PAN LOCALIZATION
NIDA EDOC Report

KHMER UNICODE WEB SITE

KHMER UNICODE CHAT

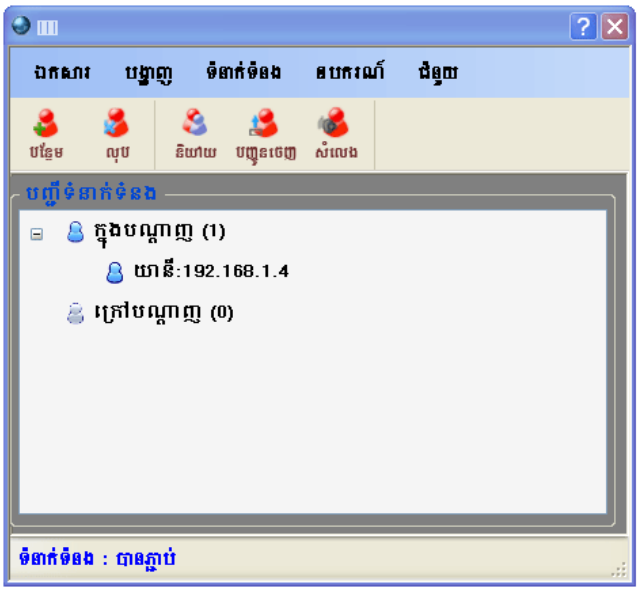


Table of Content

PAN LOCALIZATION NIDA EDOC Report 1
KHMER UNICODE WEB SITE 1
KHMER UNICODE CHAT 1
NIDA EDOC Report Monday, October 27, 2008 3
WEBSITE KHMER 3
- Screen Shots 3
- Search Page 3
- Home Page 4
- Health Page 4
- Content 5
- Issue 6
  - Content 6
  - Joomla 7
  - Cannot search many words (19/08/08) 7
  - Search many words 7
  - Textbox search and limit character to search 7
  - Filter out search terms that are too small 7
  - Limit searchword to 20 characters 7
  - Logo 7
  - SWF file 7
  - size/padding/size 7
  - To display Khmer Unicode on website without install 8
- MySQL Server 5.0 8
  - Configuration 8
  - Table Option 8
  - Column Detail 9
- Drivers 9
  - MySQL connector 5.0.9 9
  - ODBC connector 5.1 9

Khmer Chat 10
- Introduction 10
- Problem 10
- Objective 10
- Summary of requirements 10
- Use-Case Diagram 12

NIDA EDOC Report
Monday, October 27, 2008

WEBSITE KHMER

Khmer website contains of useful information for Khmer people. Plus, in this website, it has search function using Khmer Unicode.

Screen Shots

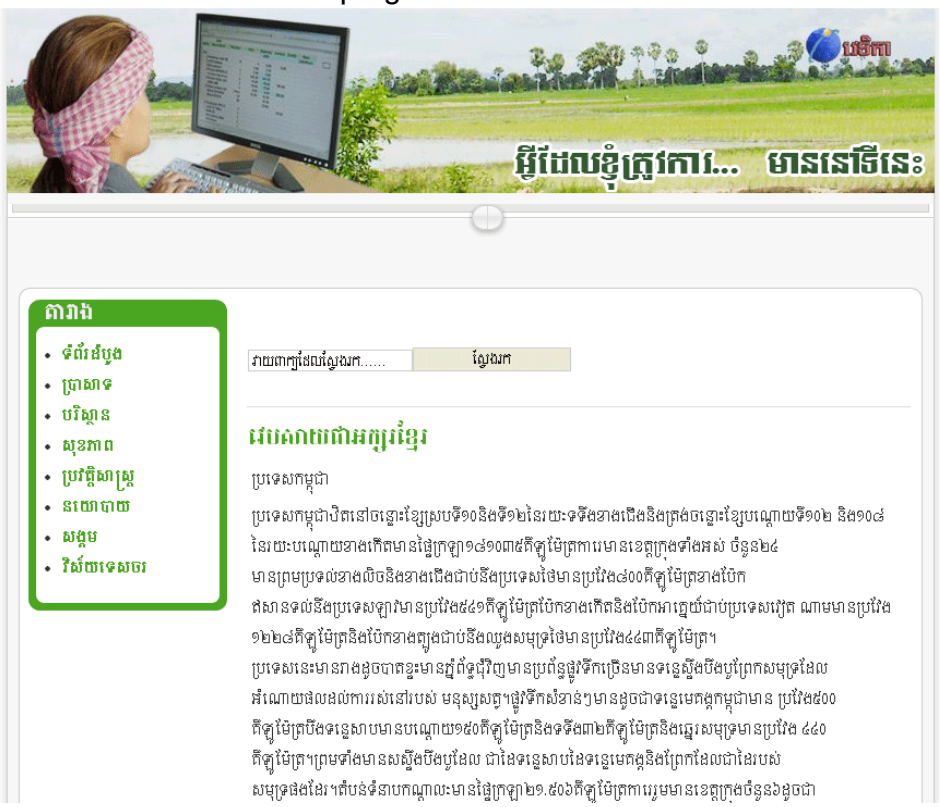
On the pages below, there are 3 sample screen shots of this Khmer website.

Search Page

This is the screen shot after we use the search function for the word “school” in Khmer language and it display results like this.

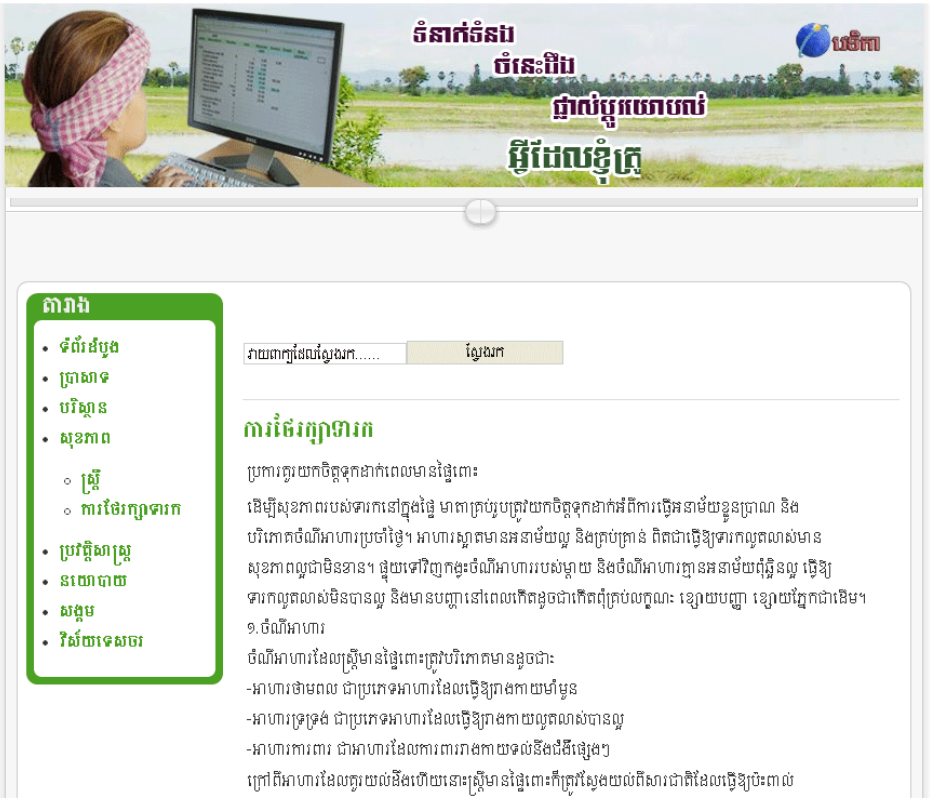
Home Page

This is the screen shot of home page.



Health Page

This is a screen shot of one of health pages.



Content

Below are the contents in Khmer language that are available in the website:

-Home Page: “ទំព័រដំបូង”

-Temple: “ប្រាសាទ”

-Environment: “បរិស្ថាន”

-Health: “សុខភាព”

- **Lady Health Care**
- **Baby Health Care**

-History: “ប្រវត្តិសាស្ត្រ”

-Politic: “នយោបាយ”

- **Cambodia Structure**
- **National Constitution**
- **Province and Town Structure**
    - **Banteay Mean Chey Province**
    - **Batthom Bong Province**
    - **Kompong Cham Province**
    - **Kompong Thom Province**
    - **Kom Pot Province**
    - **Kompong Chhnang Province**
    - **Kompong Speu Province**
    - **Kandal Province**
    - **Kep City**
    - **Koh Kong Province**
    - **Kratie Province**
    - **Mondokiri Province**
    - **Ordormean Chey Province**
    - **Pai Lin City**
    - **Phnom Penh City**
    - **Porsat Province**
    - **Preh Vihea Province**
    - **Prey Veng Province**

-Social: “សង្គម”

- **Minority People in Cambodia**

-Tourist: “ទេសចរ”

- **24 Provinces and Town**
    - **Bat Thom Bong Province**

        - **Kompong Cham Province**
        - **Kompong Chhnang Province**
        - **Kom Pot Province**
        - **Kandal Province**
        - **Kep City**
        - **Koh Kong Province**
        - **Phnom Penh City**
        - **Porsat Province**
        - **Seam Reap Province**
        - **Steng Treng Province**
        - **Takeo Province**
- **Weather in the main tourist site**
- **Other Tourist sites**
        - **Animal Guard Areas**
        - **National Park**
        - **Bird Protected Areas**

Issue

Content

Cambodia so far has lack of content. Little content is available in our content. People in NIDA do not have motivation to make content. They write very slowly and some do not have idea how to write. Starting is very difficult. When some one does not know about which topics to write, he does not produce good content.

My strategy is now changing. Content developer must be a person who really wants their work to be published. He can be journalist or in the area of tourism. Really motivated people can do this work.

Khmer Unicode has some difficulty to start. People use the traditional font called Limon that uses ANSI code with font in Khmer. By encouraging people to write directly with Khmer Unicode, every one has a reticent behavior at the beginning. As they progress and get familiar with the smart typing of Unicode, they become very keen to work with it.

There are several technical problems with Unicode on MS Windows. Office 2003 gets stuck very often. Some tools like zero space or spelling check do not work on all computers. When using MS Office 2007, the Khmer Unicode is more reliable. My team works with MS Office 2007.

We use actually Joomla to build our website. Joomla is the Content Management System. We use “MySQL” database to store all content of our website that is dynamic.

Joomla

Cannot search many words (19/08/08)

Solution: search in the Joomla's forum
Change: components/com_search/models/search.php

Search many words

```
$rows = array_merge( (array)$rows, (array)$results[$i] );
```

Textbox search and limit character to search

edit in : components/com_search/helps/search.php

Filter out search terms that are too small

```
foreach( $aterms AS $aterm ) {
        if (JString::strlen( $aterm ) < 2) {
                $search_ignore[] = $aterm;
        }
}
```

Limit searchword to 20 characters

```
if ( JString::strlen( $searchword ) > 50 ) {
        $searchword         = JString::substr( $searchword, 0, 19 );
        $restriction   = true;
}
```

Logo

We can modify in css.

SWF file

We can modify in css.

size/padding/size

We can modify in css.

To display Khmer Unicode on website without install

a. we should use embed font (Weft or CoffeeWebsiteFont3.4)
   Note: can use with IE only.

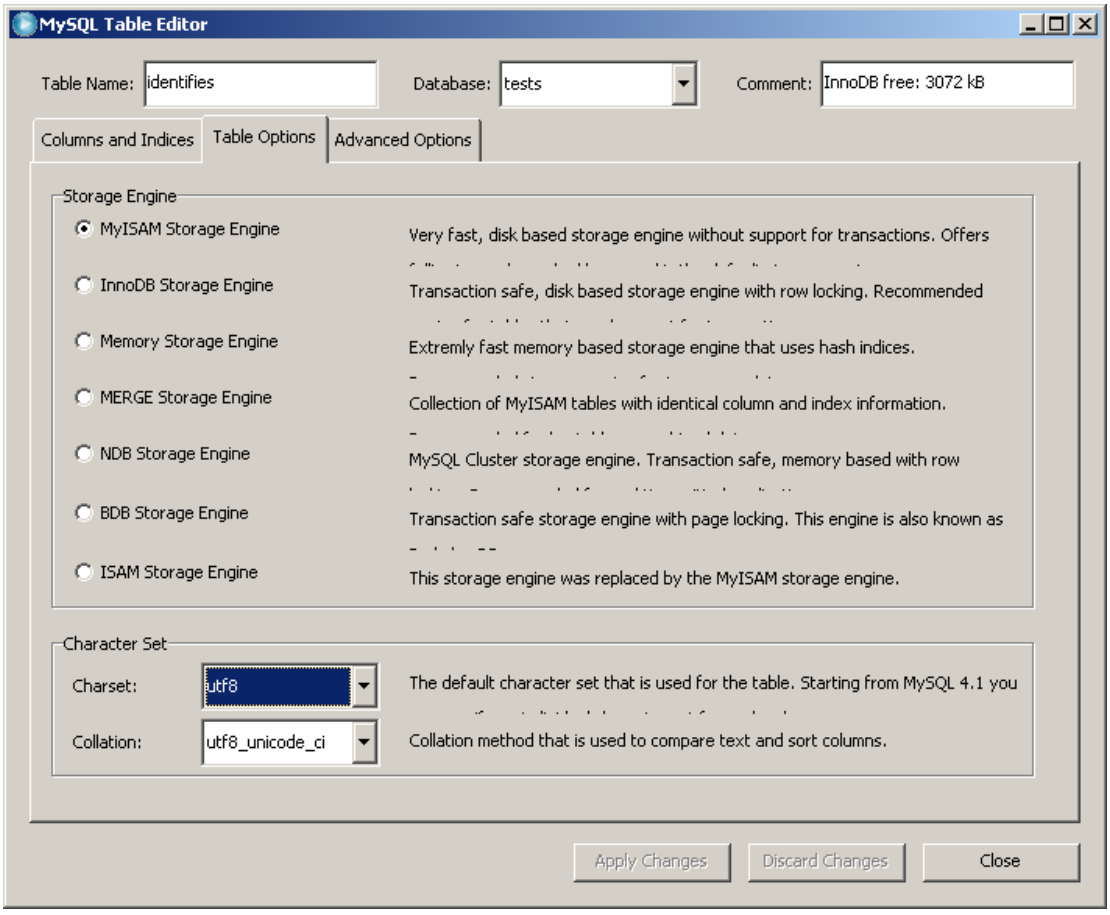
MySQL Server 5.0

Configuration

Setup the Database to support Unicode. With good configuration, data is stored in the database in Unicode format that can be searched or sorted.

Table Option

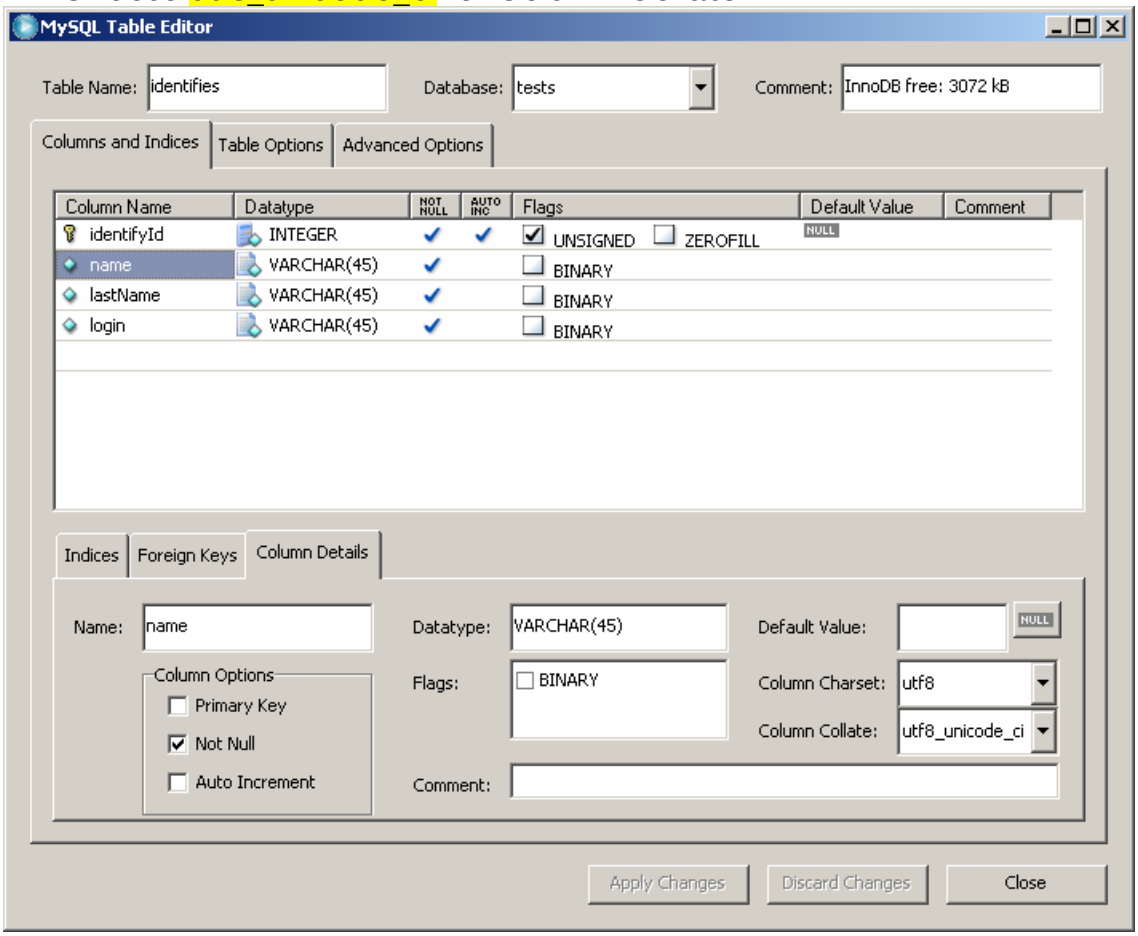
- Select on Tab Table Options
- Choose utf8 for Charset
- Choose utf8_unicode_ci for Collation (to enable the sorting)



Column Detail

Column and Indices

- Choose utf8 for Column Charset
- Choose utf8_unicode_ci for Column Collate



Drivers

We can use Khmer Unicode in MySQL by using MySQL driver or MySQL ODBC driver.

MySQL connector 5.0.9

```
string con = "Database=Tests;Data Source=localhost;User Id=root;Password=myservice;CharSet='utf8';";
```

ODBC connector 5.1

The Unicode does not support the old driver ODBC 3.51. So to enable Unicode in our application we have to download the ODBC connector 5.1

```
string con = "Driver={MySQL ODBC 5.1 Driver};Server=Sochan;Database=Tests; UID=root; Password=myservice;
Option=3;";
```

Khmer Chat

Introduction

People always have different problems and needs, especially in Cambodia. The lack of information in Khmer language is mainly considered. The NIDA EDOC chooses the livelihood subject for the project content that can be helpful for most of the Cambodian. The company wants to provide good information that is readable and understandable containing the simple words and sentences to people.

In Cambodia, the chat program in Khmer Unicode does not exist yet. The project purpose wants all Cambodian level, especially high school students, to communicate with each other in Khmer language. They can chat by using their mother tong that helps them to choose right words and catch their understanding. Oversee Cambodia can also enjoy the chat software to be close with family living in the country.

Problem

Today, we have a lot of chat applications that allow people to communicate with each other on the Internet but it is in foreign language such as English, French...etc. So the Cambodian people who don't know foreign language can't chat with each other. In order to allow them to do the communication, we have develop an application called
'Chat Khmer Unicode'.

Objective

Develop a desktop application with Khmer Unicode Interface by using the Java programming language.

Summary of requirements

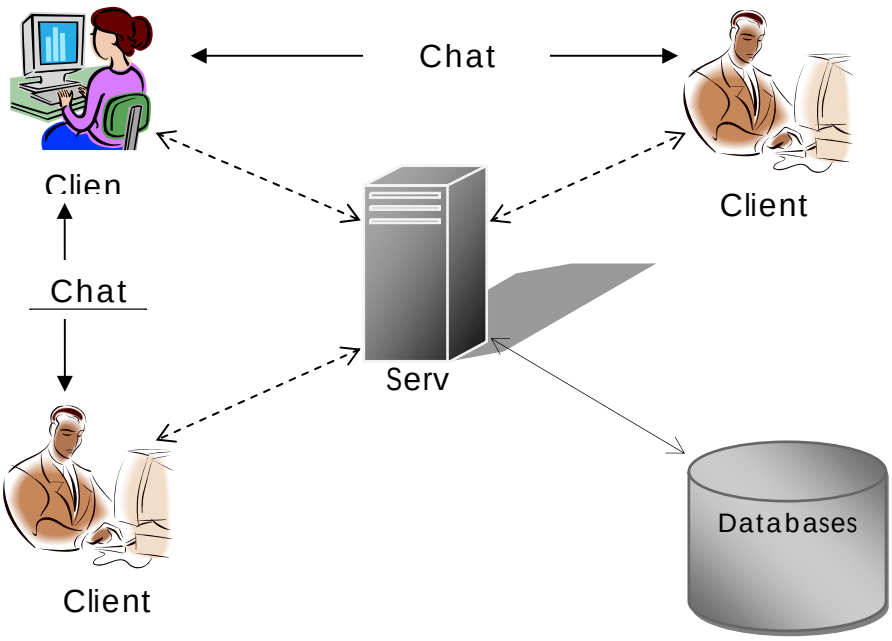
Our application is divided into two parts: Client and Server.

The client has the operations as the following:

- Chat or chat in group: allows the clients to communicate with each other.
- Transfer file: allows the clients to transfer file.
- Account modification: allows the clients to modify their account.
- Contact modification: allows the clients to modify their address book.

The server has the operations as the following:

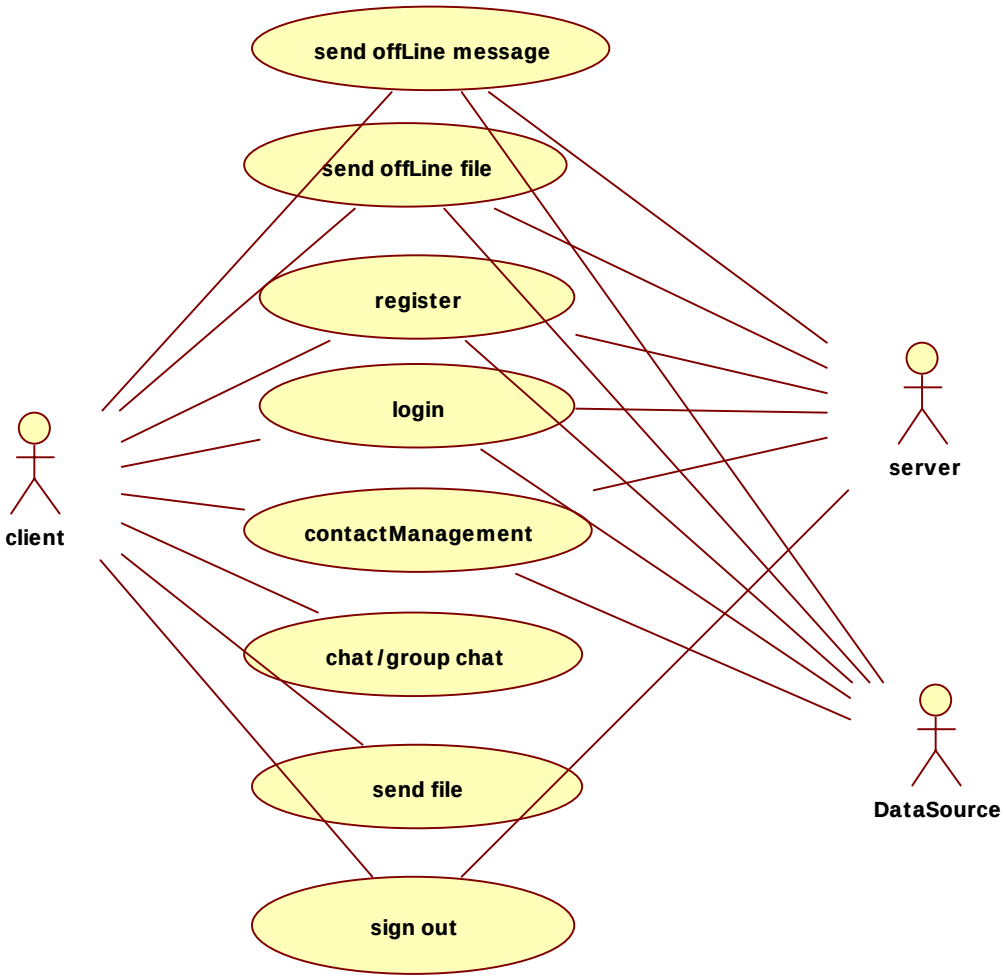
- Waiting for the requests of the clients
- Execute the requests applied by the clients (Ex: Login, Register…)



Explication: At first, the server has to run permanently before the client asks for a request. The client has to have an account in the server database before having the right to do anything. The client has to identify himself to the server. After success identification, the client can retrieve his contact list (his account) automatically and then the connection between client and server is closed, he can also chat and transfer file to another client. Assume two clients want to communicate with each other. The way that they can is the following:

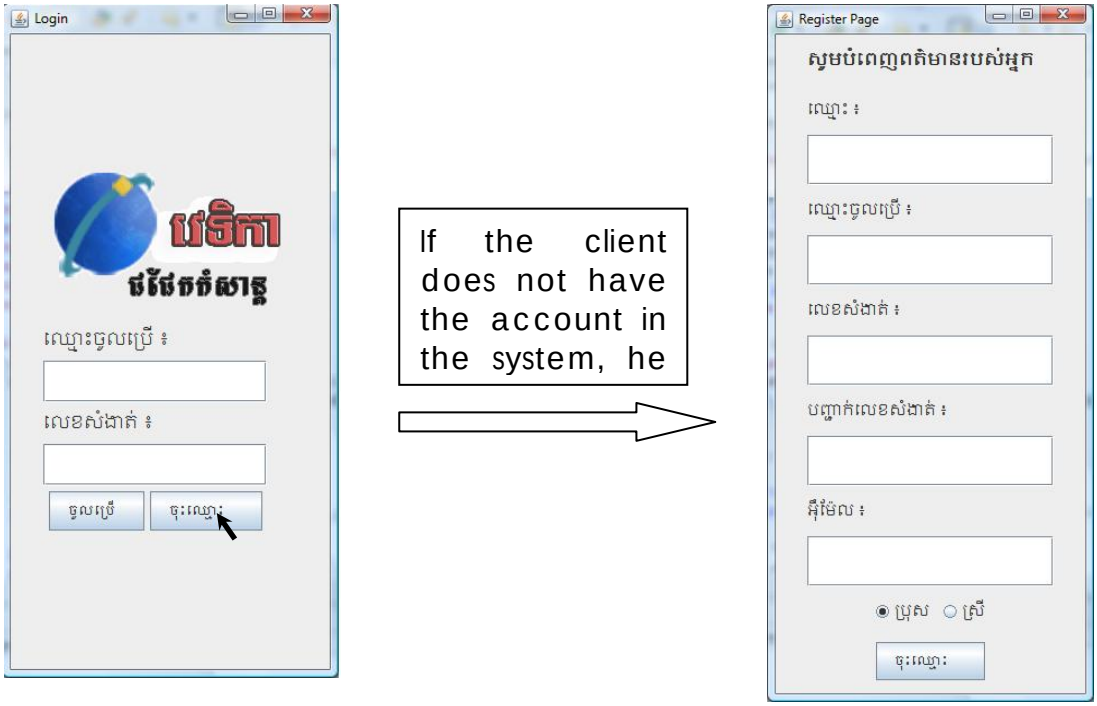
- They have to identify themselves to the server.
- After that, they can know each other.
- Then they can communicate directly without server intervention by sending the message directly (from one client to another).

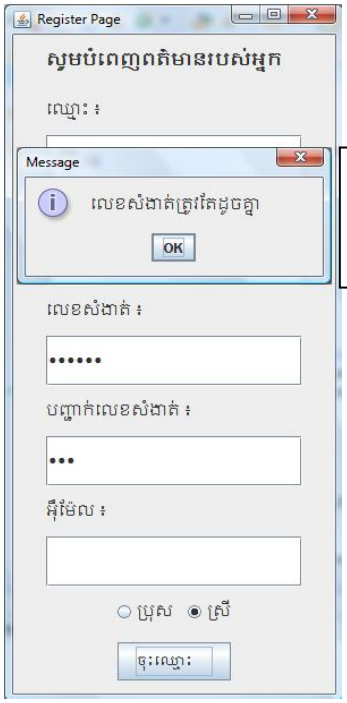
Use-Case Diagram



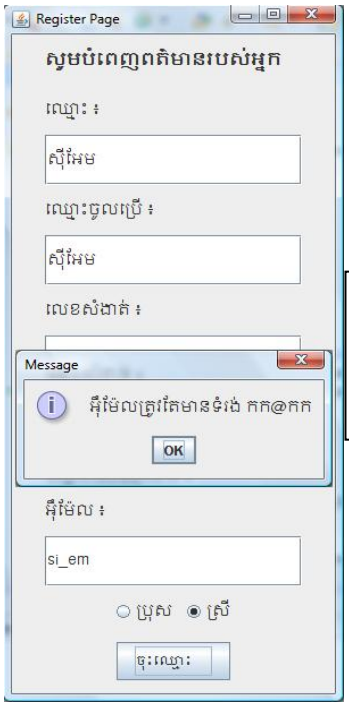
Graphical User Interface (GUI)

The first page of the Khmer unicode chat is “Login page”



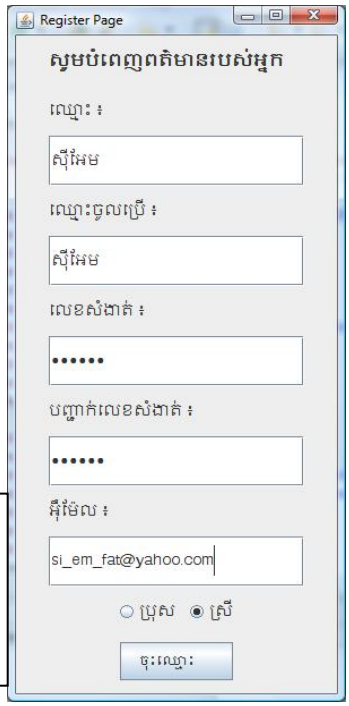


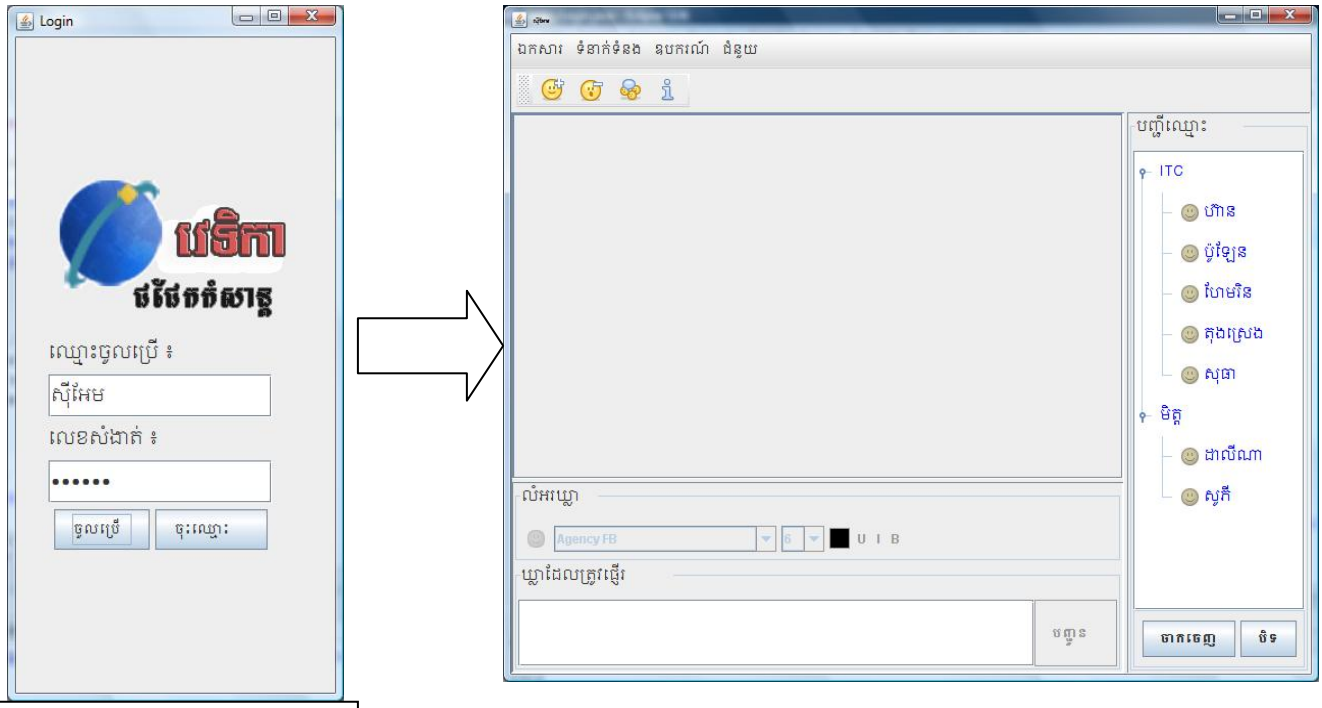
When register the information the password and the confirmation of the password must be the same. If not the same, there must be the



