A Literature Review of Artificial Intelligence

**Introduction:**
In the last few years there has been an influx of software that utilize elements of Artificial Intelligence (AI). Subfields of AI such as Machine Learning, Natural Language Processing, Image Processing, and Data Mining have become important for many of today’s tech giants. Machine Learning is actively being used in Google’s predictive search bar, in the Gmail spam filter, in Netflix’s show suggestions, and in the Cleverbot chat website. Natural Language Processing exists in Apple’s Siri and Google Voice. Image Processing is necessary for Facebook’s facial recognition tagging software and in Google’s self-driving cars. Data Mining has become a “buzz word” in the software industry due to the mass amounts of data being collected everyday. Companies like Facebook and Google collect large amounts of statistics from users every second and need a way to interpret the data they receive. Artificial Intelligence has already proven to be a useful new tool in today’s technology heavy culture.

Almost all of these technologies that have begun to implement facets of AI have only been around for a decade or less. Many of these aspects of AI have proven to be hugely helpful in industry, but these are merely applications of the technologies being researched. AI has greatly advanced in the last few years and there have been countless improvements within each subfield.
Background:
Alan Turing, the “father” of AI, wrote *Computing Machinery and Intelligence* [10] in 1950. He attempted to answer the question “Can machines think?” by developing a type of “Imitation Game” between two subjects. This game is called the Turing Test and it involves written communication between two subjects without being able to see, hear, or otherwise sense the other subject. The first subject, a human, will attempt to figure out if the second subject is a machine or another human simply from written communique. If the first subject cannot tell, or chooses incorrectly, then Turing declared that his Turing Test proved that machines can think.

However, there has been some doubt that just because a computer can respond coherently to a user questions or statement doesn’t mean the computer can actually think. Does the computer really understand the meaning behind the words, or is it simply regurgitating symbols? This was somewhat addressed in Turing’s paper, but more formally covered in a paper 30 years later by John Searle. In 1980, Searle published a thought experiment called *The Chinese Room* [7] that addressed the idea that the machine in the Turing Test is simply throwing symbols together without actually understanding the concepts. *The Chinese Room* uses the analogy of a native English speaker with no knowledge of how to speak, write, or read Chinese who is given a couple of sets of rules in English. These rules correlate input in Chinese to coherent output also in Chinese, even though the “translator” only speaks English. This question has been the topic of various research topics in Machine Learning and Natural Language Processing.
Complications in AI:
There are abundant complications when trying to create an intelligent system. Much of the old or simple AI is a list of conditions for what reaction to have based on expected stimuli. But this is arguably not intelligence, and imitating true intelligence requires an understanding of how the input relates to the output, as well as a large interdisciplinary effort among most AI subfields along with psychology and linguistics [12].

Many complications involve Human-Machine Interaction because of the complexity of human interaction. A lot of the communication that happens between humans cannot be coded facts a machine could simply recite. There are hundreds of subtle ways that humans interact with each other that affect communication. Intonations in voices, body language, responses to various stimuli, emotions, popular culture facts, and slang all affect how two people might communicate. This is hard to model in a machine that does not have a basic common sense model already in place that can learn or make inferences.

Fuzzy Logic, which is modeled after humans’ excellent ability of making approximations without any real values, poses many complications [11]. Computation, by definition, require numbers and not words or concepts.

Complications arise when trying to imitate human intuition or common sense. The amount of background information that is taken for granted by humans is immense and hard to replicate in machines [3].

There is difficulty in trying to imitate human emotion because of how complex and subjective they can be, especially when multiple emotions are expressed [1][12]. When using a Machine Learning approach, the system will process conversations that have been labeled by
humans, but these labels are not always consistent [1].

Image Processing also has complications with recognizing different locations from photos on the Internet because of the variability in images. Modeling the world from Internet photos is difficult because of how much the average Internet photo varies. Generally, image processing requires data to be somewhat consistent, but that obstacle will have to be overcome to render 3D models of popularly photographs locations on Earth [9]. Simply detecting what an image contains is tricky process.

Handling large amounts of inconsistent data is another complication, because inconsistent data is inevitable but difficult to process. But being able to take in a large amount of data and analyze the underlying concepts would be necessary to do something like summarize a novel, which is something that is currently not possible [11].

Ethical concerns arise when building a machine that can be sent into the military that could use lethal force [8]. Although this is a scary concept, it has a high priority for research by the United States government [8].

Finally, using all of the subfields of AI to develop Strong AI (Artificial Intelligence equal or better than human intelligence) is incredibly complicated. Developing a system that has sentient thought would require us to fully understand how the brain and consciousness work, which we do not.

There are a multitude of difficult complications within AI research. AI is a complex field, but much progress has been made in the last few years.
Generalizations of Research:
There is endless exciting new research in the field of AI; this review is far from a global summary of the progress made in the last ten years. There are also scores of subfields within AI, but the only topics covered here will be Machine Learning, Natural Language Processing, Knowledge Management and Fuzzy Logic, Human-Machine Interaction, and Image Processing and Computer Vision. These fields overlap and research in one field is not possible without research in another channel of AI.

Much of the research covered in this review could be applicable to developing Strong AI, but is not direct research for Strong AI. Some of the research indirectly attempts to solve Searle’s *Chinese Room* problem. Creating a machine capable of understanding the concepts behind words is important because it allows for more humanlike conversations as well as improved translation between human languages [5]. Havasi’s ConceptNet research [3] is important to the future of AI because it provides great improvements in Machine Learning, Natural Language Processing, and Fuzzy Logic by developing a type of artificial “common sense”. This graph of linked concepts allows for new information to be inferred. Research into an idea called Computing with Words [11] is exciting because it attempts to imitate humans’ innate ability to make approximations without real values. This has never been possible in machines because, by definition, computing involves crisp numbers not fuzzy approximations. There is also fascinating research into “rediscovering” basic formulas by simply observing physical models [6], research into detecting human emotion through audio and visual cues [1][12], and research into reliably removing unwanted background data from the subject in video surveillance using Machine Learning [4].
Review:

Machine Learning:
Possessing the ability to learn new information or create new relationships between known concepts is a crucial part of Artificial Intelligence. Hard coding an accurate response for every possible stimuli a machine will ever receive is nearly impossible when trying to create a machine that might indistinguishably mimic how humans interact. Because of this, there is a vast amount of research in creating efficient methods for learning from data.

Devillers is using Machine Learning techniques with audio data from call centers to detect complex human emotion [1]. Most conversations studied until this point were from recorded actors attempting to express emotion from a script [12]. The algorithm gained precision as more conversations were processed and labeled, but because of the subjective nature of emotions and inconsistent labeling the algorithm was unable to always report meaningful results [1]. It also could not detect blended emotions.

Machine Learning is also being used by Havasi to infer new information from a graph of general concepts [3]. This graph imitates common sense by linking thousands of concepts together with different types of linking edges. Each link contains information for the accuracy of each relationship [3]. Through a web page connected to this network of concepts, researchers and others on the Internet are able to see inferences being made from the existing data. Questions asked from examining holes in the data, like “What is Paprika used for?”, provide researchers with assurance that the graph is functioning correctly and actually learning. These questions also provide additional information for this artificial common sense mechanism [3].

Aside from research into creating machines with more anthropomorphic characteristics for recognizing emotions and gaining common sense, Machine Learning is being used to “rediscover”
natural laws [6]. Schmidt set up a machine to simply observe physical models such as a swinging pendulum, a double pendulum, and harmonic oscillators. It would capture position data and would analyze the data without any prior knowledge of geometry or physics. The algorithm would take all of the data, derive equations that would match the data at each point, analyze which equations became more frequently derived, and then arrive at known formulas for various physical laws. Finding formulas that were non trivially related and showed significant relationships are difficult to find, and even human researchers can have trouble distinguishing between coincidences and true relation [6]. This algorithm arrived at the equation for a circle, the Law of Conservation of Angular Momentum, and other energy conservation laws and Newtonian force laws [6].

Figure 1: This chart shows the process taken by Schmidt’s Machine Learning algorithm. Step 1 begins with the the physical model being observed and data being collected and Step 6 ends with a reliable formula becoming “discovered” to describe the data accurately.
Natural Language Processing:
Human language and conversation is complex and subjective. The current standard forms of communication with machines involve mice and keyboards, or a specific and basic set of verbal commands. This is dramatically different from how humans interact, simply because the amount of variability in human communication; red in red hair is different from red in red apple.

This fundamental problem of correctly representing concepts with symbols, or words, is greatly hindering the progression of Natural Language Processing [5]. If these challenges are overcome, systems with Natural Language Processing would have the capabilities to express beliefs they have acquired, translate languages at human translator levels, understand the difference between a red apple and red hair, and process commands like “hand me that purple thing down there” into physical action [5].

Current Natural Language Processing techniques can recognize words spoken and even do basic levels of translations. Yet, this is still primitive in terms of fully understanding the complexities of human conversation. Machines are currently unable to summarize novels [11] or perform other tasks that require moderate conceptual knowledge. Devillers’ research into detecting emotions from audio conversations [1] provides another element of understanding the meaning behind words. Havasi’s Open Mind Common Sense (OMCS) Project also provides more assistance to understanding the fundamental concepts that words represent [3]. The issue raised in Searle’s Chinese Room is a valid concern that has yet to be fully solved.

Knowledge Management and Fuzzy Logic:
Havasi’s ConceptNet is a large graph of simple concepts that link to related concepts [3]. It is a part of the OMCS Project which also includes AnalogySpace, a method for actually making sense of
all of the data [3]. This giant graph of data allows for systems using it to quickly learn from new input as well as making inferences with the existing data. This basic background knowledge mimics what humans would call common sense.

Figure 2: This figure shows a small portion of the ConceptNet graph illustrating all of the connections between “cake” and other related concepts.

Humans have an innate ability to make quick approximations without any real numbers or computations [11]. Zadeh shows that a method for performing fuzzy calculations is necessary in order to mimic this ability that humans naturally have [11]. Through computations with imprecise constraints, computing with perception (instead of measurement) becomes possible and allows for fuzzy logic. Fuzzy logic is useful when the available information is not precise enough and when threshold for error is high [11]. Computing with perception, or Computing with Words (CW) as Zadeh calls it, will allow machines to form new and more accurate inferences from natural language with an existing amount of “common sense”.
Figure 3: An example of incorrect assumptions made without fuzzy logic. The statement below each line is the assumption made from the two statements above.

Figure 4: An example of accurate assumptions made with the use of CW, or fuzzy logic, and constraint programming.

**Human-Machine Interaction:**
Most technology literate people, today, are accustomed to the idea that interacting with a computer is just different than interacting with a human. There is a push towards human-centered interfaces which emphasize removing the mechanical feeling inputs from machines and making them more
humanlike [2] [12]. This requires video input to track facial features and emotional cues, video input to track human movements and recognize actions, audio input that can detect emotions and different types of commands, and audio input that can hear and process natural language [12]. Detecting emotion allows for machines to behave in a more anthropomorphic manner because humans will recognize emotion and adjust the interaction accordingly [2]. By analyzing facial expressions, body language, conversation tones, and actual dialog, systems can anticipate human needs [12]. This would also be useful in emotional development research, tutoring, and mental disorders just to name a few [12].

Machines are already being developed for a wide range of autonomous tasks [8]. Some of these machines would be used as soldiers capable of lethal force, or as machines that can physically assist the elderly or infants [8]. Machines that are given the capacity to use lethal force or care for those that need help have a high possibility of making life changing decisions because of a lack of understanding. It is crucial for machines like these to understand the “full picture” and not only respond to simple, basic verbal commands.

*Image Processing and Computer Vision*: Human-centered design is attempting to move away from the current paradigm where a machine simply responds to given commands from a keyboard, mouse, or simple verbal commands. This shift will require an increased ability to process images and perceive information [12]. Most existing image processing software that performs facial recognition utilizes 2D spatial analysis by looking for geometric shapes and edges in the face [12]. But research suggests that the most accurate behavioral judgements of human action comes from analyzing both facial expression and body language [12].
Though not dealing with image processing of humans, Snavely has developed an algorithm for creating 3D models of areas from a series of 2D images [9]. The goal of this research is to release this algorithm “into the wild” to analyze all of the image on the Internet to create a 3D model of Earth. The algorithm is not robust enough to handle the variability of photos on the web or to consistently identify the location of the image it is processing. But if the location is set and only the images returned from a Google Image search of a specific location are used, an accurate 3D model is generated [9]. Giving a system the ability to generate a 3D model of their environment from a series of 2D images will be a large improvement to the current methods for computer vision.

Figure 5: By running all of the images returned by a Google Image search of “Yosemite Half Dome” through Snavely’s algorithm, an accurate 3D model was generated (right). Then, a famous photo by Ansel Adams (left) was analyzed with the 3D model and the model showed its best prediction for where Adams was when he took the photo of Half Dome.
Summary of Findings:
Machine Learning, Natural Language Processing, Knowledge Management, Human-Machine Interaction, and Image Processing are all interrelated and important for creating a machine that might one day be more intelligent and capable than humans. Representing the concepts behind the word is one of the main issues raised in Searle’s *Chinese Room*. The OMCS project with ConceptNet will provide a backbone with which new concepts can easily be learned. Computing with Words, in addition to ConceptNet, will increase the usefulness of the inferences made.

In addition to improvements in how a machine might “think”, there has been research into developing better techniques for detecting human emotion and analyzing images. This progression is important because it will help shift to a more human-centered paradigm, where the machine will anticipate interaction with the human instead of only responding to commands.

Looking Forward:
Artificial Intelligence has come a long way in the last decade, but there is still a large amount of work required to develop Strong AI. Giving a machine Common Sense or intuition is a critical component of allowing a machine to truly learn. Knowing how to convert the input to output appears important, but a machine that truly understands why the output relates to the input is necessary for Strong AI.

It is also necessary to further develop methods for detecting human emotions and actions. This is a multidisciplinary subject and will require advancements in Psychology, Linguistics, Machine Learning, Natural Language Processing, and Image Processing to learn how humans behave, to detect emotions, and to analyze human expressions and body language [12].
Sources:


