

UNIVERSITÀ DEGLI STUDI DI BRESCIA
FACOLTÀ DI INGEGNERIA



CORSO DI LAUREA IN INGEGNERIA INFORMATICA
TESI DI LAUREA SPECIALISTICA

***ANALISI DI FATTIBILITÀ ED IMPLEMENTAZIONE
DI UN SISTEMA DI ROBOTICA COGNITIVA
PER COMPITI DI NAVIGAZIONE***

*(Analysis of preconditions and implementation of a
Cognitive Robotic System for navigation tasks)*

Relatore:

Ch.mo Prof. Riccardo Cassinis

Correlatore:

Ch.mo Prof. Marco Ragni

Laureando:

Francesco Bonfadelli

Matricola 83174

Anno Accademico 2012-2013

Contents

1. State Of The Art	4
1.1. Cognitive Architecture	4
1.2. ACT-R	5
1.3. OpenCV	7
2. Requirements	9
2.1. Functional Requirements	9
2.2. Non Functional Requirements	9
3. Development Process	10
3.1. SCRUM	10
3.2. Working Instruments	10
4. Design	11
4.1. The Whole Software	11
4.2. The Class Hierarchy	11
5. Implementation	i
References	ii

List of Figures

1.1. <i>Structure of Act-r.</i>	7
4.1. <i>Class diagram of the whole software</i>	11
4.2. <i>Class hierarchy of the recognized objects</i>	12

1. State Of The Art

This chapter, after having introduced the concept of cognitive architecture, describes the main working instrument used for this work: ACT-R, the cognitive architecture and OpenCV, the computer vision library.

1.1. Cognitive Architecture

A cognitive architecture is the implementation through computer simulation softwares of a theory about human cognition. The theory generally relies on a wide selection of human experimental data. The design of these architectures tries to simulate human intelligence in a humanlike way.

"A cognitive architecture is not a single algorithm or method for solving a problem; rather, it is the task-independent infrastructure that brings an agent's knowledge to bear on a problem in order to produce behavior" [JEL12]. A cognitive architecture alone can not solve any problem. In order to be able to do it, it needs to be supplied with the knowledge to perform it. The combination of an architecture and the knowledge necessary to solve a specific task is called *model*. It is possible to create many models able to solve the same task. The specific knowledge is determined by the *modeler* [Sea02].

Another important feature of cognitive architectures is that they give the possibility to make comparison between the sequence of actions produced by the model and the ones produced by human beings when they try to solve the same task. In addition, the measured quantities can be not only qualitative ones, like the correctness of the goal, but also quantitative, as for example, the time necessary to complete a task. This gives the models the possibility to produce execution times, error rates and learning curves. Comparisons with human performances can be useful to evaluate the quality of model [Sea02].

As the term *infrastructure* suggests, cognitive architectures usually are built aggregating many software modules, most of which represent functions of the human brain. Anyway, there exist other modules, which coordinate the overall functioning and without which the whole architecture could not work. In the following section, which describes *ACT-R*, the

controllare
corret-
tezza
di soft-
wares

procedural module is an example of such modules. In fact, it does not represent a function of the human being, though it is necessary to coordinate the communication between all the other modules.

Some architectures, can, in addition, include some *learning mechanisms*. This can be the attempt to simulate the human memory system, thanks to which the behaviour of a human can be different after having experienced facts or consequences of a specific choice [Sea02].

1.2. ACT-R

ACT-R, that stands for *Adaptive Control of Thought-Rational*, is a cognitive architecture that implements the homonym theory developed by John Robert Anderson, professor of psychology and computer science at Carnegie Mellon University. ACT-R is a software written in Lisp and its models are written in a Lisp-like language. It is thought to have a modular structure so that it can be easily extended. The current version of the software is the 6.0.

mancano
le fonti
praticamente
ovunque

Basic concept of ACT-R's theory

In psychology, memory is defined as the processes by which information is encoded, stored, and retrieved [BEAA09].

ACT-R's most important assumption about knowledge is based on Anderson's theory about memory. Anderson, divides memory into *declarative* and *procedural*. Declarative memory refers to all the memories that can be consciously recalled. This kind of knowledge comprehends facts and knowledge that human beings explicitly know. To call back this kind of information, there must be a conscious process by the human being. For this reason, this kind of memory is also called *explicit*. In contrast, procedural memory refers to all that notions or skills that human beings have but which they learnt in an implicit way. Examples of this knowledge are, for example, driving, reading and writing. In this case, in order to call back this kind of information, the human being does not need a conscious process. That is why this kind of memory is also called *implicit* [And76].

