

# **Cloud of Reproducible Records**

## **A Unified Computational Approach to Scientific Results**

### **Reproducibility in Computational and Experimental Investigations**



**Faical Yannick P. Congo**

Supervisor:

**Prof. David Hill**

**Prof. Mamadou Traore**

**Dr. Jonathan Guyer**

Department Engineering Sciences  
Blaise Pascal University

This dissertation is submitted for the degree of  
*Doctor of Philosophy*

Engineering Doctoral School

May 2016



To my parents for all their love and support and putting through the best education possible.  
I appreciate their sacrifices and wouldn't have been able to get to this stage without them.  
To my wife Hajar for her unending support, level-headedness and love. I wouldn't have  
gotten through this doctorate if it wasn't for her.



## **Declaration**

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than xxxxx words including appendices, bibliography, footnotes, tables and equations and has fewer than yyy figures.

Faical Yannick P. Congo  
May 2016



## Acknowledgements

Mes premiers remerciements vont aux personnes sans qui je n'aurais jamais pu réaliser cette thèse, à commencer par mes deux encadrants Dr. David Hill et Dr. Mamadou Traore. Je tiens également à remercier très chaleureusement Jonathan Guyer qui m'a bien accepté d'être mon accadrant au NIST durant l'intégralité de cette thèse et sans qui cette these n'aurait pas eu quelconque valeurs que vous y trouverai. Ses conseils et sa réflexion m'ont été d'une grande aide. Il m'est nécessaire de remercier également Carelyn Campbell, mon responsable officiel au NIST, pour sa patience, sa disponibilit  et ses encouragements. Je lui dois beaucoup. Je tiens à remercier également Andrew Reid et James Warren pour leurs aides, conseils, directions et brainstorming sur les aspects de la reproducibility et de la maintenabilite des applications dans la recherche. Je me dois de remercier également, chaleureusement, les personnels administratifs qui m'ont aid  et soutenu durant ma th se : S verine Miginiac, Aur lie Bavent, B atrice Bourdieu, Socheata Sean et Nicolas Champeil. Sans eux, ma th se aurait  t  bien plus monotone, et administrativement beaucoup plus contraignante. Merci  galement   tous mes collegues du NIST qui de pres ou de loin on contribue   elargir mes horizons de penses aussi petite soit-elle : Gunay Dogan, Stephen Langer, Eric Cockain, Ronald Boisvert, Desire Banse, Martial Michel. Merci aussi   toutes les personnes que j'ai pu c toyer durant l'int gralit  de ma th se dans mes diff rentes conferences, collaborations et qui m'ont apport  une nouvelle perspective   ma comprehension : Paul Patrone, Anna Nelson, Natalie Revol, Michael Mascagni, Walid Keyrouz, ... Plus tous ceux que j'oublie. Un merci particulier   Daniel Wheeler, qui a cru en moi d s le d but.





## Abstract

Le travail accompli durant cette these a consiste en l'investigation d'une solution informatique permettant de garantir la reproductibilite de resultats scientifiques d'origine informatique ou experimentale. Lorsque nous utiliserons le mot 'reproductibilite' sans precision supplementaire nous entendons par la l'obtention de resultats dans la marge de precision acceptable du domaine scientifique concerne. Aussi nous entendons par 'garantir' ici que la solution dans le defaut de reproduire des resultats dans un environnement donne pourra proposer clairement les informations concernant l'infrastructure requise pour pouvoir realiser les resultats attendus. Au moment meme de cette redaction, en ce qui concerne la reproductibilite de resultats scientifiques de nature informatique ou experimentale contenu generalement dans une publication, plusieurs obstacles jalonnent le chemin vers leurs reconstructions. Ces obstacles sont de plusieurs nature: techniques, methodologiques et communautaires. Or la reproductibilite est la pierre angulaire de la methode scientifique parce qu'elle garantie les fondements sur lesquels de nouveaux savoirs se forment et sont valides puis accepter dans la communaute. Elle aussi le garant de la solidite des fondements sur lequel peut s'etablir avec assurance la continuite de nouvelles recherches entreprises pour l'avancement de la Science. Dans la recherche scientifique, la communaute est une ressource formidable pour collaborer, apprendre, partager, verifier, valider et contribuer a la Science. C'est le centre de toute les activites scientifiques d'un domaine precis. Par consequent, nous avons donc juge evident qu'une solution unifiante devrait commencer par federer les chercheurs en developpant un sens communautaire du partage, de la collaboration et de l'interopabilite autour de resultats reproductibles. En effet, il est important de voir que la communaute est le noyau autour duquel les resultats sont enrichis par les diverses collaborations. Et cet enrichissement beneficie les chercheurs contributeurs autant que ceux qui sont plus interesses par des procedes en amont sur ces resultats enrichis. Constat rapide de la situation et l'importance Sur la question de la reproductibilite des resultats de recherche, nous nous sommes interesses aux deux principales approches scientifiques du 21em siecle: l'experimentation pure, la modelisation informatique mais aussi leur intersection. Par experimentation nous entendons ici l'utilisation d'une machine specialisee pour produire ou analyser des donnees relative a une investigation physiquement mise en place. Aussi par modelisation informatique nous faisons references

