**Table of Content**

1. Introduction: What Exactly Is An Agent ………………………..………………………………Page 2

2. Believable Agents And What Makes Them Believable ..…..………….……………………..…Page 5

3. Giving Believable Agents Emotions ……………….……………………...……………………Page 8

3.1 Allowing Agent To Read Emotion. ………………...………………………….………Page 8

3.2 Allowing Agents to Express Emotion..…………………...…………………………..Page 10

4. Conclusion…………………………... ……………………………………………………...…Page 13

5. Bibliography……………………………………………………………………………………Page 14

Believability in Autonomous Agents

Abstract

As the field of artificial intelligence continues to grow, so does the need for intelligent autonomous agents. This paper dives into the importance of having these agents be not only intelligent, but also believable in terms of having emotional depth. Reasons for its importance as well as techniques for its implementation will be explored and explained.

**1. Introduction: What Exactly Is An Agent**

Before diving into any subject, it is always important to understand the basics, as well as give proper introduction as to the whys, whats and most importantly, the hows of the topic. So what exactly is an agent, you may ask. An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators [Russell & Norvig, 2010]. In terms of artificial intelligence, I like to see agents as basically the minds of artificial entities. They are the part that we as humans see as the driving force for the actions of ,whether it be robots or virtual persons.

These are created through the use of various algorithms that facilitate perception and reactions through the use of conditional statements. These statements, as the back-bone of the agent would take the percepts from the environment and cross reference its prior knowledge as for the appropriate response. These are referred to as agent functions, by Russell & Norvig. They ,through the illustrations below, explain how the conditions are structured based on predefined pairs of percept sequences and actions.

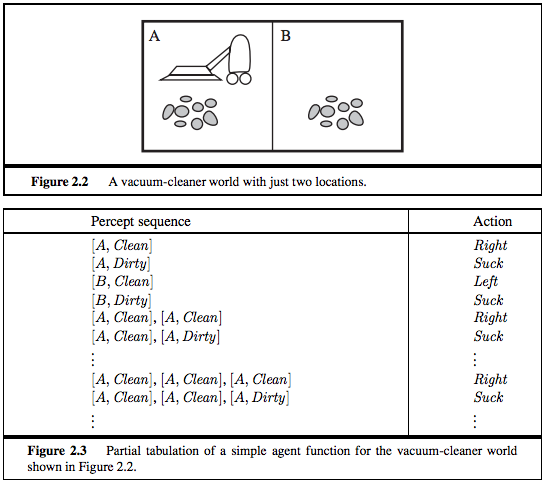


Figure 1. (Figure 2.2 & 2.3 AI A Modern Approach)

The vacuum above has the ability to perceive wether or not the area it occupies is dirty or not, and then the ability to either clean that area through the suck function or move to the next area through the left or right function. The pseudo code below shows how something like the vacuum world agent in figure 1 would be implemented in programming. The actions (suck, left, right) are placed within if statements, which are activated based on what is perceived in the environment.

place A = new place(A,B,C,D);

place B = new place(W,X,Y,Z);

vacuum vacuum = new vacuum();

while(vacuum.isFunctional() == true)

{

if(A.isDirty() == true)

{

vaccum.suck();

}

else

{

vaccum.right();

}

if(B.isDirty()==true)

{

vacuum.suck();

}

else

{

vacuum.left();

}

}

Though rather simple, the above example shows the basic workings of an agent reacting to cues from its environment. But this is only the first step in creating agents. Besides being able to interact with their environment, useful agents need to be able to act autonomously and behave believably.

**2. Believable Agents And What Makes Them Believable**

Believable agents are personality-rich autonomous agents with the powerful properties

of characters from the arts.[Loyall]. Loyall then goes on to explain that these properties are what allow for typical characters to effectively portray a personality similar to that of an organic organism with the same motivations. With this, agents are able to come across as living, breathing entities instead of just cold calculating computers.

This is basically the basis of artificial intelligence, for machines to exhibit the ability to think and be more than just calculators. Humans natively interact through the use of emotions. Facial expressions, voice intonation, body language, these are imperative to our ability to communicate with each other as well as other organism in our surroundings. It is something we have grown accustomed to seeing when interacting with other intelligent beings.

