

# Modeling Social Intelligence Through Attachment-Based Learning

Cameron Marlow, Jonah Peretti

MIT Media Laboratory

20 Ames St.

Cambridge, MA 02139

{cameron, peretti}@media.mit.edu

## I. INTRODUCTION

Augmenting social intelligence is particularly challenging because it requires a nuanced, multidisciplinary understanding of human social dynamics. Unfortunately, building a comprehensive model of human interaction is an “AI complete” problem. Nevertheless, it is possible to develop representations of particular types of interaction that can serve as a theoretical foundation for a broad range of social intelligence systems.

This paper focuses on one class of interactions: social learning. Social activity is central to learning (see Vygotski, 1980). We develop and expand a theory of social learning based on a series of seminars given by professor Marvin Minsky at MIT. Professor Minsky argues that human attachments are essential to learning, a theory he will present in his forthcoming book, *The Emotion Machine*. We position his theory here with the goal of providing social intelligence researchers with a theoretical model on which to base new system designs.

A model of attachment-based learning is particularly germane to this task because it illuminates the social interactions that promote learning. Often models of mind (see Minsky, 1985, Piaget, 1990) focus on the self-organization of the mind, and do not provide a theory of how people (teachers, parents, role-models, etc.) can influence and enhance the mental restructuring of others. Minsky's theory of attachment-based learning begins to address this problem.

## II. IMPRIMING

Minsky bases his theory of learning on the biological concept of imprinting, and refers to the human corollary “imprinting.” Imprinting explains the effect close human relationships have on learning. The basic mechanism of imprinting elevates or censors a learner's goal based on the feedback of a role model or parent. Minsky gives the example of a little girl who spends the afternoon making mud pies. She returns home covered in mud. If her mother scolds her for her dirtiness, the girl will censor her original goal of making mud pies. However, if the mother praises her creativity, the girl will elevate the value of making mud

pies. This dynamic is an example of the imprinting mechanism at work.

Minsky represents this mechanism with Boolean operators in the figure above. This AND-Gate may constitute the most basic mechanism of imprinting. As expressed in *The Emotion Machine*, imprinting can result from only three bits of input. The following Boolean values determine the status of a goal: 1) Praise/No Praise, 2) Censure/No Censure, and 3) Imprinter/No Imprinter. Despite this simplicity, in a close mentoring relationship, simple goal modification can have a far reaching effect on what a learner values.

Although Minsky's theory is an excellent start, we contend that imprinting is usually more complex than this. There are many *types* of praise and censure and each type has drastically different implications depending on the *relationship* a learner has with his imprinter. A learner and imprinter have *mental representations of each other* that influence the way that they interact. Furthermore, praise and censure do not always elevate or censor a goal, but sometimes *transform* or *create* goals. To explain attachment-based learning, we need a theory that models this complexity.

This paper is an attempt to extend Minsky's basic theory to create a more robust foundation for social intelligence research. An adequate explanation requires a whole taxonomy of learning mechanisms. We have identified several heuristics that can explain learning behaviors that are observable in the world. Furthermore, we have developed a system that models some of these heuristics. Our thought experiments and system design suggest that an expanded theory of imprinting can provide an effective explanation of the mechanisms people use to develop social intelligence.

## III. MODELING SOCIAL INTELLIGENCE

Imprinter-learner relationships can be divided into two broad categories. The first is *response imprinting* and the second is *mimicry*. The difference between these two categories is that for response imprinting to be possible the imprinter must provide feedback. In the case of mimicry, the learner ignores or does not receive feedback from the imprinter, and can only copy the imprinter's behavior and goal.

### A. Response Imprinting

Response imprinting requires the imprinter to consciously act as a teacher or mentor.