aux investigations complement simulees sur un ou un ensemble d'ordinateurs. Quant meme que le soucis de reproduire ai ete inherent chez les experimentalistes a travers des favoris comme le bloc-note de laboratoire, les notes prises etaient trop personnalisees et generalement insuffisantes ou trop endommagees pour etre d'une utilite quelconque a d'autres chercheurs utilisant la meme machine et encore moins un chercheur utilisant une autre machine. Ces experiences donnent aujourd'hui un sens important a la necessite de representations communes de l'information recolte pendant une investigation pour qu'elle soit comprehensible entre chercheurs mais egalement entre machines. Un autre objectif capital est la nature cruciale de l'automatisation de l'execution des investigations du fait du facteur manuel dans les actions des chercheurs qui pese sur l'incertitude des resultats. De plus la complexite meme grandissante des systemes de nos jours rend le bloc note obsolete parceque le nombre de parametres a ecrire et entraine une divergence plus prononcee loin de nos espoirs d'inter-operabilite entre machines. En outre, meme si le soucis de la reproductibilite des resultats informatiques aient emerges plus tardivement dans les annees 60 a travers les efforts de Chercheurs comme Jon Claerbout, la conception meme des systemes informatiques fait que certains problemes soulevés dans le monde experimental ne se posent pas ici. Neanmoins, des obstacles d'autres natures existent. La formalisation de l'information utile pour reproduire une execution conduisant aux 'memes' resultats restent un challenge du fait que l'objectif est de pouvoir reutiliser cette information pour reproduire une execution le plus simplement et automatiquement possible. Aussi la particularite virtuelle meme de la modelisation informatique laissant entrevoir par exemple un monde sans bornes physiques concretes (a l'exception des limites materiels) ouvrent la porte a un domaine supplementaire d'erreurs propres au milieu informatique. De ce fait d'autres problemes lies meme a la conception materielle et logicielle des ordinateurs ne favorisent pas la reproductibilite par defaut des executions. Et quand a la complexite des architectures nous sommes entrain aujourd'hui de concretement aller a l'ExaScale (1 milliard de milliards de calculs a la seconde). Des systemes dont la dimension et la complexite soulevent d'autres problemes en plus que ceux deja cite. Les divers defis exposes ici pour ces deux approches presentent des perspectives cruciales d'un point de vue de consensus dans la determination et la representation de l'information necessaire lors d'une investigation scientifique de maniere reproductible et aussi de formes de collaborations et de partages favorisant l'automatisation et l'interopabilite. Par ailleurs, notre recherche pour la perennite meme des propositions exposes, doit regarder au futur de l'experimentation et de l'informatique et y etre deja adapte: Les laboratoires integres et robotises et Les infrastructures ExaScale. methodology in brief Dans le cadre de cette these, l'etat de l'art s'est organise independamment des approches scientifiques exposees. Nous avons identifie trois aspects fondamentaux a l'etablissement unifie d'une solution pour la reproductibilite

