igwana: a Text-Free Search Interface

Shourov Bhattacharya  
Pongo Labs  
Melbourne, Australia  
shourov@pongolabs.com

Luke Feldman  
Pongo Labs  
Melbourne, Australia  
luke@pongolabs.com

ABSTRACT
Keyword-driven search has become an important way for people to find content on the World Wide Web. However, it is estimated that more than 700 million adults (Warrilow, 2009) lack sufficient literacy in either English or another major language to use existing search interfaces. As a result, these users are to a large extent ‘locked out’ from the Web. We seek to address this challenge through the development of a user interface (‘igwana’) to navigate effectively through large sets of content by using pictograms in place of text. This paper describes some of our preliminary research and ideas and a proposed design for the igwana system.

Author Keywords
Search, text-free interfaces, symbols, pictograms, low-literacy, semantic network, menu, search engine.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Searching for content on the World Wide Web requires a basic level of literacy in a major language. In the early days of the Web, users often used large, human-curated directories to browse through hierarchical category trees. These have now been supplanted by the use of text-based search engines which accept keyword inputs and return matched content sets. In all prevailing search interfaces, it is assumed that users are able to a) read and parse results, instructions, warnings and suggestions and b) input keywords with correct or near-correct spelling. Note that even when the end-point of a search is not text – such as videos in YouTube – the search process presupposes literacy in the interface language.

The keyword search interface is a barrier to participation for those with low literacy. Though the user base for connected devices has grown strongly in recent years, low-literacy users who are physically connected to the Web are largely shut out from social media, news, community sites, health resources and online entertainment. This is especially unfortunate as these excluded users are also often socially and economically marginalized and would derive a large marginal benefit from being online.

The ‘igwana’ project aims to encourage online participation, provide life-long learning and independence and open up new opportunities for low-literacy populations. It is motivated by the question – can low-literacy users use a simple, pictorial, low-text user interface to navigate large and complex datasets such as those found on the Web?

RELATED WORK
Low text user interfaces
Many information technology initiatives for low-literacy programs utilize basic PCs with the standard web graphic user interface (GUI). These designs are often targeted to the educated user and the literate (UNESCO, 2008). An evaluation of the usability of mobile interfaces across user bases in several countries confirmed that text-based interfaces are unusable for low-literacy users or even literate novice users (Microsoft Research).

Research in the area of low-text user interfaces is relatively sparse. A notable body of work is that of researchers at Microsoft Research Labs. Evidence suggests that audio and visual interfaces aid in information processing and can be successfully used by low-literacy users (Sherwani, 2009; Microsoft Research). Empirical work testing a job search application with low-literacy communities in India demonstrated the feasibility of using text-free user interfaces, the importance of design and difficulties overcoming a mistrust of technology (Medhi, 2007; Medhi, 2011). Other work has looked at the effectiveness of various low-text user interfaces for specific applications such as bank ATMs (Thatchera, 2006). Another technology initiative designed for the low-literate population was that of Simputer, a handheld computer designed with a simple user interface of icons based on sight, touch and audio. Although successful in allowing the population to access information, it was not widely adopted due to the need for literacy skills to read the content that was published on the internet (Chatterjee, 2004).

Pictograms and icons
There have been a number of attempts made over the last century to develop pictorial “languages” for unambiguous communication without the use of any written alphabet. Early work on picture languages was done by Neurath and Bliss (Lin, 2006); some recent examples are Zlango (Zlango website), PictNet (Takasaki, 2006) and Pictogram (Pictogram website). An extensive literature exists on the cognitive theories behind our parsing of images and the principles of good pictogram design.
Additionally, many researchers have begun to evaluate the effectiveness of pictograms for applications in various domains. Pictograms are potentially more quickly interpreted and more memorable than text, but the interpretation of even simple picture symbols can be strongly influenced by cultural factors (Clark, 1991; Tijust, 2007).

