she has drives, is autonomous in pursuing these drives, and has intelligence on how to achieve her goals. Conscious Mattie contains sensors and an action selection mechanism. She also has attention, emotions, learning, memory, and several other modules typically associated with human cognition. She fits within the framework of a well-defined high level cognitive agent.

Many of Conscious Mattie's functions may be able to be implemented more efficiently than what is found in the agent's current design. However, building this agent under the framework of Global Workspace Theory presents a new approach of viewing the functionality of her modules. Conscious Mattie is being developed with the idea of both creating a useful software paradigm and modeling the human cognitive architecture. It is the Conscious Software Research Group's hope that the implementation decisions made by the group will be studied by cognitive scientists and neuroscientists with a focus of determining if these decisions truly do model the human mind.

6.4 Future Directions

It should be stressed that one of the goals in the implementation of Conscious Mattie's Global Workspace is keeping the implementation as modular and portable as possible. This modularity principle applies to the agent's other portions as well. Development of a Conscious Software toolkit will begin approximately in the Fall of 1998. This will allow the putting together of the different Conscious Software modules together with relative ease. Each agent built via this toolkit will have a degree of learning to familiarize itself with its environment.

The Conscious Software model has applications in numerous applications. One possibility is a "conscious" travel agent. This travel agent would search the Internet to find appropriate travel information such as cost comparisons. It would also be able to make travel reservations. The agent would learn the habits of its users, such as frequently accompanying family members and preferred travel times.

Recently, in parallel to the design of Conscious Mattie, the Conscious Software Research Group has begun working on their first proof of concept project, Tutor. Tutor, whose development is supported by the National Science Foundation, will be a Conscious Software Agent which tutors students in computer literacy. A detailed description of Tutor is found at http://www.psyc.memphis.edu/trg/trg.htm (Graesser and the Tutor Research Group, 1997). Tutor will significantly extend all aspects of Conscious Mattie's architecture and functionality. The agent will extract meaning from the keyboard entries of students. Tutor will be able to formulate dialog contributions which are both conversationally and pedagogically appropriate. The agent will incorporate recent cognitive science and software advances. For example, Latent Semantic Analysis will be used to compute the quality, relevance, and truth of student contributions. Latent Semantic Analysis uses a statistical method called singular value decomposition to reduce a large corpus of texts to a one hundred to three hundred dimensional space. Production rules, sensitive to the quality, relevance, and truth of students' contributions, will be used to select Tutor's topics and the dialog moves within each topic. Tutees' inputs are classified into speech acts such as assertions or questions by language modules including a recurrent connectionist network that predicts the next speech act category and surface analysis of the input data. A talking head will be used to help convey information to students. This author is actively involved in the Tutor project, including working on the design for the agent's Global Workspace. It is hoped that Tutor will be "conscious enough" to challenge its students at the correct level.

BIBLIOGRAPHY

Bibliography

Baars, B. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press, New York.

Baars, B. (1997). *In the Theater of Consciousness*. Oxford University Press, New York.

Franklin, S. (1995). *Artifical Minds*. The MIT Press, Cambridge, Massachusetts.

Franklin, S. (1997). Autonomous agents as embodied ai. *Cybernetics and Systems*, 28(6):499–517.

Franklin, S. and Graesser, A. (1997). Is it an agent, or just a program?: A taxonomy for autonomous agents. In *Intelligent Agents III*, pages 21–35. Springer-Verlag.

Franklin, S., Graesser, A., et al. (1996). Virtual mattie — an intelligent clerical agent. *http://www.msci.memphis.edu/˜franklin/vmattie.html*.

Franklin, S. and the Conscious Software Research Group (1997). Virtual mattie — an intelligent clerical agent. Pending Publication.

Graesser, A. and the Tutor Research Group (1997). Nsf proposal for tutoring project. *http://www.psyc.memphis.edu/trg/nsfproposal.htm*.

Hofstadter, D. and Mitchell, M. (1994). The copycat project: A model of mental fluidity and analogy-making. In Holyoak, K. and Barnden, J., editors, *Advances In Connectionist And Neural Computation Theory*, volume 2: Analogical Connections, Norwood, NJ. Ablex.

Jackson, J. V. (1987). Idea for a mind. *SIGGART Newsletter*, 101:23–26.

Kanerva, P. (1988). *Sparse Distributed Memory*. MIT Press, Cambridge, MA.

Maes, P. (1990). How to do the right thing. *Connection Science*.

VITA

Myles Brandon Bogner was born in New York City on April 14, 1974. His schooling prior to high school was in Austin, TX. He attended The Westminster Schools in Atlanta, GA for high school, where he graduated Cum Laude in May, 1992. He entered Rhodes College in Memphis, TN the following August. He graduated Cum Laude from Rhodes in May, 1996 with a Bachelor of Science in Computer Science and a minor in Business Administration.

In August, 1996 Myles entered The University of Memphis, Memphis, TN as a doctorate student in Computer Science. At present Myles is at the University of Memphis. A teaching assistant for Computer Literacy his first year, Myles is currently a research assistant working on a National Science Foundation grant to develop an intelligent tutoring system. He is a member of the University of Memphis' Institute for Intelligent Systems, the Conscious Software Research Group, and the Tutor Research Group.