So what exactly makes believable agents believable? Well, in the book “Believable Agents:

Building Interactive Personalities” Loyall, proposes the following properties as necessary for an agent to be considered believable. These are Personality ,Emotion ,Self-Motivation ,Change ,Social Relationships, Consistency and the Illusion of Life. Though personally I see them all as just inter-connected facets of something larger, breaking them up like this allows for a cleaner more understandable approach for explanation. However since the paper is more so about emotion, I will tie them all together in terms of emotion and move forward from there.

**Personality -”A**ll of the particular details—especially details of behavior, thought and emotion that together define the individual.” [Loyall,1997].

As with humans, the personality of an agent would be sort of an indicator for how said agent would react to certain situations. Are they naive and typically approaching possibilities with curiosity instead of caution or are they practical and stick to the known solution to the goal instead of trying something new that might be more efficient. An agent's personality would be the driving factor for what kind of situations bring forth which emotions, for example an ill-tempered agents would more easily exhibit the emotion of anger than a calm one.

**Self-Motivation** - “The power of this requirement in autonomous agents has been repeatedly underscored for me by my experience with people interacting with the test agents the Oz group and I

have created. Repeatedly, people interacting with these agents are surprised and seem to

have a deeper respect and interest in the agents when they realize that many of the things

they do are not in response to stimuli, but rather a product of the agent’s own internal drives

and desires.”[Loyall,1997].

This quotes show how important the presence of self-motivation is for an agent to come across as a living, breathing , intelligent entity. These motivations also aid in helping to show the emotional depth of the agent to other entities observing it.

**Change** – The ability for the agent to “grow” and learn from experiences. This is something all intelligent beings do. They must be able to learn about their environment through their interactions with it and accumulate knowledge for future use. Not only that, but the ability to change, allows for agents to switch emotional states and convey feedback to those interacting with it.

**Social Relationships** - “In nearly all forms of the character-based arts, the characters interact with other characters.These interactions are influenced by whatever relationships the characters have with one another, and those interactions in turn may influence the relationships. The power that these

relationships, when done well, bring to making the character seem alive is widely agreed

upon by artists.”[Loyall,1997].

The same applies to agents. One can not come across as believable if it does not have the ability to form relationships and act accordingly based on the importance of said relationships.

**Consistency of Expression** - “Consistency of expression is one of the basic requirements for believability. Every character or agent has many avenues of expression depending on the medium in which it is expressed, for example an actor has facial expression, body posture, movement, voice intonation, etc..To be believable at every moment all of those avenues of expression must work together to convey the unified message that is appropriate for the personality, feelings, situation,

thinking, etc. of the character. Breaking this consistency, even for a moment, causes the

suspension of disbelief to be lost.”[Loyall,1997].

An agent must have all of its actuators behaving according to what emotion is trying to be conveyed. Most of us as humans would find it very unnerving, if someone exhibited all the signs of sad ness but had a huge smile on their face, well the same goes for agents. Believability cannot be achieved if behaviors out of the realm of normality are displayed.

**The Illusion of Life** – For this requirement an agent must display the existence of their goals as well as the pursuit. It must also be able to perform more than one action at a time. But most importantly, an agent must be able to react and respond to situations as they occur, and do so based on the context of the situation and resources/capabilities currently on hand.

**Emotion** – the ability to express internal mental state based on situations. But it is more than just that, it allows for the crafting of connections between intelligent entities. With the ability to both read and display emotions, an agent can more effectively interact with humans. But how exactly does one go about committing to the task of giving agents emotions and having them presented in a believable manner.

**3. Giving Believable Agents Emotions**

As mentioned in the pervious section, emotions are a necessity for creating believable agents. But to do this, one must first understand how emotions work and cause behavioral changes in intelligent creatures. And to do that the ability to both read and express emotions is important.