**Visual search**

There are several unconventional search engines that display search results in a visual format without the use of text – examples are Yometa, Bing Visual Search, Spezify and Search-Cube. However, in each of these cases, though the results are shown with little or no text, the search is initiated using keywords entered by the user. Icon-based selection systems such as the Apple OSX home screen are now ubiquitous, but we are not aware of any work that attempts to generalize these interfaces into a generic icon-based search interface suitable for use across different domains.

**IGWANA – TEXT-FREE HIERARCHICAL SEARCH**

In this section, we outline our proposed design for ‘igwana’, a text-free user interface for performing a hierarchical content search. The ‘igwana’ interface combines the use of audio and visual cues, a method proven to aid in vocabulary skills and information processing for the low-literacy population (Clark, 1991; Sarma, 2001). As an illustration, we include details of a sample ‘igwana’ implementation for accessing audio-visual information on health and wellness.

**User scenarios**

Users of igwana are typically poorly educated and lack experience using technology. They are unfamiliar with the conventions of existing user interfaces and may have a mistrust of technology. The igwana application enables these users to search and find general and practical content using touchscreen devices. Users will not use igwana to do a sophisticated or targeted search or access text-based information that is complex or technical.

**Architecture**

The igwana application consists of a) a symbol tree, content set and mapping between the two and b) a user interface.

**Data Model**

**Symbol tree**

The symbol tree is a directed, weighted graph in which a node is a pictogram with a particular meaning. Apart from the semantic content of the symbol, the meaning is also represented by attached text tags. Each edge of the graph is a path between two nodes (‘parent’ and ‘child’). Every edge in the graph has a positive weight value. The weight is a representation of the strength of association between the parent symbol and the child symbol. This will in general be a superposition of a) semantic association and b) user preference for that path.

Figure 1 shows a fragment of a sample symbol tree.

**Content set**

The content set contains items that are tagged with text. In many cases, a set of text tags is defined by metadata such as an item description or title. In the sample application considered here, the content set may be a collection of videos and photos that have been published on the Web through services such as YouTube and health networks. Content items from different sources may be aggregated into one content set, as long as each item has some text metadata associated with it.

**Mapping symbols to content**

The symbol tree is mapped to the content set using a mapping of text tags from symbol nodes to content items. The mapping may be a keyword search or a one-to-one matching of tags.

![Figure 1. Symbol tree fragment, content set fragment and mapping using keyword search](image)

**User interface**

The user interface is based on the basic standards of contextual and user-driven design (Cooper, 2003; Lee, 2005) and uses universally recognisable pictograms in a distinctive colour palette. The layout is a hierarchical menu system which is structured like a conventional navigation menu. Each menu item is a symbol from the symbol tree. Each level of the menu shows symbols from one level of the symbol tree (figures 4a – 4d) as a scrollable vertical strip.

Users are prompted to make selections from one level of the menu at a time by choosing one symbol. The next level of the menu is generated by taking all the ‘children’ of the selected symbol from the symbol tree and arranging according to the weight of the parent-child edges. Children that are connected via a higher weighting appear preferentially at the top of the next menu level.

At the bottom level of the menu, the mapping to the content set is traversed to find one or more pieces of
content as the search results, which are then displayed in an inline content viewing area.

First time user experience will be augmented by visual and audio cues.

**Modifying the symbol tree**
Unlike traditional menu systems which traverse static trees, igwana modifies the symbol tree over time using an automated alteration engine (figure 2). This engine uses user inputs to strengthen and weaken the weights of graph edges. More commonly traversed edges are strengthened by increasing their weights according to heuristic algorithms. New edges may be created between children and ancestors if the sum of weights in a path are above a threshold. Existing edges may be removed if the weight falls below a threshold.

*Figure 2. System diagram*

Over time, the effect of this modification is that more popular symbols become accessible by a) moving up within menu levels and b) jumping up to ‘higher’ menu levels. We also propose storing a local (cookie-based) copy of the symbol tree that is modified within the context of a single user. A superposition of the global and local symbol trees can then be used to render the menu. This allows the igwana interface to adapt to a particular user’s browsing preferences over time.