des resultats de recherches: La structure, le contenu et le modele d'enregistrement des informations utiles lors d'une investigation; Le partage, la collaboration, l'inter-oprabilite et l'automatisation a partir des enregistrements; L'accès, l'extraction des enregistrements enrichis par des collaborations et les differentes applications possibles. Plus en profondeurs nous exposons les problemes rencontres dans ces aspects, les solutions a ces problemes ainsi que les problemes dans ces solutions. Nous considererons l'environnement actuel des approches puis nous detaillons en perspectives nos attentes et nos predictions concernant le futur des environnements vis a vis des deux approches scientifiques. La proposition que nous exposons est une solution d'unification qui tente de resoudre le probleme de la reproductibilite dans les investigations scientifiques. Elle presente la necessite de consensus et de solutions qui tentent de resoudre les problemes rencontres aussi bien en informatique qu'en experimentation en essayant au mieux d'avoir des approches generiques unificatrices. Elle aborde chacun des trois aspects identifies en prenant en compte l'etat actuel de la technologie mais egalement la direction technologique dans les 50 prochaines annees a venir. Results of propositions \*\* Next work here \*\* Ainsi, l'implementation de notre solution a travers la proposition d'une interface standard minimal de communication (ISMC) pour les machines et un model d'enregistrement integrant jusqu'aux executions sur les clusters nous a permis de garantir un certain nombre de situations. Dans le cas de l'experimentation, l'application sur un prototype de machine integree pour des experiences en BioScience nous a permis d'assurer la repetition et de verifier la reproduction des investigations. Par ailleurs aussi dans l'aspect informatique nous avons ete capables de tester l'enregistrement, la repetition et la verification de la reproductibilite d'investigation aussi bien sur des applications sequentielles, paralleles et distribuees sur cluster a noeuds homogenes et heterogenes. Cette derniere nous permettant de garantir l'aspect ExaScale pret de notre travail. Il est fondamental egalement d'ajouter que l'application de la partie communautaire en terme de collaboration, de partage et d'enrichissement des enregistrements, nous sommes arrives a une platform cloud suffisamment generique pour etre compatible a tous les deux formes d'enregistrements possibles (experimental et informatiques). Cette platform vient egalement avec des outils permettant de connecter aussi bien les machines ayant implementees l'ISMC que les ordinateurs pour une automatisation et une inter-operabilite tres appreciable.



# Table of contents

|  |             |
|--|-------------|
| <b>List of figures</b>                                       | <b>xvii</b> |
| <b>List of tables</b>  | <b>xix</b>  |
| <b>1 On the necessity of the scientific method</b>           | <b>1</b>    |
| 1.1 Introduction . . . . .                                   | 1           |
| 1.2 Reproducibility an Overview . . . . .                    | 1           |
| 1.2.1 An introducton to the scientific method . . . . .      | 1           |
| 1.2.2 Reproducibility at the heart of science . . . . .      | 1           |
| 1.3 Terminology . . . . .                                    | 1           |
| 1.4 Conclusion . . . . .                                     | 2           |
| <b>2 On the reproducibility of scientific investigations</b> | <b>3</b>    |
| 2.1 Introduction . . . . .                                   | 3           |
| 2.2 The canonical origins of the problem . . . . .           | 3           |
| 2.2.1 Low level: Hardware and Electro-Mechanics . . . . .    | 3           |
| 2.2.2 Operating Software . . . . .                           | 3           |
| 2.2.3 Environmental . . . . .                                | 3           |
| 2.2.4 Behavioral . . . . .                                   | 3           |
| 2.2.5 Side effect of historical solutions . . . . .          | 3           |
| 2.3 Approaches and Solutions . . . . .                       | 3           |
| 2.3.1 Environment capture . . . . .                          | 3           |
| 2.3.2 Numerical garanty . . . . .                            | 4           |
| 2.3.3 Workflow capture . . . . .                             | 4           |
| 2.3.4 Data and Investigation discoverability . . . . .       | 4           |
| 2.3.5 Data Representation . . . . .                          | 4           |
| 2.4 Conclusion . . . . .                                     | 4           |