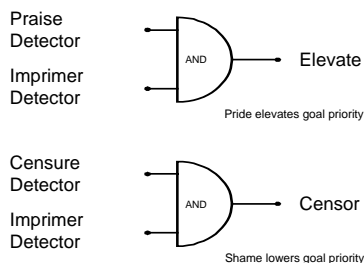


Figure 1: Minsky's model of imprinting

### 1. *Attention Getting Elevators*

The learner is only interested in getting the attention of the imprinter and does not care if the attention is positive or negative, e.g., a child throws his food on the floor to get the attention of his parents.

### 2. *Two way and Multiple Elevators*

The learner receives both praise and censure from his imprinter, causing the learner to refine his goal into several parts, e.g., a coach gives positive and negative feedback to one of his players.

### 3. *Inverted Elevators*

The learner adopts the opposite goals of the imprinter, leading to a form of negative imprinting, e.g., a teenage girl loves her father but rebels by rejecting everything he values.

## **B. Mimicry**

Mimicry works like response imprinting, but the learner does not receive or accept any direct feedback. The learner must copy the imprinter or imagine how he would respond.

### 1. *Basic Mimicry*

The learner copies an imprinter who is not attempting to be a teacher, e.g., a movie star gets a new hair style, causing legions of fans to follow suit.

### 2. *Mimicry by Analogy*

A learner copies an imprinter, but translates a goal into a new domain, e.g., a young painter forms a close bond with an aging Jazz master and decides to make paintings that are "like Jazz music".

### 3. *Reverse Mimicry*

A learner observes an imprinter and adopts the opposite goals, e.g. a girl in school is shunned by her peers, and anything she likes is seen as disgusting.

Our presentation will address these and other imprinting mechanisms. Furthermore, we will address the process through which imprinter-learning relationships change through time. A key element of social intelligence is the ability to negotiate changes in relationships. The transitions between types of imprinting are essential to this dynamic.

## **IV. THE ABLE SYSTEM**

To test our theories of social intelligence, we built the Attachment-Based Learning Environment (ABLE), a system that allows us to experiment with the concepts presented in this paper. The system uses information retrieval techniques to model the people and relations of attachment-based learning. The system is designed to dramatize the fact that people's identity is directly related to the type of mechanism they use for attachment-based learning. This is not surprising but its emphasis is

paramount: the way we learn determines what we become and how we engage in social activity.

The ABLE system represents imprinters and learners through separate corpuses of documents. The movement of documents between multiple imprinters and learners represents social interactions. These interactions provide a model of social intelligence that is grounded in the way individuals learn in social contexts.

We conducted two types of experiments. In the first, the interaction heuristic is varied in order to measure each heuristic's effect on the documents contained in the learner's collection. In this instance, several simple heuristics are used to refine a taxonomy of attachment-based learning mechanisms. This class of experiments illustrated how models of attachment-based learning shape the identity of the learner.

In the second class of experiments, the composition of the document collections were varied in order to measure the differences in how a learner interacts with a community of potential imprinters. In this instance, a complex interaction heuristic was used so that various types of interactions could emerge based on the knowledge base of the community. These types of experiments illustrated how interactions can dynamically change from one type of imprinting to another.

## **V. CONCLUSION**

The importance of human relationships to learning is often underestimated in the cognitive and learning sciences as well as within the AI community. Minsky's theory of social learning is a good first step towards correcting this oversight and we have attempted to extend his theory further by creating a taxonomy of attachment-based learning mechanisms. The ABLE system is an architecture designed to test our theories about imprinting and social intelligence. We believe that a model of imprinting is central to understanding human social intelligence and thus provides a theoretical foundation for the development of systems that support and augment social intelligence.

## **REFERENCES**

1. Vygotsky, L.S. (1980). *Mind in Society: The development of higher psychological processes*. Harvard University Press, Cambridge, MA.
2. Minsky, M. (1985). *The Society of Mind*. Simon and Schuster, New York.
3. Piaget, J (1990). *The Child's Conception of the World*. Littlefield Adams, New York.

Special thanks to Marvin Minsky for helpful comments on an earlier draft of this paper.