ADESSOWIKI: Collaborative scientific programming environment

Tutorial

Leticia Rittner – UNICAMP
André V. Saúde – UFLA
Alexandre G. Silva – UDESC
Rubens C. Machado – CTI
Mariana P. Bento – UNICAMP
Roberto A. Lotufo – UNICAMP
ADESSOWIKI: Collaborative scientific programming environment

Apresentadores

Leticia Rittner – UNICAMP
e
Mariana Pinheiro– UNICAMP
Agenda

- Part I – Introduction and Motivation
- Hands-on I – How to begin

Coffe Break

- Part II – Experiences
- Hands-on II – Writing a paper
- Part III – What comes next?
ADESSOWIKI: Introduction and Motivation

Tutorial – Part 1
Adessowiki – Definition

- Is a collaborative environment for scientific writing and programming

- Meaning of the word:
  - Adesso – Italian word for now
  - Wiki - Hawaiian word for fast
Problems to solve

- Synchronism among documentation, data, programming code and compilation tools.
  - Book on scientific programming has an accompanying CD *(usually out of sync)*

- **Literate programming** (Donald Knuth, 1984)
Problems to solve

- To reproduce a scientific paper you need:
  - Same data
  - Same code
  - Same programming environment
- The same applies to benchmark algorithms
  - Fingerprint, face recognition competition, etc.
Problems to solve

- After few days a scientific paper is written you don’t know how to regenerate the figures and tables
  - The parameters used are lost
  - Programming code is outdated or not available
  - Demos require a computer configured with many software packages

- Reproducible research
  - http://www.reproducibleresearch.net
Problems to solve

- Students deliver a programming code as a lab exercise
  - You may not have the same software environment the student has
  - It takes a lot of time to download, compile, execute and see the results

- Automatic programming assessment
Adessowiki – A solution

- Concept

- It is a wiki, where the programming code can be embedded in the text. It is executed in the server when the page is rendered updating images, figures and output text.
Wiki concept

- Wiki is a generic name for collaborative hypertext using a markup language
- Wikipedia is a good example of a wiki
- Paradigm change:
  - All users can edit wiki pages
  - Record of changes
  - Version control tools
Adessowiki – more than a Wiki

- Wiki: “the simplest online database that could possibly work”
- How to extend these concepts to computer science (programming) courses and research?
- Adessowiki
  - Wiki
  - Ability to execute code on the server as a part of the render of a wiki document
  - Execution results appear as text, images, graphics, equations and tables
  - Centralization of hardware and software configuration
Adessowiki: 3 main features

1. SaaS – Software as a Service;
2. Collaborative;
3. Executable user code.

- Documents or Applications.
Possible scenarios

- **Teaching**
  - Software support
  - Library of collaborative code and explanations

- **Research**
  - Scientific report, paper, thesis

- **Software development**
  - Client and developers in a collaborative work
Teaching

- Easy to verify if the code works
- Possible to compare different solutions to the same problem
- Prevents plagiarism (wiki log capabilities)
- Incentives collaboration (intra and inter classes)
- Solutions are easily recycled as the data, the code and its programming environment are available for all in a centralized manner
Scientific experiments can be readily reproduced
Documents contain source code, data, equations and descriptions together and synchronized
Papers, books, reports, thesis
Encyclopedia of algorithms with technical explanations
E-book of the future – text, source code and data for reader consumer and collaboration
Comparison and assessment of scientific contribution – future of paper reviewing process
Competitions for open innovation
Collaborative executable documents

- Collaborative Structured Text (like wiki)
- Allows the insertion of small parts of code (Python, C++) in the text. The graphics, plots, images are displayed as figures in the document
- Documents can be converted to:
  - PDF
  - Slide show
  - HTML
  - RTF (Word for Windows)
Adessowiki Main Page

Adessowiki - Collaborative Scientific Writing and Programming

Welcome to Adessowiki. This is a collaborative platform for scientific programming and document writing, initially dedicated to Image Processing and support for Python/C/C++ programming. We invite you to look at an overview presentation on the concepts of Adessowiki. You can also have a look at a paper presented at the WikiSym 2009, and a paper presented at WEPC/SIBGRAPI 2009.