The following example is used to explain better how these two kinds of memory work. When a person starts learning typewriting, an attempt he can make in the beginning is trying to memorize the layout of the keyboard. The knowledge of all the positions of the keys is the declarative memory. After having become a skilled typewriter, the same person will write quickly putting his fingers in the right keys and pushing them in the

correct order without thinking anymore about the positions of the keys on the keyboard. Moreover, if we ask him where the position of a certain character is on the keyboard, he will probably answer that he can not say it without looking at it. This is because, now, he is using his procedural memory [And93].

Architecture

In ACT-R, declarative memory is represented in structures, called *chunks*, and procedural memory is represented as rules, called *productions*. Chunks and productions are the basic building blocks of an ACT-R model [Bot].

The *chunks* are data structures which are defined by their *type* and their *attribute list*, that is a tuple of pairs, each of which is called *slot* and is composed by the *name* of the attribute, which is fixed for a certain chunk, and the *value* of the attribute, which, instead, can assume different values. Each chunk has also a *name* but it is not considered to be a part of the chunk itself, as it does not exist in ACT-R theory. It is used only for convenience to reference the specific chunk when writing models. The chunk-types can be organized into hierarchies [Bot].

The *productions* are the ACT-R equivalent of functions. They define sequences of actions and can be fired only if a set of preconditions is satisfied. They can be represented as if-then rules, where the *if-part* is a set of conditions that must be true for the production to apply and the *then-part* is the action of the production and consists of the operations the model should perform when the production is selected and used [Bot].

All the activities carried out by the brain, i.e. talking or moving, are performed by neurons located close together, in a well defined and limited area of the cortex. Similarly each module of ACT-R represents a function of the human brain. In the ACT-R's architecture the task of the procedural module is to coordinate all the other modules. The exchange of information between the modules is achieved through *buffers*.

Each module is independent from the others. To communicate one to each other they need to pass through the procedural module. As shown in Figure 4.2, a module can have more than one buffer; for example, the aural module has two buffers: the first is the "Aural-location Buffer", whose function is to get the focus about changes happened in the environment; the second is the "Aural Buffer", to get the real information about the environment.

The buffers are the built in interface between modules, used to exchange chunks between them. Every buffer belongs to a module, while a module can also have no buffers. Each buffer can hold one chunk at a time, which is readable by the other modules, the guidelines

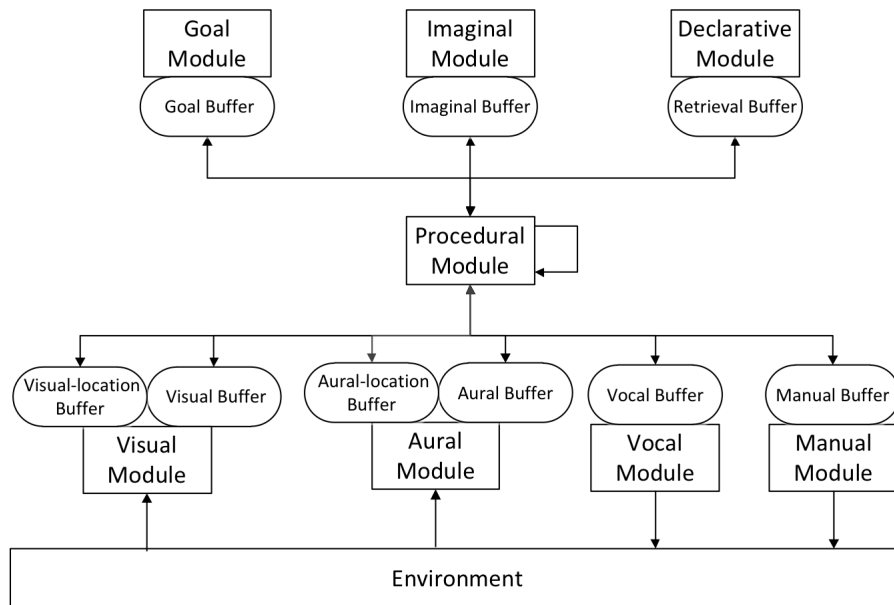


Figure 1.1.: Structure of Act-r.

say that a buffer can be written only by its owner.

The modules usually work in a parallel way, but their interactions, which are limited to a serial exchange of chunks, can be only serial. There are two reasons for this limitation: firstly the structure of the buffers can hold only one chunk; secondly only one production whose preconditions are satisfied can be fired. When the preconditions of more than one production are satisfied, the one to be fired is the one with the higher priority value, called *utility value*. This is a numeric quantity, it can be associated with each production in advance or it can be learned while the model runs.