**Sample user interface**
Symbols are presented to the user as action buttons. Touching a symbol immediately generates the next menu level by traversing to child symbols and displaying them in order of connection weighting. Figures 4a – 4d show an interface for a sample ‘igwana’ health and wellness application.

**Home screen**
The home screen (Figure 4a) shows top-level symbols. These are typically semantically broad and should segment the symbol tree in an intuitive way.

**Second level submenu**
As soon as a selection is made from the home screen, the igwana interface switches to its default layout mode (figure 4b). Each level of the symbol menu is displayed as a vertical strip, and the strips are composed left to right. The initial home screen symbols are shown in the first level menu at far left, with the current selection highlighted. Second level symbols are shown in the strip adjacent to the right, ordered top to bottom. Each strip is infinitely scrollable using mouse or touch gestures, so that the user can access more symbols than what appears on the screen.

*Figure 4b. Second level submenu*

**Third level submenu**
The third level submenu is displayed by compositing another menu strip to the right. Note that each previously selected menu strip slides up and down so that all the selected symbols are highlighted and aligned to the same horizontal line. This creates a highlighted horizontal strip that shows all selected symbols, in order from most general at left to most specific at right.

*Figure 4c. Third level submenu*

**Higher level submenus**
Higher level submenus can be added by composing successive vertical menu strips to the interface. In our experience we have found that three levels of menu with five to ten symbols per level provides a ‘sweet spot’ for many applications – enough semantic content for an effective content retrieval while keeping the user experience short and simple. Studies have shown that three to four levels with five to twelve items per level is optimal for most users (Norman, 1991).

**Content view**
The content view (figure 4d) appears after the selection has been made in the third-level submenu. It can display one or more content items to the right of the menu as a main item and a grid of thumbnails. An additional menu
strip may also show thumbnails. Touching a thumbnail loads it as the main item and plays the video.

Figure 4d. Content view

Share content and network path
An interesting igwana add-on is the ability to share a piece of content along with the path through symbol tree that was traversed to reach it. This is represented by taking the horizontal strip of previous selections and animating it “curling up” into a small “tail” which is then packed into the share link (figure 4e). The user at the target of the share will then not only see the content but also the symbols in the path by “unwrapping” the tail.

Figure 4e. Creating a symbol tail to share

CHALLENGES AND CONCLUSION
Designing pictograms for the igwana interface is a labour-intensive and critical part of the development process. An iterative design approach based on understanding users’ subjective experience is very important. Also, a short and simple user experience is key to user adoption. The igwana interface will also be augmented by audio cues in the users’ native language.

The symbol tree will be designed to have a certain initial edge configuration with weights - this will be “seeded” by representing a real ontology. We expect to design a new symbol tree from scratch for each specific domain or application using igwana.

Both the technology and the symbol trees must be validated. This will be done by field work with a representative user base, although resource constraints may make this a limited or iterative process. In particular, finding a symbol tree that is universal and unambiguous enough for a given application is a significant challenge. The more homogenous and well-understood the user base, the easier it is to design effective pictograms.

The self-modifying nature of the igwana system is also yet to be validated. In simple terms, the property that we seek for the symbol tree is that the more people use it, the more accurate it is in predicting future search paths for individual users. There may be a lengthy tuning process required to achieve that outcome.

Notwithstanding these challenges, we see great opportunity in the development of igwana. There is a vast global population that is being physically wired to the Web but is ill-served by the current state of search interfaces. We hope that the successful deployment of intuitive, language- and culture-independent interfaces such as igwana will spread the benefits of our connected world to the many millions of people who are currently excluded from it.

REFERENCES
Chatterjee, B. 2004, p. 22


Medhi, I., Patnaik, S., Brunskill, E., Gautama, S.N., Thies, W., Toyama, K. “Designing Mobile Interfaces for Novice and Low Literacy Users”, ACM Transactions on Computer-Human Interaction (TOCHI), Volume 18 Issue 1, April 2011, Article No. 2


Pictogram website http://www.pictogram.se/english/

Sarma, N., 2001, Project on Radio Education for Adult Learners: Comprehensive Review


UNESCO, 2008
