Computational Interpretation of Experience : The Foundations of Artificial Intelligence

Max Murphy

10/25/2017

The purpose of my research is to develop a means for computers to interpret experiences in largely the same way, emotionally, that humans do. My question will be what is the best methodology for simulating human experiences. This question can be broken down, further into what makes up the experiences we have, and how to do we store and recall these experiences? In addition, how are these experiences changed by stateful variables such as emotion and other traits of the self?

I became interested in Artificial Intelligence about 2 years ago. At the time I was developing an algorithm to help police departments place their units in more effective locations. Since then I have learned a lot more about AI, data science, and computer science. I have taken 2 computer science classes but I have largely self-taught myself the basic tenants of both AI and data science. I have done extensive research already into AI and this topic in general.

The main reason I'm seeking to conduct research on this specific topic is a project I had in my AP computer science principles class. The task was to encode an experience. The task seemed simple. You would encode variables that could help you tell a computer a very limited experience like a birthday party, with variables such as how many candles are the cake. For me, this wasn't enough. I had already asked myself what makes an intelligent being and how do we translate that into computer science? Within this project I had found my answer. I spent the good part of three days making a flowchart (see Document 1 and 2) that contained the essential elements of every human experience. From here, I had a strong will to conduct further research into the topic.

This field is where Psychology and Computer Science begin to align. There are a few major theories within the world of Psychology that I would like to use to help me accurately simulate different experiences. The first and by far the most important theory that I found, was the decay theory of memory. This theory states that as time goes by our brains filter out less important aspects of our memories. When we recall these memories, rather than just blanking these experiences out, we imagine the variables that were lost. This not only affects our memories, but also our social interactions. We can forgive and forget, so long as we can forget.

The second theory that is fundamental for this topic lies in our emotions. Once again, psychology helps to explain our emotions very well. There are many theories on the relationship between the conscious and the unconscious mind. The unconscious mind has just as much impact on our emotions as our conscious mind. This creates various feedback loops between the two. This can be explained easily with a bad breakup analogy. If you break up with someone, you feel sad. Under the hood, it's likely that your self-confidence will be depressed along with your confidence and security in others. This underlying impact drives future actions long after the sadness of a break up goes away.

The third theory that radically influences this topic is the theory that recall is affected by present emotions and sentiment. This theory says that when you recall memories because you

are imagining the details that were lost as your memory decayed over time. The imagination of the present must fill the gaps. Because the imaginated details didn't objectively occur, the memory can be influenced by sentiment of the present and can alter experiences in dramatic ways. A positive experience can become a negative one, when the experience is reimagined.

This research could lead to artificial general intelligence. This is a step above the artificial intelligence we use everyday, on platforms like Google. Currently, we have a variety of artificial narrow intelligence, such as Apple's Siri, that can perform well defined tasks. However, if we can recreate experiences, which are the basis of our personhood, then we can create artificial general intelligence. This iteration of AI would perform much like humans, who learn from iterative experiences. This kind of intelligence could have vast ramifications for the world at large and can help us further explore ourselves in the process.

I hope to learn more about people, and develop a rudimentary version of AGI. I hope to prove that we are not the only creatures that can actively experience, interpret, and reflect. I think that this research could be incredibly important in defining the framework in which we develop generalized artificial intelligence. I think my added perspective joining our psychology with intelligent machines, could not only bridge the gap between humans and machines, but also further the study of computer science and data science.

My research in this project will be largely experimental. The process of completing these experiments can be found later when I describe the methodology I will use. A large source of design ideas will come from research within the fields of psychology, neuroscience, and computer science. My final product should my experimentation bear results would be software that can take in objective narrative input and release a interpreted time-based output narrative. For example, I could supply this software with a series of photos, and receive out a stream of consciousness or a narrative reflection essay. The project will not be a full-fledge AGI. It will at a very basic level be a piece of software that can have variables of emotion and subconscious variables that are both affecting each other and are being affected by experiences.