|          |  |          |
|----------|--|----------|
| <b>3</b> | <b>On the contributions of this work</b>                                   | <b>5</b> |
| 3.1      | Introduction . . . . .   | 5        |
| 3.2      | Reproducibility enriched collaboration . . . . .                           | 5        |
| 3.2.1    | Introduction . . . . .   | 6        |
| 3.2.2    | Specification . . . . .  | 6        |
| 3.2.3    | Limitations . . . . .  | 6        |
| 3.2.4    | Implementation . . . . .   | 6        |
| 3.3      | Private Data and Code Discoverability . . . . .                            | 6        |
| 3.3.1    | Introduction . . . . .   | 6        |
| 3.3.2    | Specification . . . . .  | 6        |
| 3.3.3    | Limitations . . . . .  | 6        |
| 3.3.4    | Implementation . . . . .   | 6        |
| 3.4      | Investigation Recording Toolbox . . . . .                                  | 6        |
| 3.4.1    | Introduction . . . . .   | 7        |
| 3.4.2    | Specification . . . . .  | 7        |
| 3.4.3    | Limitations . . . . .  | 7        |
| 3.4.4    | Implementation . . . . .   | 7        |
| 3.5      | A Unified Computational Environment Capturing Approach . . . . .           | 7        |
| 3.5.1    | Introduction . . . . .   | 7        |
| 3.5.2    | Specification . . . . .  | 7        |
| 3.5.3    | Limitations . . . . .  | 7        |
| 3.5.4    | Implementation . . . . .   | 7        |
| 3.6      | Reproducibility focused experimental Machine Interoperability Protocol . . | 7        |
| 3.6.1    | Introduction . . . . .   | 8        |
| 3.6.2    | Specification . . . . .  | 8        |
| 3.6.3    | Limitations . . . . .  | 8        |
| 3.6.4    | Implementation . . . . .   | 8        |
| 3.7      | Numerical Caching . . . . .  | 8        |
| 3.7.1    | Introduction . . . . .   | 8        |
| 3.7.2    | Specification . . . . .  | 8        |
| 3.7.3    | Limitations . . . . .  | 8        |
| 3.7.4    | Implementation . . . . .   | 8        |
| 3.8      | Conclusion . . . . .   | 8        |
| <b>4</b> | <b>My fourth chapter</b>   | <b>9</b> |
| 4.1      | First section of the third chapter . . . . .                               | 9        |
| 4.1.1    | First subsection in the first section . . . . .                            | 9        |

|  |           |
|--|-----------|
| Table of contents  | xv        |
| 4.1.2 Second subsection in the first section . . . . .       | 9         |
| 4.1.3 Third subsection in the first section . . . . .        | 9         |
| 4.2 Second section of the third chapter . . . . .            | 10        |
| 4.3 The layout of formal tables . . . . .                    | 10        |
| <b>References</b>  | <b>13</b> |
| <b>Appendix A How to install L<sup>A</sup>T<sub>E</sub>X</b> | <b>15</b> |
| <b>Appendix B Installing the CUED class file</b>             | <b>19</b> |





## List of figures



# List of tables

|     |  |    |
|-----|--|----|
| 4.1 | A badly formatted table . . . . .                  | 11 |
| 4.2 | A nice looking table . . . . .                     | 11 |
| 4.3 | Even better looking table using booktabs . . . . . | 11 |



# **Chapter 1**

## **On the necessity of the scientific method**

### **1.1 Introduction**

Introduction here...

### **1.2 Reproducibility an Overview**

Short introduction here...

#### **1.2.1 An introduction to the scientific method**

The scientific method...

#### **1.2.2 Reproducibility at the heart of science**

content of section2.2. the role of collaboration and anything that contribute to it. how reproducibility lay the foundations for continuity and advancement which is at the heart of science....

### **1.3 Terminology**

Terminology discussion here and state of the art...

## **1.4 Conclusion**

local conclusion of the necessity of the scientific method. . .

# **Chapter 2**

## **On the reproducibility of scientific investigations**

### **2.1 Introduction**

An opening to the current issues in the reproducibility of scientific investigations...

### **2.2 The canonical origins of the problem**

An overview of the fundamental source of non reproducibility...

#### **2.2.1 Low level: Hardware and Electro-Mechanics**

#### **2.2.2 Operating Software**

#### **2.2.3 Environmental**

#### **2.2.4 Behavioral**

#### **2.2.5 Side effect of historical solutions**

### **2.3 Approaches and Solutions**

#### **2.3.1 Environment capture**

1. Principle

2. Approaches

3. Limitations

### **2.3.2 Numerical garanty**

1. Principle

2. Approaches

3. Limitations

### **2.3.3 Workflow capture**

1. Principle

2. Approaches

3. Limitations

### **2.3.4 Data and Investigation discoverability**

1. Principle

2. Approaches

3. Limitations

### **2.3.5 Data Representation**

1. Principle

2. Approaches

3. Limitations

## **2.4 Conclusion**



# **Chapter 3**

## **On the contributions of this work**

### **3.1 Introduction**

A detailed introduction presenting the motivation of this layout and work. Why do we start by collaboration first? We start very high to lower and lower level problems solving. At every level the solution proposed is open to change adaptation and high integration capability.

### **3.2 Reproducibility enriched collaboration**

Collaboration mode explained. Yet also the value of adding things on top of CoRR. How the models can evolve. Backward compatibility. API implementation for applications. Integration Examples. Toolbox on top of CoRR: Data analytics, Visualization, Interoperability broker between Applications. Fundamental similarities between machine and computer based records. Privacy issues and transition to the next section.

### **3.2.1 Introduction**

### **3.2.2 Specification**

### **3.2.3 Limitations**

### **3.2.4 Implementation**

## **3.3 Private Data and Code Discoverability**

After this point we discuss solutions on how we effectively record investigations. What we capture.

### **3.3.1 Introduction**

### **3.3.2 Specification**

### **3.3.3 Limitations**

### **3.3.4 Implementation**

## **3.4 Investigation Recording Toolbox**

A command line and gui tool to record experimental and computational investigations data and metadata. `corr-cmd`. It is also very flexible. This toolbox has an easy to extend capability that allows the integration of new recording approaches but also for interoperability integration. The submitted extension is tested then integrated and patched in the current release or the next one.

**3.4.1 Introduction****3.4.2 Specification****3.4.3 Limitations****3.4.4 Implementation****3.5 A Unified Computational Environment Capturing Approach**

We explain a critical feature in the precedent section for computational environment and the differences with machine ones. We address the fact that machine record requires strategies and leverage of a minimal protocol.

**3.5.1 Introduction****3.5.2 Specification****3.5.3 Limitations****3.5.4 Implementation****3.6 Reproducibility focused experimental Machine Interoperability Protocol**

We try to propose a protocol that allows interoperability features but at the same time a needed design to allow reproducibility: Machine computational state accessible programmatically. An example is a mechanical roller to set the temperature and a led screen display. The mechanical roll has to be infinite to allow programmatical consistency or loopback roll to the position. The best solution is having fully digital screens or softwares for User Interfaces to the machine to leverage the API.

### **3.6.1 Introduction**

### **3.6.2 Specification**

### **3.6.3 Limitations**

### **3.6.4 Implementation**

## **3.7 Numerical Caching**

We investigated here an approach of recording the investigation atomic numerics during the execution. Which results in a heavier execution and trace. For the computation it will be the ALU sys calls and for the machine it will be also the ALU but the position and status of all the sensors, motors and all electromechanical parts.

### **3.7.1 Introduction**

### **3.7.2 Specification**

### **3.7.3 Limitations**

### **3.7.4 Implementation**

## **3.8 Conclusion**

# Chapter 4

## My fourth chapter

### 4.1 First section of the third chapter

And now I begin my third chapter here ...

And now to cite some more people Read [2], Ancey et al. [1]

#### 4.1.1 First subsection in the first section

...and some more

#### 4.1.2 Second subsection in the first section

...and some more ...

##### First subsub section in the second subsection

...and some more in the first subsub section otherwise it all looks the same doesn't it? well we can add some text to it ...

#### 4.1.3 Third subsection in the first section

...and some more ...

##### First subsub section in the third subsection

...and some more in the first subsub section otherwise it all looks the same doesn't it? well we can add some text to it and some more and some more and some more and some more and some more and some more and some more ...

### Second subsub section in the third subsection

... and some more in the first subsub section otherwise it all looks the same doesn't it? well we can add some text to it ...

## 4.2 Second section of the third chapter

and here I write more ...

## 4.3 The layout of formal tables

This section has been modified from “Publication quality tables in L<sup>A</sup>T<sub>E</sub>X<sup>\*</sup>” by Simon Fear.

