

Language-Independent Communication using Icons on a PDA¹

Siska Fitrianie and Leon. J.M. Rothkrantz

Man-Machine-Interaction, Delft University of Technology
Mekelweg 4 2628CD Delft, the Netherlands
{s.fitrianie, l.j.m.rothkrantz}@ewi.tudelft.nl

Abstract. As language is fundamental to human activities, proficiency in other languages becomes important. Besides for developing abilities for communication, the knowledge is also a tool for a survival. With the introduction of computerized mobile devices, i.e. PDAs, new opportunities for communicating in other language arose. This paper describes a new communication paradigm that is language independent using icon language on a PDA. Users can create iconic messages as realization of their concepts or ideas in mind. The proof of concept tool is able to interpret and convert the messages to (natural language) text and speech in different languages. To provide faster interactions in next icon selection, the tool has icon prediction. Our user test results confirmed that using provided icons our target users could express their concepts and ideas solely using a spatial arrangement of icons.

1 Introduction

Humans communicate to share facts, feelings, and ideas among each other. The dream of being able to understand and communicate in any language has not yet been satisfied. However, there do exist signs and symbols (icons), which are understood universally. It is supported by the fact that on one hand while interacting with the environment and with others, humans form internal mental models and representation of themselves, the outside world and of things which they are interacting [19]. On the other hand, an icon is understood as a representation of a concept, i.e. an object, an action, or a relation. An icon serves a function as communication means by virtue of a resemblance between the icon and the object or the movement it stands for.

Icons also offer a direct method for conversion to other modalities. According to [6], an icon can be interpreted by its perceivable form (syntax), by the relation between its form and what it means (semantics), and by its use (pragmatics). By this

¹ The research reported here is part of the Interactive Collaborative Information Systems (ICIS) project, supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024

way, icons also form a language, where each sentence is formed by a spatial arrangement of icons [7]. The meaning of individual icons can represent a word or a phrase, which is created according to the metaphors appropriate for the context of this type of languages. Since icons are representations of models or concepts, with which humans are actually interacting, we expect this language is easy to learn. Once a set of iconic representations is established, increased usages can lead to more stylized and ultimately abstract representation, as has occurred in the evolution of writing systems, e.g. the Chinese fonts [9].

As icons offer a potential across language barriers [19], the icon language can also offer a potential as an independent human natural language like Esperanto. Thereby, any interaction using this type of languages opens opportunities for the development of different applications in different languages and domains. One might think of those where verbal communication is not possible, for example a communicator media to help people with speech disability.

Nowadays, icons are used in almost every Graphical-User-Interface-based computer software. Besides a direct manipulation on the icons allows us to have a faster interaction, as pictorial signs, they can be recognized quickly and committed to memory with surprising persistence [12]. Therefore, icons can evoke a readiness to respond for a fast exchange of information and a fast action as a result [17]. A research showed that direct manipulation with a pointer has better time performance than form filling with Soft Input Panels or handwriting recognition [15].

Despite its popularity as a personal information manager, the user interaction options for a Personal Digital Assistant (PDA) are quite limited. Like a traditional workstation, interaction with a PDA can be realized by a pointing device on a sized external keyboard. Most PDAs only contain a few physical buttons and a few of them have a small-sized keyboard. Recent researches have been done in adding multimodal capabilities to a PDA, such as a multimodal life-like character [20], fusing speech and pen input [10], speech recognition and synthesis [8], etc. However, for optimal speech recognition, the environment in which the technology is used should be the same as the training environment of the system [5], whereas PDAs are often used in various environments under various conditions. The result is misrecognition of commands, which is frustrating to the user. Since the current technology makes speech input less suitable for mobile activities, we aimed at a natural interaction style that is based on the strength of GUI. With regard to the applicability of icons as communication means and also as a natural interaction style, in this project, we developed an iconic interface that applied as a communication tool for travelers on a PDA.

2 Related Work

Icons have been used as early as the middle ages complex iconic systems such as the heraldic coats of arms, astrological signs, and the communication system of ancient Egyptian. In modern society, icons are familiar from the everyday context of living to the packaging for the latest products, for example: door signs, road signs, elec-

tronic goods manual, etc. In the computer world, they have role as a small graphical representation of a program, resource, state, option or window.

Recent researches have been done in developing computer-based iconic communication, for example: the Hotel Booking System that allows communication on a restricted domain [18], CD-Icon: a pure person-to-person communication system [2], Sanyog: a communication device for disable people in India [1], and the Elephant's memory: a computer iconic environment that allows the user to build a visual message by combining symbols from its vocabulary [14]. Most of these iconic communication systems are too complex to learn or a language-specific system. A deep research has been done by [16] on designing an iconic communication system to allow people communicate with each other when they share no common language. The system is based on the notion of simplified speech by reducing a significant complexity, such as: it has no inflection, no number, gender, or tense markers, no article and no linear order. This is possible because it was designed as a visual language in contrast to written languages, where sequencing and ordering is critical.

The icon language proposed in this paper is used to represent concepts or ideas. It was designed particularly suitable for language-independent contexts. The use of PDAs created new constraints and requirements for user interaction and the interface. Therefore, besides adopting the language simplicity principle of [16], our developed tool was designed also for a mobile use of context.

3 Corpus-based Natural Language Processing

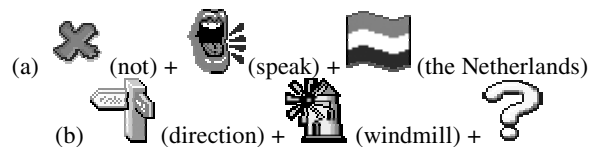






Fig. 1. Two examples of icon strings: (a) I do not speak Dutch and (b) Do you know the direction to a windmill

Icon language is a visual language where each visual sentence is a spatial arrangement of icons. Figure 1 shows two example iconic messages. This language is based upon a vocabulary of icons where each icon has a unique or multiple meanings [7]. Individual icons provide only a portion of the semantics of an iconic sentence. The meaning of individual icons with more than just one icon can still be represented by the same set of icons, but it turns out to be very difficult to determine the sentence meaning. It is due to the only thing that can be automatically derived from the semantics of the icons in an iconic sentence is a fixed word or phrase belonging to these icons. For example: “not”, “speak”, and “the Netherlands”. The meaning of an iconic sentence cannot be understood without a global semantic analysis of the sentence. This meaning is derived as a result of the combination of these icons. It is unlikely that constructing an iconic message with one to one matching of icons to words would be appropriate. Humans, on the other hands, have innate ability to see

the structure underlying a string of symbols. They are equipped with a number of innate rules that determine the possible shapes of human languages.

We conducted an experiment to have more insight of how humans express their concepts or ideas using this type of message [11]. It involved eight participants and one as a pilot. The participants were asked to form iconic sentences that represented six natural language sentences using a given set of icons. We compared the results by analyzing a large number of iconic message corpora. Although to represent a natural language sentence may use different set of icon string, the results from both studies showed a similarity. All participants tended to use icons to represent important keywords of a message. Based on this study, we developed grammar rules using Backus Naur Form. We used English as the base-language to define these rules. The terminal symbol of the grammar is a lexicon of allowable vocabulary icons. This vocabulary was grouped into its word class, such as: nouns, pronouns, proper-nouns, verbs, adjective, adverbs, prepositions, and quantifier. The grouping was done based on the metaphor represented by each icon, for example:

- noun →  (the Netherlands) |  (direction) |  (windmill)
- verb-transitive →  (speak)

These terminal symbols were combined into phrases as nonterminal symbols, such as Sentence, Noun Phrase, Verb Phrase, Prepositional Phrase and Negation Sentence. Two iconic messages in figure 1 above result the following grammar rules:

```

Noun-Phrase → noun | noun Noun-Phrase
Verb-Phrase → verb-transitive Noun-Phrase
Sentence → Noun-Phrase question-sign | Verb-Phrase
Negation-Sentence → negation Sentence

```

A parser takes an icon string and extracts each icon against the grammar rules. If the icon string is syntactically correct, the parser creates seven slots of a natural language sentence: a prefix slot for question words; a subject slot; an infix slot for a to-be or an auxiliary and a negation (“not”); a verb slot; an object slot; a preposition slot; and a suffix slot for a question mark, an exclamation mark, or an archaic word (“please”). A slot may be empty or contain more than one word. The arrangement of the seven slots depends on a sentence type. The corpus (b) in figure 1 results the seven slots depicted in table 1(a).

We developed a rule based-module to construct a complete sentence from the syntax analysis result. The rules specify conversion of an iconic sentence into a natural language sentence based on the semantic context of the iconic sentence. They derive the meanings associated with iconic sentences from the meaning associated with the individual icons forming the sentence. Based on the corpora (b) in figure 1, table 1(b) shows the results of some rules, such as: a rule for adding a subject, a rule for adding a verb, a rule for adding an article, a rule for adding an auxil-

iary, and a rule for adding a preposition. We developed these rules using a corpus based approach. Table 1(c) shows the sentence composition from the resulted slots.

Table 1. An example of an icon string conversion into seven slots (a), a construction of a complete sentence (b), and composing a natural language sentence (c).

(a) Syntax analysis result:	Prefix = [empty] Subject = [empty] Infix = [empty] Verb = [empty] Object = [noun ("direction")] Preposition = [noun ("windmill")] Suffix = [sign ("?")]
(b) Sentence construction result:	Prefix = [empty] Subject = [pronoun ("you")] Infix = [auxiliary ("do")] Verb = [verb-transitive ("know")] Object = [article ("the")] + [noun ("direction")] Preposition = [preposition ("to")] + [article ("a")] + [noun ("wind- mill")] Suffix = [sign ("?")]
(c) Sentence composition result:	Infix+Subject+Verb+Object+Preposition+Suffix Do you know the direction to a windmill?