I plan on conducting my research through experimenting with different approaches of interpretive analysis within my algorithm. I will first develop a training set of data, a series of object narratives that represent objective moments. Next, I will develop an algorithm that interprets the emotional response and subconscious response to this objective event. Finally, using the emotional response, I will develop a method for outputting the interpretation of the event. I will do this using Google's Tensor Flow library that is developed for performing machine learning operations. The text will be analyzed using a predeveloped NLP library based upon Tensor Flow to cut development time down to 3 months. I will process this using Google's Cloud Platform. This will likely be a description of the event using an emotionally-restricted corpus of words and keywords to write a block of text relating to the experience using an predeveloped NLP library.

I expect to fail constantly while performing this project. The first anticipated failure of mine is probably computing power. This project is extremely computationally intensive and it

will require efficient use of my existing recourses. This project is also very complex, so problems may arise with my oversimplification of concepts for time and complexity's sake. In addition, my experiments in creating models could fail, should the design not be congruent with the data and expectations of my development.

Sources:

Reitman, J. S. (1974). Without surreptitious rehearsal, information in short-term memory decay. *Journal of Verbal Learning and Verbal Behavior*, *13*(4), 365-377.

This finds decay within short-term memory. This study compares different methods of remembering the data and compares the difference in memory decay. Memory decay rates hovered between 33% to 42%.

Posner, M. I., & Keele, S. W. (1967). Decay of visual information from a single letter. *Science*, *158*(3797), 137-139.

This document talks about the theory of decay of nonsignificant memories. They studied what happens when memories are not assigned significance within the mind. This led to significant loss of the memory within each of the subjects involved.

Gorgoraptis, N., Catalao, R. F., Bays, P. M., & Husain, M. (2011). Dynamic updating of working memory resources for visual objects. *Journal of Neuroscience*, *31*(23), 8502-8511.

This source shows that memory that has decayed is replace by imagined details. It also studies the difference in decay rates between sequential representations of narratives and simultaneous representations. As memory load increases, memory decay increases. When items were prioritized the recall of these items was enhanced where

Thomas, D. L., & Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of personality and social psychology*, *59*(2), 291.

Tourangeau, R. (2000). Remembering what happened: Memory errors and survey reports. *The science of self-report: Implications for research and practice*, *29*, 47.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, *12*(4), 361-366.

Jacoby, L. L., Kelley, C. M., & Dywan, J. (1989). Memory attributions. *Varieties of memory and consciousness: Essays in honour of Endel Tulving*, 391-422.