This project is a collaboration between the University of Campinas (UNICAMP) and the Renato Archer Information Technology Center (CTI). The history of this project is described here.

Adessowiki is built around many freely available software such as Python, Django, LaTeX and many others. To know more about how to use this system, check the documents listed below (temporarily in Portuguese).

Warning: As parts of this contents are still in Portuguese, please use Google translator to read this material in English.

Introductory Documents

- Preparing Documents with Adessowiki
- Executing Python Code
- Toolbox Creation: Python Code
- Toolbox Creation: C/C++ Code
- Numpy Arrays

recently modified pages

- Sign
  [irittner] 2011-06-02
- MainPage
  [rubens] 2011-05-30
  [ia636:iacorr] 2011-05-29
  [luijsfsw] 2011-05-24
  [ia870:ia870] 2011-05-21
  [ia636:iavfilter] 2011-05-12
  [ia636:ianormalize] 2011-05-10
  [ia636:ianormalize] 2011-05-10
Adessowiki structure

- Adessowiki is a distributed system, based on the technologies and concepts of Web 2.0
- Kind of collaborative literate programming powered by the modern world-wide web
- Two separated web servers:
  - Wiki server, that serve information from a database and never access the local file system
  - Media server, which serve files from the server local file system
Adessowiki structure (cont.)
Wiki Markup language – **reST**

Execution of code fragments module, **XSandbox** (in a controlled environment)

**reStructuredText** extensions allows:
- Python and C/C++ code
- Images and tables from scripts
- Plots and diagrams
- Equations
- Wiki links
User creates a reST document using a web interface.

The document is parsed to create an XML tree. The Execution Sandbox executes the Python code embedded in the page and resources created by this code execution are incorporated in the XML document.

The XML document is transformed to create an HTML page that will be displayed by the Adessowiki web interface.

The same XML document can be used for the generation of other kinds of representations like, e.g. PDF documents.
Adessowiki (cont.)
Adessowiki page edition

main.Sign

Revision: 223274 (103)

1
2 Segmentation Demo
3
4 Date: 24/01/2008
5
6 Contour Detection
7
8 The original image is read and negated. The
9 contour is obtained through a morphological
10 gradient and a threshold.
11
12 .. code:: python
13   :show_code: yes
14   :show_output: yes
15   :show_images: yes
16
17   thresh = 20
18   a = mmreadgray('sign.jpg'
19   b = mmgradm(mmopen(a, mmsebox(1)))
20   mmshow(a, b>thresh, title='Contour')
21
22   print 'Image shape:', a.shape

Save and View Save Document Reload Document Verify Document
Adessowiki page edition

Segmentation Demo

Contour Detection

The original image is read and negated. The contour is obtained through a morphological gradient and a threshold.

```
1 thresh = 20
2 a = mmreadgray('sign.jpg')
3 b = mmgrad(mmopen(a, mmsebox(1)))
4 mmshow(a, b>thresh, title='Contour')
5
6 print 'Image shape:', a.shape

Image shape: (245, 326)
```
Examples

- Contour detection
- Renavam
- SUDOKU
Adessowiki

History
Version

- Redesign from previous systems such as
  - Adesso (FAPESP 1999)
  - Adessoweb (FAPESP 2002)
- Major releases
  - May 2008
  - Apr 2010 (GPU – CUDA support)
In 2011

- About 7600 pages, 234000 revisions
- 714 users
- Sandbox CUDA in operation and Sandbox Robotics in developing
- Sandboxes that interact with users
- > 20 Courses with strong use of resources
  - FEEC/UNICAMP – Computer Vision, Mathematical Morphology, Medical Imaging Processing, GPU-based Imaging Processing
  - IME/USP – Vision and Image Processing
  - UFLA – Image Processing
  - UDESC – Image Processing
Related publications

- Adessowiki -- On-line Collaborative Scientific Programming Platform.


