

Clever Texting and Panini Keypad : New technologies to support Indian languages on mobile phones.

Abhijit Bhattacharjee, Lt Gen (Retd) Prakash Gokarn

Abstract :

It is imperative for all stake holders to support regional languages of India on the phone to take the data opportunity of the mobile phone to all users. Supporting regional languages on the phone has been a challenge regarding offering a usability technology that allows easy and convenient typing on the phone. Printed characters on keypad is unsuitable in a country with many languages and scripts. And multitap support that asks the user to tap upto 9 taps to type a character is cumbersome to use. A new technology of statistical predictive texting that pre-estimates characters that the user could be typing and places them on the screen for selection by the user by single keypresses has offered a viable means to support all languages on the same phone and yet offer a means of typing that is simple to use and is also very fast. This has many advantages and has turned out to be very popular with users as well as manufacturers. This new technology called CleverTexting developed by an Indian company called Luna Ergonomics has been developed for all Indian languages into one integrated platform called Panini Keypad which allows typing in all Indian scripts and also interoperation like transliteration between them. The technology is also applicable for all global languages and has already been developed for most major global languages and is being developed for the others. It is an opportunity for an Indian innovation to determine the way the world writes on digital devices.

Background

There are 90 million people in India who understand some English. But there are already 530 million mobile phone users, and 15 million new users are being added every month. Most of the mobile users of India are unable to type on their phones and hence they cannot store their address books, cannot send an SMS, shall be out of the scope of all the VAS and information enabling services of the government and shall stay out of the the 3G,4G opportunity. Unless we build the suitable regional language ecosystem for the phone, the whole data opportunity of Indian telecom will have to be restricted to the 15% of the phone users who know some English.

The regional language ecosystem shall consist of availability of Indian fonts on the phone with accurate rendering of complex character clusters, a technology for convenient typing of all Indian languages on the phone and finally an uniform standard based encoding of Indian language data between phones and network. Of the three, the challenging was the means to support convenient typing of Indian languages on the phone. The technology has to support all Indian languages, it

should be easy to use and it should offer the means to type fast.

There was a breakthrough in this field from an Indian company called Luna Ergonomics in 2008-2009 who while exploring solutions for Indian languages, developed a brilliant new prediction technology that is based on the statistical nature of languages instead of dictionaries. The entirely new usability that was developed around this offered a means to write in all Indian languages on the same phone, needed no printed characters on keypad, was ergonomic to the thumb, was dictionaryless in implementation apart from a host of other advantages. It was one elegant solution that addressed all the industry acknowledged blindspot in one simple usability shift. The new technology developed in India was not only applicable for Indian languages, but also for all other languages of the world. It has so far been developed for all the major 11 languages of India into one integrated platform called Panini Keypad which supports Hindi, Bengali, Telugu, Marathi, Tamil, Gujarati, Kannada, Malayalam, Oriya, Punjabi and Assamese. And the global family of implementations simply called CleverTexting has been developed for 9 major global languages, namely English, Spanish, Portuguese, French, Arabic, Russian, Hebrew, Swahili and Finnish. And is now developed for many other languages of the world including Chinese, Japanese and Korean. The technologies have won many global awards and recognitions and are being actively considered by phone manufacturers.

The technologies are not only applicable for the mobile phone, but also for all other digital interfaces, be it PC/web, STB, IPTV, ATM, touchscreen kiosks, gaming consoles etc.

Texting technologies

The common handheld mobile phone offers only about 12 keys on its keypad, whereas all languages have many more characters. English has 26 and Indian languages typically have 60 or more characters that form part of the popular usage. Supporting the input of scripts on the phone has been a challenge and many different approaches have been explored for decades. One familiar approach is called multitap where several taps on a key is used to type a character. Other more intelligent approach has included a dictionary on the phone estimating the word from user inputs. These have had their advantages and disadvantages. The Indian language is even more complex to support due to the nature of the language and the above approaches have remained unsuitable. A familiar handicap with the dictionary based system of the past was the inability to support non dictionary words – names of people and places which occur so commonly in the nature of communication done on SMS.

CleverTexting

CleverTexting is a new texting technology that offers a solution with an entirely new usability. In this the phone estimates the limited possible characters that the user could be writing next and shows them on the screen. This allows the user to choose from amongst the predicted characters using a single keypress or directly from the screen in case of a touchscreen device. The predictions of CleverTexting are very accurate because they have been derived from statistical mining of the particular languages in terms of probabilities of letter combinations that actually occurred in large texts called corpora. In 90% of cases, the predictions of CleverTexting turn out

to be accurate irrespective of the language which is the key demonstrated power of the technology.

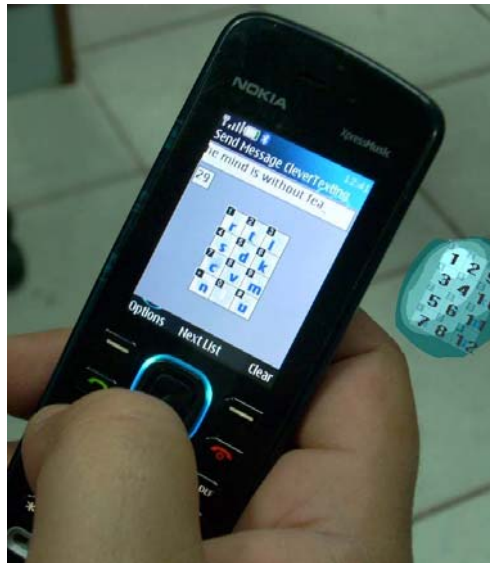


Fig 1. CleverTexting predicting characters that the user could be writing and showing them on the screen from which the user could choose using the keypad or peck directly on the touchscreen The product has been developed for all Indian languages.

Ergonomic Concerns

When the common phone is held in the right hand, all the keys of the phone are not equally accessible to the texting thumb. The key 1 would be the most accessible and the key # would be the least comfortable. Although the common phone is used by far more number of people in the world than the PC keypad, the phone keypad has not witnessed any redesign based on ergonomic considerations. You will see that the current English keypad on the phone just lays out the characters in their alphabetic order. Its called the ITUT keypad.

Since the characters are dynamically allocated in CleverTexting depending on their conditional probabilities, they are also shown on the screen according to their ergonomic ranks. Hence you would be most often pressing on keys which are most comfortable to the thumb. The character arrangements can be changed when the phone is used on the left hand.



Advantages

The approach of CleverTexting offers several different advantages both for the user as well the phone manufacturers.

Advantages to user

1. Much faster typing in terms of speed.
2. Smallest number of keytaps or keypad interactions. Almost the same number as number of characters one writes. Very little overhead.
3. Single keypress typing, no multitap. Seniors can use.
4. Seamless experience across dictionary and non dictionary words. No mode changes required.
5. Ergonomic to thumb. Pleasure to write long sentences, no blackberry thumb.
6. Easy to learn, operate for all age groups and educational profiles because only two simple rules.
7. Characters on phone keypads are normally tiny, many need glasses to read them. CleverTexting characters are on the screen and large to read.
8. Can operate in darkness (no back lit keypad required)
9. Same usability for all devices. PC, web, IPTV
10. Improves spelling habits. People will write full words due to the convenience.

11. The game like interaction with the phone is engaging and grows addictive as one improves speed.

12. The constant hand eye coordination is good for the development of the brain, to stay alert.

Advantages to the manufacturer

1. No need of printing different language characters on keypad, same phone supports all languages. Printed characters on cheap phones, also get worn out and become unreadable.

2. The ability to support a very large number of languages because support through CleverTexting for micro languages is feasible, can be delivered on schedule and cost effectively. The development is automated. And not a rule based approach. No likelihood of errors, bugs.

3. The isolation between algorithm and data. Same algo different data for a different language. The data could also be downloaded to the phone on demand or included by manufacturer based on geography. Or only included into a SIM by a telecom operator. Can be hot upgraded with newer versions.

4. The algorithm is lightweight.

5. Same usability for all languages, nothing to teach1market in terms of usability. Learn once, use everywhere.

6. Same usability applicable for touchscreen phones, nothing to teach.

Why CleverTexting works for all languages of the world.

Our languages are highly statistically correlated. In the Figure below it can be seen that in the English language, of the half a million possible four letter combinations, only about 1% make up 90% of all occurrences. The statistical curve is the same for all languages of the world and hence CleverTexting works for all languages of the world. It is the high statistical correlation which has offered great advantage to this strategy over the previous dictionary based approaches.

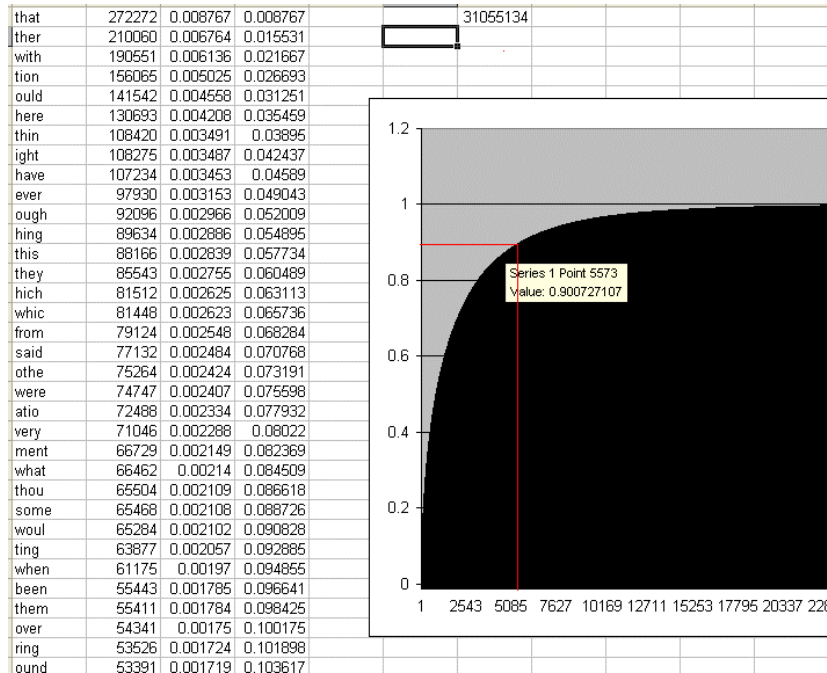


Fig 2 : Statistical correlation of the English language.

SMS compression

The technologies developed also include support for SMS compression for the first time in the world. This uses a variable bit encoding called static huffman coding which could be obtained from the statistical information of languages because all characters do not occur equally frequently in languages. Frequently occurring characters are encoded in fewer bits and rarely occurring characters are encoded in larger bits. The average number of bits per character works to about 5 bits per character which offers 300% improvement for non latin languages which uses Unicode. And about 150% improvement for latin languages.

Panini Keypad

The Panini Keypad is the brand name of the family of products for Indian languages. Named in the honour of the famous Sanskrit grammarian who lived 2500 years ago. The Panini Keypad family of products offers CleverTexting for the following Indian

languages: Hindi, Bengali, Telugu, Marathi, Tamil, Gujarati, Kannada, Malayalam, Oriya, Punjabi and Assamese as well as a few extra features available to the Indian product. This includes integrated support for all the Indian languages and interconversion via transliteration.

The Panini Keypad is fully developed and has been released to the public domain since Jan 2009. It was the winner of the Judges Choice award at the NOKIA Global Mobile Innovation contest where it was called a Mobile Necessity for Emerging Markets.

Transliteration

The Panini Keypad is an integrated platform supporting all the Indian languages, so it is possible to transliterate between the 11 languages supported. So you can write in Bengali and send in Telugu and receive a message in Gujarati and read it in Malayalam. The Panini Keypad also includes transliteration to the Roman script for phones overseas with no Indian language support.

The artificial pan indic script called Shiva.

Indian languages although most scientifically derived in the phonemic sense unfortunately use complex scripts today. These complex scripts that require advanced display systems and complex rendering engines do not lend themselves to simple implementation on limited screen capability devices. These are some drawbacks which have resisted easy adoption and support for Indian languages towards every kind of digitisation.

In the statistical work towards development of the Panini Keypad, all the languages of India were revealed in their modern forms, in terms of characters that are in current usage etc. All the Indian languages are orthographically similar and only typographically different. In other words, the word Bharat would have the same spelling in all Indian languages. The only difference being how character vary in their looks. With the above symmetry in place, an all inclusive pan Indic script system that supports all Indian languages, that has been scientifically derived with all digital considerations like segment display systems, vector rendition, OCR, readability and other factors was developed from the drawing board. A major hallmark of the script system is that all the characters are deducible based on simple rules which makes learning very easy and can help literacy. A new script system would not be easy to introduce, so the Shiva script has been unobtrusively included amongst the integrated language support within the Panini Keypad system. With the ubiquitous usage of mobile phone, people would get familiar and more comfortable to such an introduction on another day in the future when such a common script for all Indian languages would be felt necessitated.

The Sanskrit language that used the Brahmi script since the 6th century BC was a simple script that would not have posed challenges to digital implementation. It is only in the 13th century and onwards, that the Devanagari script was used by Sanskrit and other derivative scripts in the regions have followed subsequently. Therefore it would be wrong to suppose there is any sacred association between the Indian languages in their original form and the complex script that we use today. The

Shiva script resembles the original Brahmi script of the past in concept, form and simplicity. No patent application has been filed with respect to the Shiva script.

Building the regional language ecosystem for mobiles

Other impediments in enabling regional language support on phones that need to be ironed out.

1. Many major phone manufacturers have been supporting one or several Indian language fonts on the phone but there is no standardisation in it and the user is not assured of it. There are some phones sold in India which do not have support for any Indian language fonts. India has only 9 scripts and each manufacturer must include all the Indian language fonts on their phones to ensure support is provided for each language.
2. Even if the Indian language font has been provided by the manufacturer, in some cases it has been found to be of poor quality. They do not render correctly for yuktakshars or are too tiny and not readable. Quality must also be ensured for the fonts.
3. All phones in the world support all non latin languages through Unicode, both for SMS as well as for Internet data. Most phones in India also do so, but there is a miniscule number of phones that do not support Unicode encoding and decoding. This standard must be ensured by education of the stake holders for assured interoperations. Its a pleasure to see that even the inexpensive Chinese imported phones also support Hindi with Unicode support.

Applications of CleverTexting for other languages of the world.

The CleverTexting concept is applicable for every language of the world. All the languages of the world analysed by us show a statistical correlation which will lend themselves to efficient statistical predictions. The technologies have already been developed for English, Spanish, Portuguese, French, Arabic, Russian, Hebrew, Swahili, Finnish and are popular downloads all over the world as awareness of this new usability is spreading. The technologies are also being developed for all the other languages of the world including Chinese, Japanese and Korean.

The approach of CleverTexting is very productive because it is not a rule based approach and requires no linguistic support to build the product. The process of building being automated allows for the development of this technology for all the micro languages of the world. No other technology, particularly dictionary based approaches do not allow for this.

Application to other devices

The usability of CleverTexting is not only restricted to the mobile phone but is applicable for all kinds of digital interfaces. It can be used in the PC/web to write in

any Indian language. It can also be used by Set Top Boxes and in IPTV for offering a simple usability to type on the TV from the remote of the TV. It can be used in an ATM because that too is a limited keypad device with the same interface. And it can be used in touchscreen kiosks, vending machines etc.



How the technology can reach users

The technology can reach the user in the following manner.

1. Handset manufacturers implementing this technology on their phones.
2. Telecom operators offering the technology has a subscribed application to compatible phones.
3. Direct to Consumer as a software via Retail Sales and Online selling.
4. Enterprise and E-Governance solutions

Conclusion

The CleverTexting and Panini Keypad products are revolutionary in the way people can type fast on their phones and other digital interfaces. It is an approach which has addressed all the requirements of supporting multiple languages on the same phone into one elegant solution. The technologies are being studied keenly by all stake holders. It is an opportunity for an Indian innovation to define the way people write in India and different countries of the world. The Ergonomist Society of UK hailed CleverTexting as "Thumbs up to a new way of Texting" in their cover story in March 2009, describing it as powerful, promising and full of new possibilities.

References

1. CleverTexting : Thumbs up to a new way of Texting, Cover Story, Ergonomist Society, UK, Mar 09. <http://www.clevertexting.com/downloads/ErgonomistMar09p89.pdf>
2. Judges Choice award for Panini Keypad : NOKIA Global Mobile Innovation contest. Sep 09. http://www.callingallinnovators.com/judges_choice_awards.aspx
3. Whitepaper on CleverTexting

<http://www.clevertexting.com/downloads/CleverTexting&PaniniKeypad.pdf>

4. Video on CleverTexting

<http://www.youtube.com/watch?v=VJgR2nKggEI>

5. CleverTexting selected for DST – FICCI - Lockheed Martin India Innovation Growth Programme for global commercialization of Indian innovations. <http://indiainnovates.in/>

6. <http://www.PaniniKeypad.com>

7. <http://www.CleverTexting.com>

8. <http://www.PaniniKeypad.com/PaniniKeypadforweb.html>

A usability to type in all regional languages on the PC/web to fill forms in regional languages using the Panini Keypad technology.