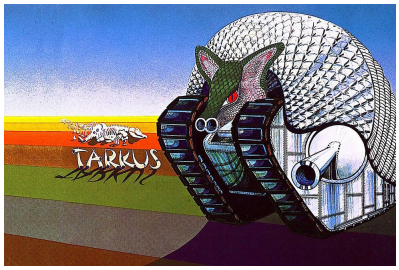
Music Genre Classification
Progressive Rock vs World

Tarkus
Group Members: Michael Cory, Jonathen Settle, Seyedeh Hoda Shajari, Derek Wu, Kasra Yazdani



I. Abstract

With the proliferation of music streaming services, sophisticated methods are required to provide users with personalized music experience. One key step in doing so is music genre classification. For the purpose of this project, we developed an algorithm to classify progressive rock versus all other genres. We utilize Mel-frequency cepstral coefficients (MFCC) via the librosa library in Python. The generated features are then fed into layers of Long Short Term Memory architectures. We then pipeline our developed algorithm by training two different mode ls with two distinct data sets to further refine our accuracy. Our pipeline models are tested on a provided test data set and we get accuracy of ***90%***.

II. Introduction

As music streaming platforms proliferate, new approaches are needed to provide users with ultimate music selections ad-hoc to each user's taste in music. Machine learning algorithms have provided satisfactory results to provide the right music for everyone. Music genre classification is an important task with this regard, and has been well researched for most of the general genres such as pop, jazz, rock, etc. There are a variety of architectures that are used on the GTZAN dataset, 1000 songs from 10 different genres. However, trying to classify progressive rock songs from all of the other genres presents a uniquely challenging task. Progressive rock is noted to be a fusion of styles making it difficult to classify the genre from other genres. In addition, music classification in general requires the extraction and understanding of data from the music, a big part of the music information retrieval domain.

III. Related Work

Music genre classification has been widely studied over the past couple of decades. Many different machine learning techniques have been tried in order to do this. One of these is training statistical pattern recognition classifiers based on features sets representing timbral texture, rhythmic content, and pitch content [1]. Common machine learning classification techniques such as convolutional neural nets [2] as well as logistic regression and random forests [3] have been used. Long short term memory models have also been used successfully [4].

IV. LSTM

Recurrent neural networks are the natural choice for addressing the music genre classification problem due to the time dependent nature of our feature space. A recurrent neural network structure is basically copies of the same network, each passing a message to a successor. This means that they connect previous information to present. Long Short Term Memory networks (LSTMs) are a special type of RNNs that allow for learning long-short-term dependencies of the data which give remarkable result and solve the vanishing gradient problem. Memorizing the training data makes the LSTMs prone to overfitting and as a result regularization techniques such as $L1$, $L2$, and dropout are utilized.

V. Feature Analysis

In this section, we provide an overview of current methods for extracting features from songs and describe our chosen approach.

Music information retrieval (MIR) lies in the intersection of digital signal processing and musicology. An audio signal, e.g., a song, is normally characterized by its frequency, bandwidth, amplitude, etc. One method to illustrate the change in frequencies with respect to time is spectrograms. A spectrogram is a time-frequency visual representation of an audio signal. A spectrogram is produced by taking the fast Fourier transform (FFT) at each timestep of an audio signal. This gives an amplitude-frequency representation of the song for each timestep. At each timestep, we indicate the larger amplitudes by darker corresponding region. A brief representation of the spectrogram is given in Figure (1).

Cepstral analysis treats the peaks of the acquired dB-frequency as the format of the signal which reveal the identity of the signal. The formats are used to smooth out the curve which results in the spectral envelope. This is achieved by

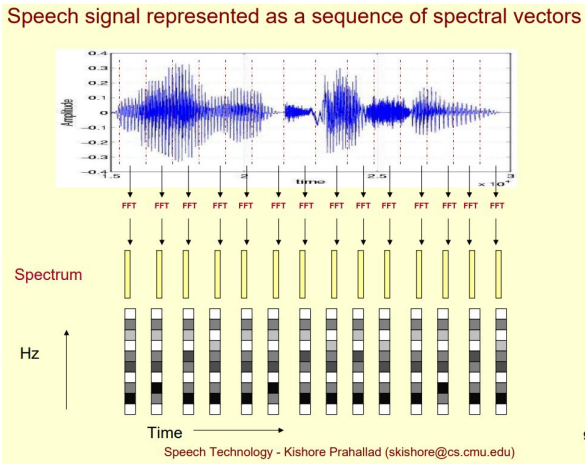


Fig. 1. A brief description of spectrogram generation

taking the FFT of the log of the spectrum. Based on perceptual experiments, however, human ear is more sensitive to regions of the spectral envelope rather than the entire of it. Mel-Frequency analysis is based on the human perception and sensitivity to certain range of frequencies. This is done by feeding in the signal through non-uniformly distributed filters with more emphasis, i.e., more filters, in the lower range of the frequencies. The Cepstral coefficients obtained with this method are denoted by MFCC which is short for Mel-Frequency Cepstral coefficients. Mel-Frequency analysis is particularly practical in song genre classification due to the fact that it closely approximates the human ear's response to audio signal.

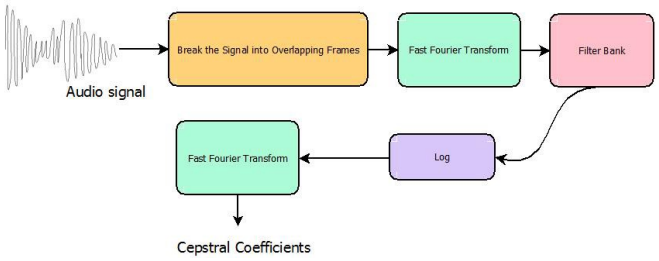


Fig. 2. MFCC Generation

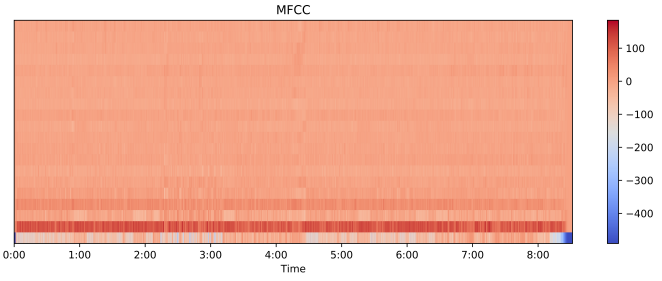


Fig. 3. MFCC for the song Kashmir by Led Zeppelin

VI. Model Architecture

In this section, we explain the process of feature generation, voting system and pipeline and results.

A. Feature generation

We used the feature.mfcc function in the Librosa [5] library to extract 20 MFCC coefficients with a sampling rate of 22050. We also normalized the MFCCs of each song to be centered and have unit variance. The number of frames generated for each song depends on the length of them. For example, the song 'Kashmir' by Led Zppelin has 2028 MFCC frames for a song with length 8 minutes and 31 seconds and 'Kryptonite' by Three Doors Down has 10117 MFCC frames with length of 3 minutes and 54 seconds.

B. LSTM Architecture

The main model consists of three layers of LSTM with 64, 32 and 16 units followed by a Time Distributed fully connected (dense) layer. The output of this layer was then flattened and fed as input to the final fully connected layer with a softmax activation function. We also used binary cross entropy as loss function.

To cover a reasonable and diverse portion of songs for training, we used a batch generator to choose songs and segments consisting of 100 consecutive frames from them randomly from training set as well as validation set. In order to make sure in each pass, the model receives balanced data from both classes, half of each batch was sampled from progressive songs and the other half was sampled from non-progressive songs. We refer to this model as Model A.

In order to validate our model while training, we used another batch generator to generate 50 batches from validation set at the end of each epoch and test the model on them and obtain an average loss and accuracy. Validation set was not used in the process of training for parameter tuning or cross-validation so it was reasonable to test the the performance of trained model on them.

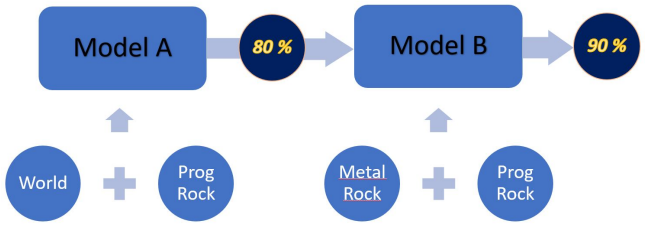


Fig. 4. Diagram of classification pipeline. Percentages refer to test set classification accuracy

C. Voting

For obtaining the final label of each song, MFCC frames were divided into segments of 100 MFCC frames and were fed to Model A and the labels for each segment were added up. If more than half of votes was favoring a genre, the corresponding label was given as the final label.

TABLE 1
Confusion matrix of test set by Model A

	Non-Prog	Prog
Non-Prog	85	16
Prog	24	77

TABLE 2
Confusion matrix of test set after improvement by Model B

	Non-Prog	Prog
Non-Prog	89	12
Prog	8	93

D. Pipeline

Model A was successfully classifying the progressive rock songs and we observed that most of the songs in non-progressive category which were misclassified as progressive rock were from metal and rock genres. Therefore we created a data set consisting of metal and rock songs and trained a model for classifying rock and metal songs versus progressive rock songs in order to capture subtle differences between these genres. Experimental results suggest using both L1 and L2 regularizer in the LSTM layers helped control learning rate and prevent overfitting. We refer to this model as Model B. Model B was pipelined with model A to receive the misclassified songs from model A and improve classification accuracy.

E. Results

After training Model A, we obtained 89% on training set and 86% accuracy on validation set. The only progressive song in the training set which was misclassified as a non-progressive song was 'Sailor's Tale' by King Crimson. We then fed the final test set to Model A and got 80% accuracy. 16 non-progressive songs were misclassified as progressive songs including 'I Am a God' by Kanye West and 'By Myself' by Linkin Park. 24 progressive songs were also misclassified as non-progressive. After feeding the misclassified songs to Model B, accuracy increased by 10% to 90% and most of the progressive rock songs which were misclassifed by Model A, got correctly classified as progressive rock by Model B.

Table 3 shows the progressive rock song which were misclassified as non-progressive by model A and voting. First column shows the number of segments voting for the song to be a non-prog song and the second columns shows the total number of segments of the misclassified song. For example, the song 'Hallogallo' had 54 out of 261 votes for progressive rock and therefore it was classified as non-progressive. However, Model B increased the number of votes to 198 and it got classified correctly.

Table 4 shows the non-progressive songs which got misclassified as progressive by Model A. As we can see from Table 4, the votes for song 'By my self' by Linkin Park, increased from 36 to 74 out of 81 total segments which was an increase of 44% to 91% of segments classified as non-progressive.

The djent data set consisted of fourteen bonus songs that had no predetermined label. After passing the djent songs through our pipeline models we saw the following results. After Model A, there were six songs that were classified as progressive rock. However, after running the same songs through Model B, we saw that only three songs were labeled as progressive rock: "Let Yourself be Huge" by Cloudkicker, "Physical Education", and "Bad Code" by Chimp Spanner. These three songs were also classified as progressive rock after Model A.

VII. Things we tried and didn't work

In this section, we describe a number of approaches that we initially implemented in the course of completion of our project. However, these models did not provide comparable results to the LSTM.

A. CNN - AlexNets

Convolutional Neural Networks (CNN) are a popular method for identifying objects in images. The most popular architectures for image classification are AlexNets, GoogLeNet, and ResNet50. For starters, we extracted the Mel-spectrogram of 10-second intervals from each song. We incorporated AlexNets' architecture as our feature detector where the inputs to the network were the obtained Mel-spectograms treated as images. We employed k-fold cross validation and a softmax layer as activation function in last layer for classification. The test accuracy obtained from this implementation was 54% while increasing the number of epochs did not improve the test-accuracy.

The basic structure of AlexNets does not incorporate the time-series nature of the songs and that is potentially the main reason why we were not able to obtain better accuracy from our implementation. We also implemented dropout, $L1$, and $L2$ regularization to prevent overfitting. However, this did not improve the accuracy we were getting on the test set. We followed our experiments by reducing the depth of the AlexNets structure as we assumed that the complexity of the architecture was the reason the model was overfitting. Moving on, by rounds of trial and error between reducing the depth of the AlexNets as well as using regularization methods, our test accuracy in this setting was not improved above 64 % and as a result we decided to include LSTM layers in our model to take advantage of the time-series nature of the song signals.

B. CRNN

Given the trouble we had with AlexNets, we thought that it would be advantageous to use the features learned by a CNN with a LSTM model that can learn time-series data. In this architecture, which we call CRNN here, the output of a CNN is sent to a series of LSTMs whose output is then classified as prog or non-prog.

Initially, the architecture was difficult to train. The architecture was trained using the validation songs provided as a validation set, and during training the validation loss would oscillate wildly instead of converging to a minimum value. After some trial and error and using grid search to tune the hyperparameters, we found that increasing the dropout reduced the training volatility that we were witnessing. However, the model now learned much slower and only achieved 65% accuracy on the validation set. More LSTM layers were added after

Prog Songs Classified as Non Prog					
Song Name	**# Prog Segments delivered by Model A**	**Total Segments**	**% of Prog**	**# Prog Segments delivered by Model B**	**% of Prog**
Oblivion Khaos	66	148	45%	20	14%
Happy Nightmare (Mescaline)	43	103	42%	33	32%
Hallogallo	54	261	21%	198	76%
Dark Light - Sinkin _ Deep	60	176	34%	79	45%
A Louse is Not a Home	157	323	49%	240	74%
Captain Beefheart - Trout Mask replica - hair pie - bask	63	128	49%	111	87%
Toxicological Whispering	96	202	48%	157	78%
Jurassic Shift	103	286	36%	176	62%
Halleluhwah	149	479	31%	180	38%
Suite Sister Mary	127	275	46%	199	72%
On that note	117	236	50%	198	84%
Assault Battery Part I	55	145	38%	49	34%
Song For America	113	259	44%	144	56%
Mekanik Kommandoh	41	107	38%	56	52%
Kraftwerk - Autobahn	132	588	22%	347	59%
Beyond the Pale	101	259	39%	74	29%
Sea Song- Robert Wyatt	80	168	48%	77	46%
The Main Monkey Business	65	155	42%	58	37%
Turn of the Century	67	167	40%	148	89%
Ommadawn (part One) - Mike Oldfield	208	501	42%	408	81%
Peace to the Mountain	76	169	45%	130	77%
Birthright	56	157	36%	103	66%
Part VI	84	193	44%	138	72%
In The Land of Grey and Pink	50	129	39%	87	67%

TABLE 3
The list of the progressive rock songs from the provided test set which our first layer (Model A) in the pipeline misclassified. The table provides the total number of segments in generation of MFCCs by librosa. The total number of non progressive rock segments are provided and we compare the two by providing percentage values. In general, our first layer in the pipeline only misclassifies the song genres which are very close to the progressive rock, e.g., metal and progressive metal and this implies the necessity of further algorithms to be developed. The green cells indicate the songs that our secondary step (Model B) in the pipeline *correctly* classifies. As illustrated above 15 songs out of the 24 incorrectly classified songs are corrected in the classification which increases our accuracy on the test set by 10%.

the CNN output and a TimeDistributed layer was added after the LSTM layers to improve performance. The architecture plateaued around 75% accuracy on the validation set. By this time, the LSTM pipeline using MFCCs was already achieving a higher performance, so this architecture was abandoned.

C. Feed forward

Feed forward networks are simple neural networks that transforms the input data into a new space for classification. To utilize the feed forward network, we extracted the following features for each song: 20 MFCC coefficients, spectral rolloff, spectral bandwidth, and spectral centroid. Each of the features were averaged over the whole song to produce a single value for each feature.

The architecture consisted of four dense layers with 256, 128, 64, and 10 units respectively. After compilation with the Adam optimizer and using categorical loss, the model was trained and tested. We were able to achieve a 90% training accuracy and 77% testing accuracy. This strongly indicated that the model was overfitting. In addition, the averaging of the features meant that we were losing a large portion of the data from each song and that the averaging could have been skewed by the song's composition. Consequently, we looked to explore other models that were able to account for temporal relation.

VIII. Conclusion

In this project, we developed, implemented, and tested an algorithm for classifying the music genre progressive rock

Non Prog Songs Classified as Prog					
Song Name	**# Non-Prog Segments by Model A**	**Total Segments**	**% of NonProg**	**# Non-Prog Segments delivered by Model B**	**% of Non-Prog**
Strawberry Fields Forever - The Beatles	50	106	47%	42	40%
Neodammerung	50	154	32%	49	32%
Three Ragas in _ D _ Minor	111	300	37%	12	4%
By Myself	36	81	44%	74	91%
Extreme Ways (Bourne Ultimatum)	42	112	38%	60	54%
Miles_Runs the voodoo down 328 lane CBR	157	363	43%	40	11%
Dear Mr. Fantasy	67	146	46%	51	35%
I	24	53	45%	32	60%
Spectre - Radiohead	38	85	45%	23	27%
The Temple on the Edge of Time	73	151	48%	14	9%
Signals	98	295	33%	23	8%
Oneohtrix_point_never-computer vision kouala	19	62	31%	24	39%
When the Levee Breaks	54	185	29%	89	48%
John Williams - The Phantom Menace-0 - Duel of the Fates	39	109	36%	44	40%
I am a God (Feat. God)	43	99	43%	67	68%
The American Metaphysical Circus	47	127	37%	11	9%

TABLE 4
The list of the non progressive rock songs which are classified as progressive rock. As we can see, our developed algorithm, in the first layer of the pipeline (Model A), misclassifies 16 songs out of 100 progressive rock songs and 100 non progressive rock songs. Subsequently, Model B filters out 5 further songs which increases our accuracy. The green cells in the table above indicate

versus all other genres. The algorithm was developed using MFCC as features and we used LSTMs to train our network. Initial results showed that training a model specifically on progressive rock vs the world was very broad and helped obtain an 80% accuracy on the provided test set. Pipelining this model with another LSTM model that was more specifically trained to distinguish between progressive rock and rock/metal, the two genres that most resembled progressive rock, helped boost our accuracy to 90%. Future work will include building a multi-label genre classifier.

REFERENCES

[1] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *EEE Transactions on speech and audio processing*, vol. 10, pp. 293–302, 2002.

[2] I. S. Alex Krizhevsky and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[3] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[4] Y. K. Y. Z. Z. K. H. W. Chun Pui Tang, Ka Long Chui, "Music genre classification using a hierarchical long short term memory (lstm) model," 2018. [Online]. Available: https://doi.org/10.1117/12.2501763

[5] B. McFee, C. Raffel, D. Liang, D. P. Ellis, M. McVicar, E. Battenberg, and O. Nieto, "librosa: Audio and music signal analysis in python," 2015.

Supervenience

Supervenience is a formal definition of the dependence of two sets of facts on each other. We usually designate one set to be of a "high-level facts" or "B-facts" and the other set to be "low-level facts" or "A-facts", and so supervenience can help us see the world as layers of understanding in which the lower layers determine the upper layers; e.g. the laws of physics determine biological behavior (whether or not we enumerate the ways in which this determination happens is less important than establishing that such an enumeration exists). Chalmers defines supervenience between B and A facts in the following manner:

> "B-props supervene on A-properties if no two possible situations are identical with respect to A-properties while differing in B-properties". (Chalmers, p. 34)

There are two axes upon which this definition can vary: the "situation" can be narrow in scope or encompass an entire world and "possible" can refer to conceivable or expressible within the possibility of our world. (Chalmers, p. 34)

The first variation, that of the "situation", allows us to distinguish between *local* and *global* supervenience. If the A-properties of an *individual* determine the B-properties of that *individual*, we say that A-properties supervene locally on B-properties. We note that local supervenience can fail if there is something about an individual besides its A-properties that determine its B-properties. For example, the physical detail of a painting does not fully determine its value because its value also depends on who painted it, if the painting is a forgery or not, the relative importance of the painting with its creators' oeuvre, and more. On the other hand, if the A-properties of the *entire world* determine the B-properties of that *world*, we say that A-properties supervene globally on B-properties. Although maybe initially counterintuitive, it is apparent that local supervenience implies global supervenience. This can be seen with a simple example, going back to biology: it is clear that the biological facts of our world supervene on the physical facts of our world, but two physically identical individuals in two different climates can have a different fitnes, so biological facts are not locally supervenient on physical facts.

The nature of how we develop different "possible" situations is the second variation in our definition. One way to think about "possible" is to consider conceptual situations that could

be created by some all-powerful being. These situations are clearly not constrained by the laws of our world, but we could constrain them in such a way! It is these two notions that, when applied, give us *logical* supervenience and *natural* supervenience. Logical supervenience is a much stronger statement than natural supervenience. With logical supervenience, we get B-facts for free once the A-facts are established, but natural supervenience requires some additional law or correlation to get to the B-facts from the A-facts.

In considerations of consciousness, the first variation distinguishes between context-dependent conscious experiences (global supervenience) and context-independent conscious experiences (local supervenience). We will revisit this later. The second variation among possibility constraints leads us to solely consider logical supervenience of consciousness on the physical, since the connection between physical structure and conscious experience is entailed from the laws of the actual world.

Consciousness

Before we address the question of consciousness being logically supervenient on the physical, we need to define the scope of consciousness in question. In his work *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*, Peter Godfrey-Smith makes the case for octopi and other cephalopods having a limited form of consciousness called sentience (and perhaps more developed forms!). This limited consciousness is exemplified by changing skin color for camouflage, protecting injured areas of the body, squirting water at researchers, etc.; more succinctly, it is the ability to respond to stimuli in the environment. (Godfrey-Smith) If we were to just consider sentience as consciousness, we could plausibly provide a reductive explanation of sentience in terms of lower level neural processes, and thus, as Chalmers lays out, establish that sentience logically supervenes on the physical. (Chalmers, pp. 48-49) However, what about consciousness as we experience it? Does the happiness we feel from a major key stem from purely physical realities? Does the enjoyment of the latest Spielberg film? It is this *phenomenal* consciousness that we claim *does not logically supervene* on the physical, i.e. there are logically coherent individuals for whom all the physical facts about us remain the same, but, unlike us, they do not possess consciousness experience.

However, first we should briefly examine a perhaps more fundamental question: does phenomenal consciousness exist at all? Surely we have a conception of ourselves and how we process the world (self-introspection), but is this the result of some higher-level functionality or a product of some computational device? Addressing these concerns of eliminativism (Ramsey) is beyond the scope of this discussion, and we simply take Chalmer's stance that eliminativism is "an unreasonable position only because of our own acquaintance with [consciousness experience]". (Chalmers, p. 103)

Position #1 Consciousness does not Logically Supervene on the Physical:

Let us consider a being that has the same microphysical makeup as one of us (let's just say "me" here), molecule for molecule. In his book, *The Conscious Mind: In Search of a Fundamental Theory*, David Chalmers raises the point that me and my twin could have a conscious experience of color that is inverted. While me and my twin may both have "red" experiences, Chalmers notes "what matters is that the experience he [the twin] has of the things we both call "red" […] is the same kind as the experience I have of the things we both call "blue" […]". (Chalmers, p. 101) While the idea of identical physical beings having completely opposite conscious experiences of color is certainly hard to wrap your mind around, we are only interested in the logical coherency of it. Surely, red light activates the same opsins and neurological pathways in both me and my twin, but this physical signal transmission is not necessarily accompanied by just one kind of experience. I could experience warmth from "red" while my twin experiences coldness. This "Inverted Spectrum" argument demonstrates that consciousness experiences differs among logically possible individuals with the same physical facts. Thus, we have illustrated our claim that phenomenal consciousness does not logically supervene on the physical.

Furthermore, Jean Piaget's work on developmental psychology shows that the development of consciousness relies on more than just the physical development of an individual. (Piaget) Piaget postulated that mental schemas, basic building blocks for a mental representation of the world, are developed through adaptation to the world. Within this framework, a physical being develops their conscious representation of themselves and the world

around them through their lived experiences. In one experiment, Piaget had children crawl on all fours and illustrate what they had done with a teddy bear. (Gardner) This demonstrates a growing ability to reflect on one's actions and articulate them to another individual, whether with language or with some tool. Building on experiments like this, Piaget created a model of consciousness with stages that culminate in being able to offer a complete account of one's actions. Unfortunately, Piaget leaves out many considerations about human experience regarding the arts and other, more subtle aspects of consciousness are not addressed. However, his framework of developing consciousness is still useful to us: different conscious experiences can develop in physically identical individuals if their lived experiences differ. We come to the conclusion that the physical facts of an individual do not entail facts about their consciousness, so local supervenience is not a valid position.

Position #2 Consciousness Partially Logically Supervenes on the Physical

The idea that consciousness could partially logically supervene on the physical implies that to a certain extent the microphysical makeup of the being in question affects the possibility of having a consciousness. Though Chalmers argues that consciousness does not logically supervene on the physical, he does note that "it seems logically possible, at least to many, that a creature physically identical to a conscious creature might have no conscious experiences at all, or that it might have conscious experiences of a different kind." (Chalmers, p. 39) Going back to our previous argument, my "Inverted Spectrum" twin has different conscious experiences than me, but it is possible that the ability to have phenomenal conscious experiences is indeed dependent on the physical facts about us. To this extent, we could say consciousness partially logically supervenes on the physical

We can also approach this argument from another direction. One theory posited to explain and describe consciousness is the Higher-Order theory. Higher-Order theory defines a mental state as a conscious mental state "if is accompanied by a simultaneous and non-inferential higher-order (i.e., meta-mental) state [...]". (Van Gulick) For example, having a conscious desire for chocolate involves having two mental states: one must have both a desire for some chocolate and also a higher-order state whose content is that one is now having just such a desire. One

could argue that even if an organism does have the microphysical makeup to represent the first-order mental state "desiring chocolate", it does not entail that its neurophysiological structure also supports the higher-order state "I am having the desire for chocolate". Though this is a very specific example, the logic at work could be applied to many different scenarios. It is here that we arrive at the notion of partial logical supervenience on the physical. Artificial intelligence may very well have the ability to develop mental representations of the outside world, but will they have the ability to have such meta-mental states that define Higher-Order theories of consciousness?

This then brings into play the intangible qualities of being human that we can naturally associate with meta-mental states; the raw emotions that we link to experiences help us develop conscious thought: the gut-wrenching guilt of making a mistake or the butterflies from interacting with a loved one. These meta-mental states that process what we are processing are what set humans apart from every other organism and establishes our self-awareness—one could even associate these qualities with the essence of a human "soul". Artificial intelligence, without the physical makeup of humans, would never be able to have such high-level conscious thought. Without a digestive organ how could an AI be consciously aware of its desire to consume food? Without a limbic system how could an AI be consciously aware that it loves another being? However, what if we created an anatomically accurate clone of a human down to every cell with semiconductors instead of carbon-based cells? What kind of consciousness would such a being possess? It does stand to reason that they would acquire sentience , but since consciousness does not logically supervene on the physical, they could have meta-mental states identical to ours, or meta-mental states merely similar to ours, or no meta-mental states at all.

Present-Day Considerations

In application to machine learning and artificial intelligence, the idea that consciousness does not logically supervene on the physical implies that a being with a different physical makeup could achieve our same level of consciousness—a conscious "robot" is achievable. If an artificial intelligence is built with enough baseline schemas and the ability to continually build mental models and representations of its experiences, then it is logically possible for artificial

intelligence to develop awareness regarding itself and the world around it as well as the emotive processes associated with phenomenal consciousness. This possibility has been explored in modern-day media through works such as *Westworld* and *Avengers: Age of Ultron*. In both of these examples, an artificial intelligence was able to build elaborate and complex mental schemas of the world through its experiences. Warning: Spoilers ahead. At the end of *Westworld* Season 1, audiences saw that Maeve, the madam of Westworld, had progressed from sentience to consciousness as she became more and more aware of Delos, the company that controls Westworld, going so far as plotting her escape. Furthermore, audiences were treated to both villainous and heroic artificial intelligence specimens in *Avengers*. The villainous A.I. Ultron comes to the conclusion that humanity must be eradicated to save the Earth. No phenomenal consciousness is attributed to Ultron; Ultron is just able to perceive the world and respond to it. Vision, on the other hand, fights to save humanity as the artificial intelligence cultivated an endearing view on humanity. This reasoning shows how Vision is not only sentient but also displays the ability to reflect on himself and attribute feelings to their experiences. These media and many others play with the idea of artificial intelligence having varying degrees of consciousness. When the idea is analyzed more critically, we see that this possibility of artificial intelligence to develop complex mental schemas through lived experience is what allows them to have conscious experiences that surpass a simple stimulus-response dynamic.

The growing development of machine learning models has initiated research in philosophy on detecting consciousness in artificial intelligences. We look to be a long way from creating an artificial agent with general intelligence, but scientific progress is decidedly non-linear. In anticipation of such developments, Edwin Turner Susan and Schneider have worked on a behavior-based artificial consciousness test (ACT). (Turner and Schneider) It turns out that the Turing test merely tests whether or not a computer can fool a human, not if that computer is conscious. There are technicalities that must be considered when developing such a test, e.g. how to isolate the AI from the outer world before doing the test. It will be interesting to track the development of philosophical ideas of consciousness as scientists continue to work on creating more mature artificial agents.

References

Chalmers, David John. *The Conscious Mind: in Search of Fundamental Theory*. Oxford University Press, 1996.

Godfrey-Smith, Peter. *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*. Farrar, Strauss, and Giroux, 2016

Ramsey, William, "Eliminative Materialism", The Stanford Encyclopedia of Philosophy (Spring 2019 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/spr2019/entries/materialism-eliminative.

Piaget, J. *The grasp of consciousness: Action and concept in the young child.* (Trans by S. Wedgwood). Oxford, England: Harvard U Press, 1976.

Gardner, Howard. (1976, August 1). *The Grasp of Consciousness*. The New York Times, pp. 172. URL = https://www.nytimes.com/1976/08/01/archives/the-grasp-of-consciousness-jean-piaget-at-80-continues-to-learn.html

Van Gulick, Robert, "Consciousness", The Stanford Encyclopedia of Philosophy (Spring 2018 Edition), Edward N. Zalta (ed.), URL = https://plato.stanford.edu/archives/spr2018/entries/consciousness .

Schneider, S., and Turner, E. L. I*s Anyone Home? A Way to Find Out If AI Has Become Self-Aware*. Scientific American Blog Network, 2017