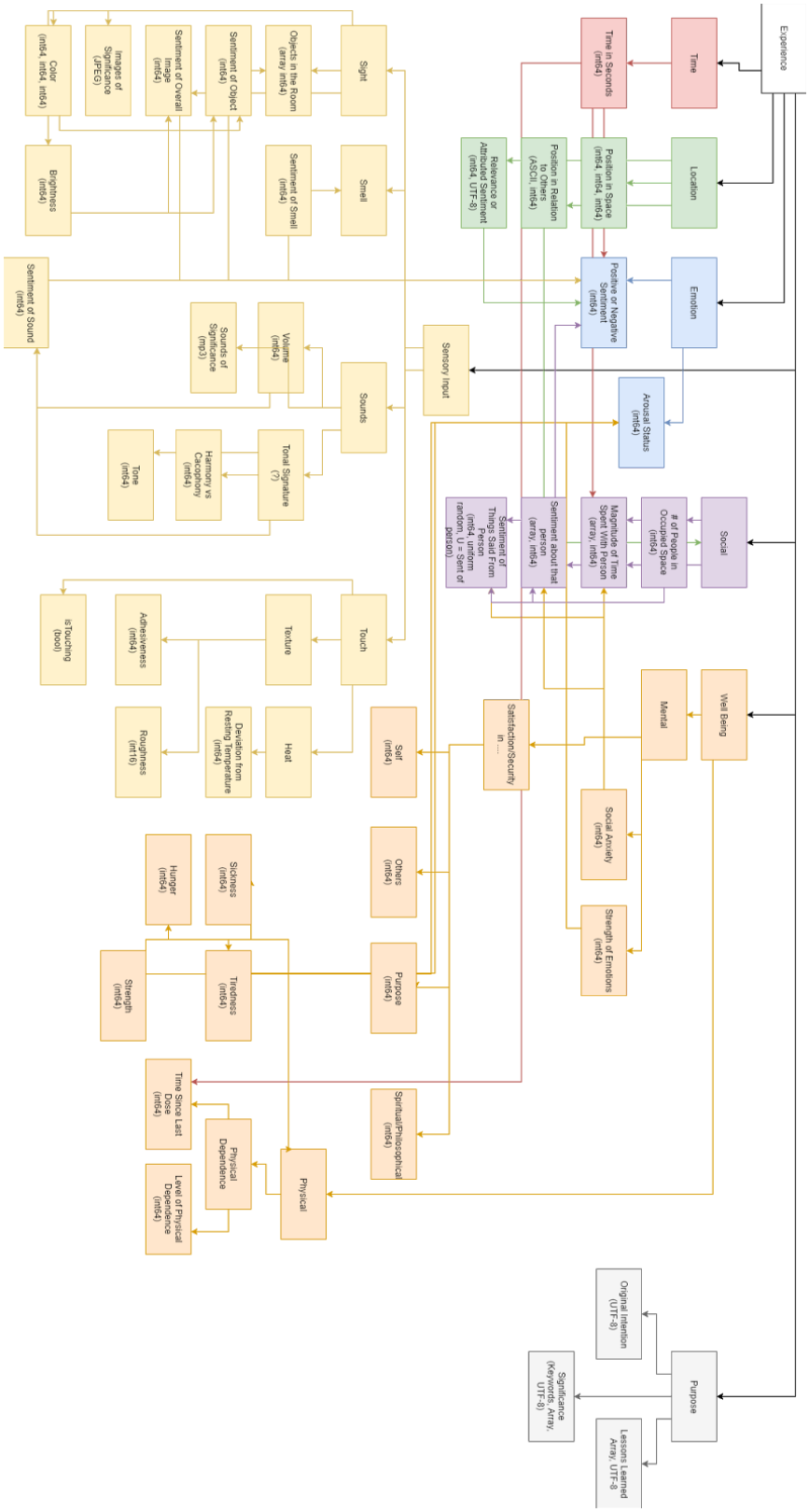
Bateman, J., & Zock, M. (2003). Natural language generation. In *The Oxford handbook of computational linguistics*.

Grishman, R. (1984). Natural language processing. *Journal of the Association for Information Science and Technology*, *35*(5), 291-296.

LeDoux, J. (1998). *The emotional brain: The mysterious underpinnings of emotional life*. Simon and Schuster.

Windmann, S., & Krüger, T. (1998). Subconscious detection of threat as reflected by an enhanced response bias. *Consciousness and Cognition*, *7*(4), 603-633.

Document 1 : My Flowchart for Experiences



Document 2 : My Explanation of the Flowchart

Experience: Humans experience things every waking moment of the day. In order for a general AI to learn and feel like a human would requires computers to quantify experiences.			
Component	**Type**	**Number of bits/range**	**Description / Comments**
Time	General Class	N/A	This is the class containing all methods relating to time in the experience
Time in Seconds	int64	64 bits	The time in seconds as the experience goes on, can be manipulated later to get to exact moments in the experience as this is a dimension element of the experience.
Location	General Class	N/A	This is the class containing all methods related to the location(s) of the experience
Position in Space	Int64, int64, int64	64 bits x3	A 3 dimensional variable relating location within the space being currently occupied (eg. Room, Hall, Street)
Position in Relation to Others	Array UTF-8 int64	64 bits x2	This is an array providing the position of the entity in relation to others within the same space. This is an array with those entities names and distances from said entities.
Relevance or Attributed Sentiment	Int64 UTF-8	64 bits x2	This is a relevance score of the strength of the sentiment of the location being occupied and the other experience that are reminiscent of that place.
Emotion	General Class	N/A	This is the class containing all methods relating to the emotion being experienced at the time of the experience. See figure 1.1
Positive or Negative Sentiment	int64	64 bits	This is an integer value from 9.223372e+18 to -9.223372e+18 that expressed the strength of the sentiment (deviation from 0, further deviation = more strength), and positive or negative elements of the emotion. *Please note this may changing rapidly in the course of a few seconds*
Arousal Status	int64	64 bits	This is an integer value containing the value of how aroused that emotion is.
Sensory Input	General Class	N/A	This is the class containing all sensory input from the experience at a given time
Sight	Sub Class	N/A	This is a subordinate class containing all visual stimuli at a given time
Objects in Space	Array int256	256 bits x len(array)	This is an array containing all the object in the space being occupied in the given time of the experience. *Please note this is largely compressed in humans after the experience and objects with less sentiment are often lost in this compression*
Sentiment of Object	int64	64 bits x len(array)	This is appended to the array of objects in space as the sentiment of each object is identified. This is a positive negative integer that contains the sentiment of each object called and inserted into the objects array. This aids in compression of object array.

Sentiment of Overall Image	int64	64 bits	This is an overall sentiment number that is compiled from the sentiment of each object in the image. *This is sent to the overall sentiment of emotion of the experience*
Images of Significance	JPEG	Variable	This is an image that is taken from the experience that has an abnormal sentiment value at a given time. Often when the experience is first felt there are many of these but as time goes by this is compressed with significant loss within images of less significance. The size of these images is tied to their significance value deviation from 0.
Color	Int64, Int64, Int64	64 bits x3	This is an RGB value of the median color of the image.
Brightness	int64	64 bits	This is a value of the average deviation from 0 of all of the RGB values.
Smell	Sub Class	N/A	This is a subclass containing all of the methods relating to the Smell of the experience at a given time.
Sentiment of Smell	int64	64 bits	This is an opinion of smell. Most people do not remember individual smells just the sentiment revolving around the smell.
Sounds	Sub Class	N/A	This is a subclass containing all of the methods relating to the sound of the experience at a given time.
Volume	int64	64 bits	This is the level of volume of the sounds being experienced at a given time
Sounds of Significance	mp3	Variable	This is a sound taken from the experience that has an abnormal sentiment value at a given time. Often when the experience is first felt there are many of these but as time passes away from the origination date the sound is compressed further and further. Often less important sounds get so compressed they get deleted. The fidelity of the sound is linked to the sounds deviation from 0.
Tonal Signature	Sub |2 Class	N/A	This is a subclass containing a tonal signature of the experience, this can be a substitute for sounds of significance in the case that they are not present.
Harmony vs Cacophony	int64	64 bits	This is a measure of whether the sound was pleasing or not. This is often how sounds are remembered.
Tonal Signature	int64	64 bits	Frequency you hear. (This gets compressed more and more as time continues)
Touch	Sub Class	N/A	This is a subclass containing other subclasses that are aspect of things you touch
isTouching	Boolean	1 Bit	This is a boolean defining whether or not something is being touched in the experience at a given time
Texture	Sub |2 Class	N/A	This is a subclass containing two aspects of a texture one feels in an experience
Adhesiveness	int64	64 bits	This is an integer from zero to 1-2^64 as a scale of the adhesiveness of what is being touched at a given time
Roughness	int64	64 bits	This is an integer from zero to 1-2^64 as a scale of the roughness of what is being touched at a given time

