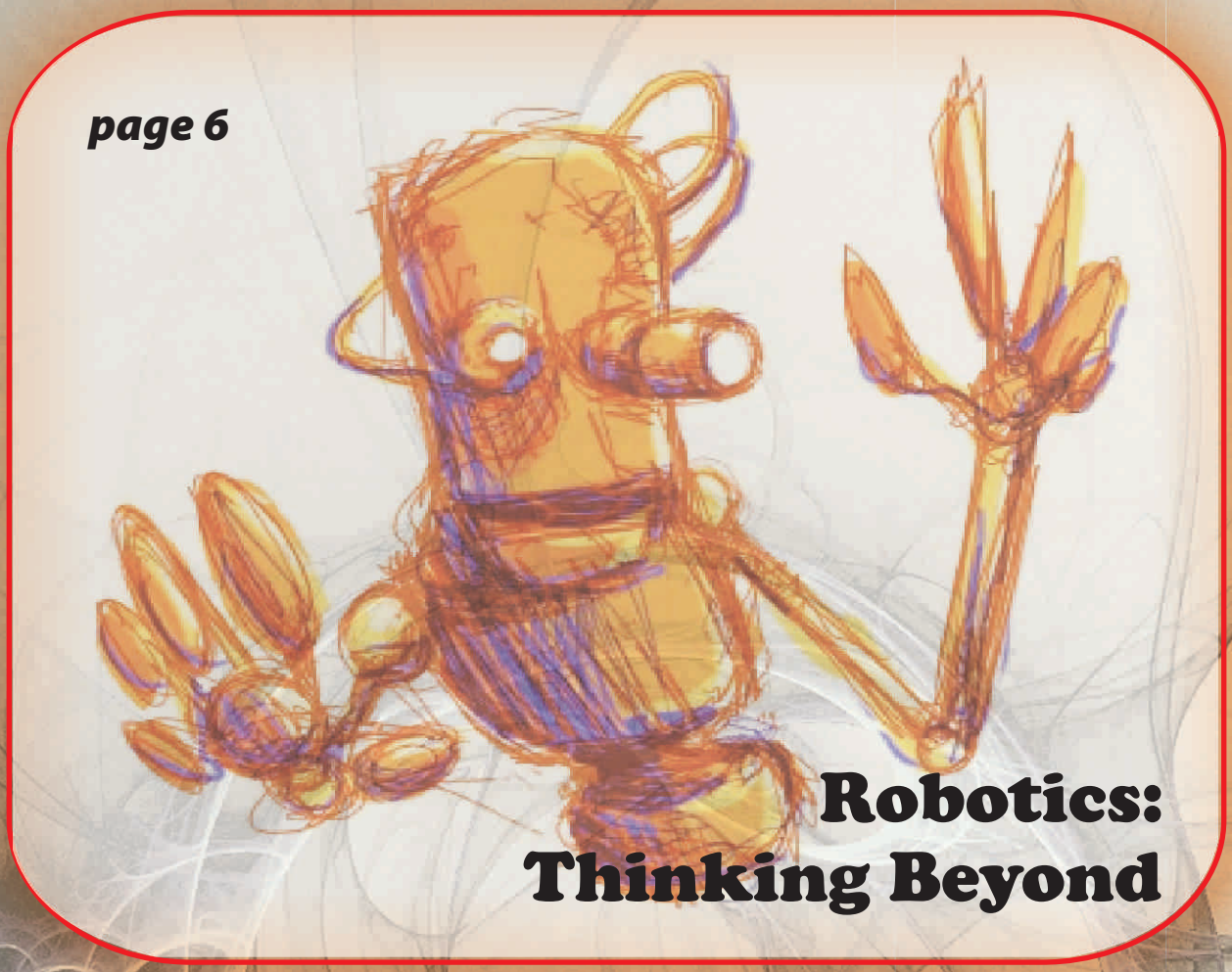
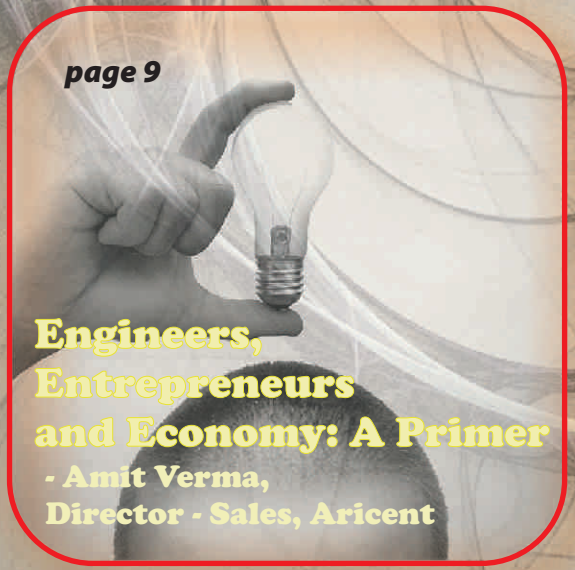
InPhase
IIT Guwahati-Cepstrum Magazine
February, 2008: Vol 2, Issue 1



page 6

Robotics: Thinking Beyond

page 9



Engineers, Entrepreneurs and Economy: A Primer

- Amit Verma, Director - Sales, Aricent

page 11

New Frontiers of Speech Recognition

- Tarun Pruthi, PhD

In Phase, February, 2008

From the editor...

Belated New Year Greetings to all the readers!

IIT Guwahati's ECE student's body Cepstrum is proud to present to you the 4th edition of its magazine In Phase. And starting this issue, we have restructured it as a bi-semesterly mouthpiece. No doubt the aim still remains the same - to connect the students, alumni, faculty and industry professionals and to bring them in-phase with what's going on in the department, in the industry and in the electronics and communications field in general Our vision for In Phase is to gradually build up as a valuable source for knowledge and information and become a respected monthly for all who are connected to electronics and communications.

This issue features an invited article from Mr. Amit Verma, Director - Sales of Aricent Inc. in which he shares details on the transitional steps from an engineer to an entrepreneur. With more and more engineers following this industrious path, it is essential for us to learn the tricks of the game first. Also featured are - Speech Processing Primer by our distinguished alumnus Tarun Pruthi, a lowdown on Wired magazine's Nextfest, a unique study of IIT Guwahati's electrical power system and an in depth analysis of the Nokia battery explosion incident. And in the cover article, we present to you a hands-on practical treatise on the fields and applications of Robotics. I hope you will find the articles interesting and useful.

And once again I'll urge you to send us your valuable feedbacks and suggestions. In Phase is still in its infancy and it needs the support of everyone to grow big and strong!

Akash Baid
Editor-in-chief
(Final Year B.Tech)

Team In Phase

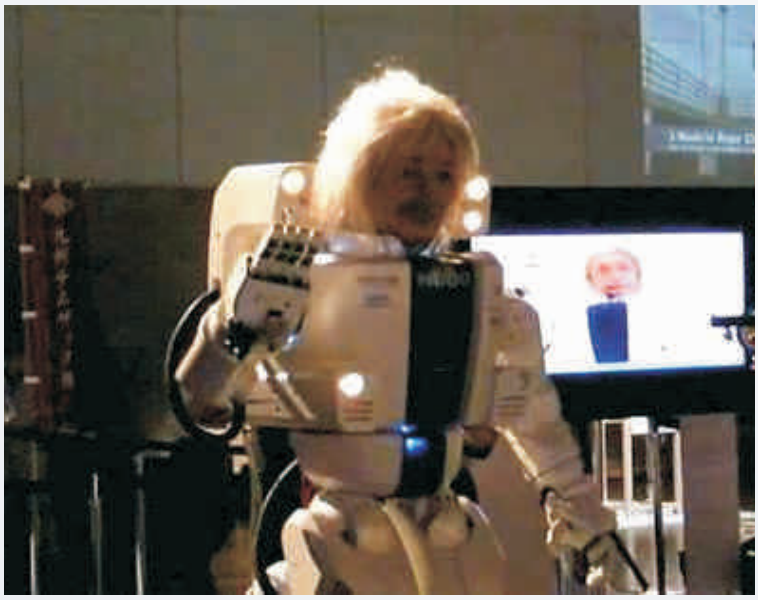
Call for Articles

We are looking for technical as well as non technical and experience sharing articles from students, alumni, faculty members and industry professionals. Articles pertaining to completed/ongoing projects, views, discussions, topic introductions, applications and professional or educational experiences are most welcome. Articles must be 1500 - 3000 words in length and should be written keeping in mind the diverse range of targeted audience including people with little as well as extensive knowledge of electronics. Please email us at cepstrum@iitg.ernet.in for any clarifications or suggestions.

VISIONS OF THE FUTURE @

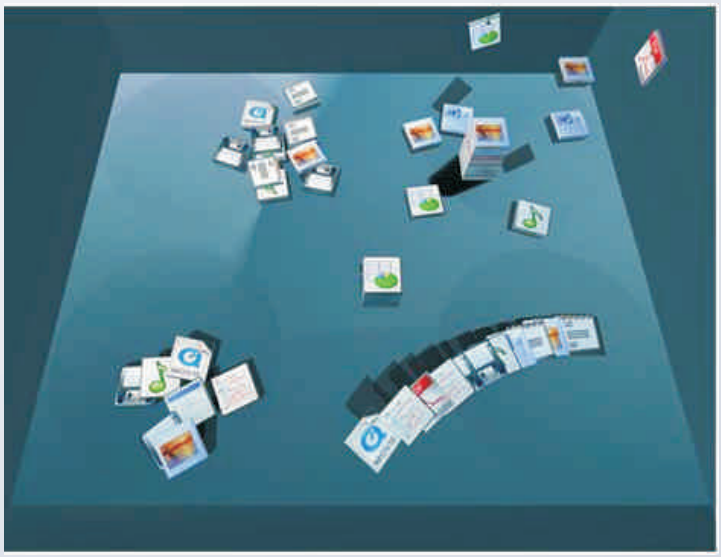
Wired NextFext'07

Have you ever wanted to be a globetrotter without moving a muscle or have you ever seen a solar-powered curtain with light show. Los Angeles had a glimpse of what our future look like at the Wired magazine's NextFest 07. Wired NextFest brought together innovations from inventors and research and development labs from all around the world, for children and adults alike. From ideas that ran from the creepy (like regenerated body parts including bladders) to conceptually fascinating (like a harp with strings made from lasers and a Lifestraw that kills sinister infectious organisms in water upon contact), over 160 exhibitors at the 4th annual Wired NextFest in Los Angeles enthusiastically revealed their visions of the future to receptive nerds of all ages. Here we give you some of the most mind boggling, fanatic innovations which featured in the fest.



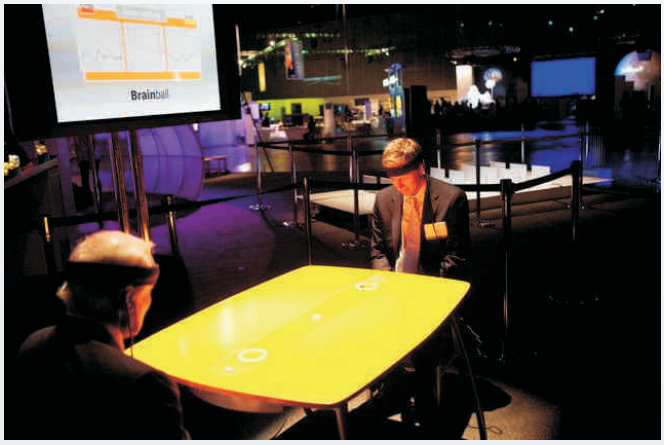
As anyone could guess from such a geeked-out expo, the star crowd-packers were the robots. The noggin of the smooth-moves Albert Hubo robot from KAIST, which holds the distinction of being the first walking robot with a face capable of showing expression, bore the eerie, exact resemblance of Albert Einstein. This humanoid is equipped with state-of-the-art voice recognition and object tracking software. HUBO can walk fluidly and move each finger individually. Another robot at the expo was SALAMANDRA ROBOTIKA, a land and water loving robot from EPFL that changes speed, direction and gait when zapped with current. This spindly, four footed yellow bot designed to mimic a salamander's spinal cord aids in biological research and may one day help us to understand the human spine.

A messy desk may signal a creative mind, but a messy computer is just a pain. Whether you are a firm believer in the power of OS X's Aqua, in love with Gnome or KDE on Linux, or simply can't wait for Vista Beta, the desktop interface plays a fundamental role in the way we use our computers. Anand Agarawala and Ravin Balakrishnan, from the University of Toronto, have developed an alternative user interface called BumpTop. BumpTop departs from the normal, flat perspective we are all used to, and tries to impart a feel of realism to the UI through the use of physics and a 2.5D perspective. In this physics-driven universe, icons look bigger and heavier depending on the file's importance, size, or last date changed. One can stack them, pile them, or crumple them up.



...Wired NextFext'07

The Invisible Train, part of Graz University of Technology's experimentation is the first "Handheld Augmented Reality" application. The Invisible Train is a mobile, collaborative multi-user Augmented Reality (AR) game, in which players control virtual trains on a real wooden miniature railroad track. These virtual trains are only visible to players through their PDA's video see-through display as they don't exist in the physical world. This type of user interface is commonly called the "magic lens metaphor". Players can interact with the game environment by operating track switches and adjusting the speed of their virtual trains. The current state of the game is synchronized between all participants via wireless networking. The common goal of the game is to prevent the virtual trains from colliding.



Another intriguing new game introduced was the "BrainBall", a game where passivity and calmness mark the winning streaks instead of activity and adrenalin. Yes, Brainball is a game that goes against the conventional competitive concept and to win the game, you must be more relaxed than your opponent. The game involves electrodes strapped to the players' heads, which can detect brainwaves that indicate the relaxation level of the player. The ball will move forward and get the player closer to victory when he or she releases alpha and theta waves, which are released only during relaxation. The winner gets the ball across the opponent's goal.

As we said, there were tens of other mind numbing exhibits but here's the final one on our list: Be a globetrotter – without moving a muscle. Built by Aksioma in collaboration with Institute for Knowledge Discovery, Laboratory of Brain-Computer Interfaces, Graz University of Technology and the Department of Information Design, FH JOANNEUM - University of Applied Sciences, Graz, Brainloop platform utilizes a Brain Computer Interface (BCI) system that lets user explore the cities and mountains of virtual google earth simply by imagining basic movements like 'move feet forward'. Such mental images trigger signals that send command to the BCI and the outside world. The result is an epic journey of mind, set to a haunting score manipulated in real time by performer Markus Rapp.

- Compiled by Amit Pal, 3rd yr. B.Tech, Dept of Ece (information and pictures courtesy www.wirednextfest.com)

How to save Rs. 1 Lakh per month !!!

In a closed community environment like our IIT Guwahati, it is intriguing to see how small numbers add up to give a huge overall effect. Power consumption for example; single users might not consume large units but when we add it all up, the numbers give us a different picture! Here we present some interesting statistics on the power consumption in IIT Guwahati. After a detailed interview of the authorities at the engineering section, we compiled the following data and analyze where and by how much can we really save.

First some numbers on the electricity bill that IIT Guwahati pays. According to authorities in the engineering section, we pay roughly Rs. 50 lakhs per month in the summer and Rs. 35 lakhs per month in the winters. Apparently the major consumers are the air-conditioning plants, the New Academic Complex and the Administrative Building. And then comes hostels. Since our main purpose here is to present some kind of figure on the amount of power that we the students waste and consume, a break up of the energy ratings of various devices in our room is given in Table 1. Typical consumption statistics for computers are in Table 2.

Appliance	Power Consumption
Tube light	35 W
Fan	60 W
Bulb	60 W
Computers	100 W

Table 1

Model	Clock Speed (Ghz)	Power (W)	Speed:Power ratio (MHz/W)
Pentium 4-C	2.4	67.6	35.5
Core-Duo	1.6-2.16	31	51.6-70
Athlon 64 3200+	-	67	-

Table 2

(The link below can be followed for details on these and other statistics:
http://en.wikipedia.org/wiki/CPU_power_dissipation #Intel_Pentium_4)

There are various types of monitors out there so we will just give you an approximate figure. CRT monitors generally consume a lot of power, figures vary around 100 W. On the other hand LCD monitors are far more efficient, generally consuming about 40 W. Again we won't go into details for speakers (as they are not pertinent to our discussion here), but they also tend to consume about 30 W.

So let's get a conservative estimate for how much power we CAN save. Even with a screensaver or with the monitor off, more than 50 Watts of electricity is used up by the CPU. At least 100 Watts is used just for start up. Printers, speakers and other ancillary devices increase the total figure even more. This is pretty much what an average student is going to use during his stay in the hostel. Some people do use small heaters, mosquito repellants or irons, but we will assume that wastage due to these devices is small enough to be ignored. Now, many of us leave the fan or the lights on at sometime or the other and simply turning them off would be a great saving in itself. But since we are making a conservative estimate we will assume that people do have the habit of flicking the switches while leaving. However, the same cannot be said about computers. Almost all of us leave our computers on when we go to sleep or when we go for our classes. Now assuming that the average student sleeps for 6 hours and goes to classes for 4 hours, we arrive at a figure of 10 hours when we absolutely do not need to leave the computer on. So adding this and some wastage due to fans, bulbs, tube lights etc, it would be safe to say that we are wasting power at the rate of around 75 W for 10 hours. This translates into wastage of 0.75 kWhr per day. Taking into account the rate of electricity, that is Rs. 3.9 per kWhr, every student is wasting around Rs. 2.9 per day.

Doesn't seem like much, huh? Read on…

No. of students in IITG with computers – 1,200
(IIT Guwahati's total student strength is 2,126 as per the academic section website but since the first year students and some others generally don't have a computer, this is a rough estimate.)

Wastage per day – Rs. 2.9 * 1,200 = Rs. 3,510
No. of days in a month – 30
Wastage per month – Rs. 3,510 * 30 = Rs. 1,05,300

Thus even by a very conservative estimate, we are wasting more than a lakh of rupees per month. And this expenditure will rise as the number of students increase gradually year by year. We usually don't pay much attention to these things since WE don't have to pay. But simple measures like switching off your computer and the lights can help save money. And all of us are aware of the dismal power situation in India. Every bit helps. Saving energy doesn't just help us to save power and hence money, it also helps us save coal, petroleum and other fossil fuels which in the long run will help us save the environment. So let's start now – Don't forget to turn off your computers when you don't need it.

(If you can find some fault with the assumptions and facts, please do inform us.)

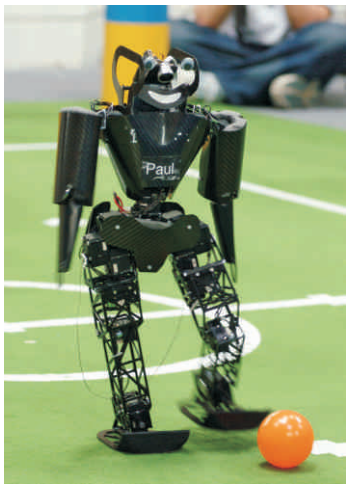
- Compiled by Rahul Sangwan, 3rd. yr. B. Tech, ECE

Robotics: Thinking Beyond

- Romesh Khaddar

By the year 2050, develop a team of fully autonomous humanoid robots that can win against the human world soccer champion team

- RoboCup



One might be tempted to infer from the statement above that a few geeks are getting ready for their new venture, but is it really just that? In a matter of words, probably yes but it's surely not just some geeky science fiction, it is increasingly becoming Reality. That scientists round the globe are trying to develop robots is a well known fact, but what kind of robots? That is the question that we are skeptical about. RoboCup is one such attempt by robotics enthusiasts from the world over to organize and channel all their resources and to develop the next level humanoids capable of playing a game which involves not only the skills of kicking the ball but also reading the opponent, organizing oneself tactically and deploying a complex strategy difficult to predict. The goal of problems like RoboCup is to foster artificial intelligence and robotics research by providing a standard problem where a wide range of technologies can be examined and integrated. So is this article is all about robocup? Not at all, of course you have to read it through to know about it.

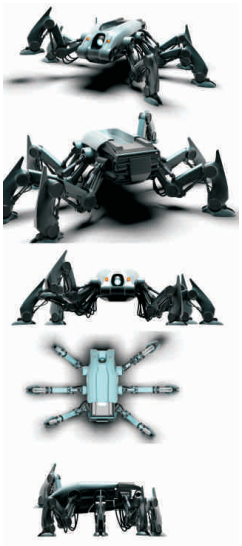
Today if we ask a person about his perception of a robot, the general answer will match to T1000 of Judgment Day or the NS-5 of I-Robot or maybe “weebo” from flubber. But that is far removed from the reality today. Even an imitation of human arm controlled by a computer or even manually with electronic control is something that can be referred to as a robot. Fundamentally thus, a robot is a mechanical system having a control unit, generally a computer, designed for a particular task or task set.

The hardware of the robot is not only what we see and feel when that mechanical devil is standing there, but a lot of other electronics and hydraulics and gears that we don't. It is the physical structure which is supposed to do work.

The first being the life of robot – a power system which one can generally think of as the famous internal combustion engines, or maybe electricity directly from the cables, or just the batteries. The kind of power system is generally decided based on the requirement. What is new in this field are the power cells, yeah not that powerful as that you saw in Rise of the Machines, but are under development and do have a capability to support a humanoid robot for days.

The second most important system of the robot is the hardware itself. From the immaculately sophisticated operating arm with multiple high torque motors or the small electromagnets to an elegant looking gyro stabilized biped, from massive big welding and cutting tools of industry for cutting through steel and concrete to small autonomous vacuum cleaners for home. Today the robots have a variety of applications defined for them and consequently there are various different mechanisms associated. Be it a complex surgery or be it a 'just for fun' environment, robots can be found anywhere and everywhere. They have started tracing the infinites of space to the minutes of human body. The hardware is one of the most challenging part in any kind of robot because of the versatility that has to be achieved.

The Third and yet another important system are the sensors. The sensors in a more crude term are the part of a robot which gathers the information about the surrounding, something similar to humans. Like we see and feel and smell and hear, we understand our environment and the basic task of our body is to get the raw data is by using them. The robots have various substitutes for our sense organs. To start with they have video camera, Infrared sensors, Ultrsonic sensors, Force and Pressure gauges (amazingly the development of substitute for human smelling capability is not that evident). Generally any single sensor is very easy to operate but the problem arises when a network is to be considered. A very simple example is a single LDR (Light Dependent Resistor) Vs Video camera; the former can be used to identify the color, but when a modern video camera is considered, it can be assumed to have millions of such LDR's (pixels) arranged in a rectangular fashion and compact manner. And it is very difficult is to extract a series of images to make a video taken into consideration of the high speed nature of our world. As the quality of video gets better, number of pixels increase and hence the need of compatible high speed to gather data from those LDR's to make a video of real world, i.e.



greater usage of sensors means more complexity.

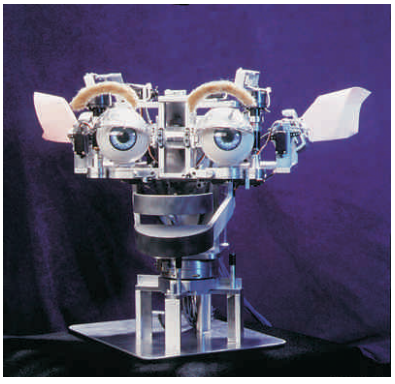
But what makes a machine different from its other counterparts, what really makes it a Robot. It is undoubtedly the intelligence associated with it, as you must have guessed by now. The fourth and most important component of any robot is the Artificial Intelligence (AI); intelligence is the ability to decide and act according to the conditions available and artificial because it is not born with the machine but has to be embedded in it by programming. The buzzword of today is AI which is actually the method to utilize the capability of any kind of hardware machine that can be built by humans. The best example for this is the Asimo robot. It can walk and climb down the stairs without needing to be programmed every now and then. But intuition says it is pretty easy to walk if not to climb down the stairs! Just analyze, a 54 kg machine that is 4.25' high which has to be balanced on two rods (legs in our case). As soon as you lift one of them, the center of mass should be in forward direction and the robot should be toppling ahead and neither sideways nor backwards (else it will fall), which is very difficult to be ensured until and unless you have a computer powerful enough to calculate all the relevant mechanics calculations involved. When the robot is moving without any guidance, the computer also has to be moved and to know about the environment, whether to climb down the stairs or to walk, a video camera is placed in the head apart from various position sensors provided at the joints to know how much angle the motor has moved. It took nearly a decade to actually build Asimo, from a mere pair of walking legs to a fully autonomous humanoid. In effect, each step of this combined entity shows that what goes behind in our brain is by no means trivial but it is on the contrary, fairly complex and the intelligence that we show and assume easy is not that simple. We might teach a child how to walk easily but in the case of a robot we have to make them walk.

Is the AI limited to walking? Of course not, AI is not only about walking but also about every trivial looking thing, whether to negotiate a maze using a camera or even without it, just checking and avoiding any obstacle in front using a set of sensors which are used in your TV remote (which is called Lower Level Vision). What about all those fancy robots from Star Wars & Terminator series? Are robots made just to do menial work for humans and nothing else?

The questions above are not as easy as pie. A lot of intellectuals and scientists around the world are trying to answer them. In fact when in one part of the world a walking robot was being developed there was research going elsewhere to actually understand and implement human behavior. To begin with, we know there are various aspects of a human being viz., thinking about solutions to complex problems, learning multiple things, organizing various sensory data, managing himself and others, showing various kinds of emotions etc. Implementing these abstract things is difficult because one can behave in strikingly different manner for two seemingly same situations. Moreover learning, which is quite natural to us is not that ordinary for machine. The general notion that 'I am teaching the robot by programming it' is actually confusion.

Beginning with learning, when we were kid, we were taught how to hold a thing by someone elder. Initially giving a bigger toy to hold and slowly in a matter of weeks thinner and thinner toys, thereby slowly finessing our grip. But given a task that using the gripper of robot, program it to hold an object, standard programming will always end up in a fixed pattern no matter however sophisticated be the hardware. As a result, if you want a screw driver to be held will not be a problem, but then the same program might break a pencil. But if we program it with a method called Neural Networking, we will be able to teach the robot about how much force to put on holding a screw driver and pencil. In other words, the master program will learn the amount of force or pressure to be applied by the gripper, just like a small child. By making a robot being able to learn, a big problem of programming every possible scenario is reduced, and the robot only learns what it is going to face. One kind of such learning programmed robot is commercially available as AIBO by Sony, the famous robotic dog that you must have come across either through mass media or in reality. The robot when bought, acts like a puppy and later learns a lot of things as the master teaches, all by simple correlation of actions and word commands, like understanding when to kick the ball and when to fetch it.

But even apart from the objective of learning, robots can do a lot more; they can show anger or laugh at your joke and talk to you just like your friend. The researchers at Massachusetts Institute of Technology have designed one such robot called Kismet that can have realistic conversations with people and has seven different facial expressions and can vary the tone of its voice as per the statement it is making. The robot can also look up to the person to whom it is speaking to. The emotions are actually the results of in depth understanding of a linguistic statement or a situation. A set of situations can be



tracked till a certain point to decide upon how to react to current situation, but to develop a system to understand language, it has to understand how we understand the language, which is not as easy as it may seem. For a proof think how hard it is sometimes even for us, the human race who devised this urbane language, to infer the alternate intended meaning of any given statement.

Asimov's Laws of Robotics:

Zeroth law

A robot may not harm humanity, or, by inaction, allow humanity to come to harm

First law

A robot may not injure a human being or, through inaction, allow a human being to come to harm except when it conflicts with the Zeroth law.

Second law

A robot must obey orders given to it by human beings except where such orders would conflict with the Zeroth or First Law.

Third law

A robot must protect its own existence as long as such protection does not conflict with the Zeroth or First or Second Law.

This brings us to a very important aspect of human imitation: that of group psychology. It is the nature of humans to live and act in unison, in a complementary manner, managing and governing themselves, with a leader and followers accomplishing various tasks, forming different teams. But envisioning a group of robots operating similarly surely looks quite impossible to a common eye. I mean unless you yourself choose a leader and say to all the robots to follow it, you are just creating havoc. The robots must be able to identify whom to follow and whom not to, i.e. they should be able to choose their leader, make a choice whether or not to follow etc. And much of this has actually been done while research is going on the more subtle issues in this regard. This particular branch of robotics is commonly referred to as Multi Agent Systems, and have a sub-branch called swarm robotics which is getting popular nowadays as it has numerous utilities in areas such as defense and task completion. Understanding of Multi agent systems is thus very important for many applications like creating a workforce for heavy industry or exploration of space and other planets, where one robot alone is not sufficient but multiple are needed to be managed by themselves to a certain extent because of various issues. Entertainment events like Robocup etc. also need this kind of robots because they need to make a strategy, choose a leader and work in a coordinated manner which maybe quite intuitive to the eyes a of common man, but ask the person who has to implement it on the robots!

Even all this doesn't means that AI is only associated with anthropomorphic and zoomorphic type of robots, It is also associated with many other kind of machines like cars, and other software like FICO which is used in the US to decide the credibility of a customer before giving loans to understand the payback capacity. So AI can help in creating robots which can actually talk, walk and behave like humans, robots which can think abstractly. A researcher from England has already made a robot which can paint like an artist, not the remakes of "Mona Lisa" but certainly original paintings which are abstract and artistic and a real work of art, admired by many in the line. That means one can create a lookalike of humans or maybe even better but shall robots be always governed by Issac Asimov's Laws of Robotics? If the robots are programmed to copy humans, they might theoretically be given the capacity to think like humans too. The objective ultimately is to create robots to do repetitive menial job, to work in unfavorable conditions for humans, and to simplify more complex tasks which will calls for a grater amount of thinking power from them. The ability to think can and will lead to the creation of an individuality of each seemingly similar robot. Each of them will start becoming unique, and they will become a part of the society. The term robots coined from Czech word robota meaning either a slave or worker doing heavy work, will start to be termed as equals. The humans will have a moral responsibility then to deal with what they have created. To some, this idea sounds comic and to some this is the bell of a great danger ahead, like the destruction attributed to Skynet in the Terminator series. And one view can be that they will be the next surviving fit species of Charles Darwin's theory who will continue to go on when even the entire human race vanishes from the earth.

Coming back to present times, there are already people working in the direction of development and formulation of moral code of conduct for dealing with robots with the capability of human intellect. But this idea brings forward a lot of questions. What exactly are we trying to achieve by incorporating intellect in these calculative machines? And if we will succeed, what are the chances that they won't start interfering with the very existence of human race? After making these kinds of machines, we will owe them a responsibility, are we willing to bear it? But above all are we trying to reinvent the wheel or are we trying to play the creator GOD?

Ah robotics, is this an engineering branch or a complete science in itself? Well that is something for you to decide.

(Romesh is a final yr. B. Tech student at the Deptartment of Electronics & Communication Engineering, IIT Guwahati)

Engineers, Entrepreneurs, and Economy: A Primer

- Amit Verma, Director - Sales, Aricent

Many of us strike upon great opportunities and business ideas but doubts, hurdles and questions come to mind every time we are faced with one. How do I start, how do I raise funds, which source of funding is good, how do I make a business plan, who is going to help me market the idea, how do I know if I have the right team, how do I manage a team of highly talented individuals, and the list goes on. This article does not cover all these questions in a lot of depth but touches upon many of them and highlights the importance of entrepreneurial thinking for today's engineers.

It is believed that entrepreneurs are good leaders, good managers of resources, problem solvers, passionate, and good risk managers. They are not necessarily among those who take risks overtly. Rather they are motivated to solve a problem and only take a calculated risk needed to solve that problem. Entrepreneurs play an important role in the growth of any economy and job creation, especially in a service and knowledge economy. While manufacturing economy favors large corporations, services economy lends itself to small businesses and continuous differentiation and hence the increased entrepreneurial activities.

Get Started The first step towards becoming an entrepreneur is to start. Sometime it is the fear of failing and sometimes it is analysis-paralysis that comes in the way. While a good idea and thorough planning is important, there is no perfect plan and no silver bullet. A simple idea if well executed can provide better returns than a poorly executed brilliant idea. McDonald's and Subway are two great example of simple ideas that are extremely well executed.

Sometime the best way to start is to not wait for the next best idea but rather improve the existing ideas. Google did not invent internet search, but definitely improved it a great deal. Ideas are important and you want to make sure that there is something unique about what you are doing but it does not have to be completely new. Does it address a real pain point in an economically viable fashion? That is the question you have to ask yourself.

Getting feedback on your idea is important. There is no point in keeping it to you. However, you may not get the best feedback by discussing your idea with just a closed circle of likeminded friends and family. They may not represent your end customer and hence their yes or no might not be very meaningful.

Ideation Ideas evolve as you work on them and take input from your potential customers. Find ways to get early feedback. Prototyping, demos, and survey are some of very important ways to gain early feedback. Customer feedback may not support your original belief and may even prove that it is not a great idea – the pain point either does not exist or there is a better alternate method for solving it. Many young entrepreneurs shy away from seeking a broad feedback due to the fear of hearing back something other than what they believe. Such a fear could lead to unnecessary investment of time and capital into an idea that should have been killed or improved at the early stage. So, seek broad and early feedback to validate your idea.

Team Formation Once you have a good idea it becomes very important to form a solid team. An experienced and talented team that adds direct value to your venture is central to the success of the venture. Starting a venture is not the best time to start pairing with your high-school buddies and start creating the “dream team” unless there are significant synergies. Most investors would like to know who all are in your team. Investors are generally more comfortable investing in people than ideas. There is a saying that the investors look at the last page of the Business Plan first. They want to know what is the background of the people in your team?, who are on your board?, who are on your advisory board?, etc.

Forming a solid team is obviously not an easy task, but perhaps the most important task. Look out for people that can add value to your idea and are passionate about it. It is no fun in working with people that don't share your passion. Make sure there is a fit and mutual respect. Once the core team of two or three is in place, then you start forming your advisory team and other functional team around that. In the early part of your venture, your best currency is your company equity so use it judicially. It should not be the case that four buddies from high-school are starting the venture and everyone is an equal member even though they may not add equal value. As you will move for funding, your team and the members that you are able to attract using equity as currency will represent the overall value of your company.

Funding

Once your core team is in place, you need to start shaping the idea further and start looking for funding. Now there are three good resources for funding early on. Friends, Family, and Fools (three Fs). So reach out to them. Get them excited about the idea. The goal is to raise enough money to develop a proof of concept, if possible, and get some market validation. It is extremely important that your idea is not just a slideware. Now, it is perhaps possible to raise funds even with slideware, but in that case the bios of the team member become the central value proposition and you still may end up giving up significant part of the equity to raise small funds. So in general it is a good idea that you take your idea as far as possible and validate the idea with user community and gather some real feedback. It is good if you can get feedback and comments from actual customers. For most investors, testimonial are more important then the product details. Nothing succeeds like success and early success are good.

Other Options

Now once you are ready for funding, you need to consider options. There is Angle money, there is VC money and there is money available from government and different large corporations. Each route has its own advantages or disadvantages. Angel money can be a good source of funding early on and can help you prove your idea and do market trial. Angel investors are individual investors and generally do not want too much control. They could be represented by a professional firm that screens Business Plans for them. They may move as a pack, so breaking into at least one of them is important.

There are also incubator funds. These funds are setup to provide seed level funding as well as also some logistic support such as space, access to VCs, talent etc. Incubators can add a lot of value to your venture and help de-risk the start. Logistical support can help you focus on developing the venture.

Venture Capital funds are more professionally managed, and hence can provide not only funding but guidance and market relationships well. Different Venture Capital focus on different stages of business (early stage, mid stage, later stage) and therefore you have to make sure you approach the right VC.

While money raising is critical to the success of your business, it is also important to know how much and when. Too much money is also not good, as it can distract you into doing things that may or may not be critical to success of your business. For example investing in a lavish office or hiring people you don't really need.

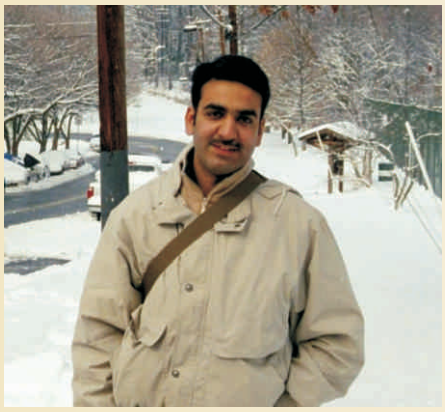
Large companies setup their own venture funds to promote innovation outside the typical corporate structure. This funding is available to both internal employees as well as external entrepreneurs. Similarly government agencies also set up funds to promote research and innovation in key areas. There may be some additional process/restrictions associated with government funding but it can also be a good source of funding.

Final Words

Never fall too much in love with your product or idea. This is one of the mistakes many entrepreneurs especially engineers make. Because your idea will evolve and will change and has be suited to changing market needs. Focus should be on solving a real problem in an economically viable manner.

Exit strategy is just as important as the entry strategy. It is good to begin with an end in mind and especially the investor would like know the exit strategy. Rather it is important that partners have clear idea how each one of them would like to exit so that there is a strategic alignment. However don't start a business with intention of flipping it. It only leads to a very narrow and short-term vision and can curtail any strategic planning.

So in summary, as problem solvers, engineers have an important role to play in the economic growth of the country and in finding economically viable solutions to today's problems. Starting a venture is not easy but true entrepreneurial spirit combined with the knowledge of how to start, create, and run a business can lead to great personal success as well as social contribution.



New Frontiers of Speech Recognition

- Tarun Pruthi

Speech recognition by machines is the process of finding the most probable sequence of words given the input signal. Speech recognition not only enhances productivity (for eg. you can dictate documents to your PC instead of typing), but is also a wonderful technology to save costs (for eg. by replacing customer service agents by automated response systems which actually talk to you instead of asking you to key in numbers), can lead to extreme miniaturization (by getting rid of all the confusing buttons) and is a dream come true for handicapped people.

Research in speech recognition has been going on for more than five decades now. Yet, the rate of adoption, a barometer for gauging the success of a technology, has been painfully slow. Despite the obvious plethora of applications, the slow rate of adoption of speech recognition speaks for itself: “There are still a lot of unresolved issues.”

According to ethnologue (www.ethnologue.com), there are 6,912 known living languages in this world! and many more dialects! Even though the makeup of the vocal tract apparatus is the same for all human beings (we all have a tongue, lips, nose, pharyngeal cavity, larynx and lungs), clearly we are able to produce a lot of different phones (or sounds) which vary from one language or dialect to another. Since the phones in one language vary widely from the phones used in another language, speakers of a particular native language introduce their own accent variations when trying to speak/learn another language. Say for example, a speaker of Hindi does not pronounce all the words in American English the same way as a native American. And that's not all! The signal produced for a particular phone by the same speaker varies depending on the word being produced, the location of the phone in the word (contextual variations), whether the speaker is speaking carefully or carelessly, whether the speaker is speaking fluently, or uttering words in isolation, whether the speaker is reading a text, or speaking spontaneously, and whether the speaker is happy or sad or angry. Even the signal for multiple productions of the same phone by the same speaker under the same circumstances is significantly variable. The presence of interfering noise in the background complicates matters even further. No wonder, that the speech recognition problem is far from being solved! Accounting for this amazing amount of variability has challenged the intellect of the brightest researchers in the world for decades, and we are still nowhere close to the performance that humans can achieve. According to a paper by Lippman (“Speech recognition by machines and humans”, 1997), Human Speech Recognition performance is an order of magnitude better than Automatic Speech Recognition. The results presented in this paper still hold true to this date since there have only been incremental improvements in machine recognition performance since then.

Research in the field has led to huge improvements in pattern recognition, signal processing, and our knowledge of human speech production and recognition. However, at the very core of it we are still using the same Hidden Markov Model (HMM) based approach with approximately similar feature vectors as we were using 30 years ago.

So, why is there so much buzz about this technology in recent years? Why should speech recognition work now? In my view, the primary reason is not the improvement in performance, but the emerging “need” for the technology and the growth of applications which can exploit it to their advantage even in its current form.

The Mobile Market

Currently there are more than two billion cell phones in this world, and the number is quickly rising. And with increasing computing power cell phones of today are used lot more than for just calling. Cell phones nowadays are used to browse the internet, send e-mails, send short textual/multimedia messages, listen to music, watch movies, click photos etc. Handheld PCs go even further by giving the power of a full PC in your palm. However, most cell phones still do not provide a full QWERTY keyboard because of limitations of size. Ones that do, have a keyboard that is too tiny even by the most modest standards. Speech recognition can come to our rescue under these conditions. Most people, even with a full size QWERTY keyboard type less than 40 words a minute, but can speak more than 120 words a minute! Imagine how easy it would be to compose e-mails, short messages and documents on the go. Further, hands-free control of cell phones is also important for safety reasons. Several states in the US have already passed laws banning the use of hand-held cell phones while driving. In such a scenario, speech recognition can easily give you the power to dial by name or number without using your hands.

Experts say that speech recognition applications for cell phones will soon flood the market. Microsoft already ships windows mobile with speech recognition built in. Nuance is moving fast to get its own version of embedded speech recognizer on cell phones. Google recently launched GOOG-411 its voice-activated search, so you can access it from any phone (mobile or land line), in any location, at any time. But this is just the beginning. Imagine the possibilities. How about telling your phone to store a

reminder by saying “Remind me to meet Neeraj in Pizza hut at 8:00 PM today”!

Other applications

Pattern recognition is not an exact science. There is always a probability associated with the correctness of the output, and given the current state of pattern recognition, the probability is far from being equal to one. True, humans are not 100% accurate either. But, the problem is that the recognition rates for machines are still below the “annoyance level”. Until the technology develops to a level where it is able to perform well under “reasonable” conditions, we need to look for applications which do not annoy people when the speech recognizer makes mistakes. Further, we need to realize that even though speech is the most natural form of communication for human beings, it may not be the most natural interface with machines. It is true that speech recognition is a highly preferable interface when hands-free control is desired, but is it always desirable? Imagine the confusion in an office where everyone is talking to their machines. Would you really choose talking to your machine in such a scenario instead of quietly typing? Besides, in some cases, it is just more convenient to type, erase and correct. This, I believe, is why dictation recognizers never took off in a big way. Therefore, we need to think of applications where errors are acceptable and the interface is natural.

In recent years people have realized this, and have started to come up with very interesting applications. Here are some examples:

Health-related applications: A wheelchair which can turn left when the user says “left”, or an Environmental Control Unit (ECU) which can turn the lights, fan, air conditioning etc. on by a simple voice command can be a potentially life-saving technology for people with severe disabilities, for eg. Quadripilegics, or ALS (Amyotrophic lateral sclerosis) patients. Further, injuries related to repetitive work (like carpal tunnel syndrome) can be avoided by creating interfaces which can be controlled through voice instead of the computer mouse.

Military applications: A soldier carrying a gun in his/her hands would very much appreciate being able to control a reconnaissance robot with his/her voice instead of a joystick. A universal translator which converts speech in one language to speech in another language can be a very useful device to interrogate prisoners in a foreign land.

Audio mining/Speech analytics is the term used to describe automatic methods of analyzing speech to extract useful information about the speech content or the speakers. For example, speech analytics may be used in a call center environment to categorize calls to identify calls from unsatisfied customers, or to analyze the recorded speech to extract critical business intelligence.

Audio/video search: Google has redefined the world through internet search. Google's mission is to organize the world's information and make it universally accessible and useful. However, most of the results it currently produces are textual in nature. Multimedia search is presently performed based only on the meta-data stored along with the media files. Clearly a method based on speech recognition is needed to look through the actual content of the media file and generate more relevant audio/video results. Any speech recognizer in its current form can be used for this application. The better the recognizer, the better the results! However, even if the recognizer is not good, it will miss some relevant links, which is fine because the recognizer can be improved incrementally to improve the results. Further, this is a particularly interesting application because it shifts the burden from the speech recognizers to the owners/creators of the audio/video file. If the creators of the file want it to be searched, they will themselves make sure that the file is recorded under conditions which are favorable for the recognizer. Companies like Blinkx have already starting offering video search based on a combination of speech recognition and video analysis.

Semi-automatic speech recognition applications: This includes all offline applications where the transcript of the speech recognizer can be corrected by a human. The ability to get a transcript with reasonable accuracy leads to significant savings in both time and cost. One example of such an application is medical transcription. Cost savings in transcription can translate into lower overhead costs and eventually cheaper healthcare for the patients. According to the AAMT (American Association for Medical Transcription), the global medical transcription pie is estimated to be anywhere between $12 billion and $20 billion, with the US being the largest market (from: http://www.acusisindia.com/AIP0302/News/Upandrunning270606.asp). Nuance already offers a speech recognition software, Dragon NaturallySpeaking Medical 9, which is specifically tuned for the purposes of medical transcription.

To conclude, speech recognition is a technology of the future and its time has come. It is a potentially life-saving technology for the handicapped and an amazing technology to improve productivity. However, there are still a large number of issues which need to be resolved before speech recognizers can be used successfully under real-world conditions. Until then, the focus should be on applications where a reasonable number of errors are acceptable, either because it doesn't matter, or because they are corrected by humans in a second pass, or because there is a strong desire for hands-free control.

(The author is a 99 batch alumni, graduating in ECE from IITG. He completed his Doctoral degree from the University of Maryland College Park in 2007 and is currently employed as a Senior Research Engineer at Think A Move, USA. His research interests include Speech Recognition, Machine Learning, Audio/Video Mining, and Digital Signal Processing.)

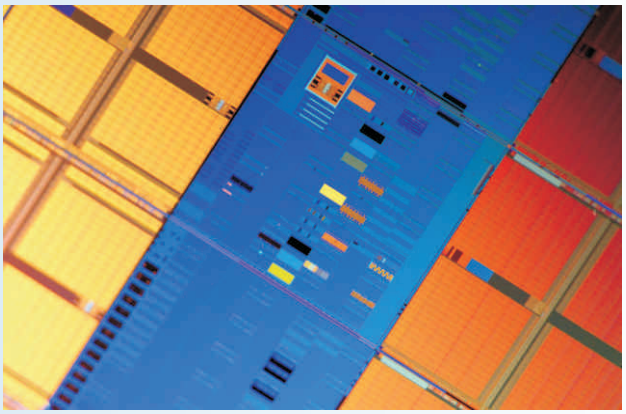
45 nm technology: What makes it so special?

- Talla Vamsi

Over the last 40 years, we have witnessed that the density of transistors on chips has been periodically doubling, as predicted by Moore's Law. Keeping up this trend Intel has launched the latest Core 2 microprocessors, code-named Penryn. The chips, based on the latest 45-nanometer CMOS process technology will have more transistors and run faster and cooler than microprocessors fabricated with the previous, 65-nm process generation. So what's so special about this technology if we have been following Moore's law for the last 40 years? Actually they are special because the chips would not have been possible without a major breakthrough in the way Intel construct a key component of the infinitesimal transistors on those chips, called the gate stack. This is the new revolution in this design and we will explore this further as to what problems Intel faced and how were they handled.

To keep on the Moore's Law curve, size of the transistors have to be halved every 24 months or so. The physics dictates that the smallest parts of those transistors have to be diminished by a factor of 0.7. But there's one critical part of the transistor that couldn't shrink anymore. It's the thin layer of silicon dioxide (SiO_2) insulation that electrically isolates the transistor's gate from the channel through which current flows when the transistor is on. That insulating layer has been slimmed and shrunk with each new generation, about tenfold since the mid-1990s alone. Two generations before Penryn, that insulation had become a scant five atoms thick.

Now, it wasn't possible to shave off even one more tenth of a nanometer—a single silicon atom is 0.26 nm in diameter. More important, at a thickness of five atoms, the insulation was already a problem, wasting power by letting electrons rain through it. Without a significant innovation, the semiconductor industry was in danger of encountering the dreaded “showstopper,” the long-awaited insurmountable problem that ends the Moore's Law era of periodic exponential performance gains in memories, microprocessors, and other chips—and the very good times that have gone with it. The solution to this latest crisis involved thickening the insulator with more atoms, but of a different type, to give it better electrical properties. This new insulator works well enough to halt the power-sucking hail of electrons that's plagued advanced chips for the past four years. As difficult as finding the new insulator was, that was only half the battle. The point of the insulator is to separate the transistor's silicon gate from the rest of the device. The trouble is, a silicon gate didn't work with the new insulator



Close-up of a 45nm SRAM Test Wafer

material. The initial transistors made with them performed worse than older transistors. The answer was to add yet another new material to the mix, swapping the silicon gate for one made of metal.

It may not seem like such a big deal to change the materials used in a transistor, but it was. A fundamental change to the composition of the transistor is pretty much unheard of. The combination of the gate and the insulator, the gate stack, hasn't changed significantly since Moore, Andrew S. Grove, and others .The problem, ultimately, is one of power. At five atoms, that sliver of SiO_2 insulation was so thin that it had begun to lose its insulating properties. Starting with the generation of chips fabricated in 2001, electrons had begun to trickle through it. In the processors made just two years later, that trickle became some 100 times as intense. All that current was a drain on power and a source of unwanted heat. The reason the gate oxide was shrunk no further is that it began to leak current. This leakage arises from quantum effects. At 1.2 nm, the quantum nature of particles starts to play a big role. The oxide layer is so narrow that the electron looks less like a ball and more like a wave. Specifically, it's a wave that defines the probability of finding the electron in a particular location. The trouble is that the wave is actually broader than the oxide layer, extending all the way to the other side and beyond. That means there is a distinct probability that an electron that should be on the gate side of the oxide can simply appear on the channel side, having “tunneled” through the energy barrier posed by the insulation rather than going over it.

The goal was to identify a gate dielectric material as a replacement for SiO_2 and also to demonstrate transistor prototypes that leaked less while at the same time driving plenty of current across the transistor channel. Intel needed a gate insulator that was thick enough to keep electrons from

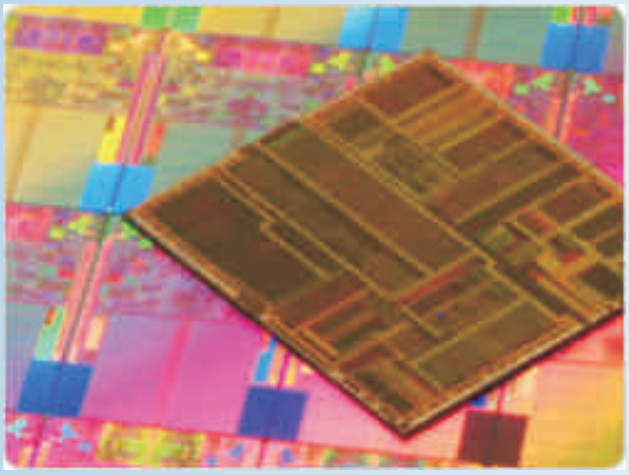
...45 nm technology: What makes it so special?

tunneling through it and yet permeable enough to let the gate's electric field into the channel so that it could turn on the transistor. In other words, the material had to be physically thick but electrically thin.

To overcome this problem, hafnium- and zirconium-based "high-k" dielectrics were used and to test capacitors were fabricated out of these. But the results were discouraging since it was found that charges got trapped at the interface between the gate electrode and the dielectric. This accumulated charge within the capacitor altered the voltage level needed to store the same amount of energy in the capacitor from one charge-discharge cycle to the next. But an innovative technology called atomic layer deposition was used for fabrication which smoothen the surface. This solved the problem of gaps and pockets in which charges could get stuck.

With two candidate materials identified, NMOS and PMOS transistors were fabricated out of them but they found some snags. For one thing, it took more voltage to turn them on than it should have—what's called Fermi-level pinning. For another, once the transistors were on, the charges moved sluggishly through them—slowing the device's switching speed. This problem is known as low charge-carrier mobility. The source of the trouble, ultimately, came down to the interaction between the polysilicon gate electrode and the new high-k dielectrics. They found out that metal gates solved the problem. Intel has not revealed it's identity due to competitiveness.

The new gate stack worked wonders in battling leakage through the gate, reducing it by more than a factor of 10. But the gate oxide is not the only source of transistor leakage chip makers have to worry about. The other significant leak is called source-to-drain or subthreshold leakage. It's a trickle of current seen even when the transistor is intended to be in the "off" state. Making transistors smaller has also meant steadily lowering the amount of voltage needed to turn them on, the threshold voltage. Unfortunately, steadily lowering the threshold voltage lets more current slip through. For many years, each new generation of transistor would increase drive current (and improve performance) by about 30 percent but would pay the price of about a threefold increase in subthreshold leakage. Leakage currents have reached levels high enough to be a noticeable portion of total microprocessor power consumption.



The Intel Xeon CPU Core

The industry is now in an era where power efficiency and low leakage are more important than raw speed increases. Compared with the previous 65-nm transistors, 45-nm high-k plus metal gate transistors provide either a 25 percent increase in drive current at the same subthreshold leakage or more than a fivefold reduction in leakage at the same drive current, or anywhere between those values. In January 2007, Intel made the first working 45-nm microprocessors using these revolutionary high-k plus metal gate transistors. One was the Penryn dual-core microprocessor, which has 410 million transistors. The quad-core version of this product will have 820 million transistors. Penryn was followed a few months later by Silverthorne, a single-core microprocessor with 47 million transistors. The invention of high-k plus metal gate transistors was an important breakthrough. Although Intel could have continued to shrink transistors to fit the dimensions needed for the 45-nm generation without this breakthrough, those transistors would not have worked much better than their predecessors, and they certainly would have expended more watts.

It is believed that this new transistor can be scaled further, and development is already well under way on the next-generation 32-nm transistors using an improved version of high-k plus metal gate technology. Whether this type of transistor structure will continue to scale to the next two generations—22 nm and 16 nm—is a question for the future. Will we need new materials and new structures again? Nobody knows for sure. But that is what makes integrated circuit research and development so exciting.

(Talla is a 3rd yr. B. Tech student at the Deptartment of Electronics & Communication Engineering, IIT Guwahati)

An Experience at IISc!

Final Year's **Prachi Singhal** shares the details of her internship at the world renowned **Indian Institute of Sciences, Bangalore**

With my internship almost confirmed at CSTR, University of Edinburgh, I was kind of relaxed with the apping work and was preparing for my trip to UK when suddenly in the month of April I got a mail from my professor saying that my internship grant was not approved and that I may not be provided the AMI scholarship. Yes, my UK internship was cancelled and I had to think of the other options in store for me. Shattered by fate, I started applying again and though I missed the deadline for the IISc fellowship programme by a good margin, I still dropped a mail to a renowned professor in my field of interest with a desire to work in a reputed university expecting some luck. Within a week's time, I received a mail from a research student working under the professor, expressing his interest in my profile and the kind of work I have been involved with in the past. He also notified me about a telephonic interview with my professor which happened to be a detailed discussion on my projects and field of interest. I was more than happy to accept this opportunity as the work involved a unique and unconventional idea.

In the Speech and Audio Lab at IISc, I found the faculty and research staff very cooperative and highly dedicated to their work. I was allotted a laboratory in the very beginning with all the essential apparatus present. My project was a research work on the idea of echo cancellation in telecommunication systems via a novel frequency-domain approach unlike the conventional time-domain adaptive methodology. Telecommunication systems essentially consist of a loudspeaker and a microphone and during any conversation, the output signal from the loudspeaker may enter through the microphone, resulting in the problem of undesired echo at far end user. Presently, the methods adopted for echo-cancellation in such systems includes using adaptive algorithms like Least Mean Square methods and its advanced versions which operate on real-time signals in time-domain and involves removal of echo signal by first detecting it and then subtracting it. My guide had already done some research in frequency-domain attributes of similarity between far-end signal (the original signal received via speakers) and echo signal. My work involved the study of these attributes and to come up with different algorithms for system design via frequency-domain approach and build a robust echo cancellation system that is as good as the time-domain echo cancellation system but also overcomes its disadvantages.

I spent my first week studying about the adaptive algorithms used for echo cancellation. These algorithms suffered with problems caused due to time-delay in far-end signal and echo signal, amplitude variations in echo signal on changing the distance between speaker and microphone, ineffective convergence of the algorithm. Due to this the system was less robust. The main motivation behind the novel frequency domain implementation was the formant matching in the far-end signal and its reverberated echo signal. This attribute can be used to differentiate the echo signal from the near-end signal and thus detect its presence automatically, rather than applying some threshold based detection for it. Another advantage was the reduced complexity of implementation in the frequency domain; the need of convolution in time-domain is just replaced by simple multiplication in frequency domain, which is further simplified to addition on logarithmic scale. The further focus of my work was to generate effective algorithms for system identification, keeping in mind the short-term analysis of signals with constrained window size. This novel approach produced better results than the adaptive methods and the research is still in progress. When implemented in practical real-world systems, this can provide robustness in all Teleconferencing system like Speakerphones, Hands-free car phone systems, Conference phones such as Polycom's Sound-station, Installed room systems which use ceiling speakers and microphones on the table etc.

The things that particularly appealed to me at IISc, were the dedication shown by all the people working there which was further encouraged by their flexible working hours and the centre being open even on weekends. The international online library of IISc is worth mentioning. The ready availability of all the needed infrastructure before-hand to everyone, be it a research scholar or an undergraduate trainee like me was very impressive. I highly enjoyed my work there and the environment further added to my motivation. It feels great to work on a technology that will make a difference in the future. It was indeed an enriching as well as a learning experience for me.

In Depth Analysis

Nokia BL-5C battery Problem

A few days back I received a call from my parents informing me about the recent problem found in Nokia BL-5C battery, the grave consequences of which could lead to an explosion in your mobile. Just like millions of Nokia users around the world, I checked the status of my phone's battery straightaway and was relieved when I found it safe. But here I was, quite bemused at what possibly could be so wrong with such a widely used technology and here's what I found out:

As per Nokia's explanation the problem was with some 46 million batteries manufactured by one of its supplier Matsushita Battery Industrial Co. Ltd. of Japan between December 2005 and November 2006.The defective batteries were getting overheated initiated by a short circuit while charging, causing the battery to dislodge. The consequences of the problem were so serious that Nokia had to exchange the batteries for free without even taking the shipping expenses.

BL-5C is a 3.7V Li-ion battery; it may seem to you that the problem was an extremely rare one as Li-ion batteries are employed in almost all modern electronics gadgets and is one of the widest used batteries. But few know that many big companies including Sony had also withdrawn their laptops and other equipments at some point of time due to the same potential risk of exploding batteries. Li-ion batteries were introduced by Sony in 1991 and had great advantages over the archaic nickel based or lead acid batteries. These batteries can hold twice the energy as the other old batteries, don't possess the memory effect i.e. easy to charge and discharge and the battery doesn't require scheduled cycling to prolong its life. It also does not suffer from sulfation problem of lead acid that occurs if the battery is not used for a long time. They have a low self-discharge and are environment friendly and easy to dispose. So where does the trouble creep in?

As the saying goes, nothing comes without a price. For obtaining high energy density, the manufacturing processes have become more critical and with thickness of only 20-25µm, any small intrusion of metallic dust particles can have devastating consequences. On rare occasions microscopic metal particles may come into contact with other parts of the battery, leading to a short circuit. Although battery manufacturers strive to minimize the presence of metallic particles, complex assembly techniques make the elimination of all metallic dust nearly impossible. A mild short will only cause an elevated self-discharge and a little heat is generated because the discharging energy is very low. If however enough microscopic metal particles converge on one spot, a major electrical short can develop and a sizable current will flow between the positive and negative plates. This causes the temperature to rise; leading to a thermal runaway i.e. temperature is increased in a positive feedback loop leading to instability.

Something similar had happened to the batteries manufactured from December 2005 and November 2006 for Nokia. Though a great risk is associated with Li-ion battery it is still believed by experts that it is among the most successful and safest batteries available and more than 2 billion batteries are produced a year for laptops, mobile phone, Digital cameras and many more. Steps are being taken by manufacturers to render them as the safest batteries, some are (1) protection to inhibit high current surges (2)Fuse to cut current flow if its temperature increases above a threshold (3) Fuse to prevent excessive charging or discharging of the battery . We can observe some safety norms by not exposing our phones to high temperature or extreme condition which are also a source of thermal runaway and using only branded batteries for these electronic equipments and most important of all not following any sophisticated tweaks mentioned on a very special forum “Orkut Power-Scrap “

Chemists are working round the clock to yield high energy density as well as safer batteries. Technologies like fuel cells and Nanotube-enhanced ultracapacitors are under development and promise to provide what we can call as an ideal battery. For instance ultracapacitor batteries would be able to achieve a power density of up to 100 kW/kg, which is three times greater than conventional Lithium-Ion batteries and it is expected to be released within the next 5 years.

- Karan Sikka, 2nd yr. B.Tech, Dept of ECE

In the Beginning

The USA has steadily lost ground in the semiconductor industry to latecomer Japan. How did Japan manage to create a knowledge intensive industry without any native seminal technology? What can we learn from Japan's success story? Let's rewind to see the beginning...

Semiconductors form the heart and soul of $ 2 trillion dollar thriving electronics industry and since the birth of transistor in 1947 at Bell Laboratories most of the technical advancements in microelectronics were made by the U.S.A. Besides the transistor, U.S. companies pioneered the integrated circuit, the dynamic memory, the microprocessor, and other critical products and processes and they continue to hold an innovative edge in many critical areas. Yet Japanese companies continue to have major chunk of the market share. In 1991 Japanese companies held 46% of the market compared to 39% held by the American companies, which was in sharp contrast to the situation in 1970's when American companies enjoyed 60% of the world market share.

Japan's success story can be traced back to post world war period i.e. 1951-1980, an era encompassing Japan's economic reconstruction and consolidation when Japanese corporations concluded more than 40,000 contracts with foreign firms, providing for the transfer of technologies deemed critical for commercial competitiveness. By eliminating the uncertainties of research and development they were able to focus on incremental growth and hence incurred huge marginal profits. This windfall of foreign technology improved and upgraded by Japanese firms paved the way for it to become a world class producer.

However transfer of foreign technology is not the only factor underlying its success story. Other nations and industries have implemented "catch-up" strategies in which the transfer of technology from abroad was a major element. What is significant about Japan's experience is the systematic, organized way in which technology was imported, adapted and upgraded.

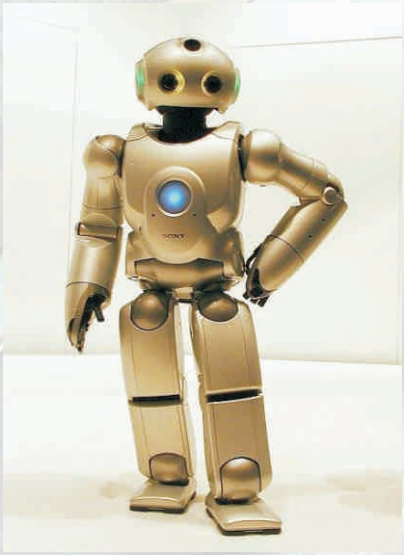
Most of the U.S companies at that time were not far –sighted. They were busy in catering to the expanding domestic market that, most of them, failed to realize the potential in other developing markets. Individual companies might have earned profits from patent revenues but as a whole it contributed to an erosion of America's industrial preeminence. Two of the world's most renowned strategists, James Abeglen and George Stalk go so far as to call the one-way outflow "disastrous". In return of cutting edge technology borrowed, Japanese firms paid a clumsy amount to U.S firms compared to what they actually generated as revenues. But looking from U.S perspective these companies had then reflected upon the circumstances prevalent at that in Japan which hindered the entrance of these companies into their domestic market.

... In the Beginning

Akio Morita
founder of Sony

This trend of world dominance was soon realized by the U.S, as 9 of the 11 DRAM producers soon left the market and EPROM producers were also incurring huge losses making Japan's dominance in the world market complete. During the 1980's it was dumping memory devices all over the U.S.A ,in retaliation to which , the U.S department of commerce filed a suit against them declaring the Japanese products to be 34.5% below the cost of production. With all the pressure on Japan a formal agreement was initiated by, Ministry of International Trade and Industry Japan, and the U.S. Trade Representative's to sort the issue of market access. The U.S Japan semiconductor agreement formally signed in 1986 met with wide spread criticism both from the EU and the American electronics industry as the prices of semiconductors began to rise significantly from their pre agreement values.

American companies were concerned as their competitive edge was getting thinner and the EU was concerned about it's declining share which had already shrunk to 1%. The agreement required Japan to stop dumping in all the markets not alone the U.S. This was a precedent setting condition as a bilateral agreement governed the behavior in a 3rd country. Also it required Japan to open up its market for American companies. As the agreement neared it's end in 1991 there was still no respite for U.S firms, hence the Semiconductor Industry Association started pressing the U.S government for another agreement with Japan. And as a result of the new agreement coming into force for the fist time after almost two decades of continuous decline in world share, the American firms started picking up. In all, six of the world's 10 largest semiconductor producers were Japanese, three were American and one was European, clearly establishing the Japanese supremacy.



SDR-4X Entertainment Robot:
A fine example of
Japan's technological mastery

The indications are clear, semiconductor is too important and lucrative an industry even for the world's only superpower to loose hold off. India, being a developing county has lost the chance to ride many waves which swept the world (only ridding high on the IT wave as it is a low investment venture) then be it the semiconductor wave or biotechnology. But we have important lessons to learn from Japan's experience with capital intensive industries that can give the right technological edge and a kind of cushion against the versatile markets.

- Rohit Bahl (3rd yr. B. Tech student at the Dept. of Electronics & Communication Engineering, IIT Guwahati)

Robust Text Independent Speaker Recognition Based on Hybrid LPC and MFCC Algorithm

In Phase - Technical Article

Sourabh Sriom, Mohit Khurana, Puneet Lakhanpal and Amit Kumar Mishra
Department of Electronics & Communication Engineering
IIT Guwahati

1. INTRODUCTION

The text independent system mentioned in this article must be able to derive its distinctive properties by the use of two characteristic features of a particular speaker`s speech, namely the MFCC [1] and LPC [3] coefficients. The training of the system can be done in any language as preferred by the user, also the conditions under which training is to be done can be non-ideal, i.e. the training sessions may be done in somewhat noisy environment. The duration of the training session is less than a minute. The same will be applicable for all the testing sessions as well, which will be shorter than the training session (about ten seconds). The features of the speakers in the database will be labeled strictly on the basis of the person's identity and independent of the text used during the training period.

This system uses the n-means clustering algorithm for modeling a particular speaker by taking the average values of the considered features from overlapping 20 ms frames of the speaker's sample recorded during the training session. The next section gives an overview of the database used for testing the classification algorithms, which is followed by an overview of the classification algorithms used in the present work.

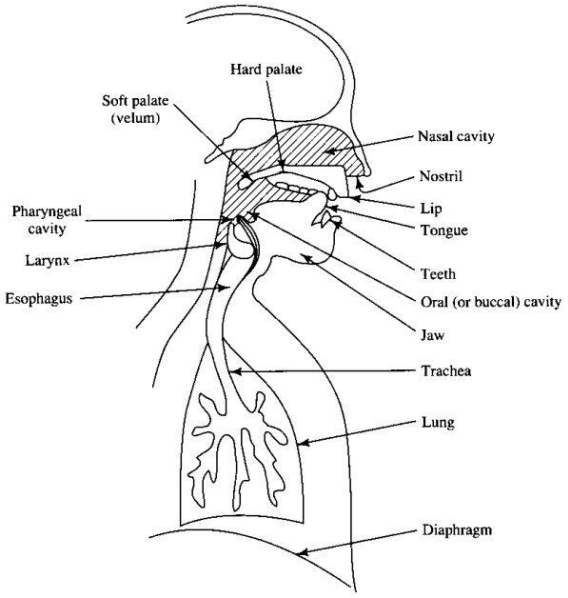
2. HUMAN MODEL OF SPEECH PRODUCTION

The differences in the speech of different speakers is attributed to the differences is the physiological and behavioral aspects of their speech production systems. The vocal tract is the major physiological reason for the differences as it is the speech production organ. The shape of the vocal tract, pitch, intonation pattern of a particular text are the characteristic features of a speaker. Analysis of these features for different speakers gives us the ability to distinguish between them. Speech signals which are nothing but acoustic waves are modified in their spectral contents as they pass through the vocal tract. The speech waves are produced when air from the lungs is carried by the trachea through the vocal folds.

Now, this source of excitation can be classified into various modes. These are known as phonation, whispering, frication, compression, vibration or a general combination of any of the above mentioned modes. The speech produced by a phonated excitation is called voiced speech. The speech produced by phonated excitation in combination with frication is called mixed voiced. Speech produced by other modes of excitation is called unvoiced.

3. DATABASE USED

The speaker recognition system experiments were performed on 9 different speakers including both males and females voices. Speech models for different speakers were modeled using arbitrary utterances generally 40 to 50 seconds. Training could be done using any thing ranging from a short paragraph to even a song. All data was recorded using a commonly available microphone under non-ideal conditions with a sampling frequency of 8kHz.



4. FEATURES USED IN DISTINCTION

4.1. Mel Frequency Cepstral Coefficients (MFCC)

These coefficients posses increasingly fine spectral information, starting from the overall spectrum shape, general formant structure to the more detailed spectral structure between the formants.[5] Assuming that *x(n)* is the input speech signal, the evaluation techniques of these coefficients can be summarized as follows[1]:

1. Calculate the energy spectrum:

$$x(k) = \sum_{n=O}^{Nw-1} x(n)W(n)\exp(-j2\Pi nk/K); \quad (1)$$

with $k \in [0, K-1]$ K is generally equal to 256 or 512;
Nw is the dimension of the Hamming window:

$$W(n) = Sc(0.5 - 0.5\cos(2\Pi n/(Nw-1))); \quad (2)$$

where Sc is a scale factor to take into account the back-ground noise. The energy spectrum is given by:

$$P(k) = |X(k)|^2 \quad (3)$$

2. Calculate the energy in each channel:

$$E = \sum_{k=0}^{5*k-1} \phi_J(k)P(k); \qquad 1 \le j \le J \qquad (4)$$

where *j* is generally equal to 24 and Φ_j is a triangular weighted function associated with the *jth* channel,

$$\sum_{k=0}^{.5*k-1} \phi_j(k) = 1 \qquad (5)$$

3. Calculate the Mel Cepstral coefficients as below:

$$c_j = A\sum_{k=0}^{J-1}\cos(j\frac{\Pi}{J}(k+.5))\log_{10}(E_{k+1}) \qquad (6)$$

The amplifying factor *A* is generally equal to 200.

When they are properly used these features give above 85% accurate results for speaker recognition systems [1].

4.2 Linear Predictive Coding (LPC)

Liner predictive coding is one of the most powerful speech analysis techniques. This technique assumes that speech is produced by a source placed at one end of a tube. The *glottis(space between the vocal chords)* acts as the source while the vocal tract acts as the tube, which is characterized by its resonances, also known as *formants.* The speech signal is characterized by its intensity(loudness) and frequency(pitch). The technique analyses the formants, separates it`s effect from the remaining signal and then analyses the intensity and frequency of the remaining signal.
The calculation of LPC coefficients is shown below.

Linear predictive coding models a signal S_n using its past values and a scaled present value [4].

$$s_n = -\sum_{k=1}^{p} a_k . s_{n-k} + G . u_n \qquad (7)$$

Here S_n is the present modeled value of the signal, S_{n-k} are the past outputs, a_k are the model parameters known as the linear predictive coefficients, *p* is the prediction order, *G* is called the scaling parameter and u_n is the present input which is generally unknown in speech models, so it is ignored while modeling. Therefore the LP approximation of
the speech signal is written as

$$\hat{s}_n = -\sum_{k=1}^{p} a_k . s_{n-k} \qquad (8)$$

i.e. it is independent of the present input.
Because of this reason we minimize the MSE between the actual output and the modeled output so that speaker dependent characteristics are not lost completely.

The prediction error is defined as difference between the actual and the modeled output and can be mathematically represented as:

$$e_n = s_n - \hat{s}_n = s_n - \sum_{k=1}^{p} a_k . s_{n-k} \qquad (9)$$

If *E* is taken to be mean square error then:

$$E = \sum_n e_n^2 = \sum_n [s_n + \sum_{k=1}^{p} a_k . s_{n-k}]^2 \qquad (10)$$

To minimize the MSE we have

$$\frac{\partial E}{\partial a_i} = 0; \quad \forall i = 1,2,3....p \qquad (11)$$

which implies:

$$\sum_{k=1}^{p} a_k . \sum_n s_{n-k} s_{n-i} = -\sum_n s_n s_{n-i} \; \forall i = 1,2,....p \qquad (12)$$

These coefficients can now be evaluated by using the 'auto-correlation' method for LP analysis. The time averaged estimates of the auto-correlation at lag τ can be expressed as:

$$R_\tau = \sum_{i=0}^{N-1-\tau} s(i).s(i+\tau) \qquad \textbf{(13)}$$

The autocorrelation method yields the system of equations named after Yule`s work on all pole modeling in sun spot analysis. The system of equations can be represented as the following matrix equation:

$$\begin{bmatrix} R_0 & R_1 & R_2 & \cdots & R_{p-1} \\ R_1 & R_0 & R_1 & \ddots & R_{p-2} \\ R_2 & R_1 & R_0 & \ddots & R_{p-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ R_{p-1} & R_{p-2} & R_{p-3} & \cdots & R_0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} = - \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_p \end{bmatrix}$$

(14)

Now on solving this matrix equation we can evaluate the LP coefficients which will be used to model speech.

5. PATTERN MATCHING

Once the database is ready and the testing speaker`s model has been created, the system now compares the model of the testing speaker with the ones in the database. In this algorithm we do so by evaluating the Euclidian distance between the corresponding characteristics. The model in the database with the least distance to the testing speaker is identified. Since here we are using a hybrid model based on both MFCC and LPC coefficients, appropriate weights were assigned to the two features according to their respective inter speaker distinction ability as indicated by the observation data.

REFERENCES

[1] Rivarol Vergin, Douglas O' Shaughnessy and Azarshid Farhat, "Generalized Mel Frequency Cepstral Coefficients for Large-Vocublary Speaker-Independent Continuous-Speech Recognition", IEEE Transactions on Speech and Audio Processing, Vol. 7, No. 5, September 1999.
[2] S. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition.", IEEE Trans. Acoust ., Speech, Signal Processing, Vol. ASSP-28, 1980.
[3] B. S. Atal, "Effectiveness of linear prediction characteristics of the speech wave for automatic speaker
identification and verification, "J. Acoust. Soc. Amer., Vol. 55, no. 6, 1974.
[4] Joseph P. Campbell, JR., Senior Member, IEEE, "Speaker Recognition: A Tutorial"
[5]Md .Khademul Islam Molla, Keikichi Hirose, "On the Effectiveness of MFCCs and their Statistical Distribution Properties in Speaker Identification." ,IEEE conference on virtual environments ,human-compute and measurement systems ,july 2004

Photonic crystal lasers move into the blue

- A brief look at the future of LASERs

When we think about LASERS, the first thing striking our minds is a long tube emitting bright light that focuses on a particular point, when aimed. Well, folks you almost got the right idea about'em with the exception that you really don't know what's going on inside the small coloured tube that you are holding in your hand. Here's a peep into the present and so called “future” of the LASER technology.

Starting from the basics, LASER stands for Light Amplification through Stimulated Emission of Radiation (LASER). The name clearly is self-explanatory. LASERS work by amplifying the light through a stimulated emission of radiated light, that we further focus onto something. Getting into the tube we come to know that a laser consists of a gain medium inside a highly reflective optical activity, as well as a means to supply energy to the gain medium. The gain medium is a material (gas, liquid, solid or free electrons) with appropriate optical properties. In its simplest form, a cavity consists of two mirrors arranged such that light bounces back and forth, each time passing through the gain medium. Typically, one of the two mirrors, the output coupler, is partially transparent. The output laser beam is emitted through this mirror.

Light of a specific wavelength that passes through the gain medium is amplified (increases in power); the surrounding mirrors ensure that most of the light makes many passes through the gain medium, stimulating the gain material continuously. Part of the light that is between the mirrors (i.e., is in the cavity) passes through the partially transparent mirror and escapes as a beam of light.

The process of supplying the energy required for the amplification is called pumping. The energy is typically supplied as an electrical current or as light at a different wavelength. A typical pump source is a flash lamp or perhaps another laser. Most practical lasers contain additional elements that affect properties such as the wavelength of the emitted light and the shape of the beam.

With its numerous application and ease of use, LASER has become a hot topic for research for the scientists and as a result these tiny light emitting things are getting better and better with each passing day. Many alterations have made LASERs more useful and easy to use and adapted according to our needs. The latest alteration was made by a group of Japanese scientists who claimed that they, for the first time produced a photonic crystal surface-emitting laser (PC-SEL) that emits in the blue-violet wavelength range. They claim that their innovative fabrication method is an important step in the development of laser sources that can be focused to a spot much less than the emitted wavelength.

They were quoted to have succeeded in developing a new method in the GaN system resulting in a current-driven PC-SEL that emits at a peak wavelength of 406.5 nm. These PC-SELs operating at blue-violet wavelengths could find uses in a variety of new areas including next-generation information storage and micro- to nano-operation in biological or medical fields. Giving an insight into the working of this amazing device, the researcher leading the team said that the difference lies in the area used for the pumping of electrons. Using a large area and the coherent oscillation produced by PC-SELs, very high-power blue-violet lasers could be achieved.

It's worth noting that until now, no fabrication techniques could form high-quality photonic crystals in the GaN system. The key to the success was the development of the air-hole retained over-growth (AROG) method which relies on the unique characteristics of GaN growth. Now taking a look at these characteristics of GaN we see that GaN growth proceeds much faster in the lateral direction than vertical growth on the crystal plane. The air holes are integrated within a layer of AlGaN situated above an active layer of GaN. A combination of nanopatterning and growth techniques is used to create a 2D GaN/air photonic crystal. The structure exhibits a photonic crystal band-edge effect sufficient for operating a current-injection surface-emitting laser.

The laser currently operates with a large threshold current however the team believes that the performance could be improved significantly by optimizing the design. Commenting on the possible ways to make the LASER more advanced than its present stage, the team said that they could improve the crystalline quality of the active layer, optimize growth conditions and semiconductor design.

Truly, technology is the keyword of 21st century and there's still more to go. We still have a long way to go and keep up our pace with the tides of time. Let's hope that some more of these advancements can be made in any field of science so that we can realize what true wonder our mother nature is.

Contributed by:
Harpreet Singh
B.Tech, 2nd Year
Dept of ECE

Crazy Gadgets

Here's our pick of some cool, crazy gadgets, that only geeks would dare possess. Certainly a lot of imagination has gone into their making and you would need a lot of that to use them as well.

- assorted from 'cyberspace'

Transparent Toaster

You love toast, but you always burn it? Then, this invention is for you. This transparent toaster allows you to see the bread while it is toasting so you just have to take it out when the colour is right. This idea is based on a transparent heating glass technology.

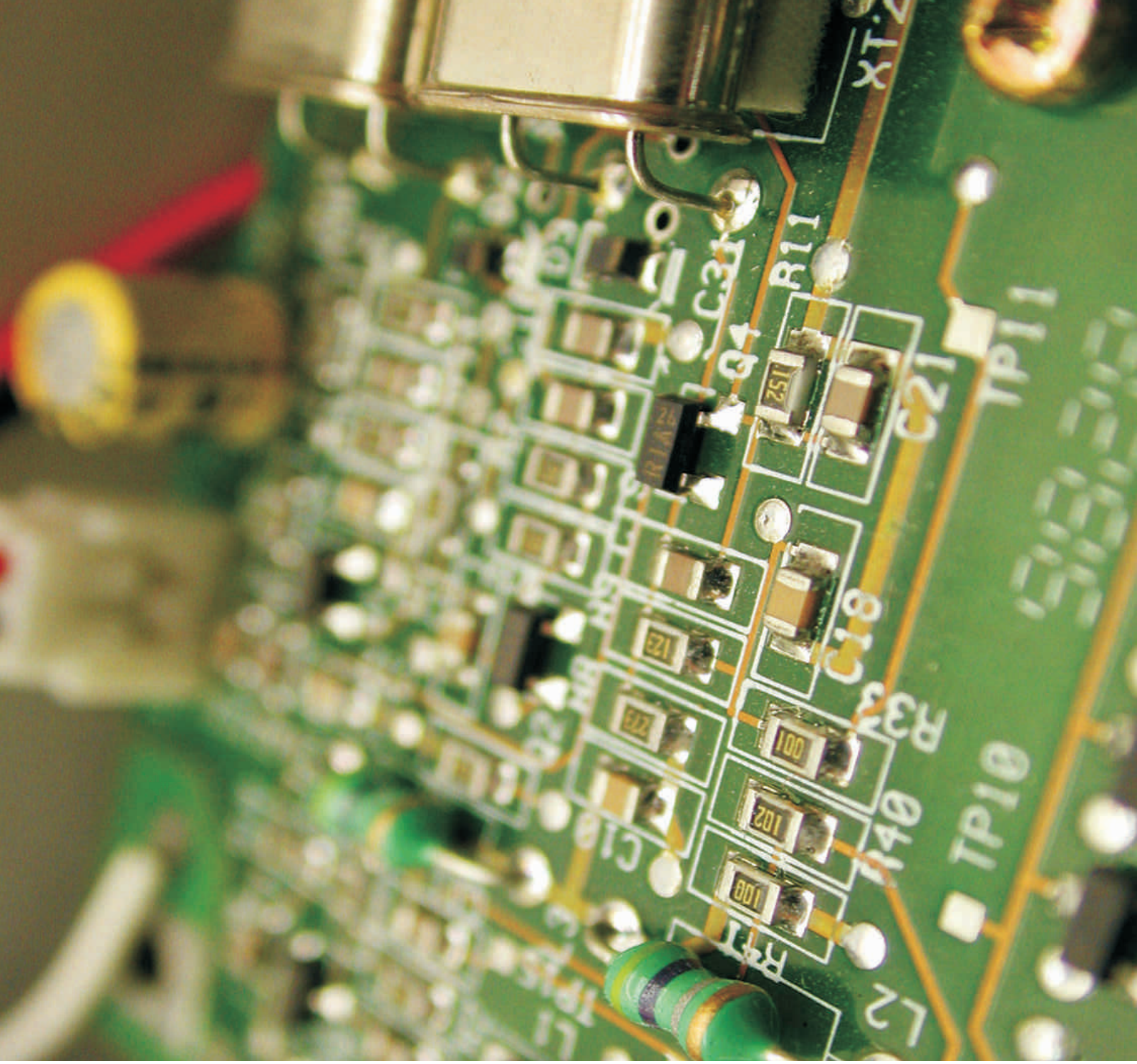
USB Vacuum Cleaner

This 20 centimeters USB Vacuum Cleaner can be plugged into your Computer USB port and help you clean your desk by removing dirt and crumbs. It has larger storage to store more dust and thr 45" cord can clean your messy desk even in your large office desk.

Laser Scissors

Cutting a straight line has never been easier. Just aim the pin-point laser and follow the line. The scissor blades are stainless steel and cut very clean with a micro serrated edge.

Passing out batch of ECE (2004-2008)



In Phase Magazine
By Cepstrum Student Society
Dept. of Electronics & Communication Engg.
Indian Institute of Technology Guwahati

Awaiting your feedback

Please fill up the online feedback form by clicking here

Cepstrum