Installation Guide for Android NDK & Tesseract Library on Windows 7 (64bit)

This guide shows the steps for the full installation of Android NDK Toolset and implementation of OCR library (Tesseract) in Android application. Tesseract is an OCR Engine that was developed at HP Labs between 1985 and 1995 and now at Google.

Author: Giorgos Papadimitriou

Publication Date: 14/08/2013

Email: gipap90@gmail.com

LinkedIn: gr.linkedin.com/pub/giorgos-papadimitriou/75/40b/b2b/

Contents

Software requirement 3
Tested on OS: 3
Java JDK 4
Android ADT Bundle for Windows (including Android SDK+Eclipse) 6
Android NDK 7
Cygwin 8
Apache - Ant 13
Tesseract Library 14
Build Android NDK and Tesseract Library using Cygwin 15
Import Tesseract library and a Simple OCR example in ADT 22
References 30

Software requirement

- Java JDK
- Android ADT Bundle for Windows (including Android SDK + Eclipse)
- Android NDK
- Cygwin
- Apache-ant

Tested on OS:

Windows 7 Professional 64-bit (Service Pack 1)

WARNING

In all installation procedure, the **directory names** and **paths** should not contain any whitespaces.

Example:

 C:\ndk-r9

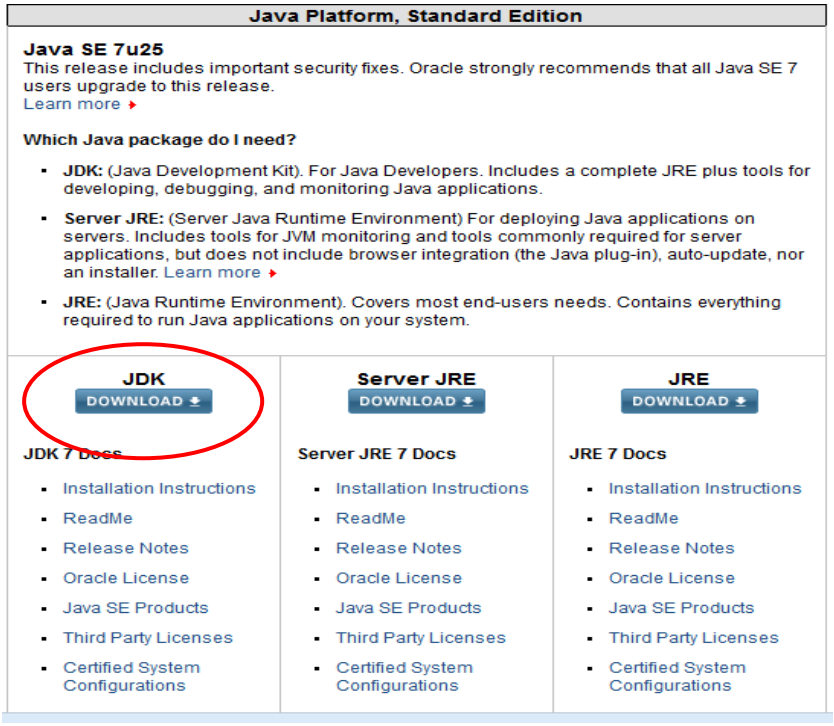
 C:\ndk r9

Java JDK

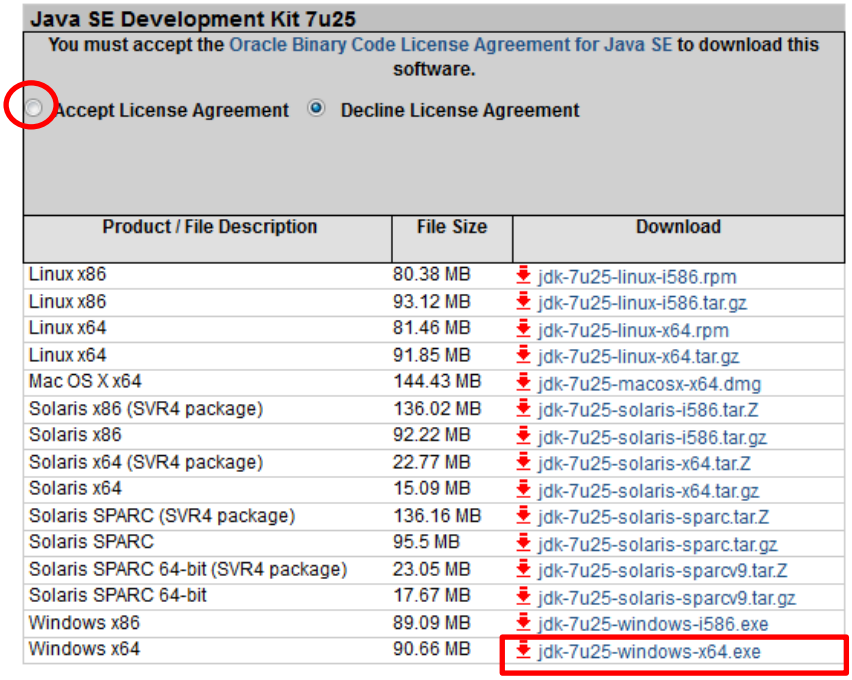
Download:

http://www.oracle.com/technetwork/java/javase/downloads/index.html

1. Click on JDK Download

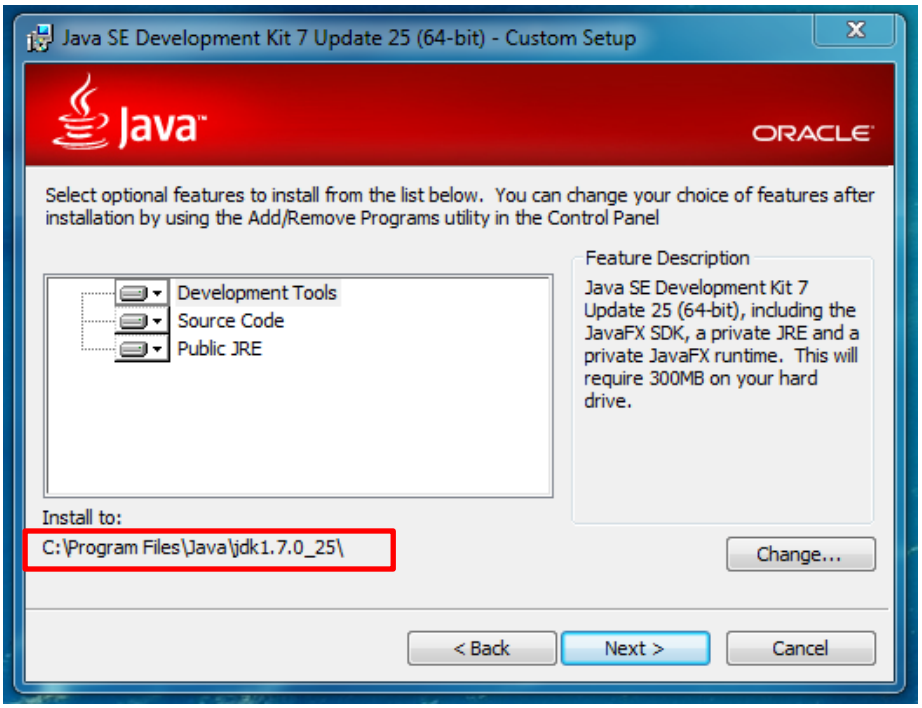


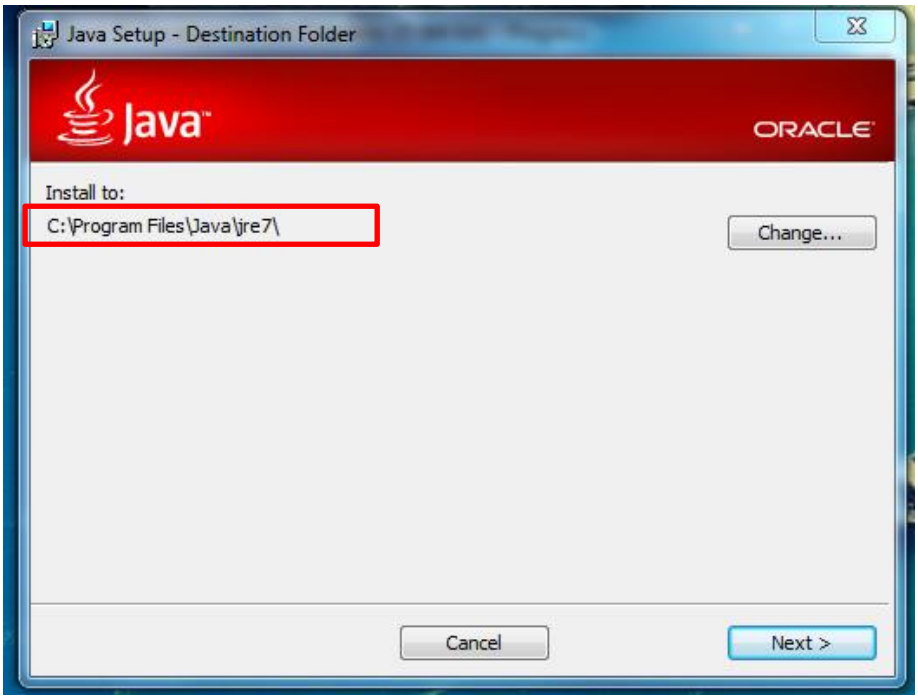
2. Click **Accept License Agreement** and then in the Download section the corresponding Windows **.exe file** version



Install:

Once the file has downloaded, install it. The installation paths should be as shown below or similar.





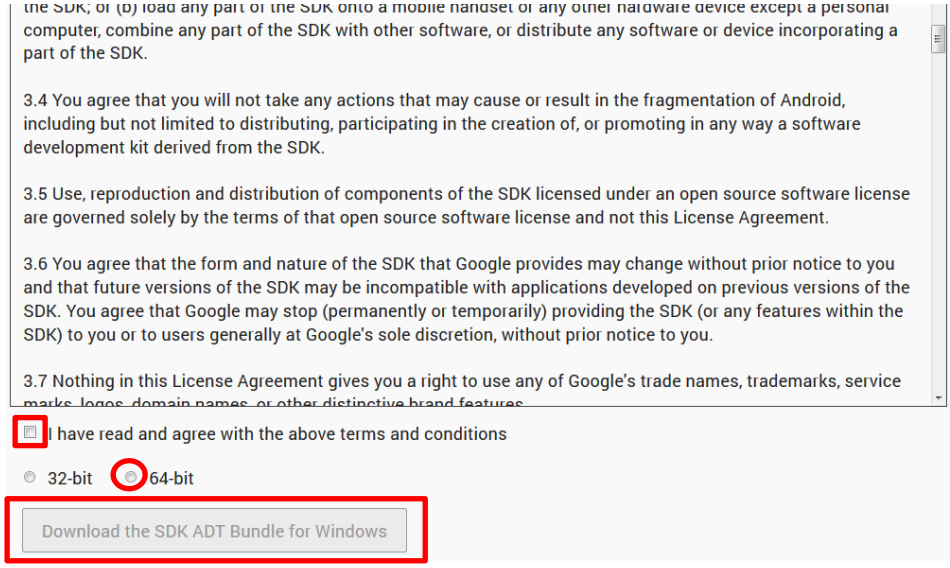
Android ADT Bundle for Windows (including Android SDK+Eclipse)

Download:

http://developer.android.com/sdk/index.html#download

1. Click Download the SDK

2. Accept the agreement, select system type and download the SDK ADT Bundle for Windows.



3. Once the file *adt-bundle-windows-x86_64-20130729.zip* downloaded, extract it on Desktop.

Android NDK

Download:

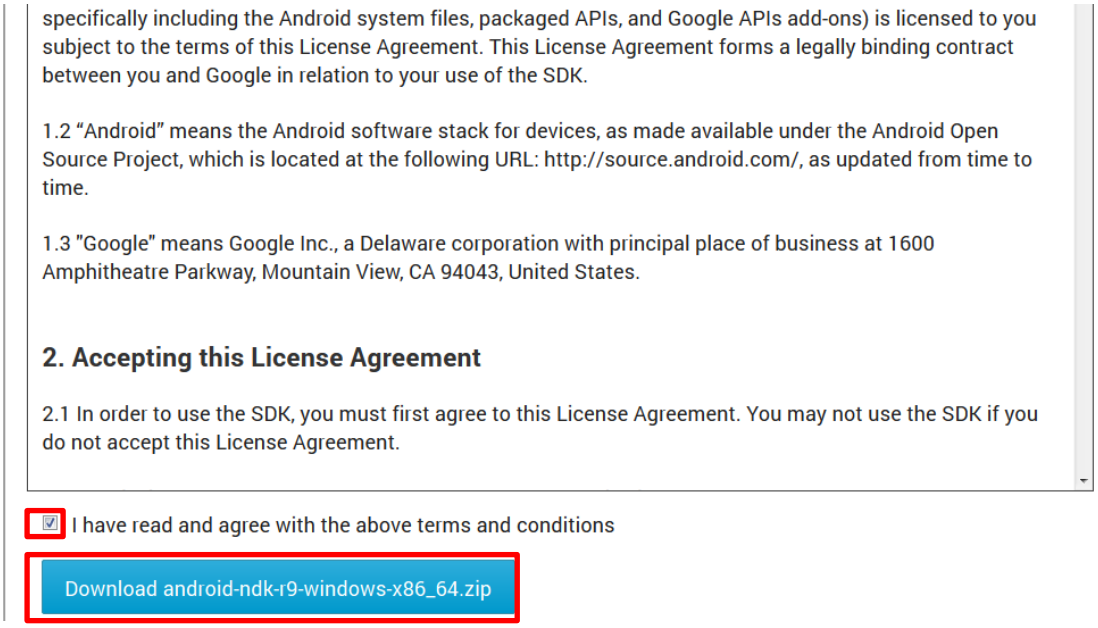
http://developer.android.com/tools/sdk/ndk/index.html

1. In the Package section click the corresponding Windows **.zip file** version.

Downloads

Platform	Package	Size (Bytes)	MD5 Checksum
Windows 32-bit	android-ndk-r9-windows-x86.zip	485200055	8895aec43f5141212c8dac6e9f07d5a8
	android-ndk-r9-windows-x86-legacy-toolchains.zip	292738221	ae3756d3773ec068fb653ff6fa411e35
Windows 64-bit	android-ndk-r9-windows-x86_64.zip	514321606	96c725d16ace7fd487bf1bc1427af3a0
	android-ndk-r9-windows-x86_64-legacy-toolchains.zip	312340413	707d1eaa6f5d427ad439c764c8bd68d2

2. Accept the agreement and download the **android-ndk-r9-windows-x86_64.zip**



3. Once the **android-ndk-r9-windows-x86_64.zip** has downloaded extract it, rename it from android-ndk-r9 to ndk-r9 and then move it on C:\

   Now the `<ndk-directory>` must be `C:\ndk-r9`

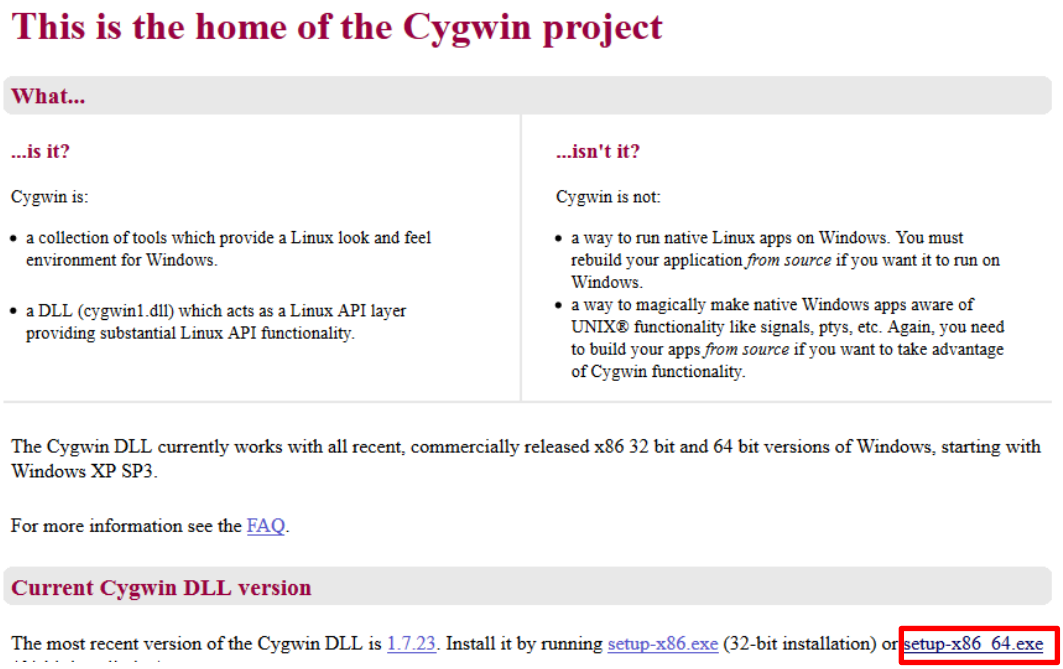
Note:
It shouldn't contain any whitespaces

Cygwin

Download:

http://www.cygwin.com/

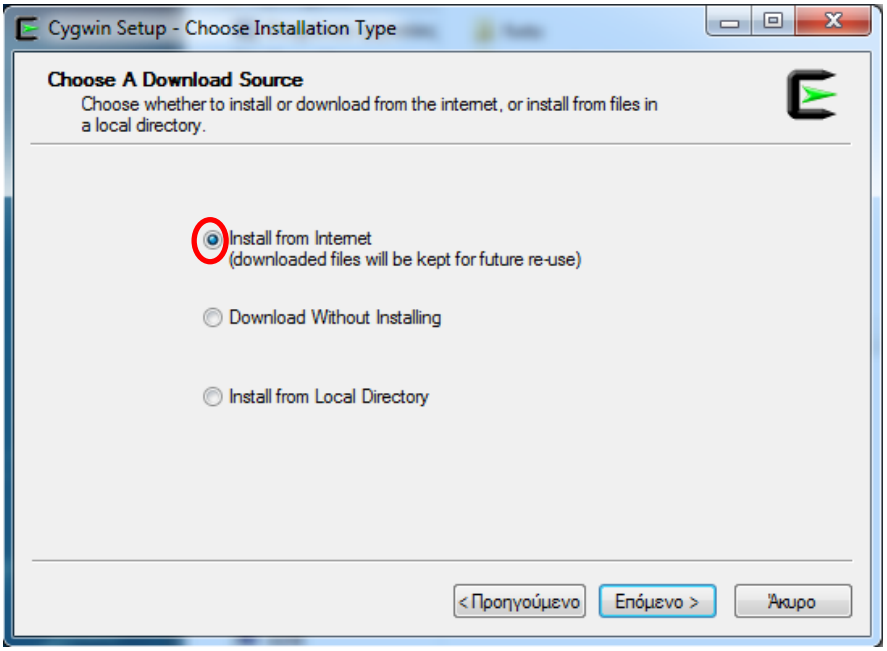
1. Click **setup-x86_64.exe**



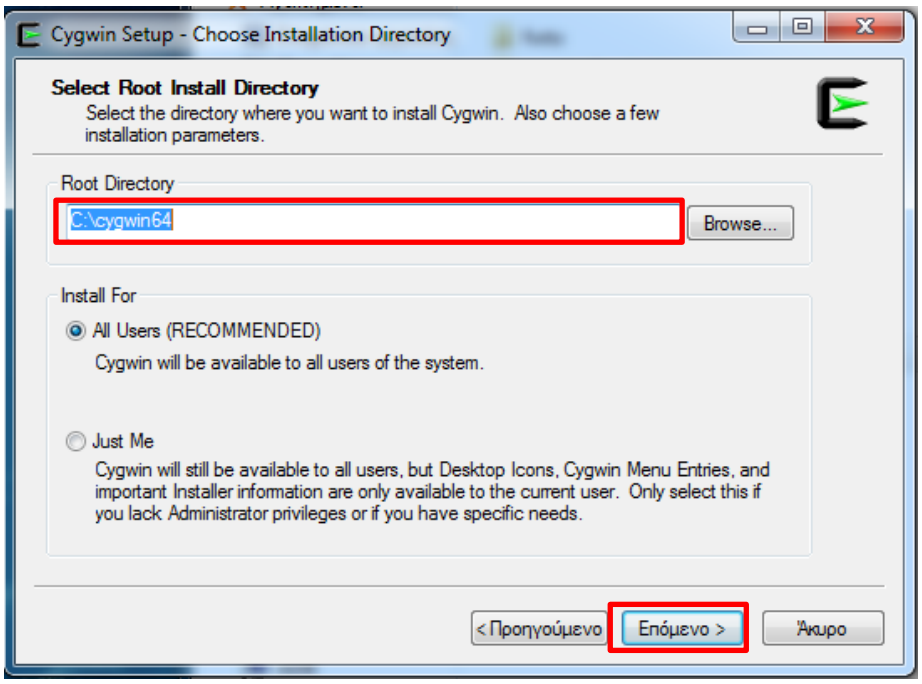
2. Once the Cygwin **setup-x86_64.exe** has downloaded, install it as shown below:

Install:

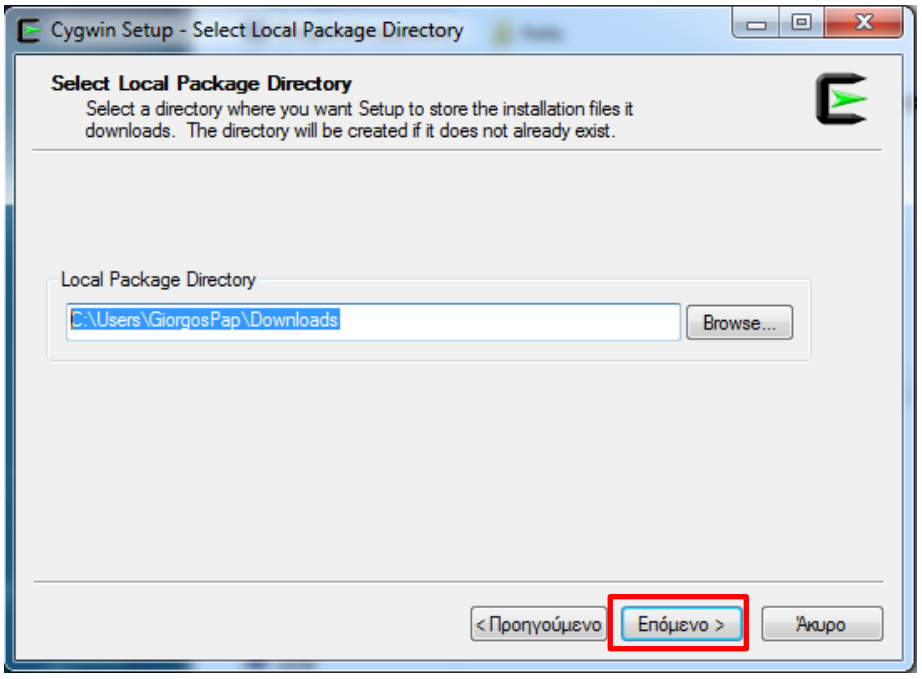
1. Choose **Install from Internet**



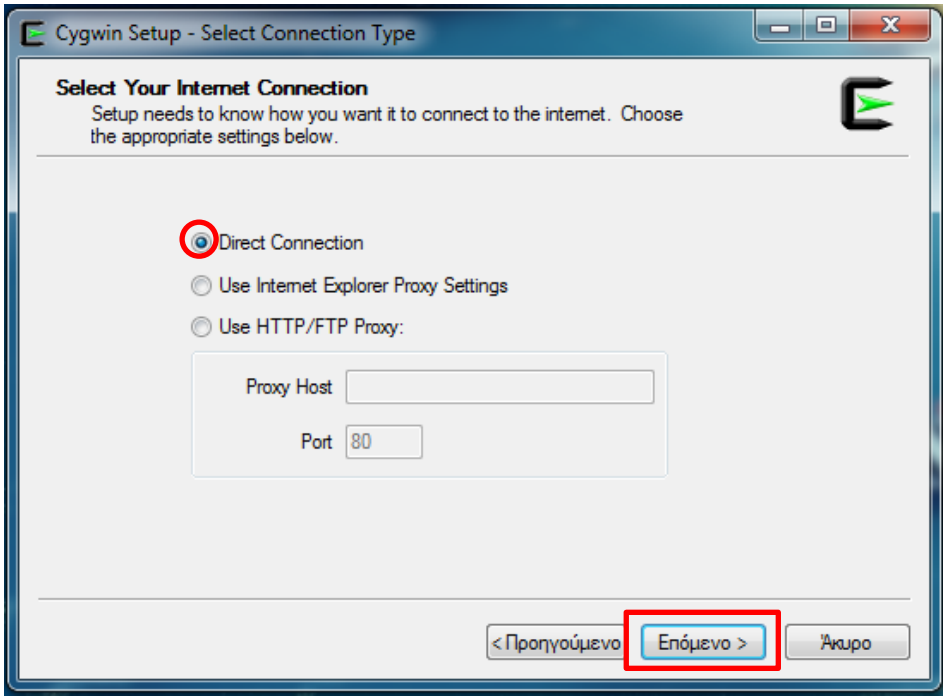
2. Leave the default installation directory and click **Next**