Applications

Up to now, it has been used in many different areas, the most important of which are problem solving, communication and perception. . For example, it has been shown that ACT-R can successfully predict *Blood Oxygen Level Dependence* of several parts of the brain, through experiments with *Functional Magnetic Resonance Imaging*.

1.3. OpenCV

OpenCV, an abbreviation that stands for *Open Source Computer Vision*, is a computer vision library that was originally developed by Intel and, later on, by Willow Garage. It

is a cross-platform library, released under a BSD license, thus it is free and open source. In the beginning it was developed in C and C++ and afterwards it was expanded by the addition of interfaces for Java and Python. OpenCV is designed for computational efficiency and with a strong focus on real-time applications. The version 2.4 has more than 2500 algorithms. The library has been used in many applications as, for example, mine inspection and robotics [Ope12b]. The following sections contain a brief history of the library and a list of its main features.

History

The OpenCV Project started in 1999 as an Intel Research initiative aimed to improve CPU intensive applications as a part of projects including real-time ray tracing and 3D display walls. The early goals of the project were developing optimized code for basic vision infrastructure, spreading this infrastructure to developers and making it portable and available for free, using a license that let the developers create both commercial and free applications.

The first alpha version was released to the public in 2000, followed by five beta versions between 2001 and 2005, which lead to version 1.0 in 2006. In 2008, the technology incubator Willow Garage begun supporting the project and, in the same year, version 1.1 was released. In October 2009, OpenCV 2.0 was released. It includes many improvements, such as a better C++ interface, more programming patterns, new functions and an optimization for multi-core architectures. According to the current OpenCV release plan, a new version of the library is delivered on a six-months basis. [Ope12a].

Main Features

OpenCV offers a wide range of possibilities. First of all, it provides an easy way to manage image and video data types. It also offers functions to load, copy, edit, convert and store images and a basic graphical user interface that lets the developers handle keyboard and mouse and display images and videos. The library lets manipulate images even with matrix and vector algebra routines. It supports the most common dynamic data structures and offers many different basic image processing functions: filtering, edge and corner detection, color conversion, sampling and interpolation, morphological operations, histograms and image pyramids. Beyond this, it integrates many functions for structural analysis of the image, camera calibration, motion analysis and object recognition. [Aga06].

2. Requirements

2.1. Functional Requirements

The final objective of this work is to make easier the creation of an ACT-R test. At the moment, a psychologist, in order to create a test, has to create two interfaces, one for the user and one for ACT-R. The user interface is quite simple, it is formed by simple shapes that can have different colours and words. The user has to read or watch the screen and then choose between a series of possibilities. In order to be analysed by ACT-R, this interface must have a corresponding one, which is simpler and contains only the most important features of the first one. This "simpler" interface is given as input to ACT-R and is written according to a fixed pattern. The figure below shows an example of the user interface of a test and of its corresponding one.

The purpose of this work is to create a module which is able to receive as input the user interface, recognize the main features in it and create the corresponding ACT-R interface, which must contain only such features. In order to do this, the software must be made by two parts. The first part will analyze the image. It will recognize and distinguish simple shapes, such as squares, rectangles, circles, stars, arrows; distinguish the colours of the objects; determine the position of each object in the image and recognize the text in the image. The second part will create the correspondent interface for ACT-R, thus...

The software is going to be tested with several ACT-R test.

2.2. Non Functional Requirements

3. Development Process

3.1. SCRUM

3.2. Working Instruments

In addition to ACT-R and OpenCV, many other tools have been used. Here follows a list of the most important ones.

- **Eclipse IDE for C/C++ Developers** as *integrated development environment* ¹;
- **Cute** as *unit testing framework* ²;
- **Mylyn** as *task and application lifecycle management* ³;
- **Trac** as *bug tracking system* ⁴;
- **Git** as *version control system* ⁵;
- **Dia** and **cpp2dia** to create UML diagrams ⁶;
- **ZBar** to read QR codes ⁷.

¹More informations at www.eclipse.org

²More informations at <http://cute-test.com>

³More informations at www.eclipse.org/mylyn

⁴More informations at trac.edgewall.org

⁵More informations at git-scm.com

⁶More informations at dia-installer.de and at cpp2dia.sourceforge.net

⁷More informations at zbar.sourceforge.net

4. Design

The following chapter describes the architecture of the software. The software design is supported by diagrams to better explain the contents and the structure of the project itself. The following sections, after having given a brief description of the whole architecture of the software, describe the class hierarchy, the .

terminare.

4.1. The Whole Software

The aim of the .