The layout of a table has been established over centuries of experience and should only be altered in extraordinary circumstances.

When formatting a table, remember two simple guidelines at all times:

1. Never, ever use vertical rules (lines).
2. Never use double rules.

These guidelines may seem extreme but I have never found a good argument in favour of breaking them. For example, if you feel that the information in the left half of a table is so different from that on the right that it needs to be separated by a vertical line, then you should use two tables instead. Not everyone follows the second guideline:

There are three further guidelines worth mentioning here as they are generally not known outside the circle of professional typesetters and subeditors:

3. Put the units in the column heading (not in the body of the table).
4. Always precede a decimal point by a digit; thus 0.1 *not* just .1.
5. Do not use ‘ditto’ signs or any other such convention to repeat a previous value. In many circumstances a blank will serve just as well. If it won't, then repeat the value.

A frequently seen mistake is to use ‘`\begin{center}`’ ... ‘`\end{center}`’ inside a figure or table environment. This center environment can cause additional vertical space. If you want to avoid that just use ‘`\centering`’

Table 4.1 A badly formatted table

|                    | Species I |      | Species II |      |
|--------------------|-----------|------|------------|------|
| Dental measurement | mean      | SD   | mean       | SD   |
| I1MD               | 6.23      | 0.91 | 5.2        | 0.7  |
| I1LL               | 7.48      | 0.56 | 8.7        | 0.71 |
| I2MD               | 3.99      | 0.63 | 4.22       | 0.54 |
| I2LL               | 6.81      | 0.02 | 6.66       | 0.01 |
| CMD                | 13.47     | 0.09 | 10.55      | 0.05 |
| CBL                | 11.88     | 0.05 | 13.11      | 0.04 |

Table 4.2 A nice looking table

| Dental measurement | Species I |      | Species II |      |
|--------------------|-----------|------|------------|------|
|                    | mean      | SD   | mean       | SD   |
| I1MD               | 6.23      | 0.91 | 5.2        | 0.7  |
| I1LL               | 7.48      | 0.56 | 8.7        | 0.71 |
| I2MD               | 3.99      | 0.63 | 4.22       | 0.54 |
| I2LL               | 6.81      | 0.02 | 6.66       | 0.01 |
| CMD                | 13.47     | 0.09 | 10.55      | 0.05 |
| CBL                | 11.88     | 0.05 | 13.11      | 0.04 |

Table 4.3 Even better looking table using booktabs

| Dental measurement | Species I |      | Species II |      |
|--------------------|-----------|------|------------|------|
|                    | mean      | SD   | mean       | SD   |
| I1MD               | 6.23      | 0.91 | 5.2        | 0.7  |
| I1LL               | 7.48      | 0.56 | 8.7        | 0.71 |
| I2MD               | 3.99      | 0.63 | 4.22       | 0.54 |
| I2LL               | 6.81      | 0.02 | 6.66       | 0.01 |
| CMD                | 13.47     | 0.09 | 10.55      | 0.05 |
| CBL                | 11.88     | 0.05 | 13.11      | 0.04 |





# References

- [1] Ancey, C., Coussot, P., and Evesque, P. (1996). Examination of the possibility of a fluid-mechanics treatment of dense granular flows. *Mechanics of Cohesive-frictional Materials*, 1(4):385–403.
- [2] Read, C. J. (1985). A solution to the invariant subspace problem on the space  $l_1$ . *Bull. London Math. Soc.*, 17:305–317.



# Appendix A

## How to install L<sup>A</sup>T<sub>E</sub>X

### Windows OS

#### TeXLive package - full version

1. Download the TeXLive ISO (2.2GB) from  
<https://www.tug.org/texlive/>
2. Download WinCDEmu (if you don't have a virtual drive) from  
<http://wincdemu.sysprogs.org/download/>
3. To install Windows CD Emulator follow the instructions at  
<http://wincdemu.sysprogs.org/tutorials/install/>
4. Right click the iso and mount it using the WinCDEmu as shown in  
<http://wincdemu.sysprogs.org/tutorials/mount/>
5. Open your virtual drive and run setup.pl

or

#### Basic MikTeX - T<sub>E</sub>X distribution

1. Download Basic-MiK<sub>T</sub>E<sub>X</sub>(32bit or 64bit) from  
<http://miktex.org/download>
2. Run the installer
3. To add a new package go to Start » All Programs » MikTeX » Maintenance (Admin)  
and choose Package Manager