3.1 Allowing Agents to Read Emotions

Firstly, reading the emotions. Humans tend to do this by observing the sentic modulations of others around them. Facial expressions, voice intonation, gestures, posture and movement, we as humans can look at each other, see these things and be able to deduce the general emotional state of one another. If agents are to effectively and believably interact with us, they should be able to do the same. However, doing so as humans does, need not be necessary. Humans as organic beings have various internal changes that occur based on varying emotional states. Blood pressure, body temperature, heart rate , these are things that are more easily picked up by the myriad of sensors we have available today. Agents can be equipped with not only sensors, but a knowledge base of results for basic emotions. This knowledge base coupled with algorithms with certainty factors would allow for the agent to deduce what the most likely emotional state is for the person interacting with it. By doing this, the agent would be able to more affectively interact with said person. It would be able to read situations and adapt to be more helpful to the user it is trying to aid. It would also open up the possibilities of fostering relationships between the human and agent. With the ability to read emotions, an agent can go from being seen as less of a piece of software doing a job, and more as an artificial assistant who is genuinely trying to aid you.

Case in point, imagine a typical calendar application on a smart watch. This applications keeps track of all the events you put into it and will remind you when the event nears. The application has no way to tell your emotional state, and as such reminds you of an upcoming event, while in a moment of emotional distress. This does not help the situation in anyway, but imagine if that same app had the ability to monitor the state of your body, and could tell you were in emotional distress. It would more effectively be able to serve you by preventing reminders until you are no longer upset or until it could hold it off no longer or even try to lighten your mood. Though simple, the example mentioned shows how, with the ability to read emotions, agents within machines would be more useful in executing their goals, by being able to tweak actions to suit the current emotional state of the user.

The ability to do this is referred to as affect recognition and is accomplished through pattern synthesis. [Picard,1997]. Picard then goes on to explain various techniques in detail as to how a computer (an agent) would be able to read expressions and emotional responses based on accumulated data on typical patterns associated with various emotional states. In the book, Affective Computing, Picard shows examples of this with computers using motion energy maps for mapping facial expressions, a vocal effects reference based on emotional states and a pitch monitor for the inspecting vocal changes. Though Picard, mentions the use of these techniques in a general sense, I personally see them as techniques more suited for personalization of an agent. Keeping track of personal facial and voice patterns allows for the agent to learn more about the main users while still having fairly accurate deduction for others. As the user and agent interact, the agent will learn to more effectively read the users emotions, which also allow for it to come across as learning and caring (showing the existence of a social relationship).

3.2 Allowing Agents to Express Emotion

But it is not enough to just be able to read emotions, agents also need the ability to provide appropriate feedback once the emotions are read, and to do this they need to be able to demonstrate emotions or emotional behavior. Picard refers to this as emotion synthesis. “If computers are ever to have emotions, then one of the things they need is the ability to synthesize or generate them.”[Picard,1997].

She them goes on to explain the 5 components of an emotional system, emotional behavior, fast primary emotions, cognitively generated emotions, emotional experience and body-mind interactions.

The importance of each component is then based on the purpose of the agent and what it needs to believably emulate.

The OCC theory is the popular choice for bringing forth a believable emotion model to agents. Though not originally intended for use in emotion synthesis, it was very useful in the synthesis of cognitive emotions.[Picard]. The model was built on the idea of “Ai systems must be able to reason about emotions”. The figure below explains how the OCC model groups emotions based on various common situations.

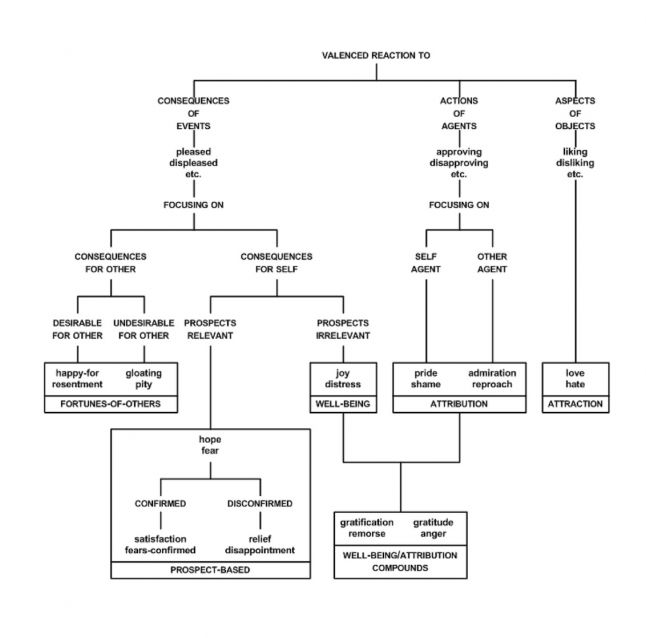
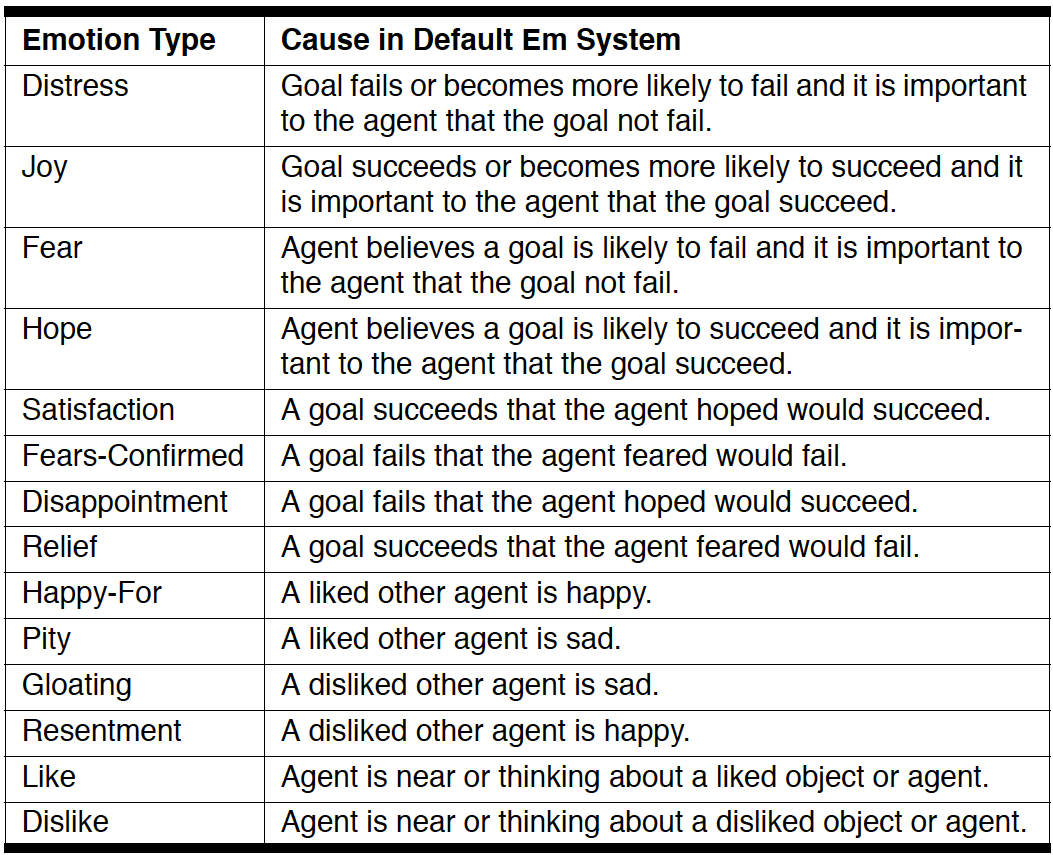


Figure 2 (Figure 7.1 in Affective Computing)

Picard then goes on to provide an example of synthesizing joy, using the OCC model. This is done by setting up conditional statements based on consequence factors for whether the event taking place is beneficial or harmful. This number is then cross-referenced with an intensity factor based on the reality and expectedness of the situation. This then creates the potential for a display of joy within the agent. Similarly a threshold amount can also be added into the mix for allowing different levels of joy, thus creating more emotional depth, and more believability within the agent.

But once in an emotional state, the agent then needs to be able to express this state outwards to other entities as well as have this emotional state effect action/re-action selection. Expression-wise, the agent may have a state monitor providing feedback to the user, but also the actions/re-actions that are executed by the agent can also be a window into its emotional state. These actions would be mapped to emotional parameters, including positivity, certainty, intensity and motivation. By having various responses mapped to one or more of these factors, the agent will be able to categorize actions just as it has emotions(sample seen below).