- Adessowiki – Collaborative platform for writing executable papers.
ADESSOWIKI: How to begin

Tutorial – Hands-on I
Hands-on I

- Click to sign in, on the superior right corner
Hands-on I

- Fill out the form
- Copy the invitation key: r/Vr9DTNnaagS0wlskzxhQ5+jgn+/zxh
ADESSOWIKI: Experiences
Tutorial – Part 3
Use cases of Adessowiki

- Collaborative research
- Scientific writing
- Code repository
- Teaching
Adessowiki

Collaborative research
Collaborative research environment

- Research groups
  - Announce opportunities for students
  - Publish research results
  - Example: Lab. Process. Inteligente de Imagens (L2I)

- Students
  - Bibliographic Revision
  - Code + Results
  - Description of meeting minutes
  - Scientific report, papers, thesis
  - Example: André Körbes
Collaborative research - strengths

- Scientific experiments can be readily reproduced
- Documents contain source code, data, equations, descriptions and compilations tools together and synchronized
- Encyclopedia of algorithms with technical explanations
- Competition for open innovation
Scientific writing
Scientific writing

- Collaborative tool for writing papers, thesis and books
- Markup text Post-processing generating HTML, PDF or LaTeX files.
- Example:
  - Image Processing book
  - Paper for IWSSIP2010
Scientific writing - strengths

- Remote collaboration with version control
- Images are automatically generated
- Comparison and assessment of scientific contribution – future of paper review process
- E-book of the future – text, source code and data for reader consumer and collaboration
44

Adessowiki

Code repository
Code repository

- Enables the creation of reusable software modules for Python and C/C++
  - Example: IA636 Toolbox

- Controlled environment for code execution. The user code can not:
  - Write to the file system
  - Open sockets directly
  - Launch processes or threads

- Data sets are also stored
  - Examples: Image dataset
Code repository - strengths

- Most commonly used algorithms available
- Possibility of comparing their algorithm with other already validated and established
  - Same hardware and same data sets
  - Run time
  - Accuracy of results
Teaching
Teaching environment

- Course page
  - General course information
  - Demonstrations as lecture notes
  - Proposed exercises
  - Students solutions for proposed exercises

- Example:
  - IA369O – Medical image processing
Teaching environment - strengths

- Easy to verify if the code works
- Possible to compare different solutions to the same problem
- Prevents plagiarism (wiki log capabilities)
- Incentives collaboration (intra and inter classes)
- Solutions are easily recycled as the data, the code and its programming environment are available for all in a centralized manner
ADESSOWIKI:
Writing a paper

Tutorial – Hands-on II
ADESSOWIKI: What comes next?

Tutorial – Conclusion
Adessowiki – Strengths

- No software installation or configuration
- The client requires just a web browser
- Collaborative as any wiki
- Synchronism between implementation and outputs (images, graphics, tables, equations)
- Easy to compare solutions (performance)
- Centralized database of images
- Centralized server and application
Acquired experience

- Writing:
  - 1 book
  - 2 book chapters
  - 1 dissertation
  - > 20 papers

- Research
  - > 15 students (Undergrad, Msc, PhD)

- Teaching
  - 5 Institutions (UNICAMP, USP, UFLA, UDESC, UFMG)
  - > 20 courses

- Toolboxes
  - IA636 – Computer Vision
  - IA870 – Mathematical Morphology
  - Watershed
Current experience

- New toolboxes:
  - Medical Imaging Processing
  - Machine learning

- Teaching:
  - New course in Pattern Recognition
Future possibilities

- Competitions
- Data repositories with annotations
- Algorithm repositories
- Collaborative research
- Collaborative writing
- New applications
- GPU application
Contacts

- Leticia Rittner
  - lrittner@gmail.com

- Mariana Pinheiro Bento
  - marianapbento@gmail.com

- André Vital Saúde
  - saude@ufla.br

- Alexandre Gonçalves Silva
  - alexandre@joinville.udesc.br

- Roberto Lotufo
  - lotufo@unicamp.br

- Rubens C. Machado
  - rubens.campos.machado@gmail.com

http://www.adessowiki.org
Thank you !