4 N-Grams Icon Prediction

To generate a smoother dialog in communication and due to the high number of presented icons, an icon prediction could help users to reduce the number of look up processes that are necessary to generate an icon string. Besides improving the input speed, the prediction also improves the quality of syntax. It predicts which icons are most likely to follow a given segment of an icon string. For this purpose, a word prediction technique was adapted. It operates by generating a list of suggestions for possible icons after the first icon is selected. The predictor only considers those next icons that can create a grammatically correct sentence. When a user selects one of the suggestions, it is automatically inserted into the string. Otherwise, the user can continue to input icons until the intended string appears.

The probability of an icon string is estimated with the use of the following Bayes rule as the product of conditional probabilities:

$$P(s) = P(w_1, w_2, \dots, w_n) = \prod_{i=1}^n P(w_i | w_1, \dots, w_{i-1}) = \prod_{i=1}^n P(w_i | h_i)$$

where h_i is the previous segment of an icon string when prediction w_i (=icon). Estimating terms of the form of $P(w|h)$ is done by assuming some generative probabilistic model using estimating conditional probabilities of n-grams type features. The n-grams language modeling aims at computing the frequency of an icon (uni-gram), icon-pairs (bigram), proposing the most likelihood icon by knowing the previous

ones (trigram), and proposing the most likelihood for the entire sentence (n-gram). As the approach needs a large amount data to compute the multi-grams model, our tool collects it from user selections during the interaction.

5 A Language Tool for Travelers

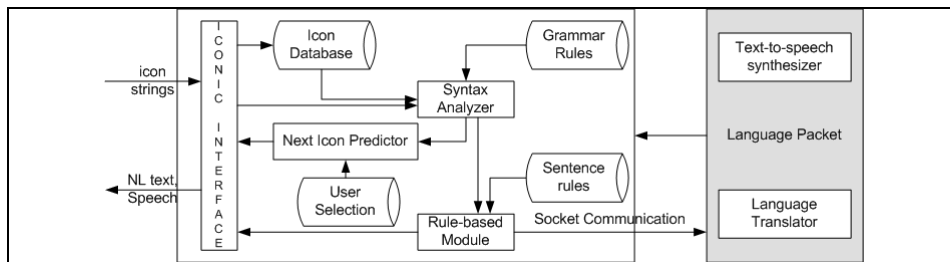


Fig. 2. The architecture of our iconic communication tool on a PDA

Our iconic communication tool was designed to play as a foreign language phrase book for travelers. A user can select a sequence of icons as a realization of his/her concepts or ideas. Besides supporting a fast interaction by converting the message into natural language, the tool also can be combined with any language translator. It also provides a speech synthesizer to read aloud the resulted translations with correct pronunciations. Figure 2 shows the architecture of our developed tool.

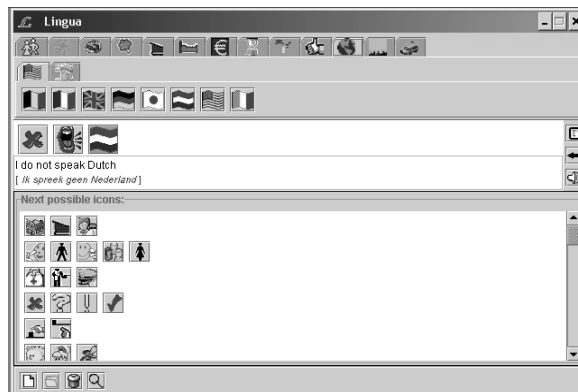


Fig. 3. The interface of our iconic communication tool on a PDA.

Figure 3 shows the current version of our language tool. On the interface, icons are grouped based on their concept. Besides supporting a compact interface, this grouping is a powerful way to hint where an icon can be found because the meaning of icons almost exists in the context of other icons [13]. The next-icon predictor ranks and groups the next possible icons based on their concept. A user can select any icon from the menu or from the prediction window. Any changing on the input,

the resulted text will be refreshed. The interface was designed to cope with the mobile users who cannot devote fully attention to operate the system.

Through a socket network, the tool is able to create a communication environment by TCP/IP connections. The socket allows the interface to communicate to other language tool (i.e. a text-to-speech synthesizer and a language translator) or to other applications if applicable (e.g. a web application or other communication system). These language tools and applications can be produced by third parties. For an experiment purpose, we used CMU's Flite [3] for reading aloud the resulted text.