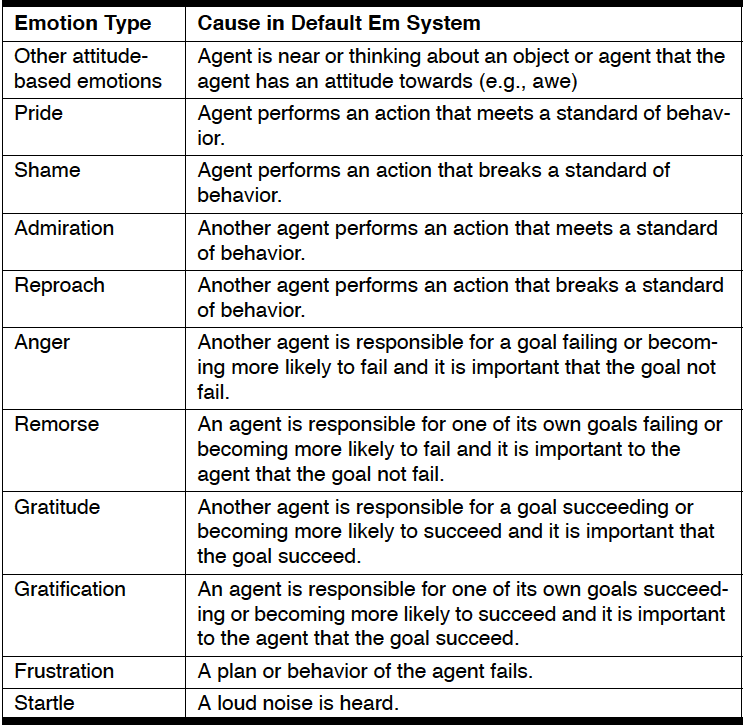


Figure 3 (Table 3.1 from Believable Social and Emotional Agents)

Emotions are fleeting however, so once expressed, they should eventually pass. Reilly as well as Loyall, mention the concept of emotional decay. That emotions should automatically decay and allow for the agent to return to its base state. However, for believability to be maintained, not all emotions can have the same decay rate. Decay rates must be calculated based on agent personality as well as emotional intensity. Special exceptions however, include fear and hope, which according to Reilly should not begin to decay until the source of said fear or hope is removed.

Once the agent is functioning emotionally, we need for it to be able to remember associations between various situations and the past emotional states (ie. The agent needs to remember what kinds of situations cause which emotions and whether or not experiencing them is necessary for the achieving the goal). A knowledge base of events and people with references to the typical emotional outcome due to interactions would be a plausible way to implement this. The agent would be able to “remember” other entities and be able to increase certainty in terms of what might happen next. This helps with the formation of agent relationships with others, thus increasing believability.

4. Conclusion

We looked at agents, perceiving, acting entities within effectible environments trying to accomplish goals. We've seen how and why we would want for agents to be more accessible to the masses through the process of making them more human like and allowing them to interact with us as we natively do with each other. We then peered into how the concepts of emotions provided us with the power to do just that.

Emotion, as a native human language, provides for a more seamless integration of agents into society as it has been built already. But not only that, emotions also provide the means for machines to understand humans and vice-versa. The mentioned techniques allow for an understanding of how it would be possible to have human interacting with machines, how said machines would be able to see, understand and in turn empathize with us as users.

Through this we would be able to not only increase efficiency, but continue to develop intelligent agents to become more and more human in behavior, which as of now is the pinnacle of believability.

Bibliography

[Loyall,1997] A. Bryan Loyall “Believable Agents: Building Interactive Personalities.”, Ph.D. thesis, School of Computer Science, Carnegie Mellon University, 1997

[Picard,1997]Rosalind W. Picard, “Affective Computing”,MIT Press Cambridge, England 1997.

[Russell et. al, 2010] Stuart Russell and Peter Norvig, “Artificial Intelligence A Modern Approach 3rd Edition” Prentice Hall Series in Artificial Intelligence, 2010.

[Reilly, 1996] W.Scott Neal Reilly, “Believable Social and Emotional Agents.”, Ph.D. thesis, School of Computer Science, Carnegie Mellon University., 1996.