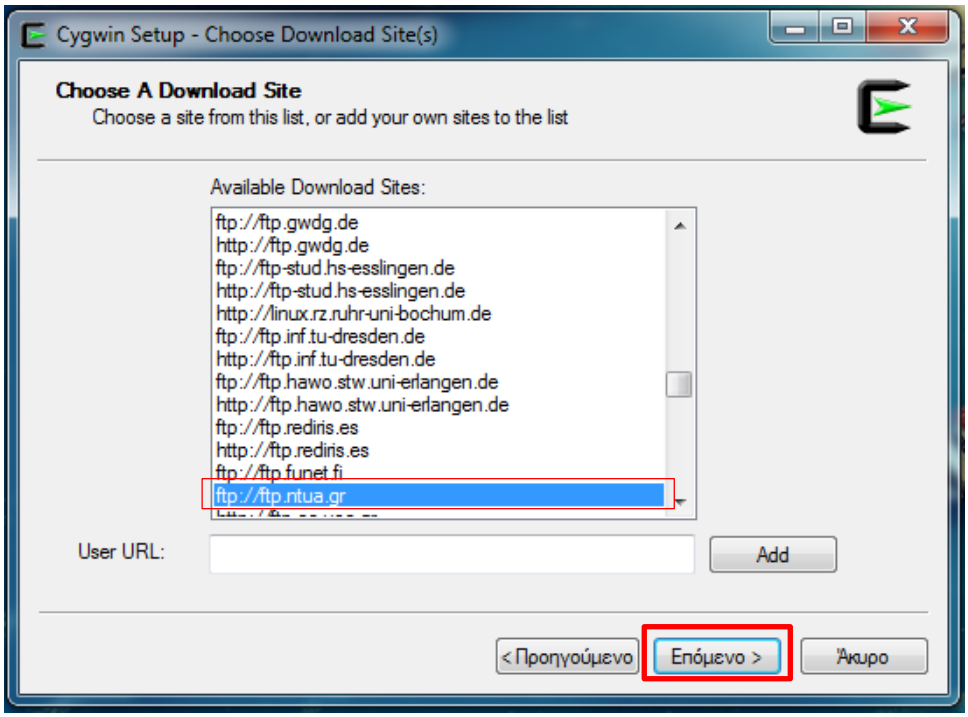
3. Leave the default Local Package Directory and click **Next**



4. Select Direct Connection and click **Next**



5. Choose a download site to download Cygwin and click **Next**

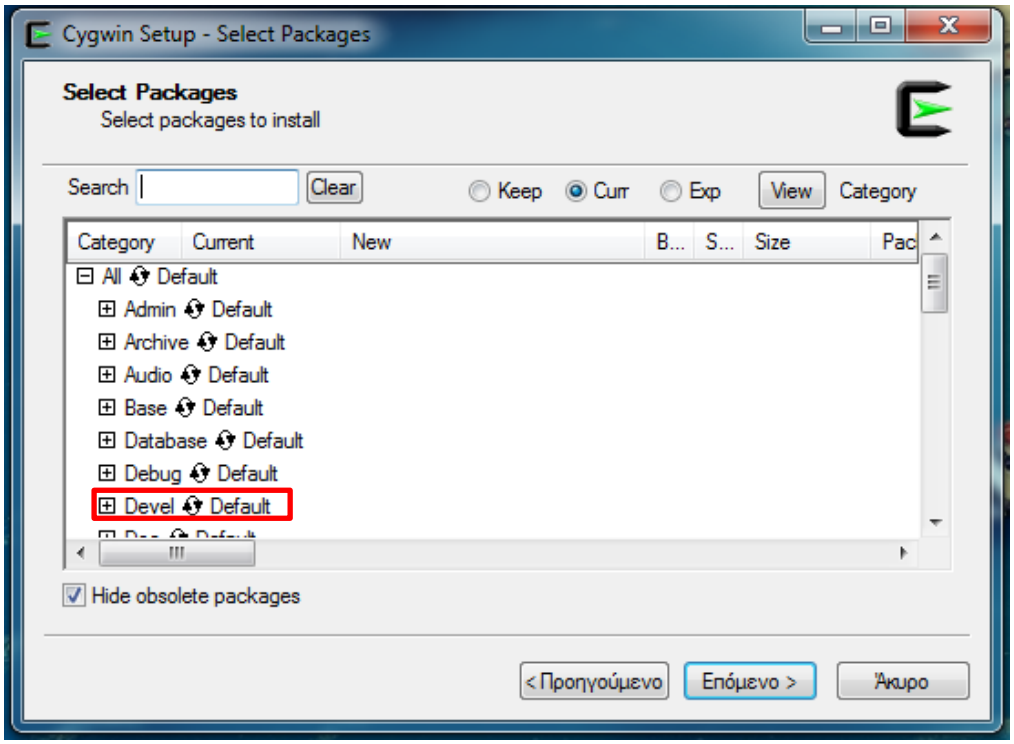


Note:

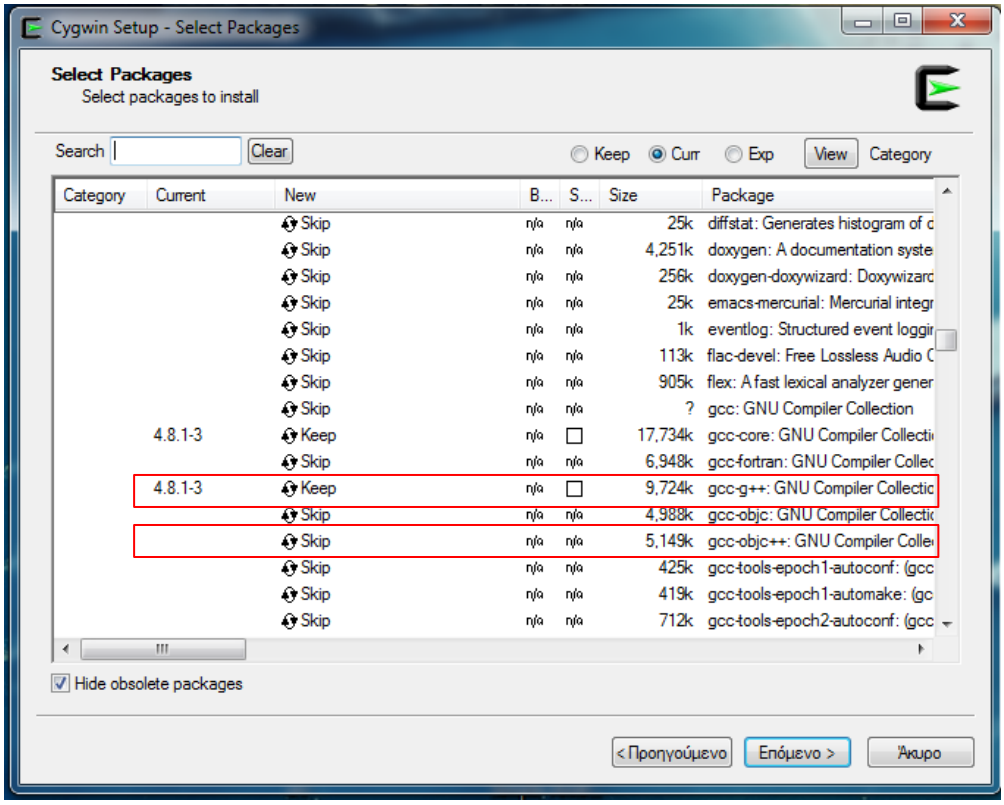
Usually we choose a site with our national TLD.

Example: Greece ➡ **.gr**

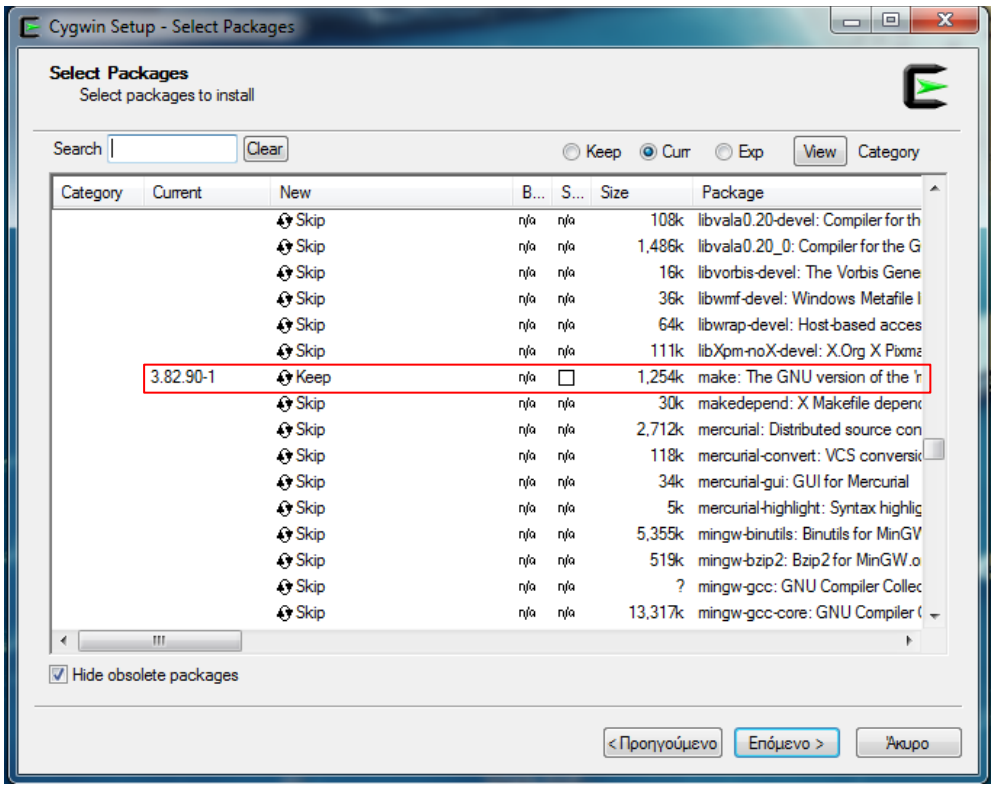
6. In the next screen we have to select the necessary packages. Click in the **Devel** selection

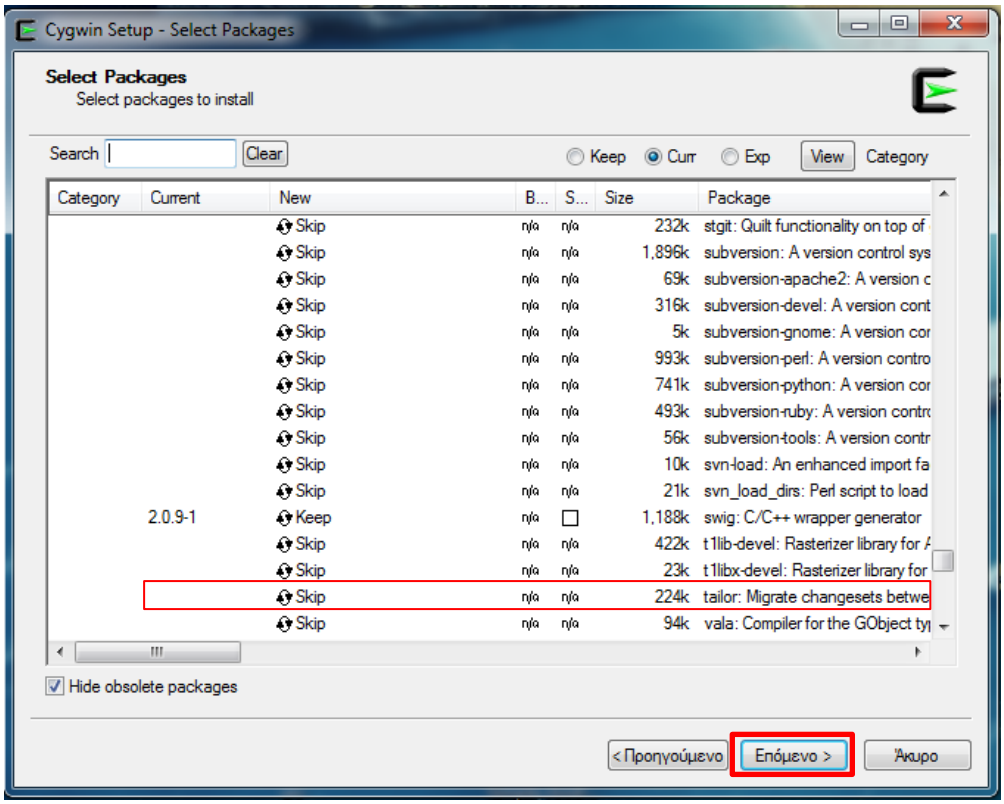


7. Scroll down until you find **gcc-core** and **gcc-g++** packages. Then click on them (once selected, the *New* section must show *Keep* label)



8. Again scroll down until you find **make** and **swig** packages. Then click on them (once selected, the *New* section must show *Keep* label). Final click **Next** to install Cygwin.