6 Evaluation

We conducted a laboratory test using the Thinking Aloud method [4]. It involved eight participants and one pilot. They were selected from different linguistics background. This test aimed at evaluating whether or not users capable to express their concepts in mind solely using a spatial arrangement of icons. The test also addressed usability issues on interacting with the communication tool. Each participant performed five tasks that were explained using cartoon-like stories. A story described a situation for travelers in which the participants should have used the tool to communicate with others. The participants were asked to create iconic messages based on the story and think aloud. There were no incorrect answers, except if a participant did not perform a task at all. All activities were recorded on a tape and all user interactions were logged for analyses purposes. At the end of the test, the participants were asked to fill their satisfaction on interacting with the tool.

From the results [11], it appeared that our target users could compose iconic messages for every given task. They understood that they could select an icon string to represent important keywords of their message. However, we also found that users tried to generate some icon strings, which were out of domain. They needed time to adapt the interface and rethink to find another relevant concept to represent their message using only the given set of icons

7 Conclusion

An experimental iconic communication tool has been developed, applied on a language tool for travelers. The developed tool was designed to be language independent with the use of icons and the connection to a socket network to communicate with other digital language aids or other applications. Thereby, it opens opportunities for the development of other application in different languages and domains.

NLP has provided a method to interpret and convert iconic messages to natural language texts/speech. We solved the problems of ambiguity and missing information that resulted by this type of messages. The BNF grammar has been used to select sentences which are grammatical correct. We also developed rules to create more natural sentences. N-grams language model was used to predict the next like-

likelihood icon given a segment of icons. Our experimental results showed that an icon interface could serve as communication mediator. However, future work is still necessary to analyze more corpora of iconic messages. Furthermore, a field study is also necessary to gather data about how people use and experience the iconic interface in real situations to cover more user requirements in mobile context use.

References

1. Basu A., Sankar S., Chakraborty K., Bhattacharya S., Choudhury M., Patel R.: Vernacula Education and Communication Tool for the People with Multiple Disabilities, Development by Design Conference, Bangalore (2002)
2. Beardon C.: CD-Icon, an Iconic Language-Based on Conceptual Dependency, *Intelligent Tutoring Media*, (1992), 3(4)
3. Black A.W., Lenzo A.K.: Flite, a Small, Fast Speech Synthesis Engine, System Documentation Edition 1.2, for Flite version 1.2, USA (2003)
4. Boren T., Ramey J.: Thinking Aloud, Reconciling Theory and Practice, *IEEE Transactions on Professional Communication*, (2000), 43-3
5. Bousquet-Vernhettes C., Privat R., Vigouroux N.: Error Handling in Spoken Dialogue Systems, Toward Corrective Dialogue, *Proc. Of ISCA'03*, USA (2003)
6. Chandler D.: *Semiotics, the Basic*, Routledge (2001)
7. Chang S.K., Polese G., Orefice S., Tucci M.: A Methodology and Interactive Environment for Iconic Language Design, *International Journal of Human Computer Studies*, (1994), 41: 683-716
8. Comerford L., Frank D., Gopalakrishnan P., Gopnanth R., Sedivy J.: The IBM Personal Speech Assistant, *Proc. of the ICASSP 2001*, USA (2001)
9. Corballis M.C.: Did language Evolve from Manual Gestures?, *3rd Conference of The Evolution of Language'00*, France (2000)
10. Dusan S., Gadbois G.J., Flanagan J.: Multimodal Interaction on PDA's integrating Speech and Pen Inputs, *Proc of EUROSPEECH'03*, Switzerland (2003)
11. Fitrianie S.: An Icon-Based Communication Tool on a PDA, Postgraduate Thesis, Eindhoven University of Technology, the Netherlands (2004)
12. Frutiger A.: *Sign and Symbols, Their Design and Meaning*, New York: van Nostrand Reinhold (1989)
13. Horton W.: *The Icon Book*, New York, John Wiley (1994)
14. Housz T.I.: *The Elephant's Memory* (1994-1996)
15. Kjeldskov J., Kolbe N.: Interaction Design for Handheld Computers, *Proc. of the 5th APCHI'02*, Science Press, China (2002)
16. Leemans N.E.M.P.: VIL, A Visual Inter Lingua, Doctoral Dissertation, Worcester Polytechnic Institute, USA (2001)
17. Littlejohn S.W.: *Theories of human communication*, 5th Edition, Wadsworth (1996)
18. Mealing S., Yazdabi M.: *Communicating Through Pictures*, Department of Computer Science, University of Exeter, England (1992)
19. Perlovsky L.I.: Emotions, Learning and Control, *Proc. of International Symposium: Intelligent Control, Intelligent Systems and Semiotics*, Cambridge, MA (1999), 131-137
20. Wahlster W., Rethinger N., Blocher A.: SmartKom, Multimodal Communication with a Life-Like Character, *Proc. Of EUROSPEECH'01*, Denmark (2001)