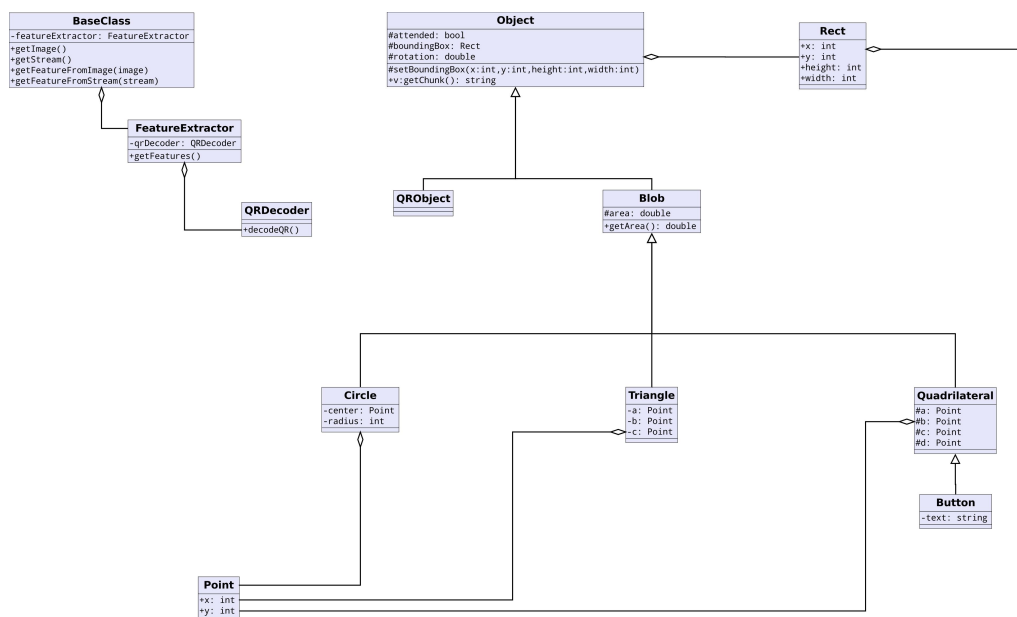


Figure 4.1.: Class diagram of the whole software

cambiare
l im-
agine
e
aggiun-
gere
la de-
scrizione...

4.2. The Class Hierarchy

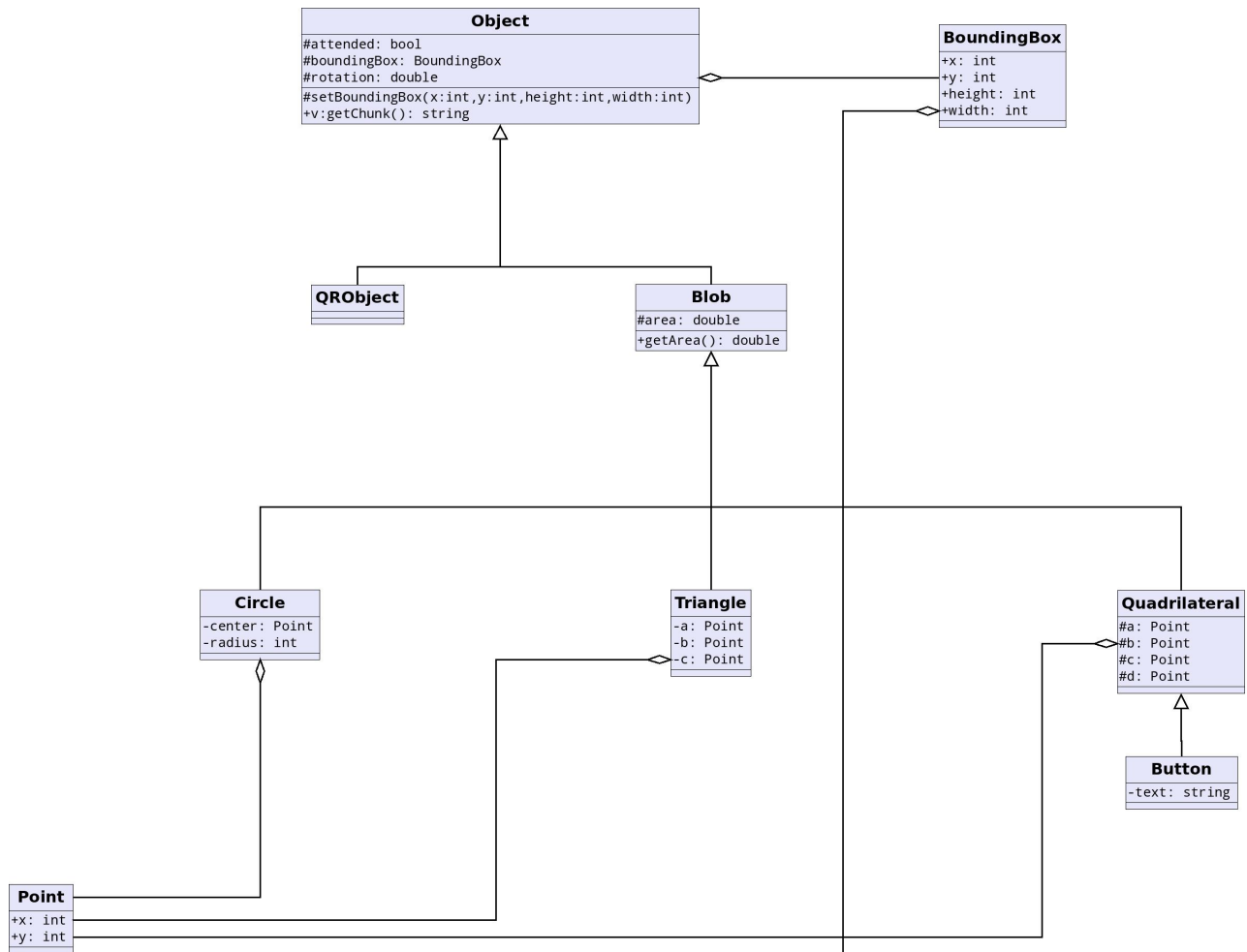


Figure 4.2.: Class hierarchy of the recognized objects

The picture above describes the hierarchy of the classes defined to contain the objects detected in the images. The *Object* class represents a generic object that can be found in an image. Thus, everything which is different from the background can be seen as an instance of the *Object* class. Every object must be bounded by a *bounding box*. This structure is represented by the *BoundingBox* class. All the bounding boxes have rectangular shape. The requirements were such that it was not necessary to define more complicate shapes. The *attended* attribute of the *Object* class is set to *true* if the object has been already returned to ACT-R, otherwise it is set to *false*. The *rotation* is a value that gives the amount of the rotation in the counterclockwise directions starting from the horizontal direction. Besides the generic objects, there are two categories of object that can be found in the processed images, the *QR codes* and the *simple shapes*. A simple shape can be, for example, a circle, a square, a rectangle or a triangle. The QR codes is represented

migliorare questa frase, fa schifo

controllare la correttezza

vedere se il metodo getChunk esiste ancora

in the hierarchy with the *QRObject* class, the generic shape with the *Blob* class. This class has the *area* parameter, which stands for the area of the object in pixel, and the *getArea* method, which returns this value. The *Blob* class is realized by three classes, *Circle*, *Triangle* and *Quadrilateral*. The first one has, as parameters, the *center* and the *radius* of the circle itself. The parameters of the *Triangle* and the *Quadrilateral* classes are respectively three and four points, which represent the vertices of the shape. Notice that the *Quadrilateral* class defines every polygon with four sides and four vertices. Thus, rectangles and squares can be instances of this class. The *Button* class is a realization of the *Quadrilateral* class. This is because, as a requirement, buttons have always rectangular or squared shapes. The additional information they add is the *text*, that is a simple message that is always present in a button. It is represented by the *text* attribute. The *Point* class identifies a generic point in the bidimensional space. Its parameters, *x* and *y*, are the two coordinates in the plane.

... ag-
giun-
gere,
correg-
gere o
finire

5. Implementation

References

- [Aga06] Gady Agam. Introduction to programming with OpenCV. Technical report, 2006.
- [And76] J.R. Anderson. *Language Memory Thought*. The Experimental Psychology Series/ Arthur W. Melton consulting ed. Taylor & Francis Group, 1976.
- [And93] J.R. Anderson. *Rules of the Mind with Mac Dis*. Taylor & Francis Group, 1993.
- [BEAA09] A.D. Baddeley, M.W. Eysenck, M.C. Anderson, and M. Anderson. *Memory*. Taylor & Francis Group, 2009.
- [Bot] Dan Bothell. *ACT-R 6 Reference Manual*.
- [JEL12] John L. Tishman John E. Laird. The soar cognitive architecture. *AISB Quarterly*, 134:1, 2012.
- [New94] Allen Newell. *Unified Theories of Cognition*. Harvard University Press, 1994.
- [Ope12a] Opencv change logs, 2012. Available on line.
- [Ope12b] Opencv web page, 2012. Available on line.
- [Sea02] Andrew Sears. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. Lawrence Erlbaum, 2002.