4. Select or search for packages to install

## **TexStudio - T<sub>E</sub>X editor**

1. Download TexStudio from  
<http://texstudio.sourceforge.net/#downloads>
2. Run the installer

## **Mac OS X**

### **MacTeX - T<sub>E</sub>X distribution**

1. Download the file from  
<https://www.tug.org/mactex/>
2. Extract and double click to run the installer. It does the entire configuration, sit back and relax.

### **TexStudio - T<sub>E</sub>X editor**

1. Download TexStudio from  
<http://texstudio.sourceforge.net/#downloads>
2. Extract and Start

## **Unix/Linux**

### **TeXLive - T<sub>E</sub>X distribution**

#### **Getting the distribution:**

1. TexLive can be downloaded from  
<http://www.tug.org/texlive/acquire-netinstall.html>.
2. TexLive is provided by most operating system you can use (rpm,apt-get or yum) to get TexLive distributions

## Installation

1. Mount the ISO file in the mnt directory

```
mount -t iso9660 -o ro,loop,noauto /your/texlive####.iso /mnt
```

2. Install wget on your OS (use rpm, apt-get or yum install)
3. Run the installer script install-tl.

```
cd /your/download/directory
./install-tl
```

4. Enter command 'i' for installation
5. Post-Installation configuration:  
<http://www.tug.org/texlive/doc/texlive-en/texlive-en.html#x1-320003.4.1>
6. Set the path for the directory of TexLive binaries in your .bashrc file

### For 32bit OS

For Bourne-compatible shells such as bash, and using Intel x86 GNU/Linux and a default directory setup as an example, the file to edit might be

```
edit ~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/i386-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;
export INFOPATH
```

### For 64bit OS

```
edit ~/.bashrc file and add following lines
PATH=/usr/local/texlive/2011/bin/x86_64-linux:$PATH;
export PATH
MANPATH=/usr/local/texlive/2011/texmf/doc/man:$MANPATH;
export MANPATH
```

```
INFOPATH=/usr/local/texlive/2011/texmf/doc/info:$INFOPATH;  
export INFOPATH
```

**Fedora/RedHat/CentOS:**

```
sudo yum install texlive  
sudo yum install psutils
```

**SUSE:**

```
sudo zypper install texlive
```

**Debian/Ubuntu:**

```
sudo apt-get install texlive texlive-latex-extra  
sudo apt-get install psutils
```

## Appendix B

### Installing the CUED class file

$\text{\LaTeX}$ .cls files can be accessed system-wide when they are placed in the  $\langle\text{texmf}\rangle/\text{tex}/\text{latex}$  directory, where  $\langle\text{texmf}\rangle$  is the root directory of the user's  $\text{\TeX}$  installation. On systems that have a local  $\text{texmf}$  tree ( $\langle\text{texmflocal}\rangle$ ), which may be named “ $\text{texmf-local}$ ” or “ $\text{localtexmf}$ ”, it may be advisable to install packages in  $\langle\text{texmflocal}\rangle$ , rather than  $\langle\text{texmf}\rangle$  as the contents of the former, unlike that of the latter, are preserved after the  $\text{\LaTeX}$  system is reinstalled and/or upgraded.

It is recommended that the user create a subdirectory  $\langle\text{texmf}\rangle/\text{tex}/\text{latex}/\text{CUED}$  for all CUED related  $\text{\LaTeX}$  class and package files. On some  $\text{\LaTeX}$  systems, the directory look-up tables will need to be refreshed after making additions or deletions to the system files. For  $\text{\TeX}$ Live systems this is accomplished via executing “ $\text{texhash}$ ” as root.  $\text{MikTeX}$  users can run “ $\text{initexmf -u}$ ” to accomplish the same thing.

Users not willing or able to install the files system-wide can install them in their personal directories, but will then have to provide the path (full or relative) in addition to the filename when referring to them in  $\text{\LaTeX}$ .