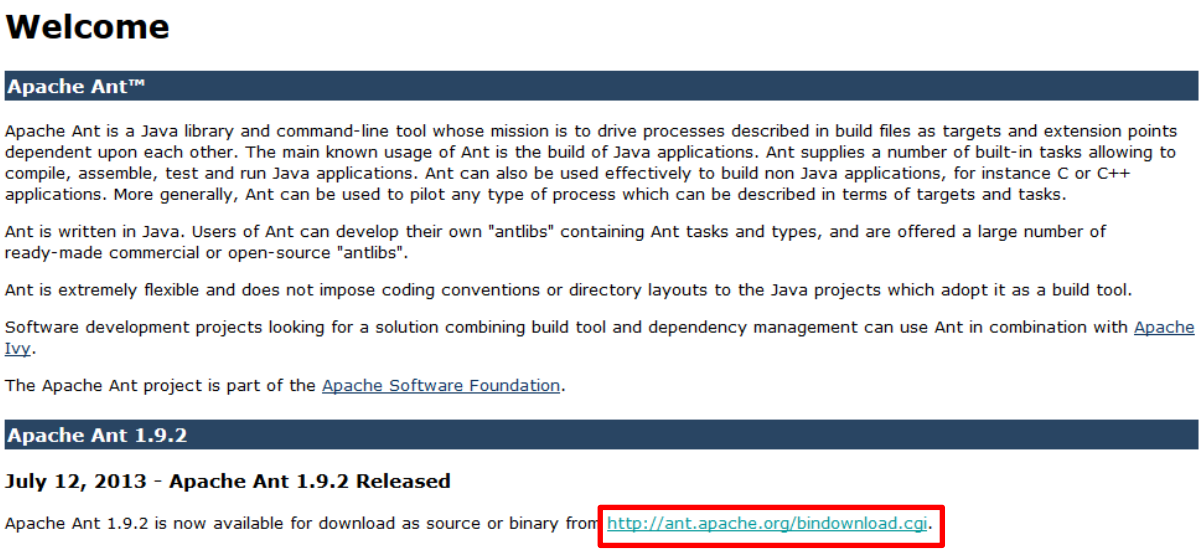


Apache - Ant

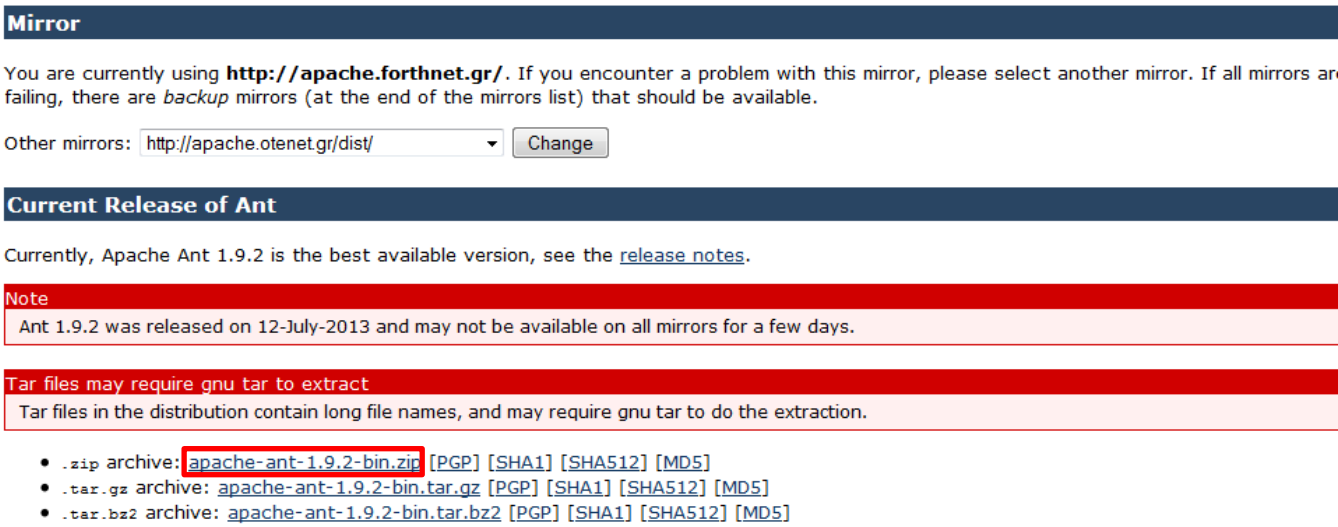
Download:

http://ant.apache.org/

1. Click on **http://ant.apache.org/bindownload.cgi.**



2. Scroll down and in the Current Release of Ant section, click on **apache-ant-1.9.2-bin.zip**



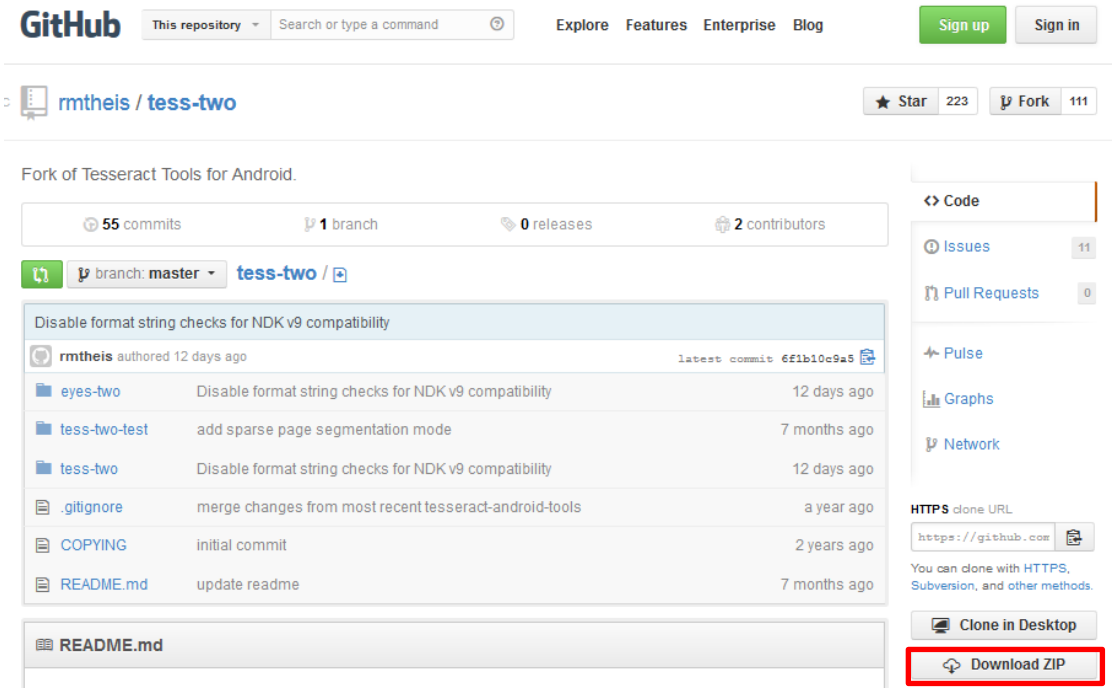
Once the **apache-ant-1.9.2-bin.zip** is downloaded extract it on C:\
Now the `<apache ant-directory>` must be `C:\apache-ant-1.9.2`

Tesseract Library

Download:

https://github.com/rmtheis/tess-two

1. Click Download ZIP

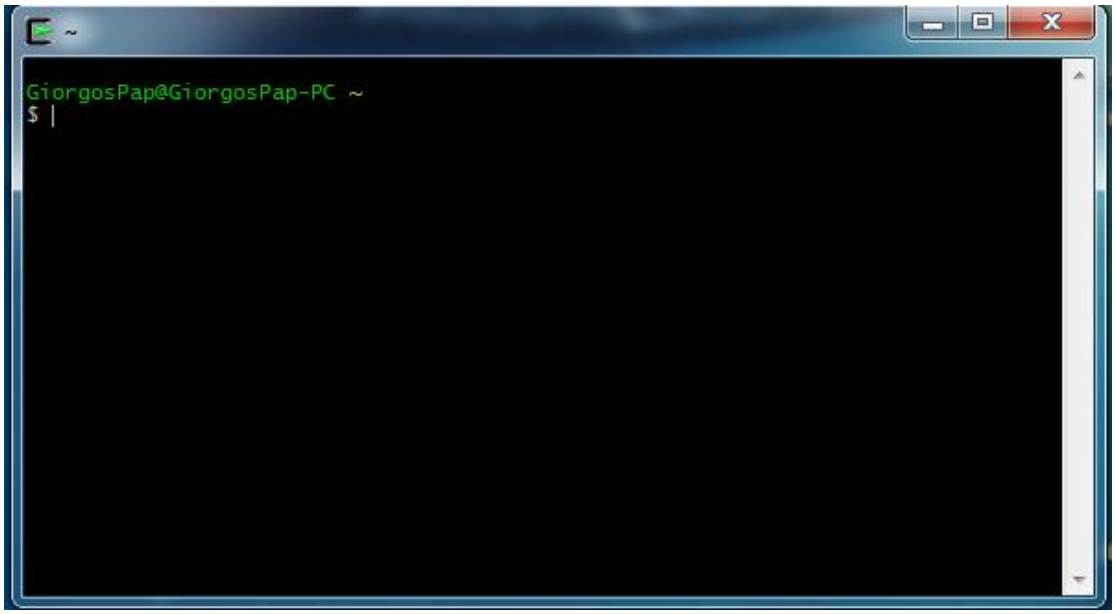


2. Once downloaded extract **tess-two-master** file on desktop.

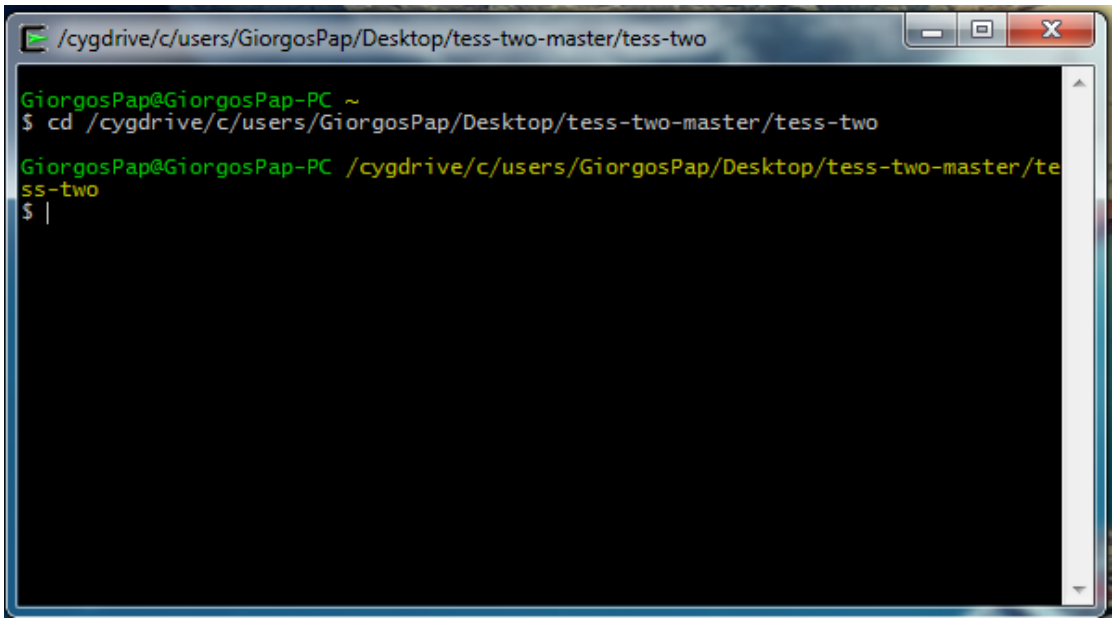
(example: C:\Users\GiorgosPap\Desktop\tess-two-master)

Build Android NDK and Tesseract Library using Cygwin

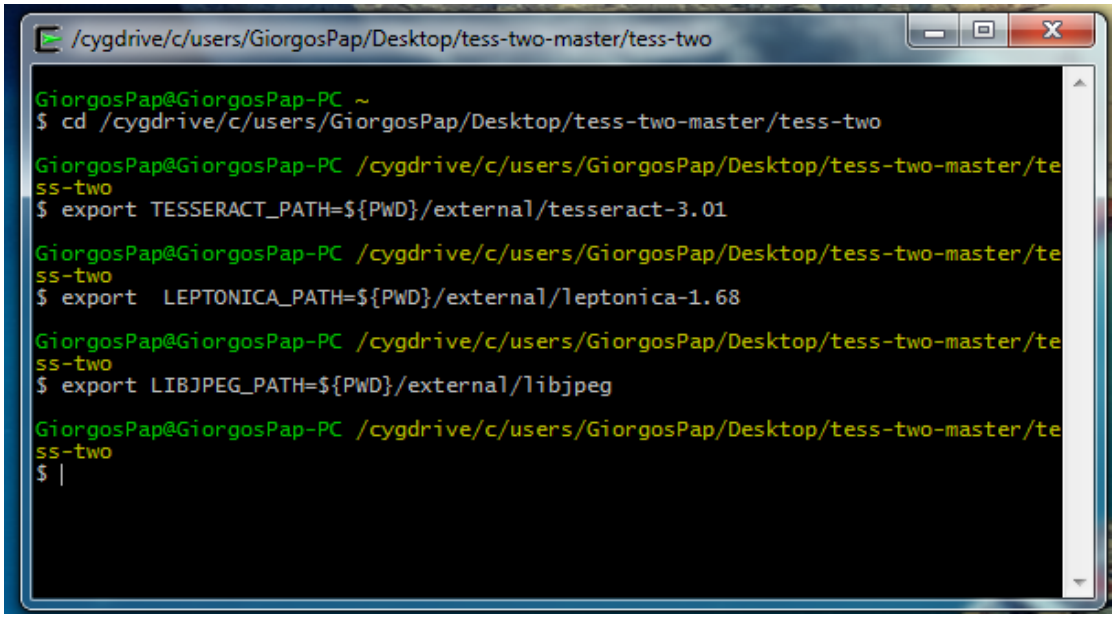
1. Run Cygwin64 Terminal:



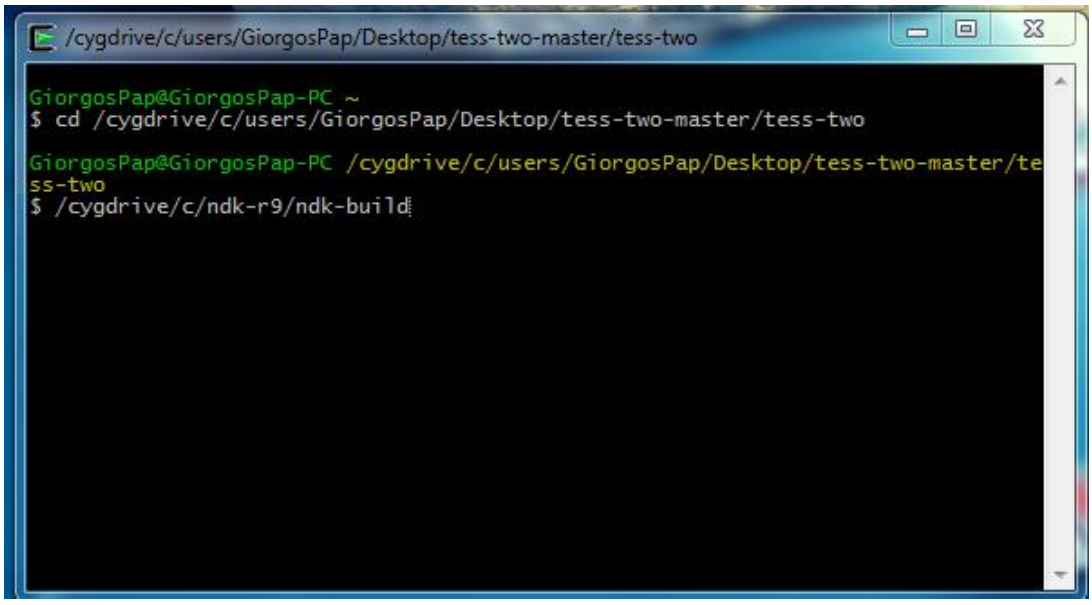
2. **cd <project-directory>/tess-two**



3. **export TESSERACT_PATH=${PWD}/external/tesseract-3.01**
4. **export LEPTONICA_PATH=${PWD}/external/leptonica-1.68**
5. **export LIBJPEG_PATH=${PWD}/external/libjpeg**



6. **/cygdrive/<ndk-directory>/ndk-build**



You have to wait about 45-60 minutes to build the ndk (depending on the computer's performance)

Note:

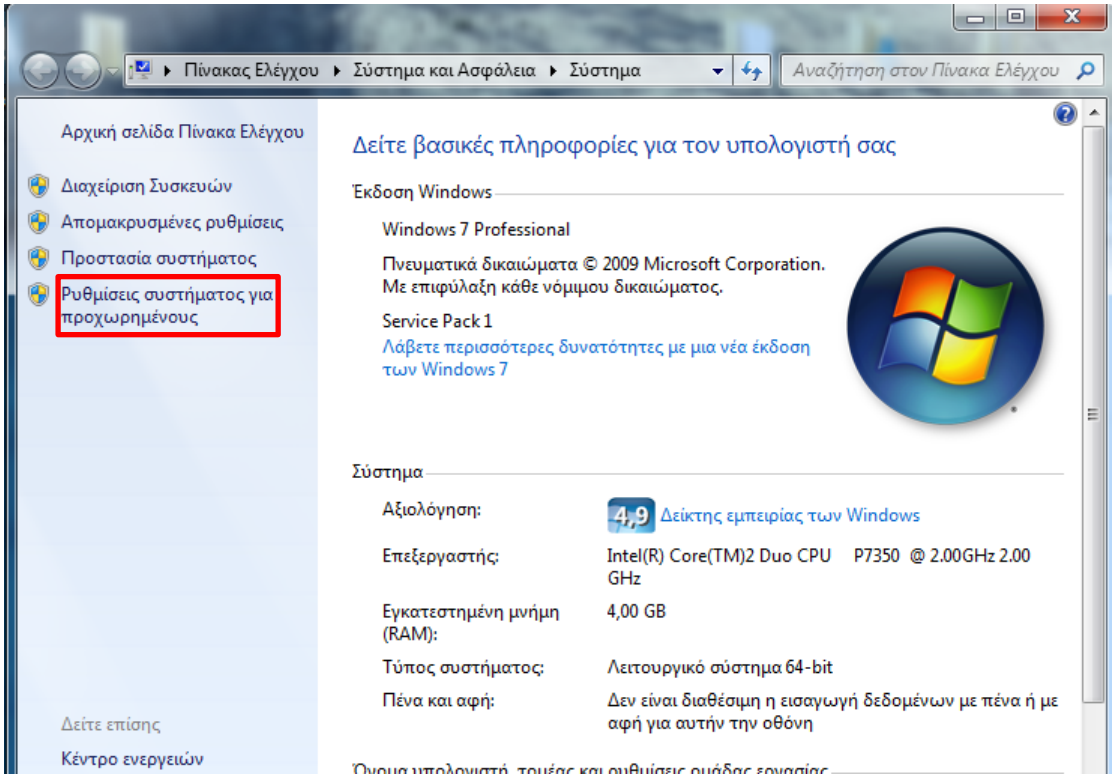
If you haven't installed additional libraries **gcc-core**, **gcc-g++**, **make**, **swig** with Cygwin then this error is shown:

```
$ ndk-build
-bash: ndk-build: command not found
```

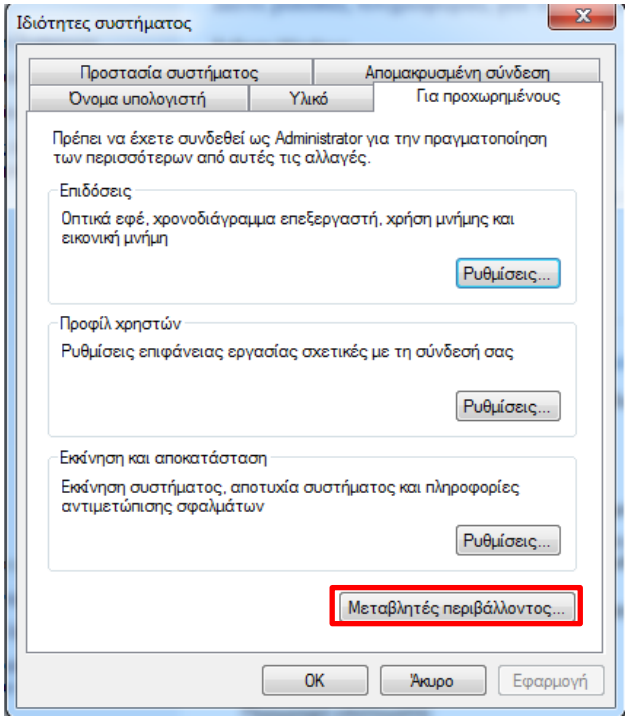
7. Create Enviroment Variables: `PATH, ANT_HOME, JAVA_HOME.`

   To Set/Create Enviroment Variable:

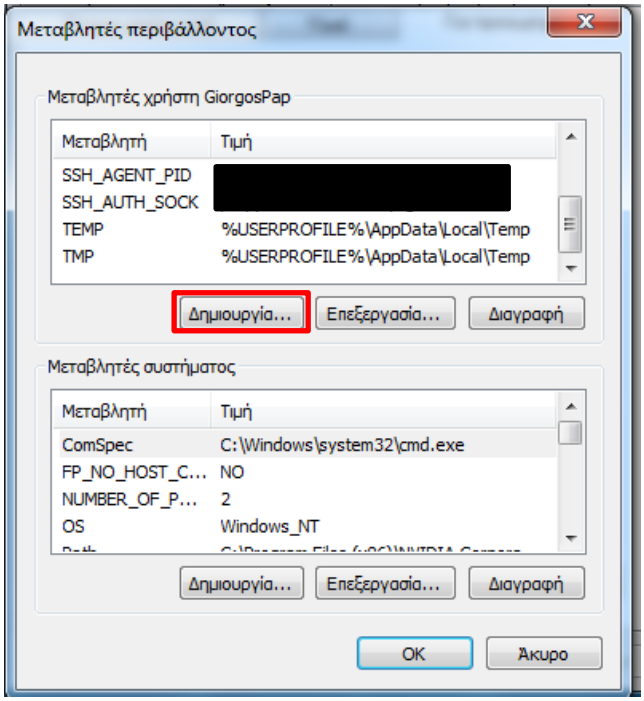
7.1 Right click on the Computer icon on your Desktop and choose **Properties** option.
7.2 Next click **Advanced system settings** on left section.



7.3 Under Advanced tab, click on **Environment Variables …**



7.4 In the next window, in *Users variables* click **New…**



7.5 Then add:

Variable name: `<VARIABLE_NAME>`

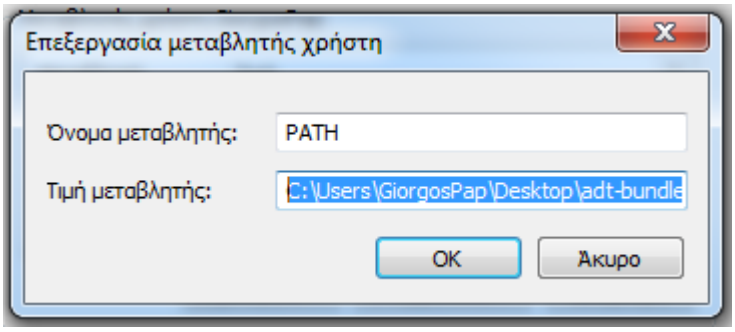
Variable value: `<program-directory>`

Note:

Make sure you separate the value with " ; "

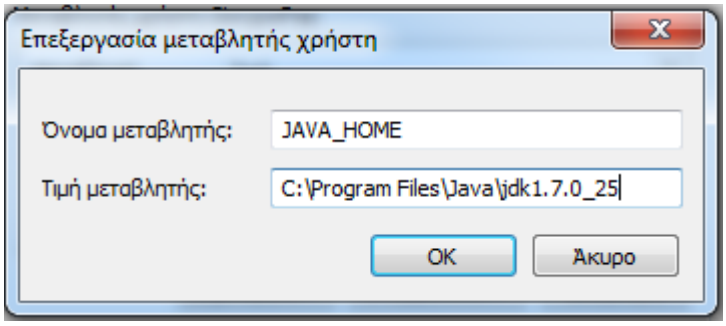
7.6 Variable name: **PATH**

Variable value: **C:\Users\GiorgosPap\Desktop\adt-bundle-windows-x86_64-20130729\sdk\tools;C:\ndk-r9**



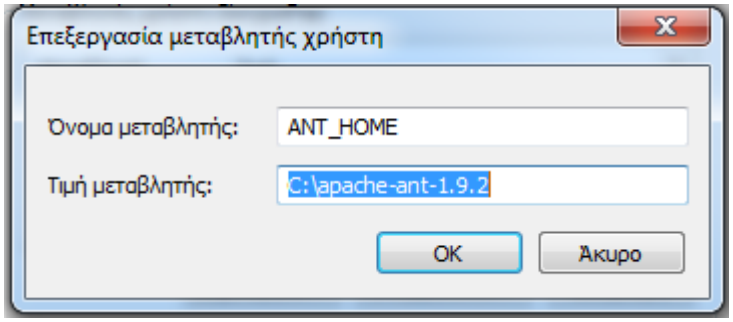
7.7 Variable name: **JAVA_HOME**

Variable value: **C:\Program Files\Java\jdk1.7.0_25**



7.8 Variable name: **ANT_HOME**

Variable value: **C:\apache-ant-1.9.2**



> **Note:**
>
> The "." after --path must be included in the command.

8. Execute the following command to update the path:

/cygdrive/c/users/GiorgosPap/Desktop/adt-bundle-windows-x86_64-20130729/sdk/tools/android.bat update project --path .

```
GiorgosPap@GiorgosPap-PC /cygdrive/c/users/GiorgosPap/Desktop/tess-two-master/tess-two
$ /cygdrive/c/users/GiorgosPap/Desktop/adt-bundle-windows-x86_64-20130729/sdk/tools/android.bat update project --path .
Error: The project either has no target set or the target is invalid.
Please provide a --target to the 'android.bat update' command.
It seems that there are sub-projects. If you want to update them
please use the --subprojects parameter.
```

If shown the above error, then you should specify a target, adding **-t <android_id>** after project as shown below:

```
GiorgosPap@GiorgosPap-PC /cygdrive/c/users/GiorgosPap/Desktop/tess-two-master/tess-two
$ /cygdrive/c/users/GiorgosPap/Desktop/adt-bundle-windows-x86_64-20130729/sdk/tools/android.bat update project -t 2 --path .
Error: Target id '2' is not valid. Use 'android.bat list targets' to get the target ids.
```

Now if the above error occurred, you should get the android target ids. To get them execute:

/cygdrive/c/users/GiorgosPap/Desktop/adt-bundle-windows-x86_64-20130729/sdk/tools/android.bat list targets

```
GiorgosPap@GiorgosPap-PC /cygdrive/c/users/GiorgosPap/Desktop/tess-two-master/tess-two
$ /cygdrive/c/users/GiorgosPap/Desktop/adt-bundle-windows-x86_64-20130729/sdk/tools/android.bat list targets
Available Android targets:
----------
id: 1 or "android-18"
     Name: Android 4.3
     Type: Platform
     API level: 18
     Revision: 1
     Skins: HVGA, QVGA, WQVGA400, WQVGA432, WSVGA, WVGA800 (default), WVGA854, WXGA720, WXGA800, WXGA800-7in
     ABIs : armeabi-v7a
```

As we see above, the available **id** is **1**. So we execute:

/cygdrive/c/users/GiorgosPap/Desktop/adt-bundle-windows-x86_64-20130729/sdk/tools/android.bat update project -t 1 --path .

9. The final command that we have to execute is:

ant release

```
GiorgosPap@GiorgosPap-PC /cygdrive/c/users/GiorgosPap/Desktop/tess-two-master/te
ss-two
$ ant release
-bash: ant: command not found
```

If the above error occurred, you should execute the 2 next commands:

export ANT_HOME=/cygdrive/c/apache-ant-1.9.2

(which assumes you've unzipped Ant into `C:\apache-ant-1.9.2`.) Then:

export PATH=$ANT_HOME/bin:$PATH

To see if all went well execute:

ant –version

It should display the version:

Apache Ant version 1.9.2 compiled on July 8 2013

```
GiorgosPap@GiorgosPap-PC /cygdrive/c/users/GiorgosPap/Desktop/tess-two-master/te
ss-two
$ export ANT_HOME=/cygdrive/c/apache-ant-1.9.2

GiorgosPap@GiorgosPap-PC /cygdrive/c/users/GiorgosPap/Desktop/tess-two-master/te
ss-two
$ export PATH=$ANT_HOME/bin:$PATH

GiorgosPap@GiorgosPap-PC /cygdrive/c/users/GiorgosPap/Desktop/tess-two-master/te
ss-two
$ ant -version
Apache Ant(TM) version 1.9.2 compiled on July 8 2013
```

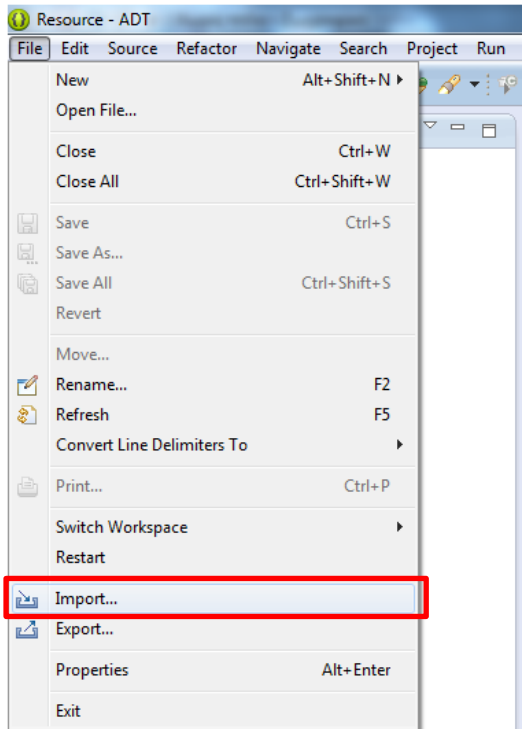
After all above type:

ant release

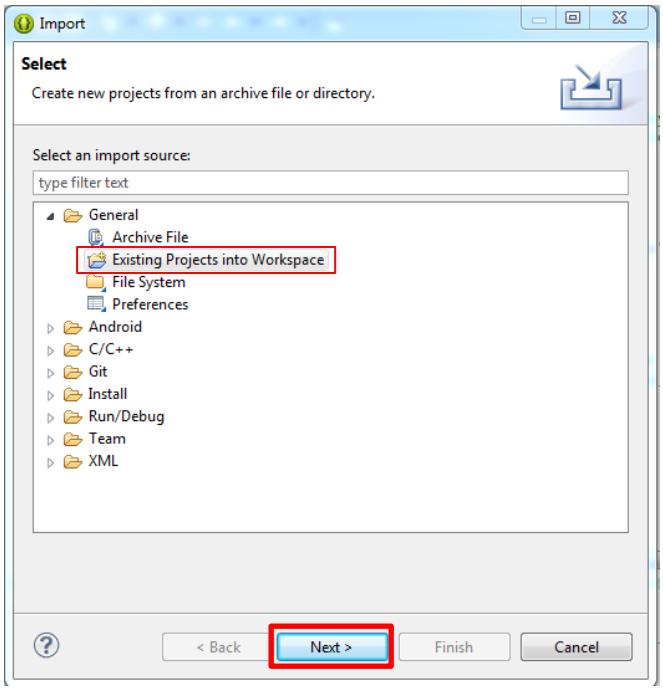
Import Tesseract library and a Simple OCR example in ADT

1. Go to Desktop >> adt-bundle-windows-x86_64-20130729 >> eclipse and run **eclipse.exe**

2. Go File >> Import…

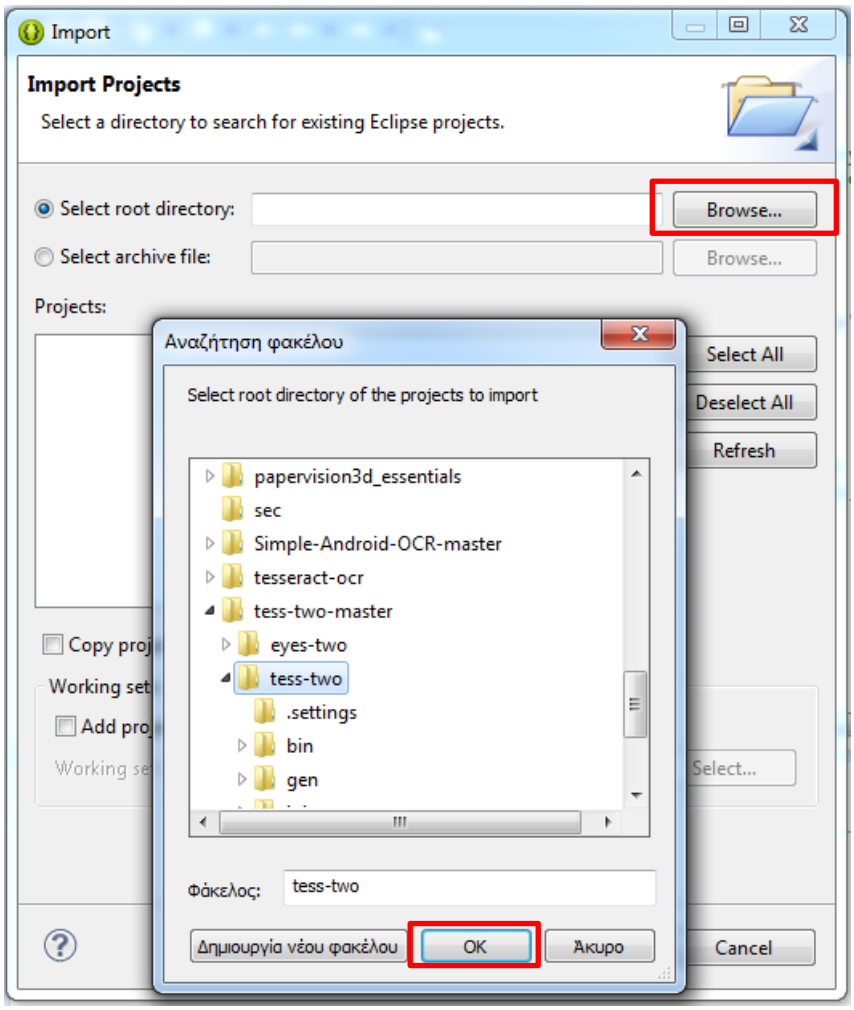


3. From file General select **Existing Project into Workspace** and click **Next**

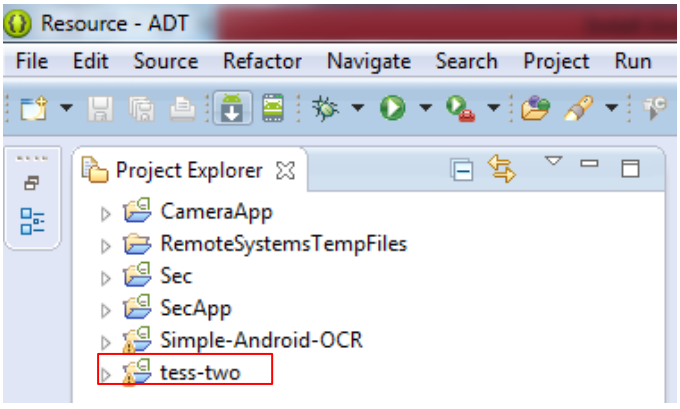


4. In the next window click **Browse…** in *Select root directory*, and then find **tess-two** file and click **OK**.

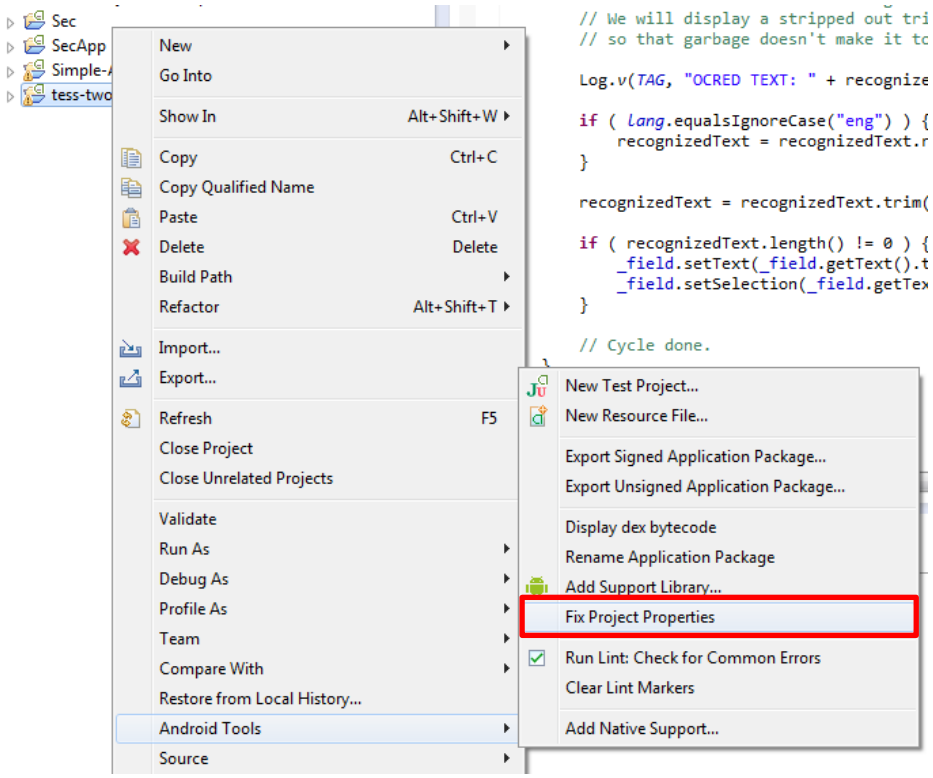
(Mine is in `C:\Users\GiorgosPap\Desktop\tess-two-master\tess-two`)



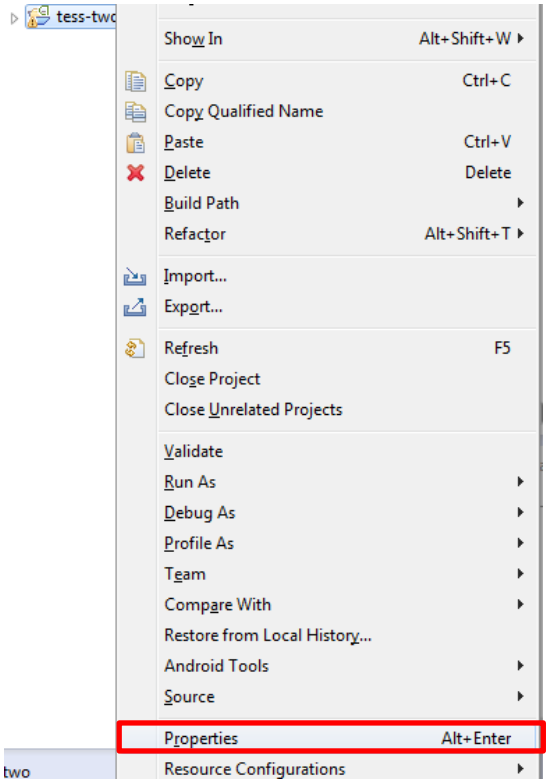
5. Click **Finish** and now you can see it on Project Explorer



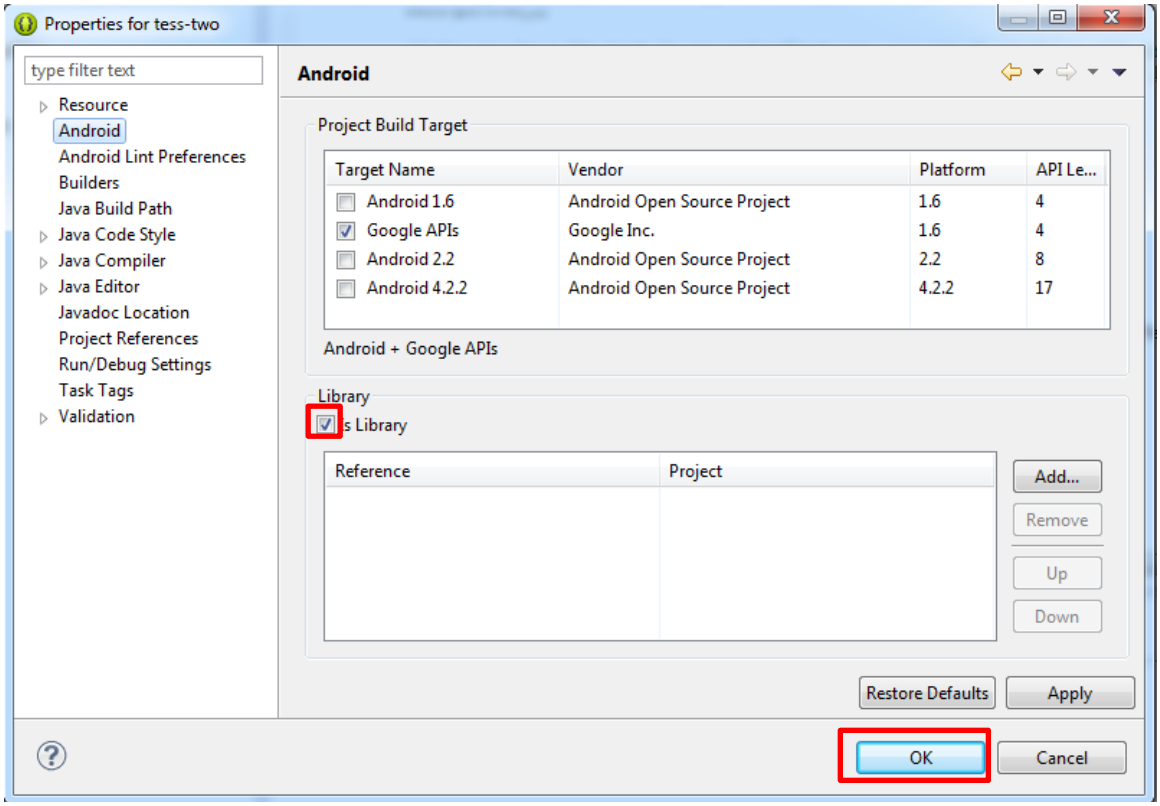
6. Right click on the **tess-two** project >> Android Tools >> Fix Project Properties.



7. Again Right click on the **tess-two** project >> Properties

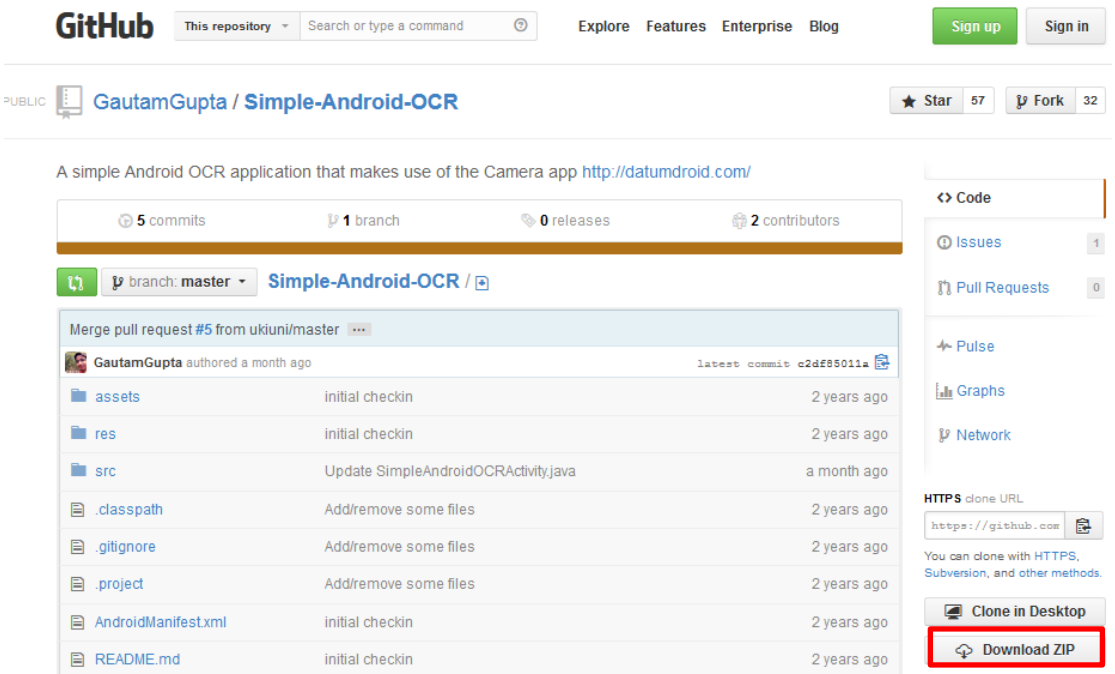


8. Select Android >> Check Is Library and click **OK**



9. Download the simple OCR android app from :

https://github.com/GautamGupta/Simple-Android-OCR



10. Extract it on desktop (Mine is in `C:\Users\GiorgosPap\Desktop\Simple-Android-OCR-master`)

11. Once extracted on desktop and you have closed ADT:

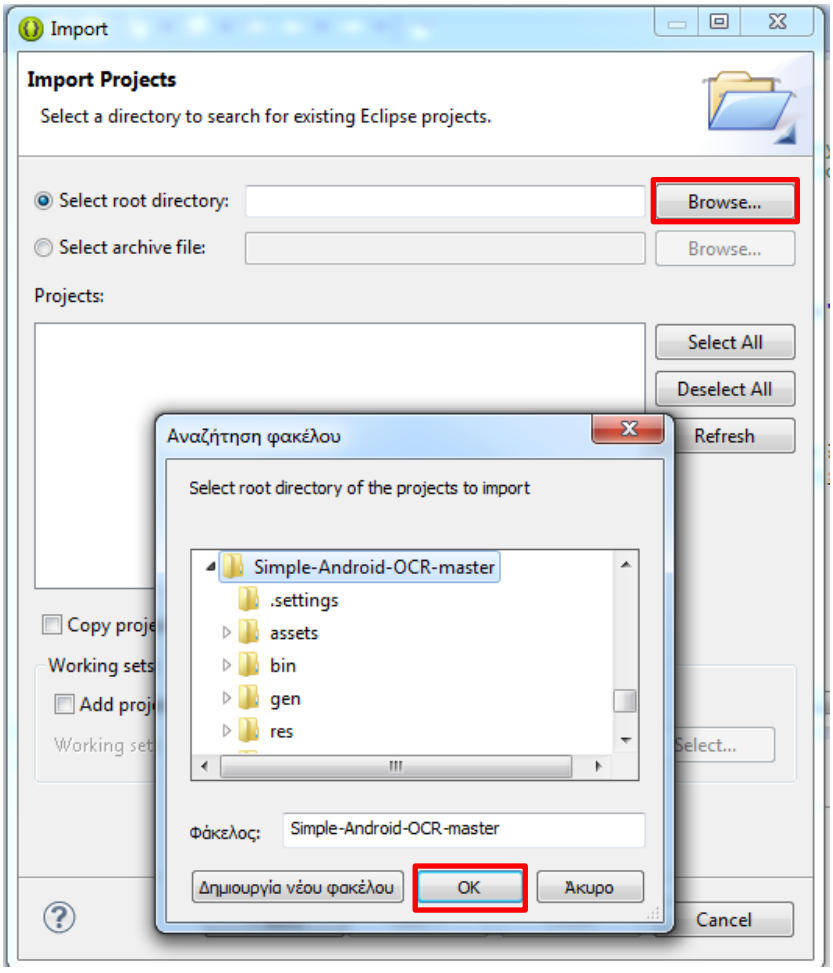
    Go again to Desktop >> adt-bundle-windows-x86_64-20130729 >> eclipse and run **eclipse.exe**

    Go File >> Import…

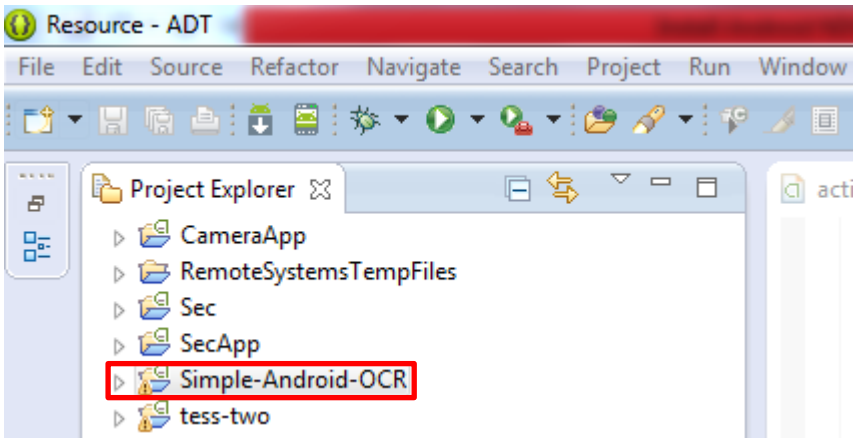
    From file General select **Existing Project into Workspace** and click **Next**

12. In the next window click **Browse…** in *Select root directory*, and then find **Simple-Android-OCR-master** file and click **OK**.

