* Project Presentation

The goal of this project was to implement a dialogue system for controlling a robot living in a virtual block world. The robot we implemented can move around objects of different forms, colours and sizes, and it can answer questions about the world and ask for clarification whenever it finds request ambiguous.

This has been performed using GrammaticaFramework (gf) for the grammar and java/prolog for the implementation. The world is described using a json file.

Below is a macro explanation of how works the global application.

On client side:

The user types a query as an input, which is handled with an ajax call in jquery. It is this particular ajax call that will execute our main function in prolog.

On the server side:

The ajax call contains several attributes: world, objects, holding, state, utterance.

All those information are stored in a JSON file that contains the world with the objects and the request. This JSON file is then read and handled using prolog.

The first step is to transform the user input into a goal that is understandable by the interpreter. This is the role of the parser.

Then the parsed query is sent to the interpreter which translates it to an understandable query by the planner we implemented.

Once the goal has been defined, we need to find a resolution plan. All the functions we need to perform object moves have been implemented in the planner. This will give us a list of all the actions we need to perform to go from the original world to the requested one.

Then this list of actions is handled by the solver which is going to produce a list of all the instructions required for the graphics rendering. The output is then sent back on the client side for the graphic animation of all the steps.

The schema below gives a macro view of the steps followed in conducting any application:



* The World

The world is represented by a floor on which several objects (of different forms, colors and sizes) can lay. The objects can stand in/on each other (if it is permitted by the world’s laws).

The goal of this project is to move around the objects, according to the request of the user, through a robot arm which can pick up and put down objects.

The floor is divided in n spaces, meaning that there is room on the floor for no more than n objects at the same time. Each space is represented as a column, so the world can be described as a list of n columns of object stake on each other.

The world we implemented can contain all the objects of different forms, colours and sizes listed on the course homepage, plus the possibility to define an object of medium size (in order to create a more complex/realistic world).

Below the list of forms, colours and sizes available in the world used in the project:

Forms: brick, plank, ball, pyramid, box, table.

Colours: red, black, blue, green, yellow, white.

Sizes: large, medium, small.



A JSON file is used to describe the world as a list of object columns.

* Improvements of the grammar

We started with the initial grammar (given at the beginning of the project) that we have completed and improved.

The major improvement we implemented is the possibility for the user to ask 3 new kinds of questions: where, what and count.

* Where

The “where” question allows the robot to answer questions such as “where is the white big ball?” Or “where are the boxes?”. The synonym for the “where” questions are the words: Find, Where is and Where are.

* What

With this “what” question, the robot can answer questions such as “what is under the red box?” Or “what are the object in the world?”. The synonym for the “what” questions are the words: What is and What are.

* Count

This last new possible question make the user able to get an answer to requests such as “count the boxes in the world.” or “how many balls are in the world?”. The synonyms for the “count” questions are: Count and How many.

* Other improvements

Another major improvement we made from the original grammar is the possibility for the user to request actions and/or ask questions about stacks instead of simply the whole world. The user can then perform request like “what are the objects in stack 2?” Or “count the small blue balls in stack 0.” Or “what are the objects on the right of stack 2?”

We also added 2 minor improvements to the original grammar: the alias “world” for “all the stacks” and the possibility for the user to put question marks at the end of his questions without having the robot answering with an error message.

Of course all those modifications regarding the grammar had to be echoed to the other layers of the application (parser and planner).

* Interpreter

After the parser (and grammar) layer we then proceed to the interpreter layer.

The interpreter aims to translate the output of the parser into something understandable by the planner. For that it uses a tree to evaluate user queries such as: "take(relative\_entity(the,object(small,black,ball),relative(leftof,basic\_entity(the(big,-,-)))

We made a set of rules to satisfy every node in the tree, e.g.

Tree = a(b(c(d)))

interp(a(X),Y):-interp(X,Y), something1(Y).

interp(b(X),Y):-interp(X,Y), something2(Y).

interp(c(X),Y):-interp(X,Y), something3(Y).

interp(d,Y)    :-, something4(Y).

Y is one which satisfies every node of the tree.

* Interpretation rules

More explicitly we have implemented the following rules:

0 interpret(object(Type,Size,Color), World, @(null), Objects, SelectedObject) :-

1 interpret(object(Type,Size,Color), World, Holding, Objects, SelectedObject) :-

2 interpret(basic\_entity(any,X), World, Holding, Objects, any(SelectedObject)) :-

3 interpret(basic\_entity(the,X), World, Holding, Objects, [SelectedObject]) :-

4 interpret(basic\_entity(all,X), World, Holding, Objects, SelectedObject) :-

5 interpret(relative\_entity(any,X, Relation), World, Holding, Objects, any(SelectedObject)) :-

6 interpret(relative\_entity(all,X, Relation), World, Holding, Objects, SelectedObject) :-

7 interpret(relative\_entity(the,X, Relation), World, Holding, Objects, [SelectedObject]) :-

8 interpret(relative(beside,X), World, Holding, Objects, SelectedObject) :-

9 interpret(relative(leftof,X), World, Holding, Objects, SelectedObject) :-

[…]

15 interpret(relative(inside,X), World, Holding, Objects, SelectedObject) :-

16 interpret(absolute(beside,basic\_stack(N)), World, Holding, Objects, SelectedObject) :-

[…]

21 interpret(absolute(inside,world), World, \_Holding, \_Objects, SelectedObject) :-

22 interpret(floor, \_World, \_Holding, \_Objects, floor).

23 interpret(take(X), World, Holding, Objects, take(SelectedObject)) :-

[…]

61 interpret(what(absolute(ontop, basic\_stack(N))), World, Holding, Objects, whatontopstack([N])).

0-1 : Get a satisfying object description

2-4 : Handle the any or all cases for basic entities

5-7 : Same as 2-4 but for relations as well

8-15 : Get an object which satisfy a relation to "object X"

16-21: Get an object which satisfy a relation to stack

22 : floor is floor

23-61: Process output to goal

Further we have some rules to see if an object is e.g. beside any other object:  
isbeside(X,Y,World) :-

    member(ColS,World),member(X,ColS), nth0(IdxS,World,ColS),

    member(ColR,World),member(Y,ColR), nth0(IdxR,World,ColR),

    (IdxS is IdxR-1;IdxS is IdxR+1).

Later we added similar rules for the questions: count where and what. We also have helper functions such as isontop(X,Y,List) which, given a list, determines the truth value of X being on top of Y.

* Quantifiers

The quantifiers allow the robot to handle query such as “put any ball in the red box”. The quantifier function uses cuts to choose one possible action, when several possibilites are available to the robot (for example, any([a,b])).

Listed below are a few examples of the cuts performed by the function:

|  |  |
| --- | --- |
| **Original query** | **Interpreted query** |
| **movebeside([a,b],[c,d])** | **movebeside([a,b],[c,d])** |
| **movebeside(any([a,b]),any([c,d]))** | **movebeside([a],[c])** |
| **movebeside([a,b],any([c,d]))** | **movebeside([a,b],[c]))** |
| **movebeside(any([a,b]),[c,d])** | **movebeside([a],[c,d])** |

Here is an example of how the cut is used in our implementation:

[…]

* Planner
* Terminal cases

The first step in our implementation was to implement terminal case, i.e. the move of an object that does not require another modification of the world. A terminal case could be “Put the black ball beside the yellow box“.

The planner takes a query as an input and build list of triplets, where each triplet represent a feasible action. There are 3 possible terminal moves, each represented by a specific triplet:

|  |  |
| --- | --- |
| **Motion** | **Triplet** |
| **Pick** | **[K1,-1,move]** |
| **Drop** | **[-1, K2,move]** |
| **Move** | **[K1,K2,move]** |

Where K1 is the position of the object to pick and K2 the position where the object has to be drop.

* Complex cases

Once the terminal moves have been implemented, we can proceed to the implementation on more complex cases. Such as moving an object that is under one or more objects.

The output of our plan function for complex cases is a list of triplets/actions (pick, drop, and move) that lead to the Goal.

In order to create this list of actions, the plan function is call recursively. The list of actions to reach the Goal is then a concatenation of all the actions to get from the original world to the requested world. This can also be view as a concatenation of all the branch of the decision tree used by prolog to find a solution to find a way to the requested state of the world (see below).

Here is an example:

The user query is “Put the blue box in the red box”.

The plan function performs an in-depth research in the prolog decision tree.

|  |  |
| --- | --- |
|  |  |

So the output of the plan looks like:

Plan = [[4,3,move],[4,2,move]]

* Heuristic

When the robot as to deal with not straight forward case, such as moving object that are not on top of a stack or moving a ball on top of a table, we need the heuristic in order to optimize handling of those case and to move objects in a better way.

The first heuristic we implemented was really simple. Basically, it was doing a depth first search and simply checking that we are not looping on an already encountered world’s state. So it was just trying to go as far as possible in a branch of the decision tree, checking if it was meeting the requirements of the query and, if not, backtracking to another branch of the tree.

To check be sure that the robot is not looping (if the resulting state of the world has not been encountered before), we incremented a list of all the encountered states of world we reached and checking if the state we are moving to is not in the list.

The new heuristic we implemented allows the robot to handle complex cases in a smarter way than the basic implementation we used.

For example, for the action “take”, the heuristic function as explain in the below scheme:



And for the move left:



* Ambiguities handling

The ambiguities handling allow the robot to handle cases where the request from the user is not clear enough. For example, in picture 1, “what is under the box?” would lead to an ambiguity, since the robot would not know if the user want to know what is under the red box or the blue box or the yellow box.

In case of ambiguity, the robot asks the user for a precision. The user then has to precise the object he is referring to. In our example, it could be “the small blue box”.

Once the user has precise the object, the robot will get this information and try to match with the entire possible goals it has identified. It then selects the unique matching solution (if it exists).

If ambiguities still occur the robot then return an error. There is no second question ask to the user since prolog does not handle while loop.

* Output

The particularity of our robot is that the possible outputs are not only an action (such has moving objects around) but also possibility some verbal information such as the number of elements stack on top of each other.

* Pierre BOUTRY contribution to the project

During this project I worked on several layers of the application. I started by writing the planner starting by the basic cases, then the complex cases and finally adding a smart heuristic to improve the treatment of robot.

I fixed a few minor bugs on the interpreter, based on the work done by Dan DOLONIUS on this part.

I added the possibility to define medium size object in JSON. This allows creating more complex and *realistic* world.

I implemented together with Julien MICHELET, the solver layer of the application. The function of the solver is to return a list understandable in Javascript. This list will be executed in Javascript on client side to perform the graphic animations of the robot as well as the verbal answers of the robot.

I also added the voice recognition brick to the application, so that the user can use a mic rather than a keyboard to send his query to the robot.

Finally, I work on how to handle “inversed list” on prolog side, since JSON and prolog seem to read list in an inversed way.

* Dan DOLONIUS contribution to the project

During this project, I was in charge of writing and implementing the interpreter layer, which is in charge of taking the output of the parser and translating it to a query understandable by the planner. The interpreter has been modified several time during the project, either to handle more complicated cases or to allow the user to perform queries of type “where”, “what” or “count”.

* Julien MICHELET contribution to the project

My first task on this project was to implement the improvement of the basic grammar. I added the possibility to ask new questions to the robot such as “what is under this object”, “where is this object” or “count the number of this specific object on this stack”.

Once it has been done, I had to improve the interpreter and planner so they can handle this new grammar and be able to answer the new type of questions offer to the user.

I implemented together with Pierre BOUTRY, the solver layer of the application. The function of the solver is to return a list understandable in Javascript. This list will be executed in Javascript on client side to perform the graphic animations of the robot as well as the verbal answers of the robot.

I also worked on the implementation of the ambiguities handling. For that I added the possibility for the robot to ask the user for complementary information regarding his query.

Finally, I implemented the last version of the CanBeOn function. This function is used to check if an object can be put on top of a stack (possibly empty).

* