

# AJSE

## American Journal of Science & Engineering

### Volume 2 Issue 3

December 2021



**American Journal of Science & Engineering (AJSE)**

Society for Makers, Artists, Researchers and Technologists (SMART)

6408 Elizabeth Ave SE, Auburn 98092, Washington, USA

ISSN: 2687-9530 (Print) and 2687-9581 (Online)

## Editor-in-Chief



### **Dr. Chuck Easttom**

University of Dallas, USA & Georgetown University, USA

**Research Interest:** Cryptography, Cyber Warfare, Engineering Processes and Digital Forensics.

Dr. Chuck Easttom is adjunct lecturer at Georgetown University and University of Dallas. He is the author of 31 books, including several on computer security, forensics, and cryptography. His books are used at over 60 universities. He has also authored scientific papers (over 70 so far) on digital forensics, machine learning/AI, cyber warfare, cryptography, bio-engineering, and applied mathematics. He is an inventor with 22 computer science patents. He holds a Doctor of Science (D.Sc.) in cyber security (dissertation topic: "A Comparative Study of Lattice Based Algorithms for Post Quantum Computing") and a Ph.D. in Technology focused on nanotechnology (dissertation topic: "The Effects of Complexity on Carbon Nanotube Failures"), as well as three master's degrees (one in applied computer science, one in education, and one in systems engineering). He is currently working on third doctorate, a Ph.D. in computer science with emphasis on applied mathematics from the University of Portsmouth (dissertation topic "On the application of algebraic graph theory to network forensics"). He is a Senior Member of the IEEE and a Senior Member of the ACM as well as a member of IACR (International Association of Cryptological Research) a member of APS (American Physical Society), and INCOSE (International Council on Systems Engineering). He is also a Distinguished Speaker of the ACM (Association of Computing Machinery). and a Distinguished Visitor of the IEEE Computer Society, and a frequent speaker at conferences. He also currently holds 55 industry certifications (CISSP, CASP, CEH, etc.) He is a member of IEEE Software & Systems Engineering Standards Committee. He has worked on the DevOps 2675 IEEE standards group 2017 to 2019 and currently a member of the IEEE Engineering in Medicine and Biology Standards Committee. Standard for a Unified Terminology for Brain-Computer Interfaces P2731.

### **From the Editor's Desk:**

**A thought about scientific rigor on research:** As scientists, we must always be striving to produce not just more research, but better quality research. A researcher should be his or her own harshest critic. Look at your own work with a skeptical eye. Could you provide clearer data? Are your references adequate and current? Is your statistical analysis appropriate and robust? Our goal as scientists is not merely to publish research, but to produce research that is truly impactful. By constantly striving to improve the quality of our own work, we improve the entire body of work in any scientific field.

### **Editorial Board:**

**Editor-in-Chief - Dr. Chuck Easttom** (University of Dallas, USA & Georgetown University, USA)

**Associate Editor - Dr. Nabeeh Kandalajt** (Grand Valley State University, USA)

### **Board Members -**

- i) **Dr. Phillip Bradford** (University of-Connecticut-Stamford, USA)
- ii) **Dr. Alex "Sandy" Antunes** (Capitol Technology University, USA)
- iii) **Dr. Izzat Alsmadi** (Texas A&M, San Antonio, USA)
- iv) **Dr. Lo'ai Tawalbeh** (Texas A&M University-San Antonio, USA)
- v) **Dr. Doina Bein** (California State University, Fullerton, USA)
- vi) **Dr. Hasan Yasar** (Carnegie Mellon University, USA)
- vii) **Dr. Moises Levy** (Florida Atlantic University, USA)
- viii) **Dr. Christian Trefftz** (Grand Valley State University, USA)
- ix) **Dr. Petros Spachos** (University of Guelph, Canada)

Page No.	CONTENT
1-8	<p><b>Bias in Cognitive Engineering for Human-Machine Teaming Literature</b></p> <p><i>Machines have a plethora of applications in spaces dangerous, hostile, or high-risk to humans. Given such applications, the fields of human-machine teaming and cognitive engineering, while nascent, are developing at a staggering rate. At the heart of these disciplines is a doctrine espousing an agent-centric concept based on common properties shared between machines and human operators. Symbiotically, the goal is to ensure machine and human mutually benefit from the teaming. Yet, the literature has been fixated on system design and machine-agent optimization for performance. Thus, the literature appears to be biased as it does not consider the cognitive requirements of the human agent. The problem then, is that recommendations from the literature focus on changing the machine to instigate less demanding human participation rather than adapting the human to induce optimal machine participation. For this reason, this work examined what characteristics of human-machine teaming literature demonstrate a statistically significant relationship with the category of focus in the same research. The characteristics-as-variables included author discipline, count of publications in the field, author affiliation, gender, and country of origin. A multinomial regression revealed a significant relationship with focus on the machine element as opposed to the human element in human-machine pairing in the cognitive engineering literature. Furthermore, author discipline, affiliation, and country of origin, demonstrated a significant bias effect towards the machine element in human-machine pairing literature</i></p> <p><b>DOI:</b> doi.org/10.15864/ajse.2301</p> <p>Jason M. Pittman(University of Maryland Global Campus, School of Cybersecurity &amp; Information Technology Largo, MD USA), Courtney Crosby (Booz Allen Hamilton, McLean, VA USA)</p>
9-14	<p><b>Shapley value in convolutional neural networks (CNNs): A Comparative Study</b></p> <p><i>Deep learning models are in dire need of training data. This need can be addressed by encouraging data holders to contribute their data for training purpose. Data valuation is a mechanism that assigns a value reflecting a number to each data instances. The SHAP Value is a method for assigning payouts to players of coalition game depending on their contribution to the total payout that entails many criteria for the notion of data value. In this paper, the value of the SHAP parameter is calculated in different convolutional neural network for varieties of image datasets. Calculated SHAP value for each data instance shows whether data is high value or low value and it is different in each model. In other words, if you have an image in the VGG model and it is high value, necessarily, it is not high value in ResNet model. The results show that the value of data varies in each dataset and model. Keywords— Deep learning, SHAP Value.</i></p> <p><b>DOI:</b> doi.org/10.15864/ajse.2302</p> <p>Seyedamir Shobeiri, Mojtaba Aajami (Department of Computer Science, Islamic Azad University of Zanjan, Zanjan, Iran.)</p>
15-23	<p><b>Moral-Driven Planning in Simple Automated Narrative Generator</b></p> <p><i>A moral is a vital part of the story and is vital for both the author and the audience. However, moral-based story planning is relatively underexplored. This paper proposes a planning-based Simple Automated Narrative Generator framework that explicitly considers the moral input and plans the story events around it. To achieve that, we determine the character's emotional arc based on the polarity of the input moral and then plan the events that adhere to the emotional arc. In the experiment of 103 human evaluators and 35 generated narratives for 7 morals, we find encouraging normalized confidence of 0.86 for intended moral being conveyed by the narratives and reasonable confidence of 0.71 for selecting the narratives with morals versus without morals.</i></p> <p><b>DOI:</b> doi.org/10.15864/ajse.2303</p> <p>Sanjay Garg(Indrashil University), Purvish Khalpada(Nirma University)</p>
24-31	<p><b>An Approach to E-Commerce with Voice Recognition</b></p> <p><i>E-Commerce or Electronic Commerce is process of doing business through computer networks virtually. A person sitting in front of his computer can access all facilities to buy and sell products. E-commerce has reduced human effort by reducing physical work and saving time of the seller and buyer both.The main advantage of E-commerce is that the user can browse online shops and compare prices of products sitting at home in their comfort. The buyer/consumer surfs through the internet to the seller's website and views the products sold by them. There the buyer compares the price and selects items that he/she wants to buy and then proceeds with transaction.</i></p> <p><i>Most E-commerce websites are limiting consumer's senses to only viewing. Shopping can be made more convenient if the buyer could communicate with the website just by verbally asking for the product and receiving audio + visuals as a result. Here comes the role of AI powered voice recognition. This is the problem we aim to solve to make the user experience better by combining two very popular technologies that are Reactjs and Alan AI.The power of React makes a website faster. React allows developers to create large web applications that can change data, without re-loading the page. The main purpose of React is to be fast, scalable, and simple. Alan AI is an advanced Voice AI Platform that allows us to add a voice interface to our app without overhead.</i></p> <p><b>DOI:</b> doi.org/10.15864/ajse.2304</p> <p>Debangi Choudhury, Riya Saha, Anushka Dey, Tiyasha Dey, Srilekha Mukherjee(UEM, Kolkata)</p>
32-37	<p><b>Study on the limitations of Stacking Technique for Bandwidth Improvement of Microstrip Patch Antennas</b></p> <p><i>Stacking of parasitic patch layers on rectangular micro strip patch antennas and its effect on the bandwidth and gain of the antenna are analyzed. For that two different frequencies 2.4 GHz and 11.5 GHz have been chosen from S band and X band respectively and stacking of patch layers have been done up to a certain limit assuming the minimum requirement of gain of 10 dB for S band antenna array and 12 dB for X band antenna array. So in both the cases, at first a 2x2 rectangular patch antenna arrays have been optimized to reach up to the gain requirement and then the stacking of parasitic layers have been done and results are checked and analyzed one by one.</i></p> <p><b>DOI:</b> doi.org/10.15864/ajse.2305</p> <p>Dipankar Saha, Mandar Chakrabarti, Abir Chattopadhyay(University of Engineering &amp; Management, Kolkata)</p>



# Bias in Cognitive Engineering for Human-Machine Teaming Literature

Jason M. Pittman  
 University of Maryland Global Campus  
 School of Cybersecurity & Information Technology  
 Largo, MD USA  
 jason.pittman@umgc.edu  
 ORCID: 0000-0002-5198-8157

Courtney Crosby  
 Booz Allen Hamilton  
 McLean, VA USA  
 ORCID: 0000-0002-9956-5962

**Abstract - Machines have a plethora of applications in spaces dangerous, hostile, or high-risk to humans. Given such applications, the fields of human-machine teaming and cognitive engineering, while nascent, are developing at a staggering rate. At the heart of these disciplines is a doctrine espousing an agent-centric concept based on common properties shared between machines and human operators. Symbiotically, the goal is to ensure machine and human mutually benefit from the teaming. Yet, the literature has been fixated on system design and machine-agent optimization for performance. Thus, the literature appears to be biased as it does not consider the cognitive requirements of the human agent. The problem then, is that recommendations from the literature focus on changing the machine to instigate less demanding human participation rather than adapting the human to induce optimal machine participation. For this reason, this work examined what characteristics of human-machine teaming literature demonstrate a statistically significant relationship with the category of focus in the same research. The characteristics-as-variables included author discipline, count of publications in the field, author affiliation, gender, and country of origin. A multinomial regression revealed a significant relationship with focus on the machine element as opposed to the human element in human-machine pairing in the cognitive engineering literature. Furthermore, author discipline, affiliation, and country of origin, demonstrated a significant bias effect towards the machine element in human-machine pairing literature.**

**Keywords:** *cognitive engineering, human-machine teaming, bias, literature characteristics, regression*

## I. INTRODUCTION

The human-machine teaming (HMT) paradigm implies a dynamic whereby human and machine agents work toward a common end state (Ryan, 2018). Yet historically, anthropomorphic assumptions framed HMT research. That is, machines existed in such a framework to assist humans, rather than partner with them (McDermott et al., 2017; Hancock et al., 2011; Schaefer et al., 2017). As a result, the critical lens centered on system design. However, while system design as a technical engineering discipline improves machine performance, it does not inherently address the cognitive requirements of the human participant. The problem then, is that recommendations from HMT literature focus on changing the machine to instigate less demanding human participation rather than adapting the human to induce optimal machine participation.

We suggest the problem is worse because of the lack of comprehensive HMT research in the military, where the generic human agent in human-machine teaming becomes an individual Soldier agent (Ryan, 2018). In such scenarios, non-specific domain research is an abstraction at most. Thus, this study aims to disambiguate universal HMT literature by untangling bias. Specifically, the purpose of this study was to investigate the potential relationship between characteristics of existing HMT literature that may lead to bias toward optimizing the machine over the Soldier.

## II. BACKGROUND

### A. Human-Machine Teaming

Fundamentally, human-machine teaming is an agent-centric concept based on symbiotic properties.



Symbiosis, by definition, is achieved when both agents mutually benefit through one or more exchanges (Schaefer et al., 2014). Thus, an optimal state occurs when both agents accomplish intended incremental objective towards a common end state. Yet, HMT research biases toward the machine; the human is present in the equation only insofar as they are a recipient of the exchange. To that end, qualitative teaming analysis has limited itself to themes related to design content and concerns only the manipulation of the machine element (McDermott et al., 2017). In other instances, quantitative research minimized the role of human-related factors altogether (Hancock et al., 2011). Further, a recent meta-analysis highlighted the limitations of historical literature. Specifically, a lack of empirical research on human-machine interaction and subfield specificity constrains efforts to leverage the work against modern applications (Schaefer et al., 2017).

Accordingly, the updated descriptive model for HMT presented by Schaefer et al. (2017) is a more exact representation of the partnering dynamic natural in human-machine teaming arrangements. Such a model echoes complementary research seeking to expand conventional frameworks and consider such elements as intelligence, cognition, and autonomy, which are critical to machine intelligence advancement (Fong et al., 2001; Yagoda, 2012; UK Ministry of Defence, 2018). Yet, while Schaefer et al. (2017) made critical contributions to distinguish between human- and machine-focused characteristics and expand the role of the human in the interaction, the model remained binary and categorical. That is, demographic traits are expressly nominal; and states, cognitive factors, and emotive factors are all ordinal variables. In short, the findings do not capture the complexity and dynamism of modern, or future, human-machine teaming.

### *B. Considering Transhumanistic Philosophy*

Given the bias emphasizing the machine portion of HMT, it naturally follows that research into human engineering for HMT is limited. There are a multitude of explanations for such bias. However, we feel the philosophical framework underlying HMT and human cognitive engineering (CE) is materially different and calls for examination as a first-cause principle. Interestingly, the potential philosophical frameworks align with the operational components in HMT.

Cognitive engineering implies a deliberate manipulation of human cognitive patterns for some specific benefit. As such, this study will consider cognitive engineering transhumanistically. Transhumanism rejects the concept of finality as it relates to human development (More, 2013). Specifically, transhumanists view technology as the primary means of progressing the human species along natural and artificial world continua (Bostrom,

2005). Thus, the philosophical framework of transhumanism is tightly coupled to human cognitive engineering.

Conversely, it appears that the bias towards the machine in general HMT is more closely associated with post-humanistic thought. Where transhumanism holds the transcendence of humans through technology as a central tenet, posthumanism is a philosophical correction to such anthropocentrism. In other words, posthumanism holds non-humans as capable of cognition as humans (Wolfe, 2009). Certainly, the focus on the machine- at a minimum, even as a mediating construct- in HMT literature associates strongly with posthumanism philosophy.

Accordingly, this paper adopts Woods and Roth's (1988) definition of cognitive engineering as a creative discipline based in applied science. That is, cognitive engineering is human behavioral design aimed at improving participation in complex systems (Woods & Roth, 1988). Consequently, the transhumanistic approach to HMT places the onus of active participation as much on the human as the machine (insofar as it is the addition of the machine that engenders the expansion of human ability). Such a concept is essential to the idea that successful human-machine teaming is not necessarily a technical re-engineering of the machine but rather a cognitive re-engineering of the human vis-à-vis technology.

### *C. Principles and Military Applications*

The theoretical architecture of cognitive engineering describes the mechanisms that bound the application of cognitive science in a given domain. Specifically, we inferred from Woods and Roth (1988) that the machine is part of the domain world, not a tool within it. As such, "one must understand how representations of the world interact with different cognitive demands imposed by the application world in question and with characteristics of the cognitive agents, both for existing and prospective changes in the world" (p. 424). Here, the benefit of cognitive engineering in the military HMT space is in its ecology (Vicente, 2002). That is, it applies specifically to the "multidimensional, open world" inherent in military operations whereby multiple cognitive agents work in tandem (Woods and Roth, 1988, p. 418).

Moreover, CE emphasizes contextual knowledge acquisition, meaning that it accounts for the semiotic and semantic distinctions intrinsic to a given domain (Woods & Roth, 1988). For example, military terrain analysis and tactical communications act as unique delivery systems for domain-specific semiotics. Similarly, plain English code words hold one meaning to the agent, analyst, or operator while conferring a different meaning to the civilian. What makes this contribution particularly important to military HMT is



that it alleviates the focus on individual objects in the space and instead focuses on “changing behavior [or] performance in that [space]” (Woods & Roth, 1988, p. 419).

By adopting Woods and Roth’s (1988) framework, the main differentiate between machine-centric approaches to HMT engineering and the transhumanistic one becomes introspection. That is, the ability of the participant to understand their own active and passive participation in the world (Halff, Hollan, & Hutchins, 1986; National Research Council, 2008; Sherlin et al. 2011). Such a concept echoes the Army’s recent adoption of learner-centric environments and upholds existing cognitive training doctrine (TRADOC, 2017). Thus, the goal of cognitive engineering is to achieve the mental functions and expertise outlined within the HMT world: situational recollection, holistic recognition, intuitive decision making, and absorbed awareness.

In a model where the human and machine are cohesive domain actors in the same world, the symbiotic end state is one in which humans and machines execute intuitive functions naturally, while simultaneously working cooperatively toward logical functions. To carry out such parallel and convergent task orientation, the military adopted the Boyd decision making model, known colloquially as the *OODA loop* (Clarke & Knudson, 2018). The OODA model’s integration into the military organizational system rests in its alignment to military measures of effectiveness (MOE). That is, the goal of both OODA and MOE is to manipulate the predicted future world toward a specific end state, just as cognitive engineering is founded on the principle of changing behavior (Woods & Roth, 1988; Clarke & Knudson, 2018). When taken together, the resulting theoretical framework already supports CE for military applications.

#### D. Military Cognition

Thus, imperative to the practical application of CE for HMT are (a) viewing both the machine and human as domain agents of the same world; (b) acknowledging HMT as a continuous endeavor toward a given end state; and (c) recognizing end state relevancy as the sum of its parts. In fact, existing literature draws direct links between CE and successful military training (Halff, Hollan, & Hutchins, 1986; Noble, 1989; Blacker et al., 2018). Further still, cognitive science is regularly applied to training domains outside the scope of military intervention (Simons et al. 2016; Strobach & Karbach 2016). Notwithstanding, Blacker et al. (2018) differentiate military CE from other domains, noting, “outcomes for multifaceted skills [are especially important], as military operations inevitably require coordination of multiple cognitive abilities. For example, there is not one specific laboratory-based

task that encompasses all components of room clearing, piloting a military aircraft, or navigating a ship” (p. 3). Thus, the cognitive requirements in a military population are innately distinguishable from other populations (Blacker et al., 2018). Accordingly, military cognitive engineering must account for the unique demands of service members.

### III. METHOD

We sought to answer a single research question: what characteristics of HMT literature (e.g., author, scientific discipline, etc.) demonstrate a statistically significant relationship with the type of HMT research (e.g., machine-focused). To that end, we operationalized six characteristics of HMT literature as variables. Moreover, we grounded the characteristics by adopting those illustrated by Elamrani and Yampolskiy (2019) and used by Norris (1997) to relate those characteristics to bias. Such characteristics include author discipline, count of publications in the field, author affiliation, gender, and country of origin. Where one or more characteristics are unknown, it is indicated.

#### A. Coding

*Author discipline* was coded by the author’s self-identified area of research in available profiling, including either within the publication or available on their public page. *Count of publications in the field* was calculated by counting all publications associated with the author’s name, regardless of author position. *Author affiliation* was determined first by the affiliation listed on the publication; when no affiliation was listed on the publication or by the publishing authority, affiliation was determined by the author’s current professional or academic position or role. *Gender* was coded as binary and represented only male or female. *Country of origin* was based on the origin of the author’s affiliation.

Furthermore, we identified three types of research: *Machine-focused*, *Human-focused*, and *Hybrid*. To identify these types, we measured the frequency of related keywords in each piece of literature. We interpreted a frequency of greater than 51% to indicate majority and thus type. A frequency in which machine and human keywords were evenly split was a hybrid type. For example, keywords were associated with each code (machine-focused, human-focused, and hybrid), such that *machine*, *computer*, *system*, *robot* represented machine-focused terminology while human encompassed terms such as *human*, *person*, *people*, *Soldier*.

#### B. Population and Sample

The initial sample of relevant HMT literature consisted of the entire online archive of EBSCO studies specific to eight search queries (“human machine teaming”, “human robot teaming”, “human

machine systems”, “human machine trust”, “human automation trust”, “human machine cognitive engineering”, “human machine design”, and “human robot interaction and behavior”). We selected only those studies which were published in an academic journal and had relevant search engine ranking. When search engine ranking was taken into account, the population size decreased from 1,051 to 201. In addition to the ESCBO search, the review also included the entire online archive of Google Scholar results as of January 1, 2014 (6,300) specific to the search query “human machine teaming”. The results were further refined by search engine ranking. As such, our analysis was restricted to 118 studies or 47% of the relevant population. We excluded foreign language literature as language translation would have been outside the scope of this study and, without such, we would not have been able to properly code for “type of research”.

### C. Analysis

Once the studies were coded, we generated descriptive statistics and followed with a multinomial logistic regression (Gayle, Lambert, & Davies, 2009) as a predictive analysis technique to determine what characteristics of HMT literature (e.g., author, scientific discipline, etc.) demonstrated a statistically significant relationship with the type of HMT research (e.g., machine-focused). Such a technique was most appropriate given the type of dependent variable in the study. Additionally, the use of multinomial logistic regression supported the complex, psychological nature of the research question.

Since this study posits that there is underlying bias in the literature, logistic analysis appropriately disambiguates the relationship between the dependent variable (bias) and its associated indicator variables (author discipline, count of publications in the field, author affiliation, gender, and country of origin). Further, because we identified five potential indicators, a multinomial logistic regression specifically fit our data analysis needs. Coded data were imported into IBM SPSS and the appropriate function executed. Then, the bias prediction results were assessed to be significant only when  $p \leq 0.05$  (bias exists or does not exist) for both the overall model and for individual indicator contribution to the model.

## IV. RESULTS

The general question we sought to answer was whether author discipline, publication count, institutional affiliation, gender, and country of origin might be predictors for the apparent bias of published cognitive engineering literature focusing on the as opposed to the human element in human-machine pairing. Ultimately, we found our overall model fit our conjectured outcome wherein three of the five

individual factors demonstrated statistical significance. The following sections present the descriptive and quantitative findings.

### A. Description of Literature and Factors

The literature we collected focused on topics that related human cognitive engineering and machines. That is, the research examined only those publications that sought to explore the relationship between organic and artificial agents. Once the sample was selected, factors were extracted from the literature and processed per our research protocol. The results are as follows.

#### 1) Author Discipline

We found a total of 73 author disciplines in the sample of literature. The percentage of authors associated with these fields ranged from 0.2 percent (one author) to 17% (73 authors). There were 31 disciplines showing a 0.2 percent representation whereas there were 21 disciplines with one percent or greater representation.

#### 2) Publication Count

Next, we observed author *publication count* (Table 1) as being diverse ( $M = 70.4$ ,  $SD = 121.67$ ) with the minimum number of publications being one and the maximum being 1,190. Further, the majority of authors had publication counts in the one to one-hundred range (80%) out of 421 total authors. The median was 25 and the mode was one.

TABLE I. NUMBER OF AUTHORS WITH PUBLICATION COUNTS CATEGORIZED BY HUNDREDS

	1-100	101-200	201-300	301-400	401-500	501+
<i>Machine</i>	164	24	7	4	5	2
<i>Human</i>	100	8	4	2	2	1
<i>Hybrid</i>	73	10	6	4	4	1
<i>Totals</i>	337	42	17	10	11	4

#### 3) Affiliation

Following publication count, we examined the *authors' affiliation*. The distribution of this variable favored University affiliations. That is, the majority of authors were publishing for a University or as representatives of a University (%). This was followed by U.S. military agents (%) and Industry (%). Industry reflects any commercial or private, non-research-based entity.

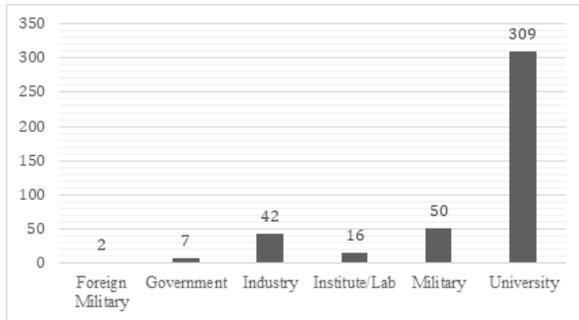


Fig. 1. The distribution of author affiliation by professional category

4) Gender

Gender is represented as a binary categorical variable. Of known author gender, there were 223 males and 115 females. As such, men represented 66% of authors in this space. While both men and women favored machine-centric research, the distribution within each gender for human- and hybrid-focused research differed. Specifically, the male variance between human- and hybrid-focused research was negligible (56 compared to 57, respectively). Comparatively, women were more likely to publish human-focused literature (38 authors compared to 27).

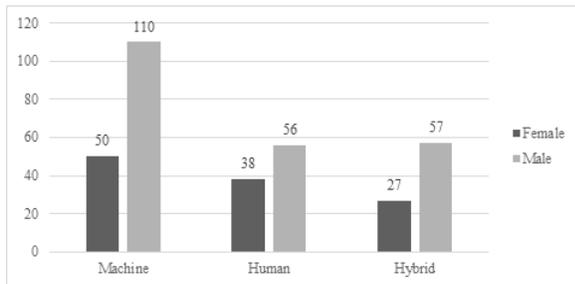


Fig. 2. The distribution of author gender categorized by bias coding

5) Country

Country of origin was based on the origin of the author’s affiliation. That is, the country factor is representative of the strength of the author’s professional or academic ties (i.e. affiliation), rather than their birth place. The decision to code country based on this methodology was to prevent the disclosure of PII related to an author’s birth record. The decision also mitigated validity risks since birthplace identification is not a readily publicized or available piece of information.

There were 427 counts of Country. The United States was represented in a plurality of the literature, with 203 associations (48%). This was followed by China (11%), Italy (6%), Japan (6%), Australia (4%), the United Kingdom (4%), and Germany (3%). The remaining countries accounted for less than 10 codes each.

6) Factors and Relationships

We identified five potential factors which we operationalized as predictor variables: author discipline, publication count (in the field), institutional affiliation, gender, and country of origin. None of the variables were transformed during the logistic regression process. Procedurally, we first analyzed the overall model and then proceeded to examine individual variables. The results are detailed in the following subsections.

7) Overall model evaluation

The overall model appeared to fit our conjecture, based on the research methodology and analysis protocol (Table 2). That is, the identified predictors fit within the stated model more positively compared to the null model. Granted, the ultimate validity of the model was limited by the sample size; however, the results exhibit a foundation for inference insofar as there is a measured effect at work given the selected variables. Determining which specific variables were involved constituted the next step.

TABLE II. MULTINOMIAL LOGISTIC REGRESSION FOR OVERALL MODEL

Model	-2 Log Likelihood	Chi-square	df	Sig.
<i>Intercept Only</i>	841.110			
<i>Final</i>	137.240	703.870	482	.000

8) Predictor variables

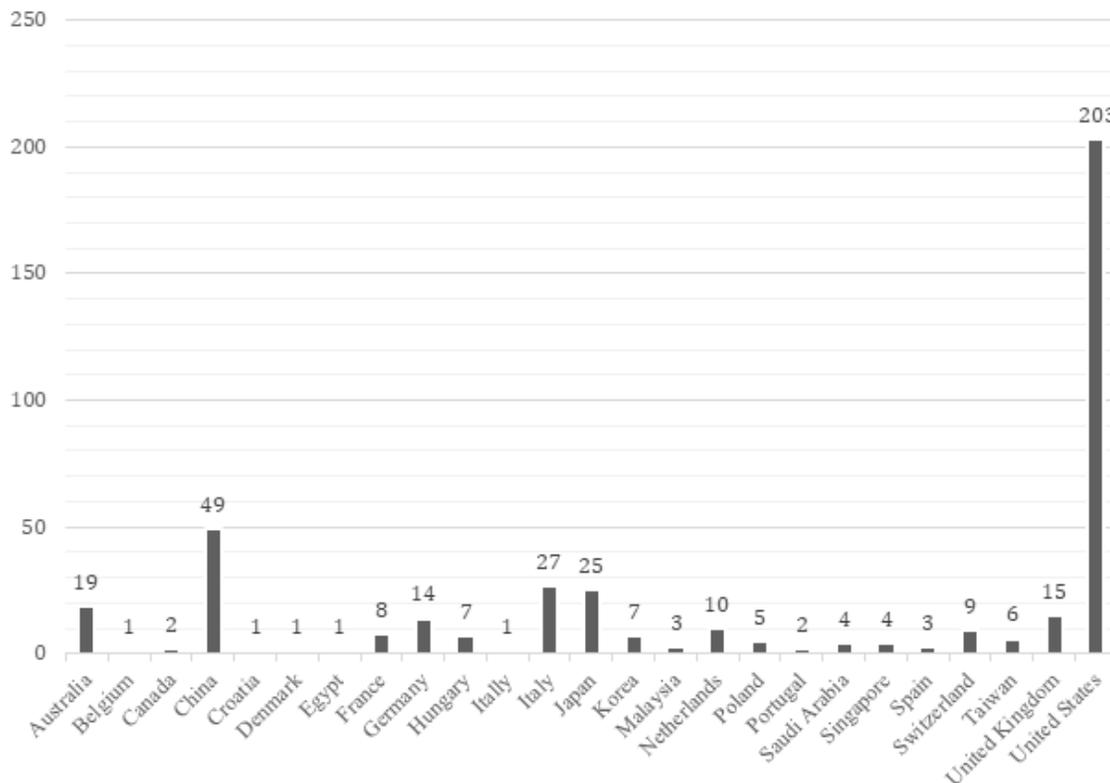
The multinomial logistic regression of the variables (Table 3) revealed three of the five with a significance less than 0.05. More specifically, the discipline, affiliation, and country of origin associated with authors demonstrated significant effect on the outcome of the analysis. Within the context of this work, the results point at a positive bias towards the machine element in human-machine pairing literature.

TABLE III. MULTINOMIAL LOGISTIC REGRESSION OF PREDICTORS VARIABLES

Effect	-2 Log Likelihood	Chi-square	df	Sig.
<i>Intercept</i>	137.240	.000		
<i>Discipline</i>	351.879	214.639	132	.000
<i>Pub Count</i>	368.754	231.514	268	.948
<i>Affiliation</i>	163.794	36.553	16	.047
<i>Gender</i>	139.268	2.027	4	.731
<i>Country</i>	204.970	67.730	40	.004

V. CONCLUSIONS

The Human Machine Teaming domain is researched and applied through an anthropomorphic lens that emphasizes system design over cognitive engineering (McDermott et al., 2017; Hancock et al., 2011; Schaefer et al., 2017). The subsequent focus on using machines as a tool to help humans, without the requisite reengineering of the human cognitive space



to effectively partner with machines, is further exacerbated in the military space. When examining various agent actors, such as Soldiers, the specificity of the environment must be considered.

To that end, the importance of human-machine teaming and cognitive engineering should not be underestimated. Ensuring machine and human agents mutually benefit from teaming is not just a hope but a necessity for the nascent disciplines to solidify the

Fig. 3. The frequency of authors country of origin based on the author's affiliation.

underlying doctrines. However, the literature has fixated on system design and machine-agent optimization for performance in lieu of maintaining a holistic perspective.

As a means to bolster such a perspective, we sought to answer a single research question: *what characteristics of HMT literature (e.g., author, scientific discipline, etc.) demonstrate a statistically significant relationship with the type of HMT research (e.g., machine-focused.* To that end, we operationalized six characteristics of HMT literature as variables: author discipline, count of publications in the field, author affiliation, gender, and country of origin. We then identified three types of research: *machine-focused, human-focused, and hybrid.* Once our research methodology and protocol were established, we conducted a multinomial logistic regression. The model was then formulated with the identified dependent and indicator variables and fit

using maximum likelihood estimation within IBM SPSS. Results show that discipline, affiliation, and country of origin associated with authors demonstrated significant effect on the outcome of the analysis. Within the context of this work, the results point at a positive bias towards the machine element in human-machine pairing literature.

As such, we posited that the machine is part of the domain world, not a tool within it (Woods & Roth, 1988). Therefore, research bias inconsistent with this position is detrimental to the effective advancement and execution of HMT. Yet, our model demonstrates that bias does exist in the HMT literature. What is not clear in these findings is what the potential effects of such bias may be to ongoing research and field applications.

Certainly, bias toward machine-centric research, at the expense of efforts to augment human cognition,



limits the possibility of meaningful advancement in HMT. Specifically, the current paradigm places undue emphasis on machine engineering or human engineering. Consequently, the transactional framing of machines solely as a tool for humans creates a ceiling whereby machine-human trust is constrained by the limits of human anthropocentrism. Rejecting such posthumanistic views will bring human-centric HMT research into parity. The resulting shift will empower developers to move beyond system design and machine-agent optimization to unlock beyond the horizon advancements in the field.

Looking beyond, approaching parity in the human versus machine research point of view may empower additional but related research pathways. For example, machine learning has long been haunted by a black box decision making problem. That is, existing machine learning systems are unable to express why a decision is made within the set of plausible decisions. Whereas common techniques are geared towards having machines explain internal decisions, perhaps more success might be had if a teamed human were able to directly query the machine.

On a similar note, because of the possible existential risks posed by advanced machine intelligence, containment or boxing is a crucial topic. Understandably machine intelligence is central to this research at the cost of exploring how humans interface with the containment. However, containment by its nature limits or prevents interactions between machine and human. In addressing the bias towards machine-centric cognitive engineering, some answers for how humans and machine intelligence interact across containment may become clearer.

#### REFERENCES

- Blacker, K. J., Hamilton, J., Roush, G., Pettijohn, K. A., & Biggs, A. T. (2019). Cognitive training for military application: a review of the literature and practical guide. *Journal of cognitive enhancement*, 3(1), 30-51.
- Bostrom, N. (2005). A history of transhumanist thought. *Journal of evolution and technology*, 14(1).
- Clarke, A. J., & Knudson, D. I. (2018). *Examination Of Cognitive Load In The Human Machine Teaming Context*. Naval Postgraduate School Monterey United States.
- Fong, T., Thorpe, C., & Baur, C. (2001). *Collaborative control: A robot-centric model for vehicle teleoperation* (Vol. 1). Pittsburgh: Carnegie Mellon University, The Robotics Institute.
- Gayle, V., Lambert, P., & Davies, R. B. (2009). Logistic regression models in sociological research. *University of Stirling, Technical Paper, 1*.
- Half, H. M., Hollan, J. D., & Hutchins, E. L. (1986). Cognitive science and military training. *American Psychologist*, 41(10), 1131.
- Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, 53(5), 517-527.
- McDermott, P. L., Walker, K. E., Dominguez, C. O., Nelson, A., & Kasdaglis, N. (2017, June). Quenching the thirst for human-machine teaming guidance: Helping military systems acquisition leverage cognitive engineering research. In *13th International Conference on Naturalistic Decision Making* (pp. 236-240).
- More, M. (2013). The philosophy of transhumanism. *The transhumanist reader*, 8.
- National Research Council. (2008). Potential Intelligence and Military Applications of Cognitive Neuroscience and Related Technologies. In *Emerging Cognitive Neuroscience and Related Technologies*. National Academies Press (US).
- Noble, D. D. (1989). Cockpit cognition: Education, the military and cognitive engineering. *Ai & Society*, 3(4), 271-296.
- Norris, N. (1997). Error, bias and validity in qualitative research. *Educational action research*, 5(1), 172-176.
- Elamrani, A., & Yampolskiy, R. V. (2019). Reviewing tests for machine consciousness. *Journal of Consciousness Studies*, 26(5-6), 35-64.
- Ryan, M. (2018). *Human-Machine Teaming For Future Ground Forces*. Washington, DC: Center for Strategic and Budgetary Assessments.
- Schaefer, K. E., Billings, D. R., Szalma, J. L., Adams, J. K., Sanders, T. L., Chen, J. Y., & Hancock, P. A. (2014). *A meta-analysis of factors influencing the development of trust in automation: Implications for human-robot interaction*. Army Research Lab Aberdeen Proving Ground Md Human Research And Engineering Directorate.
- Sherlin, L. H., Arns, M., Lubar, J., Heinrich, H., Kerson, C., Strehl, U., & Serman, M. B. (2011). Neurofeedback and basic learning theory: implications for research and practice. *Journal of Neurotherapy*, 15(4), 292-304.



Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. (2016). Do “brain-training” programs work?. *Psychological Science in the Public Interest*, 17(3), 103-186.

Strobach, T., & Karbach, J. (2016). *Cognitive training*. New York, NY: Springer.

TRADOC (2017). Robotic and autonomous systems strategy. *Army capabilities Integration Centre*, Retrieved from [http://www.tradoc.army.mil/FrontPageContent/Docs/RAS\\_Strategy.pdf](http://www.tradoc.army.mil/FrontPageContent/Docs/RAS_Strategy.pdf).

UK Ministry of Defence. (2018). Joint Concept Note 1/18

Human-Machine Teaming. Development, Concepts and Doctrine Centre. Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/709359/20180517-concepts\\_uk\\_human\\_machine\\_teaming\\_jcn\\_1\\_18.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/709359/20180517-concepts_uk_human_machine_teaming_jcn_1_18.pdf)

Wolfe, C. (2009). Human, all too human: “Animal studies” and the humanities. *PMLA/Publications of the Modern Language Association of America*, 124(2), 564-575.

Woods, D. D., & Roth, E. M. (1988). Cognitive engineering: Human problem solving with tools. *Human factors*, 30(4), 415-430.

Yagoda, R. E., & Coover, M. D. (2012). How to work and play with robots: an approach to modeling human-robot interaction. *Computers in human behavior*, 28(1), 60-68.



# Shapley value in convolutional neural networks (CNNs): A Comparative Study

*Seyedamir Shobeiri*

Department of Computer Science, Islamic Azad University of Zanjan, Zanjan, Iran.

seyedamir.shobeiri@iauz.ac.ir

*Mojtaba Aajami*

Department of Computer Science, Islamic Azad University of Zanjan, Zanjan, Iran.

aajami@iauz.ac.ir

**Abstract**—Deep learning models are in dire need of training data. This need can be addressed by encouraging data holders to contribute their data for training purpose. Data valuation is a mechanism that assigns a value reflecting a number to each data instances. The SHAP Value is a method for assigning payouts to players of coalition game depending on their contribution to the total payout that entails many criteria for the notion of data value. In this paper, the value of the SHAP parameter is calculated in different convolutional neural network for varieties of image datasets. Calculated SHAP value for each data instance shows whether data is high value or low value and it is different in each model. In other words, if you have an image in the VGG model and it is high value, necessarily, it is not high value in ResNet model. The results show that the value of data varies in each dataset and model.  
**Keywords**— *Deep learning, SHAP Value.*

## I. INTRODUCTION

A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$ , if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ . The machine learning can be split into different classes of problems, Supervised learning and Unsupervised learning. In supervised learning, these algorithms, also called classifiers, receive data that has already been labeled. In fact, the data sent to the classification algorithms tells these algorithms which category each data should be placed in.

We do not tell the algorithm which category each data should fit into. In fact, we have no presumption about what category or group the existing data falls into, and the algorithm automatically detects the data classification. This is why these types of algorithms are called unsupervised learning algorithms.

The third type of algorithm, which can perhaps be classified as unsupervised algorithms, is a group called reinforced learning. In this type of algorithm, a machine (actually its controller program) is trained to make a specific decision, and the machine is based on its current position (set of available variables) and permissible actions (e.g. moving forward, moving Back and forth...) makes a decision that for the first time, this decision can be completely random and for each action or behavior that occurs, the system gives him a points from this feedback , The machine realizes whether it has made the right decision or not to repeat the same action the next time in that situation or try another action and behavior.

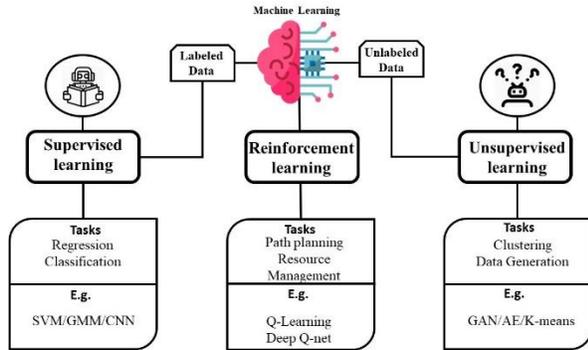


Fig. 1 Machine learning overview.

Deep Learning refers to a machine learning technique that establish artificial neural networks (ANNs) to imitate the structure and function of the human brain. Indeed, an artificial neural network, or neural network (NN), is a collection of connected computational units or nodes called neurons arranged in multiple computational layers that map a data input into a desired output. Each neuron applies a linear function to its inputs that sums up the products of weights and inputs. Subsequently, the output of this function is passed through an activation function. NN generates the desired output via feed-forward data flow and then updates the weights of each neuron by backpropagation of errors during the training phase.

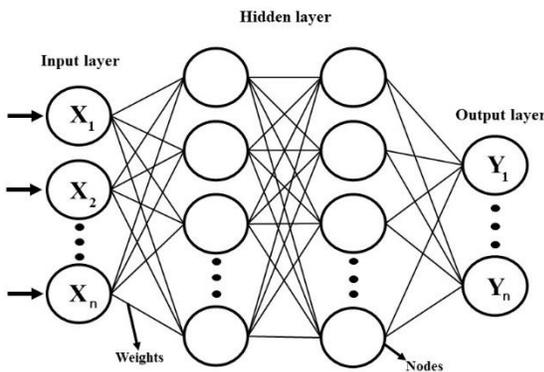


Fig. 2 Neural Network Overview.

In general, according to the employed NN architecture supervised deep learning can be categorized to the Recurrent neural networks (RNN) and Convolutional neural networks (CNN). RNN is a type of artificial neural network commonly used in speech recognition and natural language processing. Recurrent neural networks recognize data's sequential characteristics and use patterns to predict the next likely scenario. CNN is a type of artificial

neural network used in image recognition and processing that is specifically designed to process pixel data.

A Recurrent Neural Network is a type of neural network that contains loops, allowing information to be stored within the network. In short, Recurrent Neural Networks use their reasoning from previous experiences to inform the upcoming events. Recurrent models are valuable in their ability to sequence vectors, which opens up the API to performing more complicated tasks.

Convolutional neural network, is a deep learning neural network designed for processing structured arrays of data such as images. Convolutional neural networks are widely used in computer vision and have become the state of the art for many visual applications such as image classification, and have also found success in natural language processing for text classification. CNNs are very good at picking up on patterns in the input image, such as lines, gradients, circles, or even eyes and faces. It is this property that makes convolutional neural networks so powerful for computer vision. Unlike earlier computer vision algorithms, convolutional neural networks can operate directly on a raw image and do not need any preprocessing. A convolutional neural network is a feed-forward neural network, often with up to 20 or 30 layers. The power of a convolutional neural network comes from a special kind of layer called the convolutional layer. Convolutional neural networks contain many convolutional layers stacked on top of each other, each one capable of recognizing more sophisticated shapes. With three or four convolutional layers it is possible to recognize handwritten digits and with 25 layers it is possible to distinguish human faces. The usage of convolutional layers in a convolutional neural network mirrors the structure of the human visual cortex, where a series of layers process an incoming image and identify progressively more complex features. The architecture of a convolutional neural network is a multi-layered feed-forward neural network, made by stacking many hidden layers on top of each other in sequence. It is this sequential design that allows convolutional neural networks to learn hierarchical features. The hidden layers are typically convolutional layers followed by activation layers, some of them followed by pooling layers. A simple convolutional neural

network that aids understanding of the core design principles is the early convolutional neural network LeNet-5, published by Yann LeCun in 1998. LeNet is capable of recognizing handwritten characters.

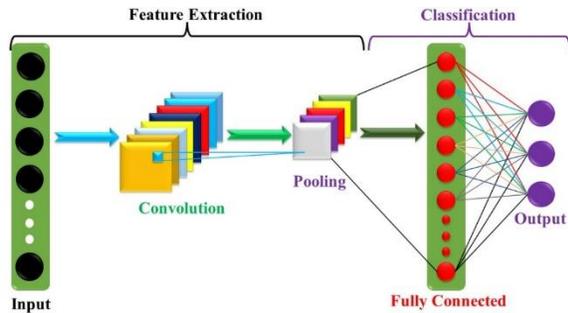


Fig. 3 Convolutional neural networks (CNN) Overview.

Machine learning models in general and deep learning in particular need to be trained with a wide variety of data in order to have acceptable performance. Hence there is a high demand from developers of machine learning models for training data. On the other hand, this data is owned by sectors that often do not voluntarily provide their data to model developers. In addition, we are interested in using data to training a model that will improve the efficiency of the model. Given this, there is a need for a mechanism that can evaluate the data needed and used to training the model. By using this method to provide incentives to data owners to share their data or to evaluate the data that is available for free before using it for training in terms of quality. In the seminal work, a mechanism for data valuation is provided. Inspired by Shapley Value method, this mechanism has presented a formula that can be used to numerically express the value of a specific data in improving the performance of a specific model. In this article, we have selected several famous CNN models that are widely used in various image recognition applications. we have calculated Shapley Value for two datasets, ImageNet and 10 Monkey Species, in each of these models. The rest of paper is organized as follows. Section II describes the architecture of several well-known CNNs.

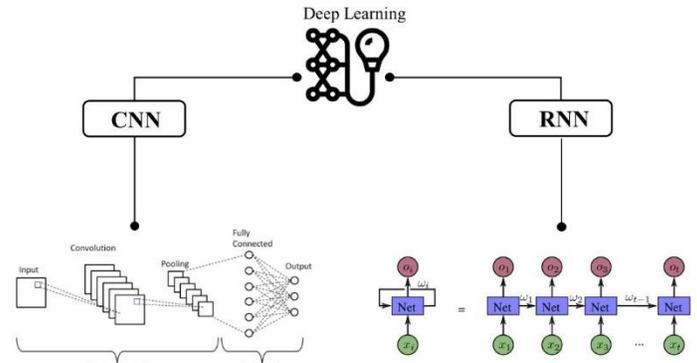


Fig. 4 NN architecture supervised deep learning Overview.

## II. prolific CNN architectures

This paper seeks to empirically establish a correspondence between the shapely value and the components of a CNN based learning systems including model, training data set. For this purpose, we selected some CNNs as the representatives of the models that are widely used in practice.

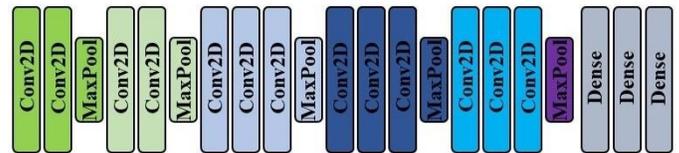


Fig. 5 vgg16 architecture.

Fixed size 224 x 224 RGB image is input of cov1 layer. The image through a convolutional.3\*3 layers, which is the smallest size to capture the notion of left/right, up/down, center. it also utilizes 1x1 convolution filters, which as a linear transformation of the input channels. The convolution stride is fixed to 1 pixel; the spatial padding of convolution layer input is such that the spatial resolution is preserved after convolution, i.e., the padding is 1-pixel for 3x3 convolution layers. Spatial pooling is carried out by five max-pooling layers, which follow some of the convolution. Max-pooling is performed over a 2x2-pixel window, with stride 2. Three Fully-Connected (FC) layers follow a stack of convolutional layers (which has a different depth in different architectures): the first two have 4096 channels each, the third performs 1000-way ILSVRC classification and thus contains 1000 channels .The final layer is





Inception-V3, at first, it introduced for the ImageNet Recognition Challenge. Inception assists classification of objects in the computer vision.

$$\phi_i = \sum_{S \subseteq D \setminus \{x_i\}} \frac{V(S \cup \{x_i\}) - V(S)}{\binom{|D|-1}{|S|}}$$

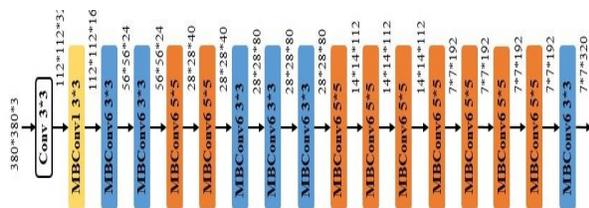


Fig. 10 EfficientNet B0 architecture.

EfficientNet-b0 is a CNN that is trained on more than a million images from the ImageNet database. The model can classify images into 1000 object categories, such as camera, boat, bird. So, the network learned rich features. The input of network is 224\*224. EfficientNet-B0 the developed by AutoML MNAS.

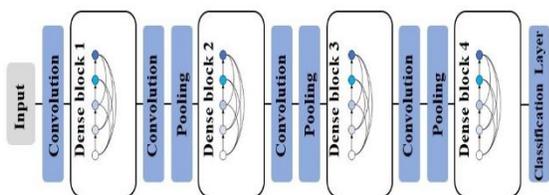


Fig. 11 DenseNet169 architecture.

### III. Shapley Value

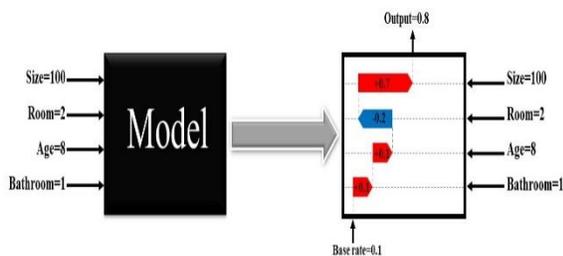


Fig. 12 Shapley Value example.

Shap Value performs the equitable data valuation in supervised machine learning. For a given set of training data points D and a performance metric, the “Shap Value” value  $\phi_i$  of a data point  $x_i \in D$  is defined as:

Where  $V(S)$  is the performance of the model trained on subset S of the data.  $V(S)$  is the prediction accuracy on the validation set. Intuitively, the Shapley value of a data point is a weighted average of its marginal contribution to subsets of the rest of the dataset. As a result, it can be used as a measure of data quality: a data point with a high Shapley value is one that improves the model’s performance if we add it to most subsets of the data, while a data point with a negative value on average hurts the performance of the model. Exact computation of Eq. requires an exponential number of computations in the size of the dataset, which is infeasible in most realistic settings. In fact, High value indicate high quality of image and correct label while low value represents low quality of image and incorrect label.

Finally, SHAP Value has three outputs, which are value, growth rate, and main data, respectively. Value: An array that each cell represents a pixel, each cell of the array contains another array that contains three cells and represents the RGB effect as shown: Value = [ [ R, G, B], [ R, G, B], [ R, G, B], ... ] and the growth rate, which is a base number, and our main data, which is the original values of our image. How to calculate the value of an image is as follows:

$$Value = \left( \sum_{i=0}^n ([R_i + G_i + B_i]) \right) + Base$$

According to the above formula  $i = 0$  because the array cells start from zero and N is the number of pixels, R represents the effect of red, G represents the effect of green, B represents the effect of blue, which indicates each of these pixel colors. How effective the image has been



in our model is that the sum of the effects of red, green, and blue with the base, which represents the growth rate, reflects the value of image. next, two datasets are tested that whether the value of an image is always the same or depends on another factor.

#### IV. Evolution and results

Due to the large volume of images, we randomly selected 50 images from each dataset and created two datasets smaller than ImageNet and 10Monkey Species. Images in the datasets are numbered from 0 to 49, and to refer to each image, Refer to the relevant number. In this article, we use the first formula, we obtained the value of each color of each pixel (RGB), and then use the second formula, we obtained the value of each image on different Convolutional neural networks, include VGG16, ResNet50, DenseNet169, MobileNet, EfficientNetB0, MobileNetV2. The table below has two rows that represent high quality and low-quality images and six columns that represent different architectures. The numbers you see in the tables indicate the number of images in each dataset. The numbers in the first row indicate the high-quality data of datasets in the architecture and the second row indicates low quality data in the same architecture.

Table 1: ImageNet Datasets

	VGG16	ResNet50	DenseNet169	MobileNet	EfficientNetB0	MobileNetV2
High Value	3	3	3	3	11	31
Low Value	34	25	18	32	25	18

Table 2: 10 Monkey Species Datasets

	VGG16	ResNet50	DenseNet169	MobileNet	EfficientNetB0	MobileNetV2
High Value	34	30	19	25	32	31
Low Value	46	43	36	38	46	37

#### V. Conclusions

Training data is used for training deep learning models. This need can be addressed by encouraging data holders to contribute their data

for training purpose. We need a method for numerical evaluation of data that the Data valuation method is very efficient. The Shap Value is a way to evaluate each data relative to its contribution to deep learning models.

According to the above experiments, we realized that the value of each image is very different from the network architecture and the value of each image is directly related to the architecture of that network.

#### REFERENCES

- Ghorbani, A., & Zou, J. (2019, May). Data shapley: Equitable valuation of data for machine learning. In *International Conference on Machine Learning* (pp. 2242-2251). PMLR.
- Lundberg, S. (2018). *SHAP documentation*. SHAP. <https://shap.readthedocs.io/en/latest/index.html>

# Moral-Driven Planning in Simple Automated Narrative Generator

Sanjay Garg  
Indrashil University  
gargsv@gmail.com

Purvish Khalpada  
Nirma University  
16ftphde19@nirmauni.ac.in

**Abstract**— A moral is a vital part of the story and is vital for both the author and the audience. However, moral-based story planning is relatively underexplored. This paper proposes a planning-based Simple Automated Narrative Generator framework that explicitly considers the moral input and plans the story events around it. To achieve that, we determine the character's emotional arc based on the polarity of the input moral and then plan the events that adhere to the emotional arc. In the experiment of 103 human evaluators and 35 generated narratives for 7 morals, we find encouraging normalized confidence of 0.86 for intended moral being conveyed by the narratives and reasonable confidence of 0.71 for selecting the narratives with morals versus without morals.

**Keywords**—computational storytelling, narrative generation, computational narratology, story planning, automatic story generation, moral based narrative generation, moral driven planning, moral of the story

## I. INTRODUCTION

Storytelling is an integral part of human civilization (Herman, 2013). The concept of storytelling is as old as the cave paintings, even older than language development itself. Storytelling, being with humanity for so long, has affected the evolution of the human brain to the extent that neuroscientists and evolutionary biologists believe that storytelling is what makes us human (Gottschall, 2012). With its roots so deep, storytelling can influence a person, causing the metamorphosis of Mister Gandhi to Gandhi, for example, and the society and the whole generation. The part of the story that makes the most impact is the story's message or moral (Turner, 1993).

Although computational storytelling has been a field of research for some time, the significant contributions have been towards generating and identifying the set of coherent events, addressing the creativity problem, and how the interaction of the audience can affect the development of the story. Explicitly incorporating morals in computationally generated narratives has been an underexposed element of computational storytelling. So, this paper proposes a moral-driven planner to lead event planning in Simple Automated Narrative Generator (SANG), a framework based on (Khalpada & Garg, 2021).

One factor that makes planning a story is the breadth of the exploration space of the possible events at each planning turn. To constrain the exploration, SANG uses the structure of an

emotional arc, which depends on the polarity of the moral input. SANG uses a Directed Causal Graph (DCG) similar to (Khalpada & Garg, 2019) and a knowledge base to plan the events around the determined emotional arc. This flow is abstractly depicted in fig. 1.

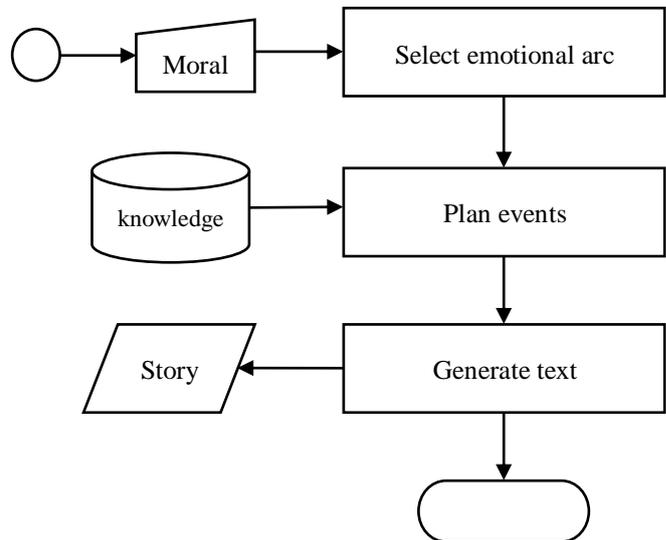


fig. 1 System overview

SANG takes moral input in higher-level language similar to predicate logic, with at least two atomic formulas connected with a causal relation. The last atomic formula determines the polarity of the moral. For example, the input  $forgive(PersonX, PersonY) \rightarrow reward(PersonX)$  is encouraging the forgiving nature. On the other hand, the input  $steal(PersonX, PersonY, car) \rightarrow punish(PersonX)$  discourages the theft. fig. 2 contains an example story for the moral 'be forgiving'.

Section V discusses the story generation and analysis in section VI. Before that discussion, Section II discusses the prevalent narrative generation techniques and why planning is relevant, more than ever, even with the rise of end-to-end narrative generative models. Section III explores plot structure and emotional arc, and section IV discusses character modelling. Finally, section VII concludes the paper with thoughts on future work.

## II. STORY GENERATION STRATEGIES

Even with the advancement in artificial intelligence and natural language processing, narrative generation is still an open task. One of the earliest approaches was developed by Lakoff (Lakoff, 1972), who used rewrite rules to create a grammar of Russian folk tales based on Propp's (Propp, 2010) findings. The story grammar consists of expandable rules coded by the author in the system. In the story grammar, the non-terminals consisted of the chunk of text representing a respective story fragment, with variables for attributes like character names. As it was computationally less expensive, and it could incorporate a story structure, many approaches like (Bringsjord & Ferrucci, 2000)(Machado et al., 2001)(Spierling et al., 2002)(Onuczko et al., 2008)(Pemberton, 1989)(Kybartas & Verbrugge, 2014) used story grammar as a foundation to generate stories. Because the selection of expansion rules was random, and there was hardly any contextual information passing, often the stories turned out to be very incoherent (Black & Wilensky, 1979). Hence, the researchers considered generating stories as problem-solving, to an extent that Meehan (Meehan, 1977) argued that storytelling can be more, but it is at-least problem-solving.

Input: forgive(x,y) → reward(x)

*Output story: Once, there lived two kings Josh and Jacob, with neighbouring kingdoms. The kingdom ruled by Josh was one of the most prosperous kingdoms. So, Jacob envied Josh very much. He used to create troubles for Josh. One day, being tired of him, Josh declared war against him. Rich army of Josh badly defeated Jacob and Josh disarmed him. Josh had an opportunity to kill his troublemaker. However, he did not make the kill as Jacob apologised to him. He forgave Jacob, and they returned to their kingdom.*

*After a few years, a new bloodthirsty king, Lian, rose to power in a far foreign land. Lian was very greedy and wanted the prosperous kingdom of Josh. He was mighty and could beat Josh. So, he attacked Josh. Lian was almost winning the war. But Jacob came to help Josh. Lian could not defeat Josh and Jacob together. They together defeated Lian. Josh was delighted to get help from Jacob. He understood how forgiving Jacob had saved his life and his kingdom. He decided to be more forgiving towards others' mistakes.*

fig. 2 Generated story for the moral 'be forgiving'

Translating the narrative generation into a problem-solving task requires an initial story world state, a goal state, and a knowledge base of events that could transform one state into another. Because the narrative generator had a goal state input from the human author and used a knowledge base crafted by the human author, the narratives are usually coherent and sensible. It is one of the most widely used techniques. Approaches like (Meehan, 1977)(Turner, 1993)(Skorupski & Mateas, 2010)(Gervás et al., 2005)(Lebowitz, 1985)(Pérez y Pérez, 1999)(Bae & Young, 2014)(Barros & Musse, 2008)(Charles et al., 2011)(Riedl & León, 2008)(Peinado et al., 2004) use problem-solving. As the selection of events primarily relied on the goal state and current state, the resultant stories often had inconsistent characters. Hence, researchers considered using autonomous

agents as story characters and building stories from their actions.

Apart from proposing problem-solving, Meehan (Meehan, 1977) was one of the earliest to use agents to build story generation. With a rise in the research of autonomous agents in artificial intelligence and the better applicability of agent-based story generation in interactive environments like games, the story generation based on autonomous agents became popular soon. Approaches like (Meehan, 1977)(Theune et al., 2003)(Fairclough & Cunningham, 2003)(Cai et al., 2010)(Shim & Kim, 2002) used autonomous agents. Often the approaches based on autonomous agents assign a set of goals to the agents and use planning to build the series of actions to achieve the goal, requiring a similarly extensive knowledge base and high computational complexity.

One way to dodge those requirements is by searching in the corpus of stories with techniques like case-based reasoning. However, that often requires exhaustive conditioning on cases to avoid nonsensical generalization and adaptations. Due to advances in neural networks and machine learning, another approach is training and using a neural network to model narrative generation. One of the most intuitive networks for the task would be an already established sequence-to-sequence (seq2seq) neural network(Sutskever et al., 2014). It has attained a significant improvement in multiple language generation fields, like the generation of Wikipedia articles (P. J. Liu et al., 2018), poems (Zhang & Lapata, 2014), including generating narratives from a set of incoherent sentences (Jain et al., 2017). However, it is learning a language model, and it does not differentiate between story content and representation, unlike Chatman's taxonomy (Chatman, 1980). Hence, it has inherent difficulty in capturing the foundational characteristics (D. Liu et al., 2020) like character personalities (Bahamón & Young, 2017), story conflicts (Ware & Young, 2014), or action intents (Riedl & Young, 2010) like their planning counterparts. Also, as it does not keep track of the story world or pass contextual information, most of the stories become incoherent (Fan et al., 2018), have an unexplainable and irrelevant set of events (Yao et al., 2019).

To partially overcome these difficulties, Jian et al. (Jain et al., 2017) used incremental encoding and common-sense graph to generate a logical ending to a story of a few sentences. This strategy takes the previous set of events of the existing story as a context and predicts the next event. Although it only generated a single event, the idea of considering the set of events to make stories more consistent and coherent was intriguing. Later, Fan et al. (Fan et al., 2018) proposed to use a prompt to take a set of words or sentences as an input and generate stories in the context of the prompt. It demonstrated that having a set of plot points increases the probability of the narratives adhering to the original intention. Yao et al. (Yao et al., 2019) went further. They planned a storyline from the given title and used it to generate the story. This strategy provided even better coherence, as the events in the storyline shared context. Lara et al. (Martin et al., 2018) used a similar concept of adopting and integrating the event planning with the seq2seq based model, although they used network instead of planning. They proposed using one seq2seq based recurrent multi-layer encoder-decoder network Event2Event that learns and predicts the events and using another neural network Event2Sentence that generated natural language text for the event generated by Event2Event. These

approaches addressed the event dependency and coherence, while Liu et al. (D. Liu et al., 2020) addressed the issue of character consistency by learning character embeddings like (Oraby et al., 2019). Like Lara et al. (Martin et al., 2018), they built story generation in two phases, action selection, constrained by the learned character embedding, and sentence generation. However, they still suffer from repetition, inconsistencies, coherence problems, and the resultant stories usually do not have a structure and do not make sense as a whole story. Guan et al. (Guan et al., 2020) speculate that this is mainly because of a lack of common-sense knowledge, understanding of causal and temporal relationships between events. They proposed an extended GPT-2 based model that incorporates common-sense knowledge from ATOMIC (Sap et al., 2019) and ConceptNet (Speer et al., 2017) further blurring the line with the conventional story generation approach. The **Error! Reference source not found.** shows an example story generated by their model conditioned to ROCStories Corpus (Mostafazadeh et al., 2016).

Input: [MALE] has been married to his wife for 20 years.

Output story: he was very nervous about his upcoming reunion. he decided to go to an italian restaurant and get some food. the food was good but he could n't decide which dish he wanted. he decided to try a pork dish instead.

fig. 3 Example story from Guan et al. (Guan et al., 2020)

The story in fig. 3 is an excellent example for demonstrating the weakness of the discussed learning models. The story is logical and coherent, yet it lacks a purpose, central message, or moral. Also, the approaches that use neural networks adapt various aspects of planning or autonomous agents (or character profiles) to derive better content. This, with the rise of automatic knowledgebase construction (Bosselut et al., 2020), we believe we are in a situation to afford a more extensive knowledge base to use planning and address underexplored parts of the story, like the moral.

### III. BACKGROUND

#### A. Story Taxonomy

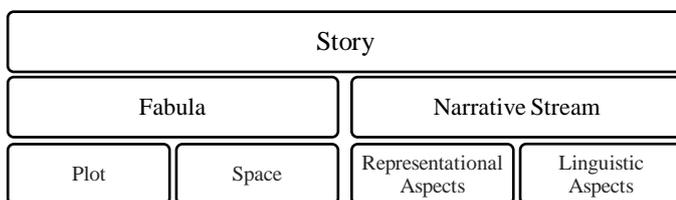


fig. 4 Story taxonomy

Inspired by Chatman's proposal (Chatman, 1980), we define story taxonomy as shown in fig. 4. In the taxonomy, space consists of every element of the story world. It includes all the characters, objects, locations, atmosphere, etcetera. The plot is the action part of the story, where all events exist. Together plot and space create the fabula, which is the main content of the story. On the other hand, the narrative stream defines how the fabula unfolds and is represented to the audience. It includes additional aspects of the story, like representational order of the events, length of the description

of the space elements and events, which events are to be omitted, from whose perspective will the story be told, etcetera. One of the benefits of separating narrative stream from fabula is that by varying narrative streams, many stories can be generated for the same fabula (Callaway & Lester, 2002). Additionally, it also allows the process of fabula generation to be linguistically invariant. SANG considers the taxonomy and generates the fabula in internal representation form. Then, SANG converts fabula to the natural language story. The resultant stories are in simple English as of now. However, the language generation system can be replaced to generate the story in any other language without affecting other parts of the SANG.

#### B. Plot Structure

Aristotle was one of the first to formally study and identify the plot structure in a drama (Butcher & others, 1907). Since then, litterateurs have considered the plot structure an essential part of the plot (Theune et al., 2003). Many plot structures have been proposed and used in the literature. One of the famous plot structures is Freytag's pyramid (Freytag, 1908) as shown in fig. 5.

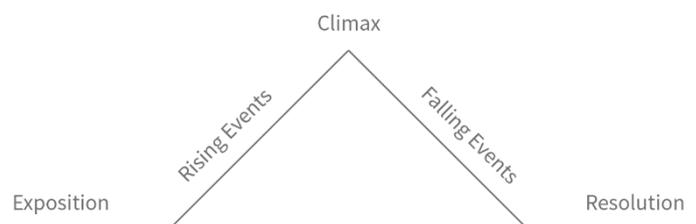


fig. 5 Freytag's plot structure

Irrespective of the shape and position of the components, most plot structures consist of five elements like Freytag's pyramid. The exposition holds the events related to the introduction of the space. It sets up the story world before the plot's main events. On the other hand, resolution pictures the story world after the plot's main events. The climax is where the main events are. It defines the essence of the story theme. Rising and falling events connect exposition with a climax and climax with resolution.

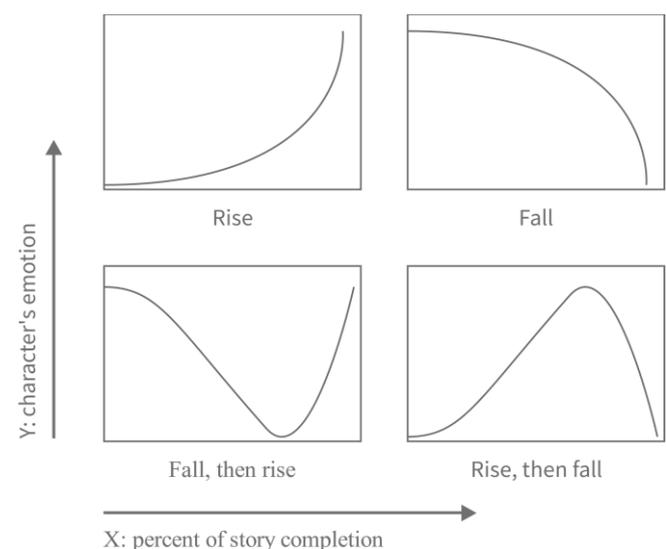


fig. 6 Emotion arcs

Freytag's pyramid may look like a graph, and we can interpret it better as a part of story completion versus tension

it builds in the audience's mind. So, the plot should first introduce the character and then build the tension up gradually, reaching the peak of climax and gradually relieving tension towards resolution. Now, predicting the audience's tension can be complex, as it requires modelling of the audience and their cognitive processes. Instead, we can reduce the problem of predicting the audience's tension into the problem of tracking the protagonist's struggle or emotions. The more the protagonist struggles, the more the tension in the audience's mind (Komeda & Kusumi, 2006). We call the resultant graph emotion arc, inspired by the proposals of (Reagan et al., 2016). Four emotion arcs are shown in fig. 6. At a given time  $\tau$  in the story-world, we compute the character emotions as the signed summation of all the emotion the character is having at  $\tau$ , where positive emotions like joy carry the positive sign, while negative emotions like sorrow carry the negative sign.

#### IV. CHARACTER MODELLING

Characters are a vital part of the stories. The reason why character modelling is so crucial for story development is empathy. Characters in the story allow the audience to see the story world through their eyes. With empathy to one of the characters (usually the protagonist), the audience lives the story, and the events substantially impact the audience's mind. An author can convey the emotions, situations, perspectives, beliefs, etcetera, with the help of the characters, which would have been difficult otherwise.

SANG uses a library of pre-built characters. After planning the story events, it uses the strategy of building prime sets (Khalpada & Garg, 2019) and selecting the character that can satisfy the requirements of the prime trait set. If no such character exists in the library, SANG uses inverse planning to mutate the character from the closest match into a full match.

Often, characters are introduced by SANG to resolve a planning lock, a situation where the planner cannot plan the next event because the requirements of any event cannot be fulfilled with the current story world. Character aspects are maintained in the memory and updated as the story progresses. These aspects are as below.

##### A. Basic Character Data

It includes basic information about a character, like a name, age, health, etcetera. The majority of these parameters remain static during the narrative generation process. The only fields which may require updating during the story generation are the location and health of the character at a particular instance in the story world.

##### B. Emotions

A story with emotionless characters would be more of a report describing events, even if the events are perfectly arranged logically. Modelling emotions can be a complicated job. Rather than exploring the depths and breadths of the complex study of emotions, SANG follows a simple classification of emotions inspired by Plutchik's wheel (Plutchik, 2001). Emotions are stored in a way represented by the fig. 7. (Consider \* as Kleene's star from regular expressions, meaning 0 or more occurrences of the entity.)

<id, emotion, subject, object\*, strength, causal id>

fig. 7 Internal representation of character emotions

Emotions are not independent of each other. It is hard to imagine a person feeling rage and bliss simultaneously. As SANG has a common-sense knowledge base, SANG understands these dependencies. In SANG, the following scenarios can cause a change in the emotion of a character.

- Actions that the character performs. For example, if a person commits a crime, he may regret it and become sad later.
- Actions that other characters perform. For example, if a princess commits a crime, the king may become sad. It depends on the emotional change that action may cause and the emotional bond between the characters. For example, the king is happy with happiness value 2, his emotional bondage with the princess is 8, and the killing action causes sadness of 9. The king is now sad by the value 6. The strength of emotion varies in the range [0-10].
- State of the other character. For example, if a princess dies, the king may become sad. It depends on the property value and the emotional bonding between the characters. It is calculated similarly as in the previous case.
- Change in relationships. For example, a prince may become happier getting married to a princess. It depends on the emotional bond and his emotional value, that is, the emotional bonding of the prince towards the princess and the prince's happiness before the marriage.

##### C. Relationships

Like emotions, relationships between characters are also crucial for a narrative to interest the reader. The different types of relationships are defined in the database of relationship types in the knowledge base. SANG assigns the relationship to the characters and specialize them, i.e., converting them from gender neutralized form to gender-specific form based on the sex of the characters. For example, relation (parent, child) is converted to (father, daughter).

#### V. STORY GENERATION

The overview of the system flow is shown in **Error! Reference source not found.** As discussed in section II, we use planning to generate events. To use planning, we need an initial state of the story-world  $I$  and the final state of the story-world  $F$ . Instead of feeding in the real initial and final story world, we instead use the story-world at the beginning of rising events as  $I$  and end of falling events as  $F$ . This offers finer control over the generation of exposition and resolution and ensures that the resultant story has an exposition and resolution part. If we had opted for the other way, the planner would often omit these parts by considering the subsequent possible causal events.

To derive the pilot events, and hence the story-world,  $I$  and  $F$ , we first determine the emotion arc along with  $I$  and  $F$  from the moral. We then use



- planning to build fabula between them,
- backward chaining to build exposition, and
- forward chaining to build the resolution.

**A. Deriving Pilot Events**

If we classify morals by polarity, we have encouraging and discouraging morals. Encouraging morals motivates positive actions and values, for example, "be forgiving," while discouraging morals demotivates negative actions and values, for example, "do not judge a book by its cover." Each moral can be conveyed in two ways:

- The character exhibiting that trait meets the respective consequence. For example, a person may judge another person based on appearance and lose something important because of such judgment.
- The character not exhibiting that trait misses the respective consequence. For example, a person does not judge another person based on appearance and is rewarded with something important because of his choice of not judging.

Psychologists suggest that appreciation and inspiration are the best way to encourage a person (Dweck, 2008) to act or behave in a certain way. Hence, we associated encouraging morals with the arcs ending on higher emotion while we associate discouraging morals with arcs that end on the lower arc. We then use the emotional arc to guide the planning of events.

The initial point of rising action, or *I*, in the story would be the first time the character shows the value or performs an action encouraged or discouraged by the moral. For example, if the user inputs a discouraging moral like "do not break the rules",

- the *I* would be the character's introduction to the situation where the character is tempted to break the rule.
- the *F* would be character meeting the consequences of breaking the rule.

On the other hand, if the user inputs an encouraging moral like "obey the rules", the *I* would be the same as above, but *F* would change to rewarding the character for obeying the rules. We then use the *I*, the *F*, and the character database to plan a set of events between exposition and resolution.

**B. Planning Events**

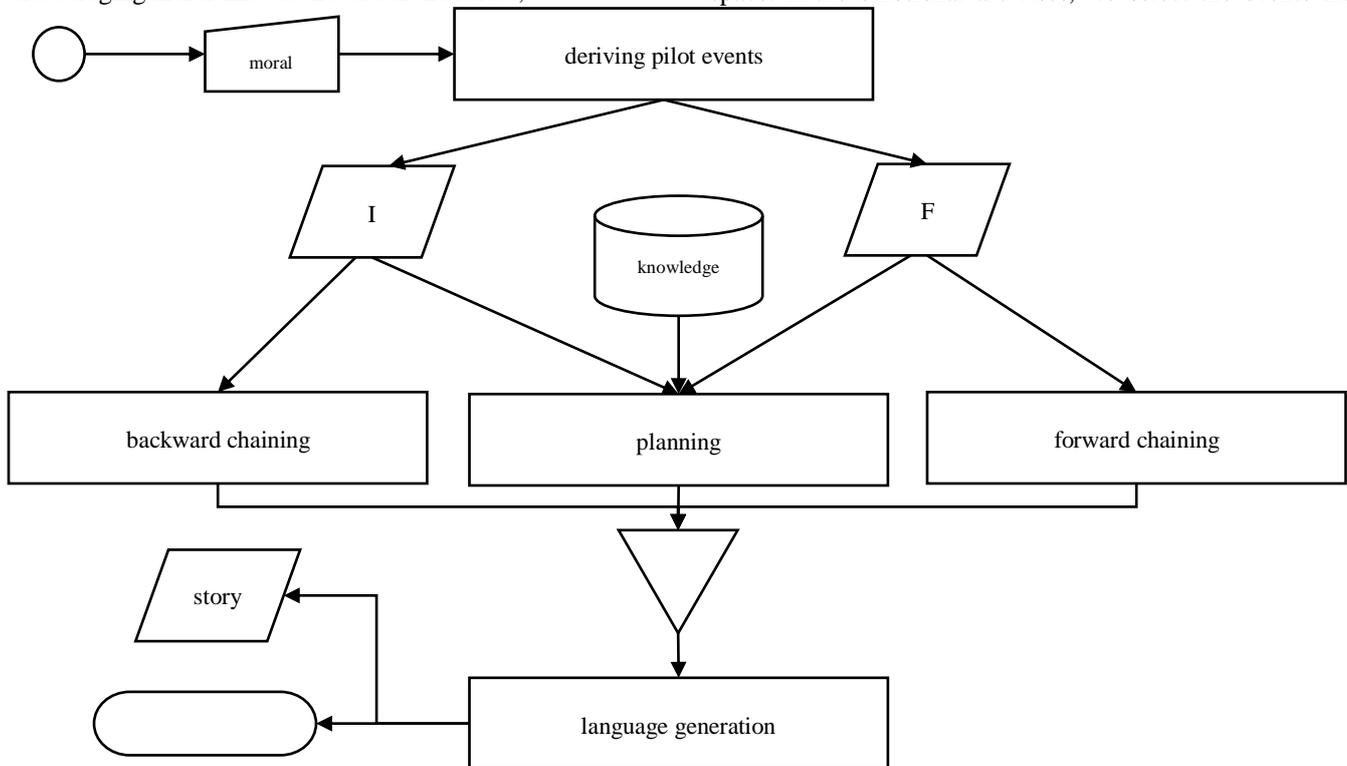
To plan the events, we use Directed Causal Graph (D.C.G.) like in (Khalpada & Garg, 2019), where nodes *n* represent the plot events and edges *e* represent the causal connections between the events. In addition to ConceptNet (Speer et al., 2017), we use CoMET (Bosselut et al., 2020) trained over ATOMIC2020 (Hwang et al., 2021) dataset to supplement the knowledge base crafted in first-order logic.

To ease the process of language generation, we represent the nodes in an internal representation form as shown in fig. 9. (Consider \* as Kleene's star from regular expressions, meaning 0 or more occurrences of the entity.)

```
<id, event, subject*, object*, medium*, strength*, location, causal id>
```

fig. 9 Internal representation of events

While planning, we often have multiple choices for the next possible event that meet the story-world criteria at each node. We use the emotional arc to constraint the exploration space. If the emotional arc rises, we select the events that



trigger or strengthen positive emotions in character. Similarly, if the emotional arc falls, we select the events that trigger or

strengthen negative emotions. Although planning with the emotional arc significantly reduces the search space, it often gets stuck in a planning lock.

A planning lock is a condition when none of the subsequent possible events meets the requirements (story-world criteria or emotion arc requirement). To overcome this, we use an inference mechanism to develop a subplot or prequel to change the current story-world context or the emotional arc to the matching one. For example, if we want Jill to date Jack but cannot meet the requirement of "Jill intending to date Jack", we can use the example inference rules from fig. 10. The result of the rules would be an introduction of a friend of Jill, who would then persuade Jill to date Jack.

<p><b>Problem:</b> ?y.intend(date, ?z)</p> <p><b>Rules:</b> ?y.be(friend, ?x) → ?y.capable(persuade, ?x) ?y.persuade(\$activity, ?x) → ?x.intend(\$activity)</p>
--

fig. 10 Example of an inference rule

### C. Avoiding Repetitive Stories

Another issue that might arise from limited exploration space is repetitive story generation. To avoid that, we remember the edges we traversed for a given story-world context and events and try to avoid that edge next time when an analogous situation arises. Let us assume that, given a story-world context, we are at an event  $e_i$  and we have a set of subsequent possible events  $\{n_1, n_2, n_3, n_4\}$  from which  $\eta = \{n_1, n_2, n_3\}$  that meet the story-world criteria. Suppose planner selects  $n_3$  as the next event, then we remember that for a given story-world context, on event  $e_i$  we selected  $n_3$ . We also maintain the times an edge was selected as  $x$ . For our example, we would increase the  $x$  for  $n_3$  while decreasing the  $x$  for  $n_1$  and  $n_2$  reflecting they both were possible subsequent events that met the story-world criteria, but the planner used  $n_1$  instead. Note that the  $x$  for  $n_4$  would be unaffected because although it was a possible next event, it had not met the story-world criteria, so the planner could not have used it. We then update the likelihoods on the edges  $e_{(n_i, n_k)}$  of all events  $\forall n_k \in \eta$  connecting to  $n_i$  as

$$\rho_{e_{(n_i, n_k)}} = \rho_{e_{(n_i, n_k)}} - \left( \gamma \times \tanh\left(\frac{x_{e_{(n_i, n_k)}}}{|\eta|}\right) \right)$$

where,  $|\eta|$  is the cardinality of the set of subsequent possible events that meet the story-word criteria,  $\gamma$  is the user tunable effective rate, defining how much of the consecutive use of the edge  $e_{(n_i, n_k)}$  would affect the weight. We use  $\tanh$  function to increase the penalty for subsequent edge uses. If any  $n \in \eta$  that is not preferred for a while,  $x$  of it may become negative, resulting in increase of the likelihood of the edge.

### D. Language Generation

SANG uses a rule-based system to generate English sentences from the internal representation format. It first uses context-sensitive grammar-based associations to generate simple sentences. Table 1 shows an example snippet of a fabula in internal representation form and simple story sentences derived from it.

Table 1 Example of simple sentence generation

Events in internal format	Simple sentence
<e1101, see, c1720, c1386.trait[2], l2710>	Jack was in a garden. Jill was in a garden. Jill was beautiful. Jack saw the beauty of Jill.
<f1602, love, c1720, c1386, 8, e1101 & c1720.prefTrait[0]>	Jack liked beauty. So, Jack strongly fell in love with Jill.
<e1102, propose, c1720, c1386, l2710, f1602>	Jack proposed to Jill in a garden because Jack fell in love with Jill.
<e1103, refuse, c1386, e1102, f1530>	Jill refused to Jack proposing to Jill because Jill liked Raj.
<f1603, sad, c1720, 5, e1103>	Jack became sad because Jill refused to Jack proposing to Jill.

SANG then uses language protocols defined in first-order logic from knowledgebase to generate compound and complex sentences. One such example is shown in figure 11.

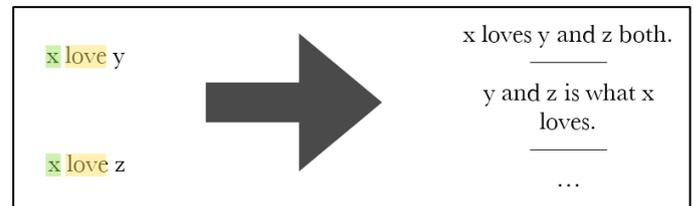


figure 11 An Example of Compound Rule

Table 2 shows an example of applying language protocols on simple sentences from Table 1.

Table 2 Example sentences after applying language protocols

Simple sentences	Rewritten surface sentences
Jack was in a garden. Jill was in a garden. Jill was beautiful. Jack saw the beauty of Jill.	Jack went to a garden, where he saw a beautiful girl, Jill.
Jack liked beauty. So, Jack strongly fell in love with Jill.	Because he always liked beauty, he deeply fell in love with Jill.
Jack proposed to Jill in a garden because Jack fell in love with Jill.	So, he proposed to her.
Jill refused to Jack proposing to Jill because Jill liked Raj.	But, she liked Raj, so she refused.
Jack became sad because Jill refused to Jack proposing to Jill.	Jack became sad.

## VI. EXPERIMENT

To evaluate the narrative generation, we considered the following aspects.

- Does the narrative convey the expected moral?
- How well do the stories with moral fare against baseline stories without moral?

We generated 5 narratives for 7 different morals each to evaluate the narratives. We added 5 random stories generated using (Khalpada & Garg, 2019) without morals to eliminate outliers. We then asked 103 human evaluators to list a maximum of 3 morals they think the narrative conveys. To avoid bias and one-to-one relation, we provided a list of 21 morals and a none option to assign each narrative its closest moral.

Table 3 Human evaluation

Moral	Inter-evaluator agreement	Moral confidence in top 3	Moral as prime (Normalised)
Be forgiving	0.91	0.74	0.81
One should obey the rules.	0.83	0.85	0.89
It is okay to be different.	<b>0.71</b>	0.68	0.74
Do not judge a book by its cover.	0.86	<b>0.43</b>	<b>0.69</b>
One should not lie.	0.94	<b>0.96</b>	0.98
It is wrong to hurt others.	0.91	0.83	0.91
Hard work always pays off.	<b>0.97</b>	0.91	<b>0.99</b>
	<b>0.87</b>	<b>0.77</b>	<b>0.86</b>

Table 3 shows the result of manual evaluation. A strong agreement average of 0.87 between the evaluators imply that they strongly agree with the result. A good average of 0.77 for moral confidence shows that the generated narratives nicely convey the intended moral. This is a good result, even for cherrypicked morals, especially as there are hardly any baseline contenders. To better understand the confidence of the moral in the narrative, we filter the positive samples and examine how many of them have the intended moral with the highest confidence (on top of the list of 3 morals). An average of 0.86 shows strong confidence in the intended moral, suggesting that if an evaluator has associated the intended moral with the narrative, it is highly likely that the intended moral would have the highest confidence.

Inspecting the individual results, we realized that the narratives for the complex morals like 'do not judge a book by its cover' have lower confidence. We suspect this is mainly because of the higher order of exploration the planner may require.  $observe(PersonX, PersonY.attribute1) \rightarrow assume(PersonX, PersonY.attribute2) \rightarrow punish(PersonX)$  requires exploration of the plot events and character attributes that convey the moral. In one generated narrative, because one character had a motorcycle, the protagonist assumed that the character would be short! Although being short and possessing a motorcycle are valid attributes for a character, they do not contrast enough to highlight a stereotype in society.

To understand how well the narratives with morals fare against those without morals, we cherrypicked 5 narratives from the proposed approach and paired them with the other 5 narratives generated using (Khalpada & Garg, 2019). We asked the same 103 evaluators to select a story from the pair they are likely to tell kids. Narratives with the proposed

approach were selected with encouraging confidence of  $\approx 0.71$ .

## VII. CONCLUSION

We proposed a planning-based approach to generate narratives for a given moral. A moral is a vital part of the story. The second experiment shows that narratives with morals are highly likely to be shared with kids. The first experiment shows that the proposed system generates the narrative for intended morals with a high confidence.

## REFERENCES

- Bae, B. C., & Young, R. M. (2014). A computational model of narrative generation for surprise arousal. *IEEE Transactions on Computational Intelligence and AI in Games*, 6(2), 131–143. <https://doi.org/10.1109/TCIAIG.2013.2290330>
- Bahamón, J. C., & Young, R. M. (2017). An empirical evaluation of a generative method for the expression of personality traits through action choice. *Proceedings of the 13th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, AIIDE 2017*, 144–150.
- Barros, L. M., & Musse, S. R. (2008). Towards consistency in interactive storytelling: Tension arcs and dead-ends. *Computers in Entertainment*, 6(3), 7–9. <https://doi.org/10.1145/1394021.1394036>
- Black, J. B., & Wilensky, R. (1979). An evaluation of story grammars. *Cognitive Science*, 3(3), 213–229. [https://doi.org/10.1016/S0364-0213\(79\)80007-5](https://doi.org/10.1016/S0364-0213(79)80007-5)
- Bosselut, A., Rashkin, H., Sap, M., Malaviya, C., Celikyilmaz, A., & Choi, Y. (2020). CoMET: Commonsense transformers for automatic knowledge graph construction. *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, 4762–4779. <https://doi.org/10.18653/v1/p19-1470>
- Bringsjord, S., & Ferrucci, D. a. (2000). Artificial Intelligence and Literary Creativity: Inside the Mind of BRUTUS, a Storytelling Machine. *Computational Linguistics*, 26(4), 642–647. <https://doi.org/10.1162/coli.2000.26.4.642>
- Butcher, S. H., & others. (1907). *The poetics of Aristotle*. Macmillan.
- Cai, Y., Shen, Z., Miao, C., & Tan, A. H. (2010). DIRACT: Agent-based interactive storytelling. *Proceedings - 2010 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2010*, 2(May 2016), 273–276. <https://doi.org/10.1109/WI-IAT.2010.89>
- Callaway, C. B., & Lester, J. C. (2002). Narrative prose generation. *Artificial Intelligence*, 139(2), 213–252. [https://doi.org/10.1016/S0004-3702\(02\)00230-8](https://doi.org/10.1016/S0004-3702(02)00230-8)
- Charles, F., Porteous, J., Teutenberg, J., & Cavazza, M. (2011). Timeline-based Navigation for Interactive Narratives Categories and Subject Descriptors. *8th Conference on the Advances of Computer Entertainment Technology*. <http://www.ace2011.org>

- Chatman, S. B. (1980). *Story and discourse: Narrative structure in fiction and film*. Cornell University Press.
- Dweck, C. S. (2008). *Mindset: The new psychology of success*. Random House Digital, Inc.
- Fairclough, C., & Cunningham, P. (2003). A Multiplayer Case Based Story Engine. *Game-On, Bremond*, 41-.
- Fan, A., Lewis, M., & Dauphin, Y. (2018). Hierarchical neural story generation. *ACL 2018 - 56th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference (Long Papers), 1*, 889–898. <https://doi.org/10.18653/v1/p18-1082>
- Freytag, G. (1908). *Freytag's technique of the drama: an exposition of dramatic composition and art*. Scott, Foresman and Company.
- Gervás, P., Díaz-Agudo, B., Peinado, F., & Hervás, R. (2005). Story plot generation based on CBR. *Knowledge-Based Systems, 18*(4–5), 235–242. <https://doi.org/10.1016/j.knosys.2004.10.011>
- Gottschall, J. (2012). *The storytelling animal: How stories make us human*. Houghton Mifflin Harcourt.
- Guan, J., Huang, F., Zhao, Z., Zhu, X., & Huang, M. (2020). A Knowledge-Enhanced Pretraining Model for Commonsense Story Generation. *Transactions of the Association for Computational Linguistics, 8*, 93–108. [https://doi.org/10.1162/tacl\\_a\\_00302](https://doi.org/10.1162/tacl_a_00302)
- Herman, D. (2013). *Storytelling and the Sciences of Mind*. MIT press.
- Hwang, J. D., Bhagavatula, C., Le Bras, R., Da, J., Sakaguchi, K., Bosselut, A., & Choi, Y. (2021). (Comet-) Atomic 2020: On Symbolic and Neural Commonsense Knowledge Graphs. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(7), 6384–6392.
- Jain, P., Agrawal, P., Mishra, A., Sukhwani, M., Laha, A., & Sankaranarayanan, K. (2017). *Story Generation from Sequence of Independent Short Descriptions*. <https://doi.org/10.475/123>
- Khalpada, P., & Garg, S. (2019). Balancing Consistency and Plot Structure in Computational Storytelling. *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2019*, 550–555. <https://doi.org/10.1109/IEMCON.2019.8936138>
- Khalpada, P., & Garg, S. (2021). Simple Automated Narrative Generator (SANG). *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, 906–914.
- Komeda, H., & Kusumi, T. (2006). The effect of a protagonist's emotional shift on situation model construction. *Memory & Cognition, 34*(7), 1548–1556.
- Kybartas, B., & Verbrugge, C. (2014). Analysis of ReGEN as a graph-rewriting system for quest generation. *IEEE Transactions on Computational Intelligence and AI in Games, 6*(2), 228–242. <https://doi.org/10.1109/TCIAIG.2013.2290088>
- Lakoff, G. P. (1972). Structural Complexity in Fairy Tales. In *The Study of Man* (Vol. 1, pp. 128–150).
- Lebowitz, M. (1985). Story Telling as Planning and Learning. *Poetics, 14*(6), 483–502.
- Liu, D., Li, J., Yu, M.-H., Huang, Z., Liu, G., Zhao, D., & Yan, R. (2020). A Character-Centric Neural Model for Automated Story Generation. *Proceedings of the AAAI Conference on Artificial Intelligence*. [www.aaai.org](http://www.aaai.org)
- Liu, P. J., Saleh, M., Pot, E., Goodrich, B., Sepassi, R., Kaiser, L., & Shazeer, N. (2018). Generating wikipedia by summarizing long sequences. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*, 1–18.
- Machado, I., Paiva, A., & Brna, P. (2001). Real characters in virtual stories promoting interactive story-creation activities. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2197*, 127–134. [https://doi.org/10.1007/3-540-45420-9\\_14](https://doi.org/10.1007/3-540-45420-9_14)
- Martin, L. J., Ammanabrolu, P., Wang, X., Hancock, W., Singh, S., Harrison, B., & Riedl, M. O. (2018). Event representations for automated story generation with deep neural nets. *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, 868–875.
- Meehan, J. (1977). TALE-SPIN, An Interactive Program that Writes Stories. *IJCAI, 77*, 91–98.
- Mostafazadeh, N., Chambers, N., He, X., Parikh, D., Batra, D., Vanderwende, L., Kohli, P., & Allen, J. (2016). A corpus and cloze evaluation for deeper understanding of commonsense stories. *2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT 2016 - Proceedings of the Conference*, 839–849. <https://doi.org/10.18653/v1/n16-1098>
- Onuczko, C., Szafron, D., & Schaeffer, J. (2008). Stop Getting Side-Tracked. *AI Game Programming Wisdom, 4*, 513–528.
- Oraby, S., Reed, L., Tandon, S., T.S., S., Lukin, S., & Walker, M. (2019). *Controlling Personality-Based Stylistic Variation with Neural Natural Language Generators*. July, 180–190. <https://doi.org/10.18653/v1/w18-5019>
- Peinado, F., Gervás, P., & Díaz-Agudo, B. (2004). A Description Logic Ontology for Fairy Tale Generation. *International Conference on Language Resources and Evaluation: Workshop on Language Resources for Linguistic Creativity*, 56–61.
- Pemberton, L. (1989). A modular approach to story generation. *Fourth Conference of the European Chapter of the Association for Computational Linguistics*, 217–224. <https://doi.org/10.3115/976815.976845>
- Pérez y Pérez, R. (1999). *MEXICA: A Computer Model of Creativity in Writing*. 245.
- Plutchik, R. (2001). The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical

- practice. *American Scientist*, 89(4), 344–350.
- Propp, V. (2010). *Morphology of the Folktale*. University of Texas Press.
- Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1), 1–12.
- Riedl, M. O., & León, C. (2008). Toward Vignette-Based Story Generation for Drama Management Systems. *Workshop on Integrating Technologies for Interactive Stories*, 23–28.
- Riedl, M. O., & Young, R. M. (2010). Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39, 217–268. <https://doi.org/10.1613/jair.2989>
- Sap, M., Le Bras, R., Allaway, E., Bhagavatula, C., Lourie, N., Rashkin, H., Roof, B., Smith, N. A., & Choi, Y. (2019). ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 3027–3035. <https://doi.org/10.1609/aaai.v33i01.33013027>
- Shim, Y., & Kim, M. (2002). Automatic short story generator based on autonomous agents. *Pacific Rim International Workshop on Multi-Agents*, 151–162.
- Skorupski, J., & Mateas, M. (2010). Novice-friendly authoring of plan-based interactive storyboards. *Proceedings of the 6th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment, AIIDE 2010*, 174–179.
- Speer, R., Chin, J., & Havasi, C. (2017). Conceptnet 5.5: An open multilingual graph of general knowledge. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1).
- Spielring, U., Grasbon, D., Braun, N., & Iurgel, I. (2002). Setting the scene: Playing digital director in interactive storytelling and creation. *Computers and Graphics (Pergamon)*, 26(1), 31–44. [https://doi.org/10.1016/S0097-8493\(01\)00176-5](https://doi.org/10.1016/S0097-8493(01)00176-5)
- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, 4(January), 3104–3112.
- Theune, M., Faas, S., Faas, E., Nijholt, A., & Heylen, D. (2003). The Virtual Storyteller: Story Creation by Intelligent Agents. *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference*, 204–215. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.12.5222>
- Turner, S. R. (1993). MINSTREL: A computer model of creativity and storytelling. In *ProQuest Dissertations and Theses* (pp. 841-841 p.). <http://search.proquest.com/docview/304049508?accountid=43623>
- Ware, S., & Young, R. M. (2014). Glaive: a state-space narrative planner supporting intentionality and conflict. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 10(1).
- Yao, L., Peng, N., Weischedel, R., Knight, K., Zhao, D., & Yan, R. (2019). Plan-and-Write: Towards Better Automatic Storytelling. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33, 7378–7385. <https://doi.org/10.1609/aaai.v33i01.33017378>
- Zhang, X., & Lapata, M. (2014). Chinese poetry generation with recurrent neural networks. *EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 670–680. <https://doi.org/10.3115/v1/d14-1074>



# An Approach to E-Commerce with Voice Recognition

DEBANGI CHOUDHURY,  
UEM Kolkata,  
Email: debangiofficial@gmail.com

RIYA SAHA,  
UEM Kolkata,  
Email: riya02saha@gmail.com

ANUSHKA DEY,  
UEM Kolkata,  
Email:  
anushkadeyapd@gmail.com

TIYASHA DEY,  
UEM Kolkata,  
Email: Tiyashadey199@gmail.com

SRILEKHA MUKHERJEE,  
UEM Kolkata,  
Email: srilekha.mukherjee@uem.edu.in

**Abstract:** E-Commerce or Electronic Commerce is process of doing business through computer networks virtually. A person sitting in front of his computer can access all facilities to buy and sell products. E-commerce has reduced human effort by reducing physical work and saving time of the seller and buyer both. The main advantage of E-commerce is that the user can browse online shops and compare prices of products sitting at home in their comfort. The buyer/consumer surfs through the internet to the seller's website and views the products sold by them. There the buyer compares the price and selects items that he/she wants to buy and then proceeds with transaction.

Most E-commerce websites are limiting consumer's senses to only viewing. Shopping can be made more convenient if the buyer could communicate with the website just by verbally asking for the product and receiving audio + visuals as a result. Here comes the role of AI powered voice recognition. This is the problem we aim to solve to make the user experience better by combining two very popular technologies that are Reactjs and Alan AI. The power of React makes a website faster. React allows developers to create large web applications that can change data, without re-loading the page. The main purpose of React is to be fast, scalable, and simple. Alan AI is an advanced Voice AI Platform that allows us to add a voice interface to our app without overhead.

## INTRODUCTION:

Developing a web site for the Internet (World Wide Web) or an intranet (a private network) is called Web Development [1]. Web development ranges from developing a simple static page to complex Web applications. Web development also commonly refers to Web design, Web engineering, client liaison, Web content development, client-side/server-side [2] scripting, network security configuration, Web server and e-commerce development. Web development is basically the collaboration between departments and not the domain of a designated department. Three kinds of Web developer specialization include front-end developer, back-end developer and full-stack developer. Responsibility of the Front-End Developers is behaviour and visuals that run in the user browser, while back-end developers are responsible to deal with the servers. Since the time the web is commercialised, Web development has been a rapid growing industry. Development industry is growing rapidly because every business wishes to use their Website for advertising and selling their products and services to customers.

Some of the many open source tools for Web development are GlassFish, BerkeleyDB, stack, LAMP (Linux, Apache, MySQL, PHP) and Perl/Plack. Cost of learning Web development is minimum because of these. Easy-to-use WYSIWYG (What You See Is What You Get) Web-development software, such as BlueGriffon, Adobe Dreamweaver and Microsoft Visual Studio are another contributing factor to the



growth of the industry. It is still required to have knowledge of Hypertext Mark-up Language (HTML) [3] or of programming languages to use such software, but the basics can be learned and implemented quickly.

Online shopping [4] could be a sort of e-commerce that permits customers to directly purchase product or services from any vendor over the net employing a mobile app or an online Browser. customers notice a product of interest by visiting the web site of the distributor directly or by looking among multiple sellers employing a program, that displays the a product's convenience and valuation at completely different e-retailers. As of 2020, customers will search on-line employing a vary of various computers and devices, as well as desktop computers, laptops, pill computers and smartphones. Internet search evokes the physical analogy of searching for product or services at an everyday "bricks-and-mortar" distributor or searching centre; the method is named business-to- consumer (B2C) online shopping. Once an internet store is ready up to modify businesses to shop for from another business, the method is named business-to-business (B2B) [5] on-line searching. A typical on-line store allows the client to browse the firm's vary of product and services, read photos or pictures of the product, alongside info regarding the merchandise specifications, options and costs.

Online stores sometimes alter shoppers to use "search" options to search out specific models, brands or things. Online customers should have access to the net and a legitimate methodology of payment so as to finish dealings, like a master card, AN Interact-enabled revolving credit, or a service like PayPal. For physical product (e.g., paperback books or clothes), the e-tailer ships the product to the customer; for digital product, like digital audio files of songs or computer code, the e-tailer sometimes sends the file to the client over the net. The biggest of those on-line selling companies square measure Alibaba, Amazon.com, and eBay.

## LITERATURE SURVEY:

Online shopping is basically electronic commerce to purchasing products or services directly from the seller through internet. With time, more people than before are using online shopping for an extensive variety of items, from household essentials to designer dresses and shoes to airplane tickets. Now purchasers have multiple options to choose their products and services while they are shopping through online platform.

Online searching has several distinctive characteristics. Huseynov and Yildurm (2014) [6] highlighted that the shortage of physical interaction tends to be the important obstacle in on-line retail sales followed by the privacy of individual info and security of monetary transactions over the net. Demangeot and Broderick (2010) [7] conjointly discovered that perceived simple use doesn't have an effect on the behavioural pattern during this case rather influenced by security and privacy problems. No relationship is constructed between the client and therefore the on-line look within the presence of perceived on-line risk although a client spent hours on the net Zuroni & Goh (2012) [8]. Day-by-day style, preference and selections are varied relating to various factors like the net emergence. However, this development desires some a lot of understanding associated with the consumer's behaviour. shopper behaviour analysis identifies a general model of shopping for behaviour that depicts the processes employed by customers in creating a buying deal call (Vrender, 2016) [9]. Those styles are preponderant to the merchandiser as they'll make a case for and predict shopper purchase behaviour.

Jarvenpaa, and Todd (1997) [10] projected a model of perspective, behaviour, and searching intention towards net searching normally. The look includes many indicators classified into four broad classes like product price, quality services offered through the web site, the searching expertise, and also the risk perception of the web searching. Chang, Cheung, and Lai (2005) [11] studied classes of variables, that drive on-line searching activity. In their study, they divided the options into 3 broad



classes. Perceived characteristics of the online sale channel square measure the primary one which incorporates risk, on-line searching experiences, advantage, service quality, trust. The second class may be a web site and merchandise options that square measure risk reduction measures, website options, and merchandise characteristics; and also the last cluster is client characteristics. Varied styles of options, demographic variables, client searching orientations, client originality and psychological variables, computer, net information, and usages drives client characteristics.

Riley, Scarpi, and Manaresi (2005) [12] projected that consumer-related factors have an effect on on-line getting and therefore the ensuing implications for on-line retailers. The authors give sensible suggestions for retailers to cut back or overcome a number of the barriers that stop customers from increasing the number of product purchased on-line. Their work provides a radical analysis of 3 key factors which will influence shopper purchase behaviour on the net. These factors embody product-related factors (e.g., product kind, brand name, etc.), consumer-related factors (e.g., shopper experience, attitudes, risk perceptions, searching orientation, etc.), and retail-related factors (e.g., ways and tactics). Finally, the paper mentioned however e-retailers selling efforts (retailer factors) may be accustomed overcome the barriers to net getting ensuing from specific product and shopper connected characteristics. Clear and simple to implement recommendations to managers offered.

Rahman, et.al (2018) [13] explored the most factors, that issues the on-line customers to buy on-line and the way these factors influence the customers, once looking on-line. The paper emphasised on why customers hesitate or delay whereas looking on-line. The study known worth, confidence, security, convenience, time, when sale service and discounted deals because the main influencing factors. The paper known anxiety of sharing personal and money info collectively of the prime obstacles on the means of on-line looking. One study has known the impact of demographic factors (age, gender,

education, income) on on-line looking behaviour (Possession of net, Frequency of on-line purchase, Motivation drives for on-line purchase) of customers. The paper concludes that gender has important impact on on-line looking behaviour.

Another work (2012) [14] analyzed the factors moving on on-line searching behaviour of shoppers which may be one amongst the foremost vital problems with e-commerce and selling field. It absolutely was complete that age and education have sturdy correlational statistics with perspective to on-line searching. The paper considers web site design/features because the most influencing factors. Also, demographic variables like age, gender, and level of financial gain play a facilitating role as a result of they influence shopper perception and shopper behaviour that drives them towards on-line searching Kim, Zhao, & Yang, 2008; Laforet & Li, 2005; Sabbir Rahman (2012) [15]. In China, on-line searching intention depends on consumers' age, income, and education additionally as legal status most significantly their perceived utility Gong, Stump, & Maddox (2013) [16].

All the views given by the analyst above mentioned shows how it can expand and become more profitable as it holds a definitive power over the consumers which influence the market.

## BACKGROUND STUDY:

Problems faced in e-commerce websites and companies:

- **It is all Visual**

E-commerce websites are mostly visual in nature, which can be a disadvantage for visually unstable eyes or old people who have trouble viewing the product and trouble in reading and writing in digital devices. Since the popularity of E-Commerce website is increasing day by day, we should look into making it feasible for all spectrums of people.



- **An absence of online identity verification**

When a customer visits an e-commerce website, the portal has no clue about the customer and the information they have entered. The customer's information provided is real or not remains doubtful. Ordering something on Cash on Delivery (COD) payment mode using fake phone numbers or non-existing addresses can lead to large revenue losses.

- **Customer experience on delivering Omni channel**

Nowadays, customers can reach out to us through many numbers of times. They can visit to our website, contact with a agent, drop a message on your facebook/instagram or any other page or connect with you through a live chat session.

- **Competitor Analysis**

In today's competitive world, all other shopping aids will offer the same things and services as us. If we don't have any great strategy, it will be difficult for us to survive.

- **Stuck in at the old-school way of approach to selling**

The reason several e-commerce firms realize on-line marketing therefore troublesome is that they're, ironically, stuck within the past. Most of them lack the required insight into client behaviour and shopping for patterns, knowledge which may facilitate them thrive within the current e-commerce atmosphere.

- **Shopping cart abandonment**

Shopping cart abandonment could be a vast issue. Even e-commerce giants aren't proof against this drawback. For example, once brick Associate in nursing mortar heavyweight Nordstrom started an e-commerce portal, they witnessed massive losses from abandoned carts. The tedious and bug-filled checkout method was inflicting customers to escape in their droves. Nordstrom had to come back up

with a brand new checkout style, turning it into a ballroom dancing method.

- **Maintaining customer loyalty**

Even with the best-designed web site out there, while not client trust and loyalty, the business is absolute to struggle.

Creating new customers then maintaining them needs a vast effort. One in all the explanations e-commerce corporations particularly faces a challenge in building client trust and loyalty is that the merchandiser and client don't understand one another. Nor will they see one another. Thus, the client is robbed of the senses they'd usually suppose in face to face transactions. This could solely be created up for through time and energy. Across multiple transactions, eventually, the corporate will build thistrust and loyalty.

- **The headache of product return and refund**

When a product is came back, whether or not thanks to a disgruntled client or broken product, the business suffers an important loss in cargo and name. Logistical and shipping prices have perpetually been problematic to e-commerce sellers delivering their product for gratis.

- **The struggle of competing on price and shipping**

Online merchants oftentimes vie on value. Lots of sellers list constant merchandise on their sites, however the costs could also be completely different. Price cutting war notably affects little e-commerce businesses, as mid-sized and enormous competitors will typically supply merchandise for fewer. Mix this with free shipping, and smaller corporations merely can't afford to vie on value.

- **A problem of data security**

Security problems will result in nightmare eventualities. Fraudsters post spam and attack the online host server, infecting the websites with viruses. They'll doubtless gain access to confidential information



concerning your customer's phone numbers, card details, etc.

## **PROPOSED SOLUTION:**

### ▪ **Introducing Voice Interface**

As we discussed in problem statement of the background study, E-commerce has to be made feasible for a large spectrum of people including the visually impaired or old people who have troubled eyesight, and also face problem in reading and writing in digital devices.

We introduce something very interesting that is Voice interface. Although many e-commerce websites already have Voice Interface but we are unique in a way that we are not only searching products through vocal instructions but also listening to the product description, and listening reviews and even writing reviews all through voice.

### ▪ **Solution for an absence of online identity verification**

First of all, look out for signs of suspicious activity. This might take the shape of significantly high worth or giant orders, determine pretend phone numbers and email addresses, check whether or not nothing codes match with the state/city. Send a verification link once a client signs up, via text message or email, to validate the client is real. With COD purchases, an automatic decision might even dial bent the client, asking them to validate the delivery address.

### ▪ **Delivering an Omni channel customer experience**

Make sure to equip your team with the correct technology. Up to date, visual engagement tools change your organization to serve customers across all bit points, channels, and journeys. Identify the key channels. Verify that channels square measure most vital to your customers. Client support employees ought to contact customers via their most well-liked channels, phone, email, live chat, video call, on-line facilitate centres or in-app electronic messaging. Integrate and optimize those channels, adding personalised messages and providing matched interaction with live chat or video job. Maintain the context. Direct the conversations supported a user's previous response. Keeps a track of client

conversations victimization parameters like user profile. This way, you'll be able to invariably respond contextually, regardless of the channels they used.

### ▪ **Competitor Analysis**

Conduct thorough analysis into competitors. Place stress on developing a method enabling you to shine brighter than your competitors. Use social media platforms and blogs for promoting merchandise. Invest in promotional offers to assist produce additional net presence and thus additional customers.

### ▪ **Stuck in at the old-school way of approach to selling**

Offer product in distinguished marketplaces like Amazon, eBay, etc.

These e-commerce sites have already got an existing network of consumers thus pitching and stigmatization your product becomes comparatively simple.

### ▪ **Segment visitors**

Visitor segmentation permits e-commerce firms to spot and communicate with guests supported their customer journey, past conversations, geographical location, browsing behaviour, referral page, and far additional. In turn, e-commerce firms will anticipate customer desires, giving them a customized service, like (free delivery or promo codes), making a additional economical stream of conversions.

### ▪ **Shopping cart abandonment**

Consider redesigning your go-cart, ensuring there aren't any bugs or an unnecessarily long and frustrating type filling method. Live chat solutions scale back go-cart abandonment instances. Combining this with options like traveller segmentation, it's doable to proactively reach resolute customers throughout key stages of their journey, providing the chance to interact with customers World Health Organization would otherwise drop out while not getting. Visual tools will facilitate address client queries throughout the checkout method.

### ▪ **Maintaining customer loyalty**

You must check that the client is happy with



the complete method, from ordering on-line to shipping. many on-line retailers is also commercialism a similar product, therefore you need to determine competitive advantage and nurture your client service consequently.

Increase trust with guests by:

Displaying your address, signalling, photos of employees, client testimonials and believability badges on your web site.

Add a live chat choice to the web site.

Create blogs, they assist build trust. Create client service a priority over profit. Remember, it's easier to take care of Associate in nursing existing client than to seek out a replacement one.

Create loyalty programs. The points can't be transferred to alternative firms, that the client can get topurchase from you.

- **The headache of product return and refund**

Return and refund also are a part of nice client service.

It would be an enormous mistake to forget that. the most effective issue you'll do is build a robust returns policy. Take into account the subsequent once planning it:

Be clear. Never hide your policy.

Use plain English, intelligible to everybody. Not all of your customers are clued up with the word.

Try avoiding phrases like, "you must" 'and "you square measure required", which can appear harsh and suspend potential customers.

Outline what they will expect from you. Offer completely different choices for payments and shipping.

Educate employees regarding your come policy, in order that they will assist customers effectively.

Be ready to face the value of your mistakes.

If the merchandise is shipped wrong, take additional effort to stay the client happy.

- **The struggle of competing on price and shipping**

To survive in a competitive market, e-commerce companies need to distribute their inventory to fulfilment warehouses, become extremely resourceful shippers, or find some unique products to minimize this problem. Ultimately every online shopper expects free and fast shipping at the lowest price.

- **Competing against retailers and manufacturers**

Stopping makers marketing product on to customers might not be potential, however there ar a couple of ways to attenuate the matter.

Give priority to makers less probably to sell on to customers.

Offer the merchandise at a cheaper price or with extra advantages to extend sales.

Restrict the manufacturer from marketing the merchandise on to customers by setting this call at the contract. It'll be insufferable for each manufacturer, however you'll be able to work with smaller makers.

- **A problem of data security**

Manage your own servers.

It is suggested not to use common FTP to transfer files.

If the patron desires something, they're going to initial begin looking on the web once a client desires one thing, their initial port of decision is that the net. Though E-commerce transactions have enlarged grownup at a crazy pace at an improbable rate since its birth, the competition to put your product on the highest and to create a gentle sale has additionally enlarged then has the competition to create the simplest use of it.



## **TECHNOLOGY REQUIREMENT**

*Technologies that may be required are:*

- MERN Stack
- Alan AI

*Machine Learning components that may be used are:*

- Recommendation System
- Sentiment Analysis
- Voice Interface

**Recommendation system** sometimes build use of either or each cooperative filtering and content- based filtering (also called the personality-based approach),as well as alternative systems like knowledge-based systems. cooperative filtering approaches build a model from a user's past behavior (items antecedently purchased or hand-picked and/or numerical ratings given to those items) additionally as similar selections created by alternative users. This model is then went to predict things (or ratings for items) that the user might have AN interest in. Content-based filtering approaches utilize a series of distinct, pre-tagged characteristics of AN item so as to advocate extra things with similar properties. Current recommender systems generally mix one or additional approaches into a hybrid system.

**Sentiment analysis** (also called opinion mining or feeling AI) refers to the employment of language process, text analysis, linguistics, and statistics to consistently establish, extract, quantify, and study emotional states and subjective data. Sentiment analysis is wide applied to voice of the client materials like reviews and survey responses, on-line and social media, and care materials for applications that vary from selling to client service to clinical drugs.

**Voice Interface** helps the consumer by making the e-commerce experience easier by voice searches and adding items to cart, learning more about the item, listening to product reviews, just by vocal instructions.

## **CONCLUSION**

In this present work, our study on most common online user activities showed that most of the respondents use Internet mainly for communication, social networking and multimedia and entertainment reasons, while shopping online is still finds a lower rank in the list. This result is quite similar with various published statistics in this regard, clearly showing, that even if online shopping is seeing rapid growth in India, it is still to become the most favourite among majority of the online user. More focus in building trust in the area of order fulfilment and quality assurance will definitely help to build customer confidence on online sellers and will boost its growth further. In this context cash- on-delivery payment should be provided in more variety of products with greater geographical coverage, as it has already given huge success to number of online sellers in Indian market.

We introduce something very interesting that is Voice interface. Although many e-commerce websites already have Voice Interface but we are unique in a way that we are not only searching products through vocal instructions but also listening to the product description, and listening reviews and even writing reviews all through voice. It was also found that online shoppers are often worried about security concerns while making payments through their debit / credit cards or net banking facilities, as the security of their accounts may be compromised. But newer technologies like digital wallet or e-cash services are having fewer risks involved, since bank account details of the customers are not used every time they make a payment online. Though internet is available to a major part of the country's mobile users, its speed is still a major setback, while performing online transactions. To overcome such setback the e commerce web sites and shopping apps should be designed to work in low bandwidth network also.



## REFERENCES

1. Ceri, S.; Daniel, F.; Matera, M. & Facca, F. M. (2007), 'Model-driven development of context-aware Web applications', *ACM Trans. Internet Technol.* 7, 1 - 33.
2. Das, O. & Das, A. (2008), Performability Evaluation of Mobile Client-server Systems, in 'Proceedings of the 2008 ACM Symposium on Applied Computing', ACM, New York, NY, USA ,pp. 2197--2201.
3. van Kesteren, A. & Stachowiak, M. (2007), 'HTML Design Principles', World Wide Web Consortium, Working Draft WD-html-design-principles-20071126 .
4. (2018), Online Shopping., in Reda Alhadj & Jon G. Rokne, ed., 'Encyclopedia of Social Network Analysis and Mining. 2nd Ed.', Springer, .
5. Cullen, A. J. & Webster, M. (2007), 'A model of B2B e-commerce, based on connectivity and purpose', *International Journal of Operations & Production Management* 27 (2), 205--225.
6. Huseynov, F. & Yld r m, S.. (2014). Internet users' attitudes toward business-to-consumer onlineshopping: A survey. *Information Development.* 32. 10.1177/0266666914554812.
7. Demangeot, Catherine & Broderick, Amanda. (2010). Consumer Perceptions of Online Shopping Environments: A Gestalt Approach. *Psychology and Marketing.* 27. 117 - 140. 10.1002/mar.20323.
8. Zuroni, M. J., & Goh, H. L. (2012). Factors influencing consumers' attitude towards e-commerce purchases through online shopping. *International Journal of Humanities and Social Science*, 2(4), 223--230
9. Vrender. (2016). Importance online shopping. Retrieved May 17, 2016, from <http://www.scooparticles.com/shopping-articles/clothing-articles/importanceonline-shopping-1495828.html>
10. Jarvenpaa, S. L., Todd, P. A., Jarvenpaa, S. L., & Todd, P. A. (1997). Consumer reactions to electronic shopping on the world wide web
11. Chang, M. K., Cheung, W., & Lai, V. S. (2005). Literature derived reference models for the adoption of online shopping. *Information and Management*
12. Riley, Scarpi, Manaresi (2005). Drivers and Barriers to Online Shopping: The Interaction of Product, Consumer, and Retailer Factors, In I Clark III, & T.B. Flaherty. *Advances in electronic marketing* (pp 45-66). Hershey PA: Idea Group Publishing.
13. Rahman, Mohammad Anisur & Islam, Md. Aminul & Esha, Bushra & Sultana, Nahida & Chakravorty, Sujan. (2018). Consumer buying behavior towards online shopping: An empirical study on Dhaka City, Bangladesh. 5. 1-22. 10.1080/23311975.2018.1514940.
14. Liao, S. H., Chu, P. H., Chen, Y. J., & Chang, C. C. (2012). Mining customer knowledge for exploring online group buying behavior. *Expert Systems with Applications*, 39(3), 3708--3716. doi:10.1016/j.eswa.2011.09.066
15. Sabbir Rahman, M. (2012). Dynamics of consumers' perception, demographic characteristics and consumers' behavior towards selection of a restaurant: An exploratory study on Dhaka city consumers. *Business Strategy Series*, 13(2), 75--88. doi:10.1108/17515631211205488
16. Gong, Wen & Stump, Rodney & Maddox, Lynda. (2013). Factors influencing consumers' online shopping in China. *Journal of Asia Business Studies.* 7. 10.1108/JABS-02-2013-0006.

# Study on the limitations of Stacking Technique for Bandwidth Improvement of Microstrip Patch Antennas

Dipankar Saha  
University of Engineering & Management, Kolkata

Mandar Chakrabarti\*  
University of Engineering & Management, Kolkata  
mandarc02@gmail.com

Abir Chattopadhyay  
University of Engineering & Management, Kolkata

**Abstract**—Stacking of parasitic patch layers on rectangular micro strip patch antennas and its effect on the bandwidth and gain of the antenna are analyzed. For that two different frequencies 2.4 GHz and 11.5 GHz have been chosen from S band and X band respectively and stacking of patch layers have been done up to a certain limit assuming the minimum requirement of gain of 10 dB for S band antenna array and 12 dB for X band antenna array. So in both the cases, at first a 2x2 rectangular patch antenna arrays have been optimized to reach up to the gain requirement and then the stacking of parasitic layers have been done and results are checked and analyzed one by one.

**Keywords**— Patch antenna, Micro strip antenna, Bandwidth Improvement, Stacking of Patches, S band , X band , Antenna array, Parasitic patch layer.

## 1. INTRODUCTION

Patch antenna is one of the most popular and most commonly used micro strip antennas. It is widely used in portable wireless devices for the ease of fabricating it on printed circuit boards. The micro strip is actually a very thin metallic strip placed on a ground plane with a dielectric material in between patch and ground plane. When the antenna is excited, the electromagnetic waves generated within the dielectric undergo reflections and the energy is radiated from the edges of the metal patch. These multiple patch antennas can be used to make a high gain antenna array. Micro strip patch antenna has lower fabrication cost, smaller size, lower mass and they are even capable of dual or tri band operations. But using micro strip patch antenna provides also some challenges towards us as a form of its lower gain, lower power handling capability, inherently lower band width, lower efficiency due to dielectric loss and conductor loss etc.

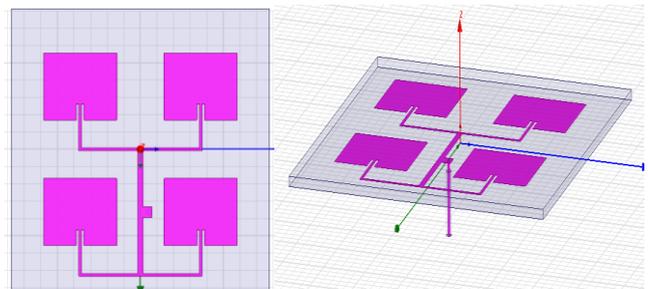
There are different bandwidth improvement techniques already available to us to improve the bandwidth of a microstrip patch antenna. Some of the popular bandwidth improvement techniques are aperture coupled, sequentially rotated array, stacking of patches, proximity coupled etc. Aperture coupled and proximity coupled are actually two different feeding techniques to the microstrip patch antenna. They are also called non-contacting feeding because in those cases the feed is not directly connected with the patch itself or via any usual microstrip line. In case of sequentially rotated array, the array elements are nothing but just rotated at a particular angle value to enhance the bandwidth of the antenna array though it is not a very effective way of

bandwidth improvement. And stacking of patches is what the domain of our interest here is to work on. It is a layer of just metallic patches, printed on a substrate layer and those are not connected to any of the feed. Each of those patches are positioned aligned with antenna axis for each of the radiating antenna elements of its lower substrate level. These types of patches are also known as parasitic patches.

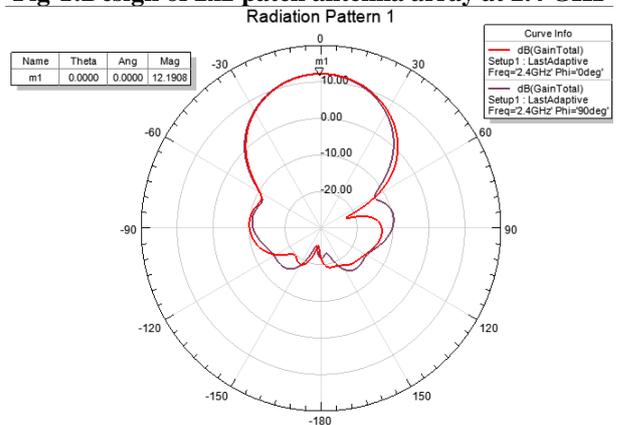
So, using stacking technique, the bandwidth of the antenna arrays will be observed and analyzed here and it will be an investigation to find out the limitations of the bandwidth improvement techniques by patch stacking.

## 2. ANALYSIS OF DESIGNS AT 2.4 GHz

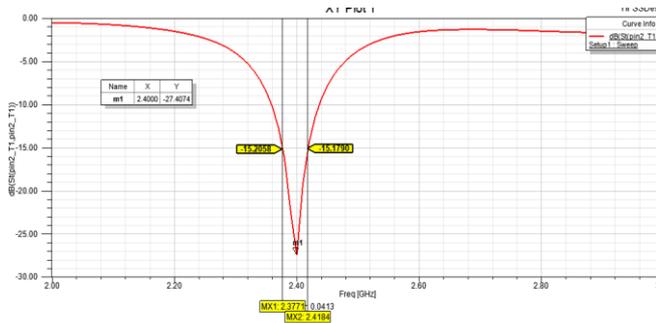
At first case, the S band frequency 2.4 GHz is considered to go ahead with. So, the baseline is established using 2x2 array of single layer conventional rectangular patch antennas at 2.4 GHz, designed using Rogers 5880. Its patch length 38.5mm , patch width 44mm and substrate height 3.2mm , total ground plane length for array 162mm , total ground plane width for array 155mm.



**Fig-1: Design of 2x2 patch antenna array at 2.4 GHz**

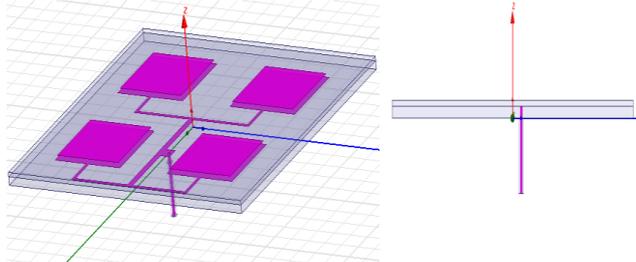


**Fig-2: Gain = 12.19 dB.**



**Fig-3: Return loss = -27.4 dB , Bandwidth = 0.0413 GHz = 41.3 MHz**

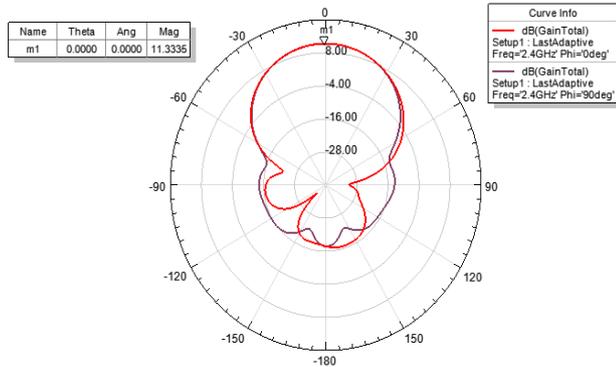
**2.1 Stack layer 1:**



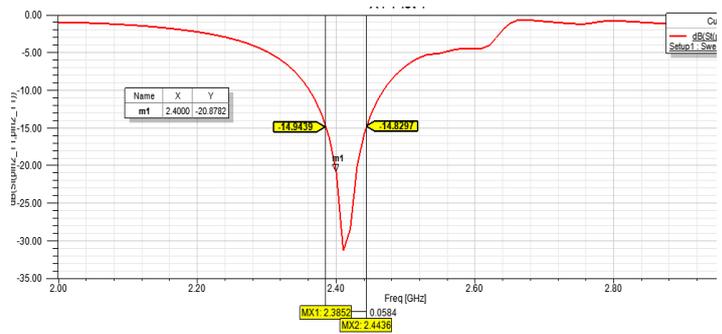
**Fig-4: Design of antenna array at 2.4 GHz with 1 layer**

Design Parameters	Values
Lambda	125 mm
Ground plane length	145 mm
Ground plane width	140 mm
Substrate height	3.2 mm
Patch width	39 mm
Patch length	37 mm
Air gap height	0 mm
Spacing between patches	0.54 lambda
Parasitic patch length	35.5 mm
Parasitic patch width	36.5Mm

**Table -1**  
Radiation Pattern 1

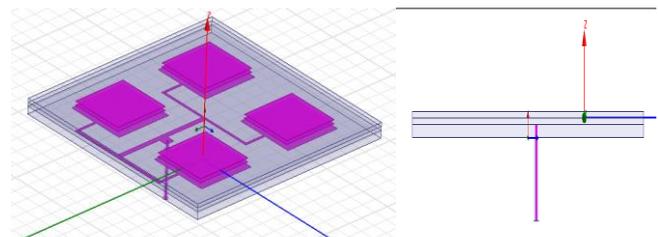


**Fig-5: Gain = 11.33 dB.**



**Fig-6: Return loss = -20.87 dB , Bandwidth = 0.0584 GHz = 58.4 MHz**

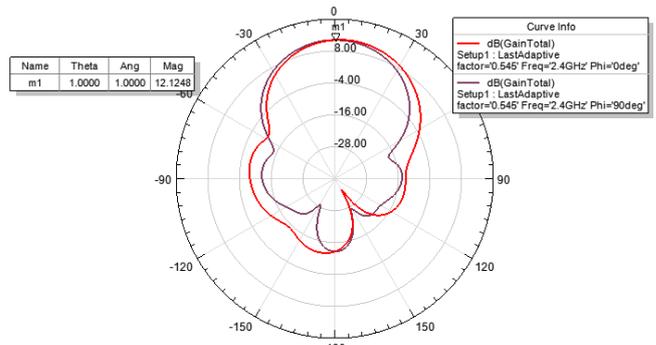
**2.2 Stack layer 2:**



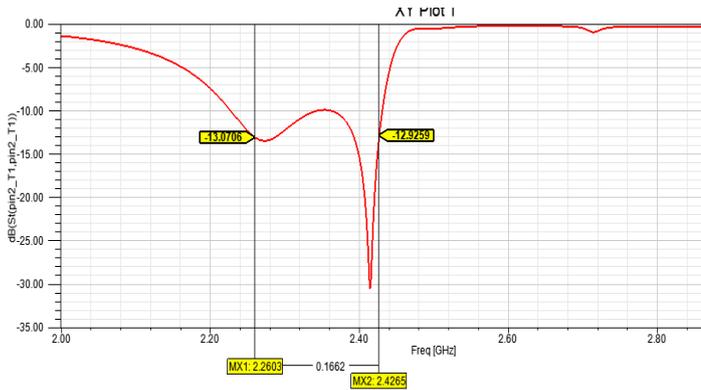
**Fig-7: Design of antenna array at 2.4 GHz with 2 layers**

Design Parameters	Values
Lambda	125 mm
Ground plane length	145 mm
Ground plane width	140 mm
Substrate height	3.2 mm
Patch width	39 mm
Patch length	39 mm
Air gap height	0 mm
Spacing between patches	0.54 lambda
Parasitic patch length	37.5 mm
Parasitic patch width	35 mm

**Table -2**  
Radiation Pattern 1



**Fig-8: Gain = 12.12 dB.**



**Fig-9: Return loss = -30 dB , Bandwidth = 0.1662 GHz = 166.2 MHz**

**2.3 Stack Layer 3:**

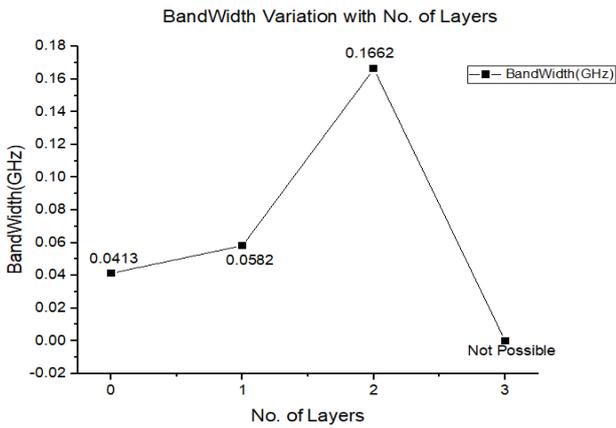
Further stacking is not possible because the gain curve is getting disturbed and even the minimum 10 dB gain is not reached and because of the presence of this multiple layers of Duroid and patches within the near field region of the antenna array, the return loss is also very affected for this case.

**2.4 Percentage Change in Bandwidth:**

No. of layers	Percentage change in BW
1	+39.7%
2	+184.5%

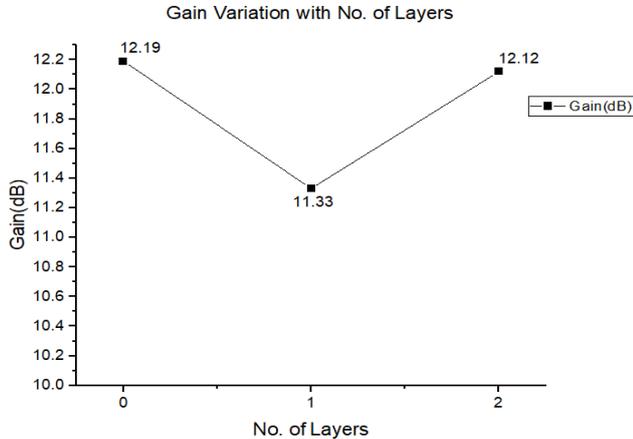
**Table -3**

**2.5 Bandwidth Variation Over No. of Layers:**



**Fig-10**

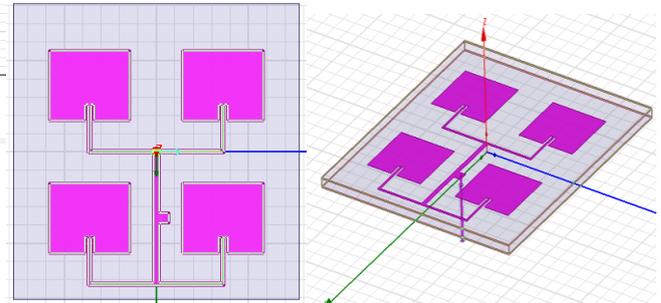
**2.6 Gain Variation over NO. of Layers:**



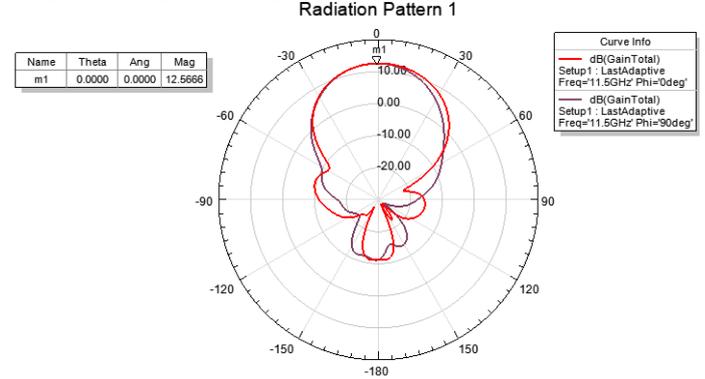
**Fig-11**

**3. Analysis of Designs at 11.5 GHz:**

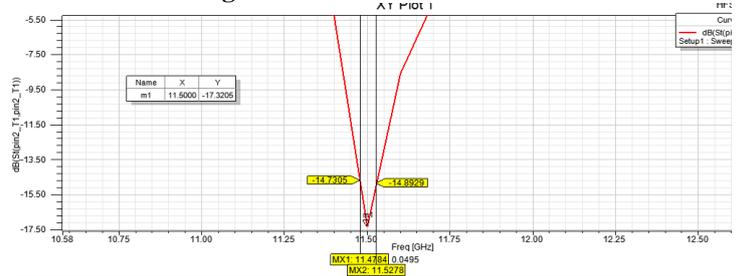
Now, the X band frequency 11.5 GHz is considered to go ahead with. So, now the baseline is established using 2x2 array of single layer conventional rectangular patch antennas at 11.5 GHz, designed using Rogers 5880. Its patch length 8.3mm , patch width 10mm and substrate height 0.254mm , total ground plane length for array 35mm , total ground plane width for array 32mm.



**Fig-12: Design of 2x2 patch antenna array at 11.5 GHz**

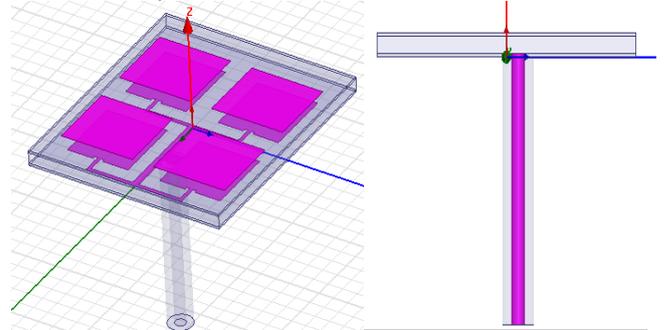


**Fig-13: Gain = 12.56 dB.**



**Fig-14: Return loss = -17.32 dB , Bandwidth = 0.0495 GHz = 49.5 MHz**

**3.1 Stack layer 1:**



**Fig-15**

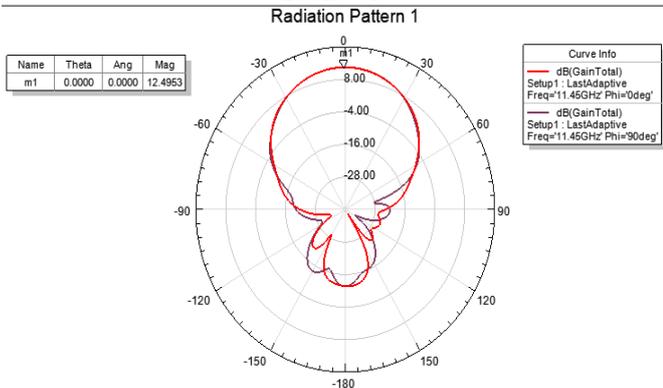


Design Parameters	Values
Lambda	26.2 mm
Ground plane length	30 mm
Ground plane width	26 mm
Substrate height	0.254 mm
Patch width	8.32 mm
Patch length	8.62 mm
Air gap height	1.3 mm
Spacing between patches	0.485 lambda
Parasitic patch length	8.83 mm
Parasitic patch width	9.8 mm

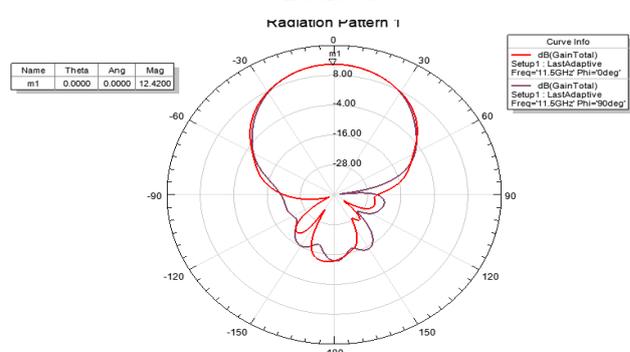
**Table - 4**

Design Parameters	Values
Lambda	26.2 mm
Ground plane length	30 mm
Ground plane width	26 mm
Substrate height	0.254 mm
Patch width	8.32 mm
Patch length	8.62 mm
Air gap height	0.5 mm
Spacing between patches	0.485 lambda
Parasitic patch length	8.5 mm
Parasitic patch width	9.5 mm

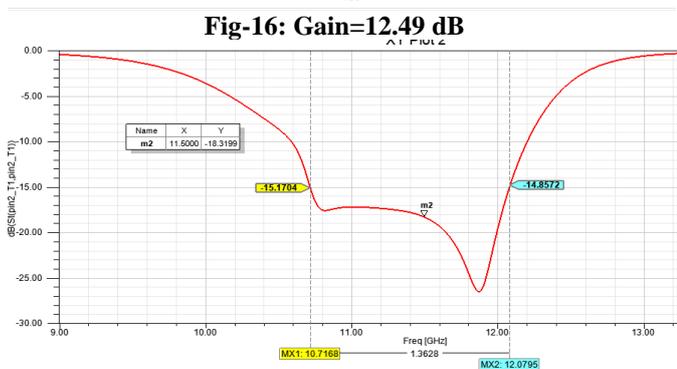
**Table - 5**



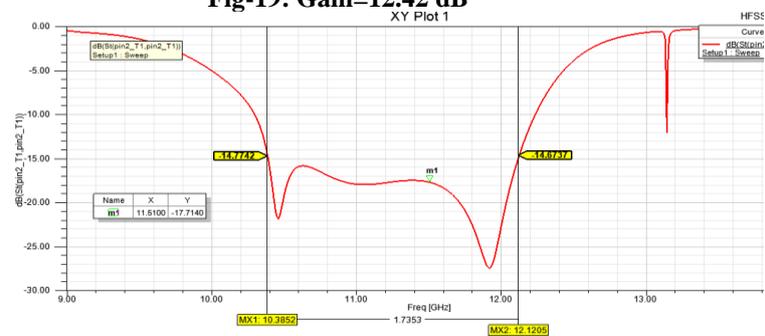
**Fig-16: Gain=12.49 dB**



**Fig-19: Gain=12.42 dB**

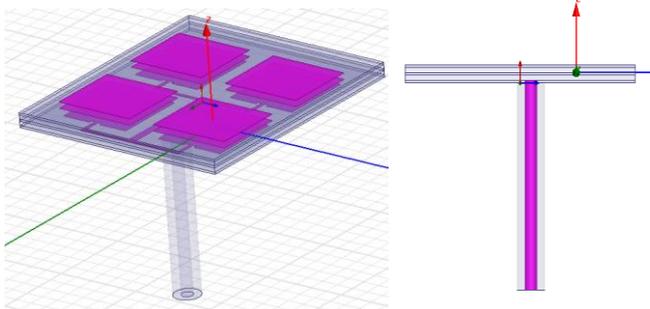


**Fig-17: Return Loss= -18.31 dB , Bandwidth = 1.36 GHz**



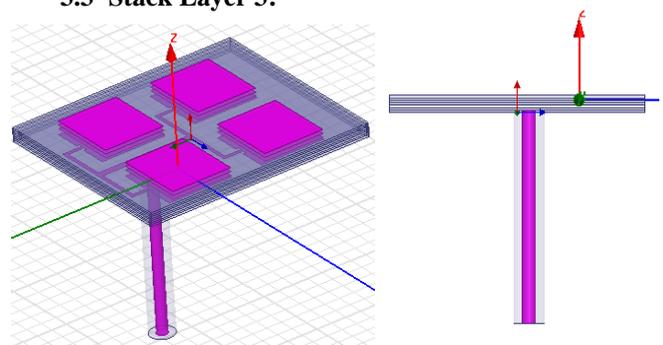
**Fig-20: Return Loss= -17.71 dB , Bandwidth = 1.735 GHz**

**3.2 Stack Layer 2:**



**Fig-18**

**3.3 Stack Layer 3:**

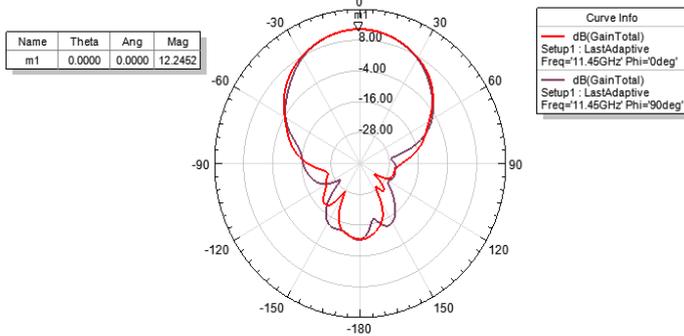


**Fig-21**

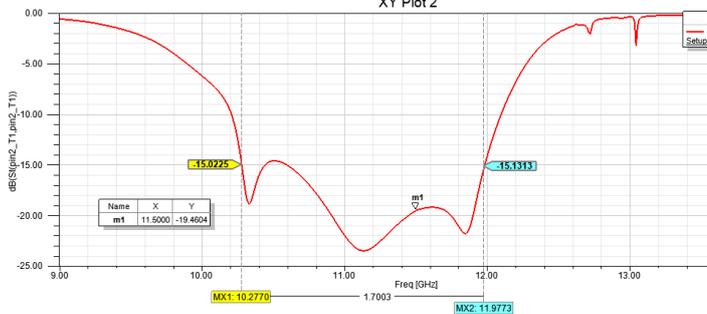


Design Parameters	Values
Lambda	26.2 mm
Ground plane length	30 mm
Ground plane width	26 mm
Substrate height	0.254 mm
Patch width	8.32 mm
Patch length	8.62 mm
Air gap height	0.2 mm
Spacing between patches	0.485 lambda
Parasitic patch length	8 mm
Parasitic patch width	9 mm

**Table - 6**  
Radiation Pattern 1



**Fig-22: Gain=12.24 dB**  
XY Plot 2



**Fig-23: Return Loss = -19.46 dB , Bandwidth = 1.700 GHz**

**3.4 Stack Layer 4:**

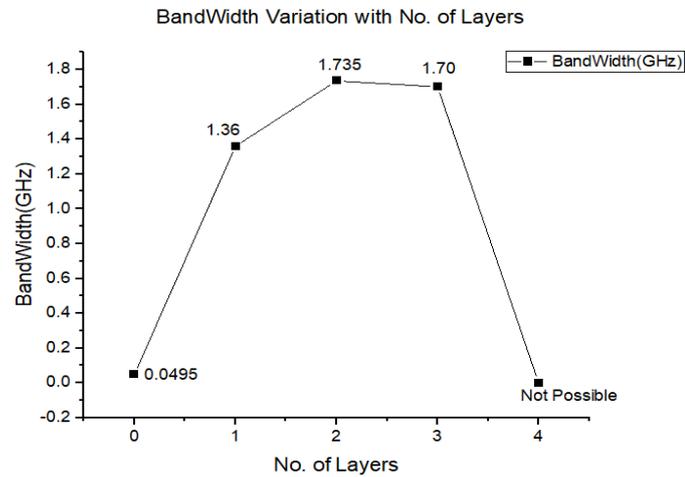
Further stacking is not possible because the gain curve is getting disturbed and even the minimum 12 dB gain is not reached and because of the presence of this multiple layers of duroid and patches within the near field region of the antenna array, the return loss is also very affected for this case.

**3.5 Percentage Change in Bandwidth:**

No. of Layers	Increment in BW
1	+2647%
2	+27.57%
3	-2%

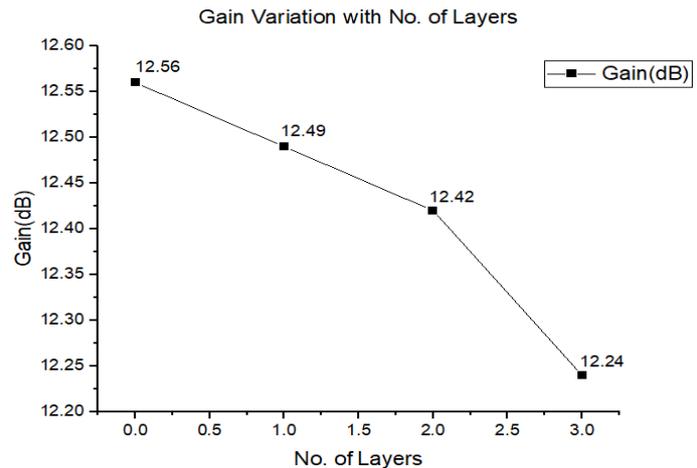
**Table - 7**

**3.6 Bandwidth Variation Over no. of Layers:**



**Fig-24**

**3.7 Gain Variation Over No. of Layers**



**Fig-25**

**4. CONCLUSION**

Stacking techniques provides good method for bandwidth improvement but it has some limitations. As some substrates layers and metallic patch surfaces are placed within the reactive near field region of the primary antenna element so it hampers the gain and the return loss of the antenna by a very significant fraction so, we need to have a very good amount of gain and need to have result loss as low as possible before adding those extra layers. And on the other side, it is not that the number of parasitic patch layers can be increased up to as many number as we want. It has certain limits after which if we want to add an extra layer, it can be seen that the gain and return loss are highly affected which is not acceptable otherwise there must be some physical restriction which will not allow us to do so.

### References

1. Rathi V, Kumar G, Ray KP. Improved coupling for aperture coupled microstrip antennas. *IEEE Transactions on Antennas and Propagation*. 1996;44(8):1196-1198
2. Kumar G, Gupta KC. Nonradiating edges and four edges gap-coupled with multiple resonator, broad band microstrip antennas. *IEEE Transactions on Antennas and Propagation*. 1985;33:173-178
3. A New Approach for Bandwidth Enhancement Technique in Microstrip Antenna for Wireless Applications D. Yoharaaj1, Raja Syamsul Azmir2 and Alyani Ismail3 Department of Computer and Communication System Engineering, Faculty of Engineering, Universiti Putra
4. Kumar G, Ray KP. *Broadband Microstrip Antennas*. Artech House; Boston; 2003
5. Targonski SD, Waterhouse RB, Pozar DM. Design of wideband aperture stacked patch microstrip antenna. *IEEE Transactions on Antennas and Propagation*. 1998;46(9):1245-1251
6. Gain and Bandwidth Enhancement Techniques in Microstrip Patch Antennas - A Review by Alok Kumar Research Scholar CT Group of Institutions Jalandhar Punjab , Nancy Gupta Assistant Professor CT Group of Institutions Jalandhar Punjab , P.C. Gautam Scientist 'E' DRDO Chandigarh Punjab.
7. Reddy KTV, Kumar G. Stacked square microstrip antennas for wideband circular polarization. In: *National Conf. on Communications*, NCC. Kanpur: Indian Institute of Technology; 2001. p. 125-128
8. Chew WC. A broadband annular ring microstrip antenna. *IEEE Transactions on Antennas and Propagation*. 1982;30:918-922
9. Vlasits T. Performance of a cross aperture coupled single feed circularly polarized patch antenna. *Electronic Letters*. 1996;32(7):612-613
10. Awida MH, Suleiman SH, Fathy AE. Substrate-integrated cavity-backed patch arrays: A low-cost approach for bandwidth enhancement. *IEEE Transactions on Antennas and Propagation*. 2011;59(4):1155-1163
11. Latif SI, Shafai L, Sharma SK. Bandwidth enhancement and size reduction of microstrip slot antennas. *IEEE Transactions on Antennas and Propagation*. 2005;53(3):994-1003
12. Arya AK, Kartikeyan MV, Patnaik A. Efficiency enhancement of microstrip patch antenna with defected ground structure. In: *International Conference on Recent Advances in Microwave Theory and Applications*, Microwave; 2008.