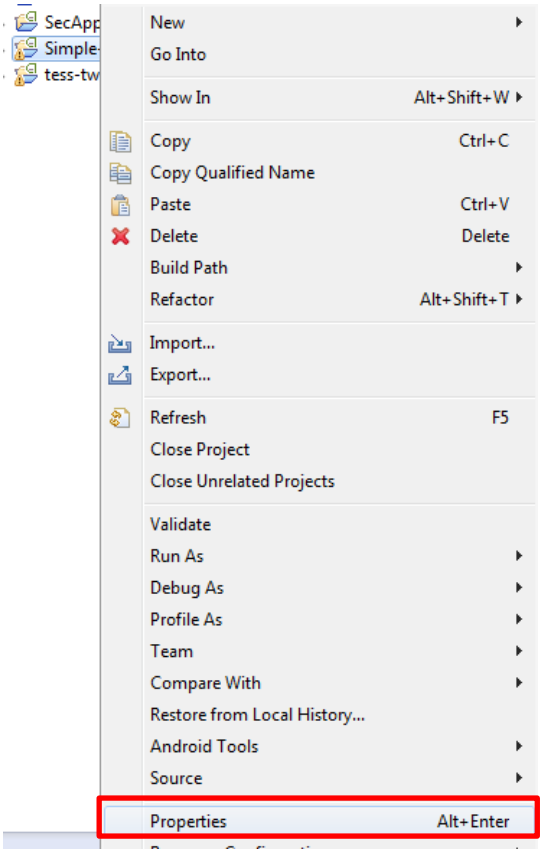
(Mine is in `C:\Users\GiorgosPap\Desktop\Simple-Android-OCR-master`)



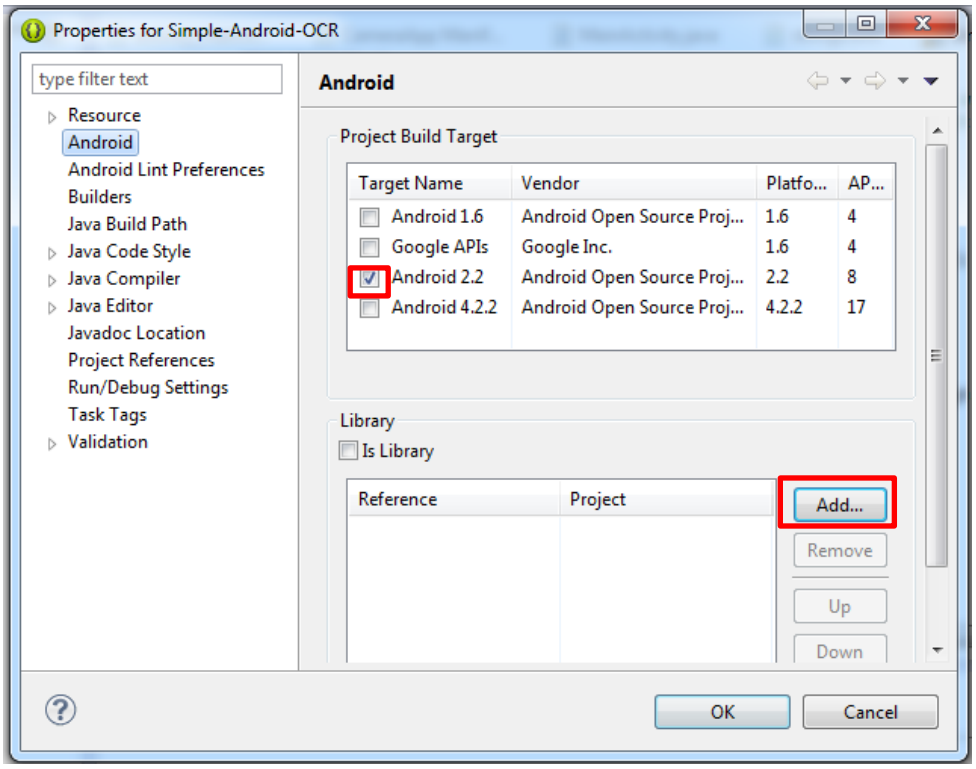
13. Click **Finish** and now you can see it on Project Explorer



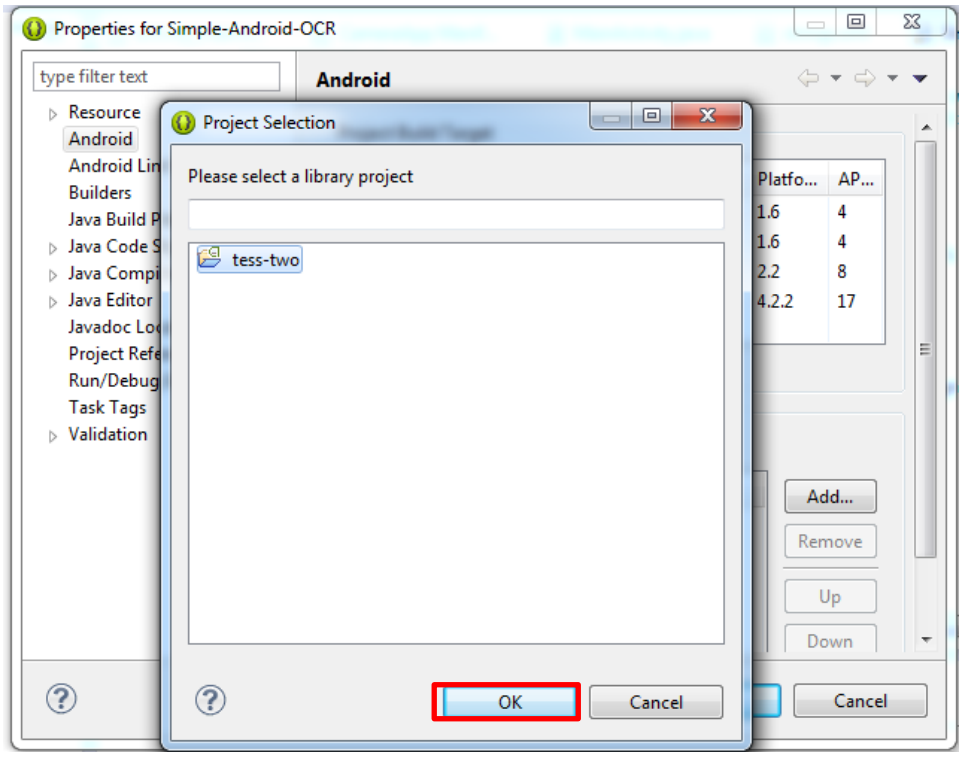
14. Right click on the **Simple-Android-OCR** project >> Properties



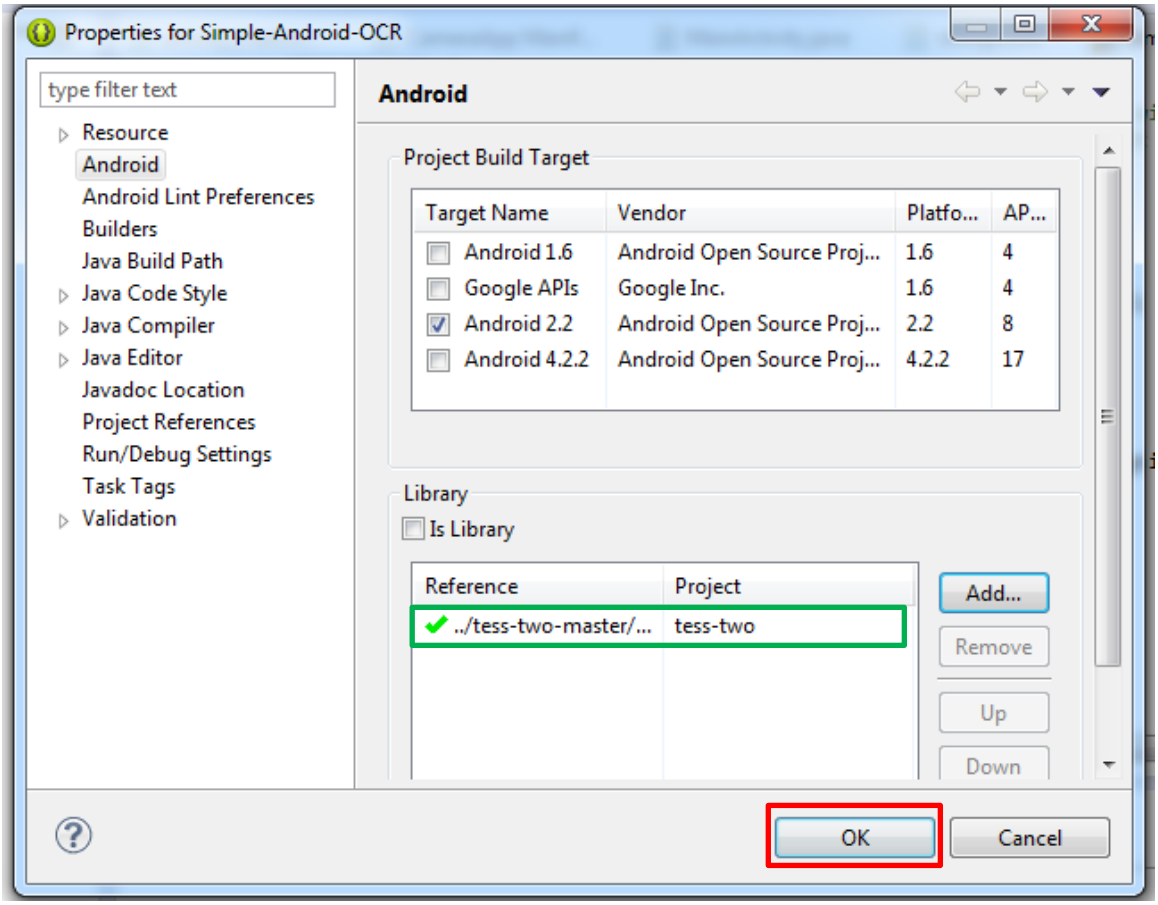
15. Select Android, check Project Build Target (I chosed Android 2.2, because i tested on an Android 2.2 smart-phone) and click **Add…**



16. In the *Project Selection* window select **tess-two** library and click **OK**



17. Now you can see the **imported library**. Click **OK**



18. Finally you can run the App that uses the **Tesseract** library for the OCR operation.

References

http://code.google.com/p/tesseract-ocr/downloads/list

http://www8.hp.com/gr/el/home.html

http://gaut.am/making-an-ocr-android-app-using-tesseract/

http://kurup87.blogspot.gr/2012/03/android-ocr-tutorial-image-to-text.html

http://www.oracle.com/technetwork/java/javase/downloads/index.html

http://developer.android.com/tools/sdk/ndk/index.html

http://www.cygwin.com/

http://ant.apache.org/

http://stackoverflow.com/questions/3858732/installing-ant-on-cygwin

http://developer.android.com/training/basics/firstapp/index.html

http://stackoverflow.com/questions/5025986/ndk-cygwin-path-specific-issues

http://stackoverflow.com/questions/11370220/cygwin-cannot-find-android

http://stackoverflow.com/questions/12749382/android-ndk-android-command-not-found