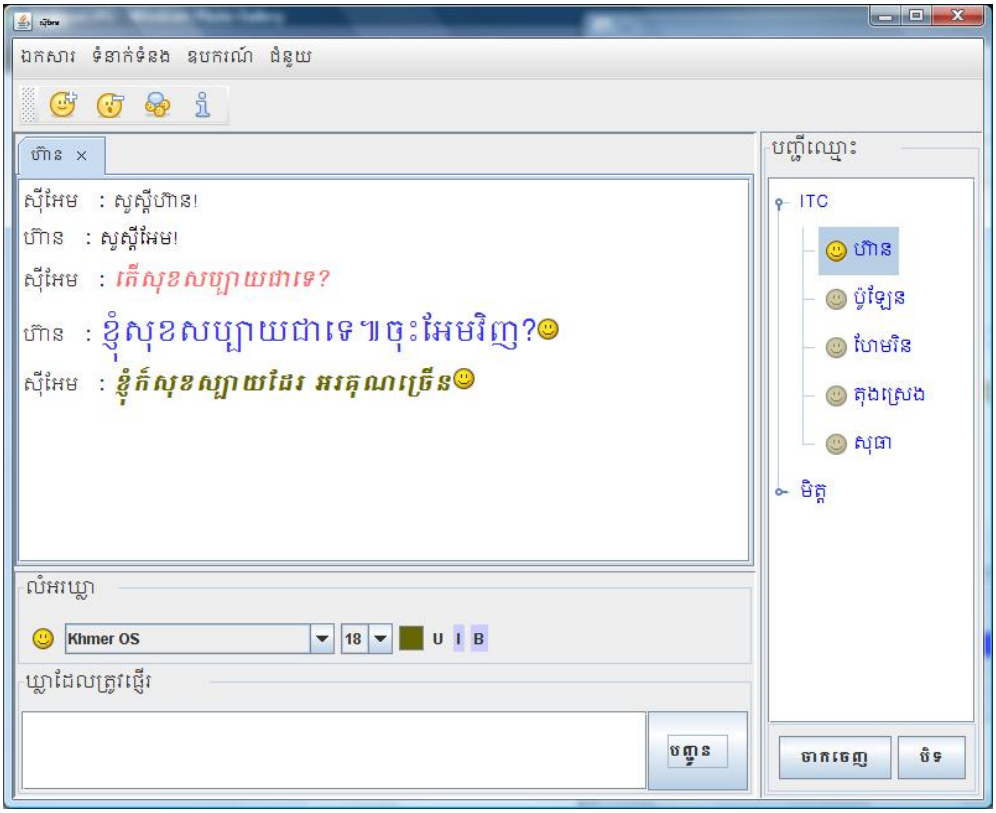
The email must have the form such as aaa@aaa.com

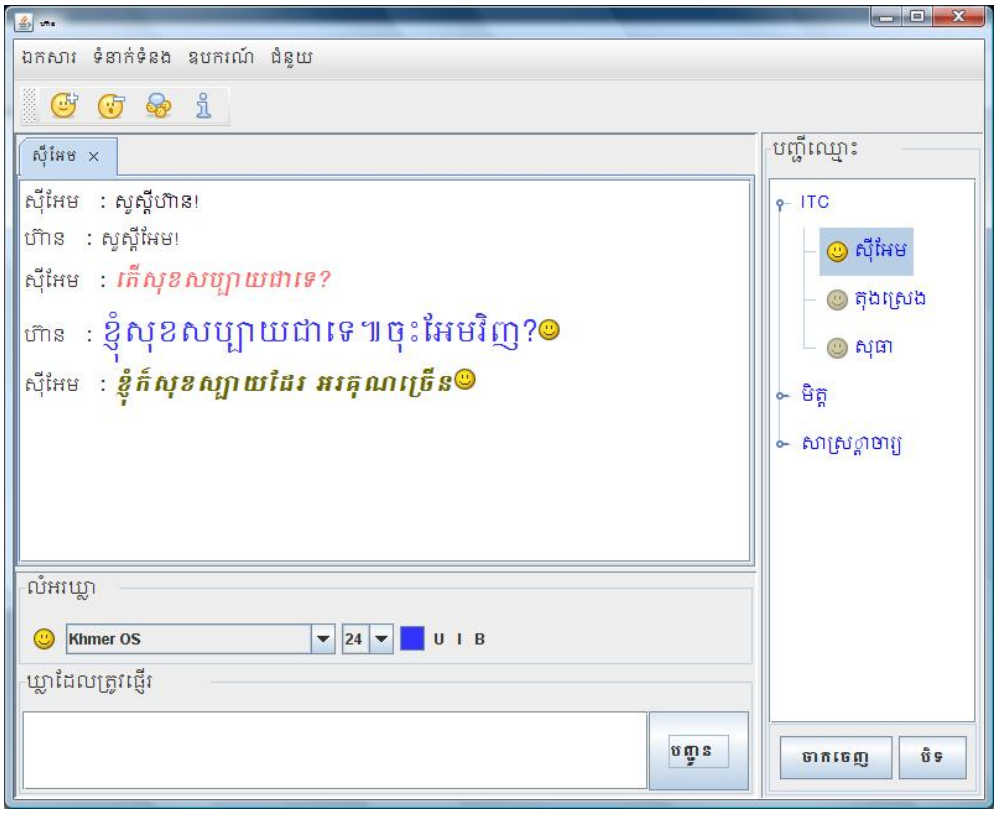
This is the final page of the register. The information must be completed all part





When login, the client have to fill userName and password




ឯកសារ ទំនាក់ទំនង ឧបករណ៍ ជំនួយ
ស៊ីអែម ×
ស៊ីអែម : សួស្តីហ៊ាន!
ហ៊ាន : សួស្តីអែម!
ស៊ីអែម : តើសុខសប្បាយជាទេ?
ហ៊ាន : ខ្ញុំសុខសប្បាយជាទេ។ចុះអែមវិញ?
ស៊ីអែម : ខ្ញុំក៏សុខស្បាយដែរ អរគុណច្រើន
បញ្ជីឈ្មោះ
ITC
ស៊ីអែម
តុងស្រេង
សុធា
មិត្ត
សាស្ត្រាចារ្យ
លំអរឃ្លា
Khmer OS
24
U I B
ឃ្លាដែលត្រូវផ្ញើរ
បញ្ជូន
ចាកចេញ
ចំទ