Heat	Sub |2 Class	N/A	This is a subclass containing aspects of heat of an object being touched at a given time
Deviation from Resting Temperature	int64	64 bits	This is an integer that shows the deviation of the self from the resting temperature of the thing that is experiencing at a given time
Well Being	General Class	N/A	This is a class containing all of the details of well being of the self at a given time
Mental	Sub Class	N/A	This is a subclass containing the mental aspects of the well being of the self at a given time
Social Anxiety	int64	64 bits	This is actually a very good measure of introvertedness vs extrovertedness, this is an integer describing the anxiety of the person in social situations. 0 = neutral -1(2^64-1)/2 = Extreme introvertedness (2^64-1)/2 = Extreme extrovertedness
Strength of Emotions	int64	64 bits	The strength the emotions being felt, this affects all sentiment values based on the emotion variables, and can alter sentiment values of certain details of the experience 0 = neutral 2^64-1 = Emotions have taken over all mental processes
Satisfaction/ Security in ...	Sub |2 Class	N/A	This is a subclass containing stability variables that can affect strength of emotions, sentiment values, and social anxiety values
Self	int64	64 bits	This is an integer value containing the value of self satisfaction this is fed by the other three securities and the average of the sentiment values of the previous experiences that recently occurred
Others	int64	64 bits	This is an integer value based on social anxiety and sentiment values from the social general class, this feeds social anxiety variables
Purpose	int64	64 bits	This is an integer of the average sentiment values of the work experienced previously, this is often compressed and placed in this variable as the experiences of working break down. This affects general outlook on life and therefore are the basis of the sentiment variable found in emotion, and can affect reflection upon previous memories.
Spiritual/ Philosophical	int64	64 bits	This is an integer of the security within believed place in the universe, this can be any philosophy. If you are questioning your philosophical beliefs or are acting against believed philosophy this value goes down. This is tied to purpose of experience and believed philosophy.
Physical	Sub Class	N/A	This is a subclass containing the physical aspects of the well being of the self at a given time, it is important to note that this can have the same or greater effects on emotions as the mental well being variables on the emotions
Sickness	int64	64 bits	An integer value of sickness felt at a given time. This is directly linked to sentiment values overall and sentiment values of the entire experience, smell, and sight. These sentiment values are inversely related as the sickness value rises the sentiment values drop.
Hunger	int64	64 bits	This is an integer value of the hunger value of the experience at a given time, this is generally used to redirect the purpose of the next experience and if this value gets to high sickness begins to rise as well. This value is compressed immediately unless it deviates abnormally from 0.

Tiredness	int64	64 bits	This is an integer value of the tiredness of the experience at a given time, this has a large effect on the logic of Natural Language Processing of the being. This raises the social anxiety variables and lowers the strength variables. Arousal status is lowered significantly by this value raising. This vaguely lowers overall sentiment values.
Physical Dependance	Sub |2 Class	N/A	This is subclass of physical dependance of addictive substances. This encompasses all addictive material.
Time since last dose	int64	64 bits	This is a value that increments along with the increment of time. This gets set to 0 at the use of addictive substance. This is multiplied by level of dependence and raise social anxiety, sickness, tiredness, and strength of emotion values
Level of Physical Dependence	int64	64 bits	This is a value that is actually a 1 - 0 scale. This is a decimal point value not an actual integer. The level of physical dependence can rise dependant on the strength of addictive substance. Most start at 0. Complete addiction is 1. This can lower with less use of the substance, and other chemical means. More detail later.
Purpose	General Class	N/A	This is a class containing the purpose of the experience what its significance originates from, and how this affects the being that experienced it.
Original Intent	UTF-8	Variable	This is a description of the original intent of the experience, or *how the experience started*. This can be compressed and manipulated by emotion and well being variables upon recall.
Lessons Learned	Array UTF-8	Variable	This is an array of lessons that were learned, this includes the class of lesson(details later) and what was added. This is added to the knowledge base of the being experiencing the experience.
Significance	Array Keywords UTF-8	Variable	This is just a list of keywords to help index the experiences within the larger being. The list are just things associated with the experience.

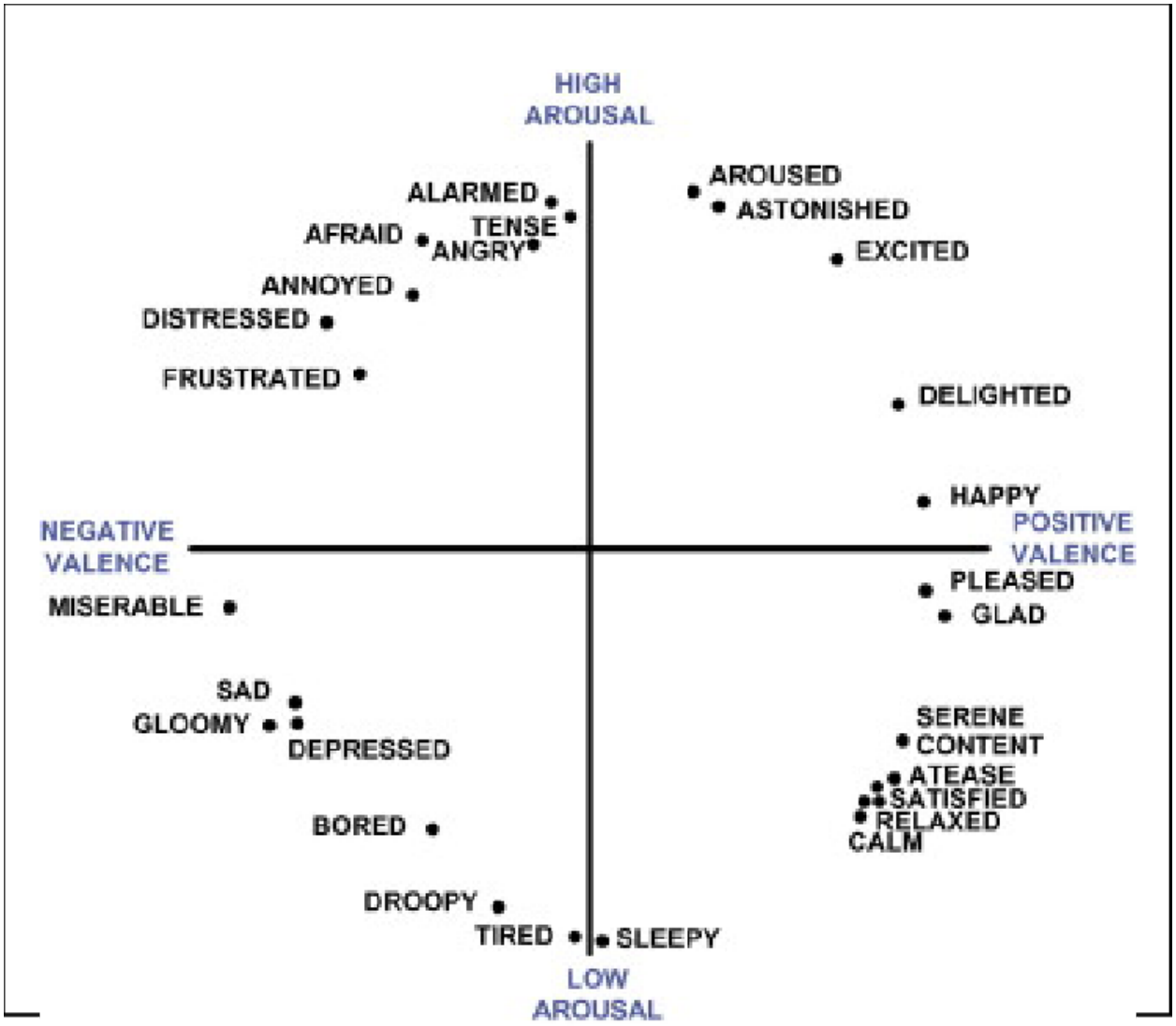

HIGH
AROUSAL
ALARMED
TENSE
AFRAID
ANGRY
ANNOYED
DISTRESSED
FRUSTRATED
AROUSED
ASTONISHED
EXCITED
DELIGHTED
HAPPY
NEGATIVE
VALENCE
POSITIVE
VALENCE
PLEASED
GLAD
MISERABLE
SAD
GLOOMY
DEPRESSED
SERENE
CONTENT
ATEASE
SATISFIED
RELAXED
CALM
BORED
DROOPY
TIRED
SLEEPY
LOW
AROUSAL

Figure 1.1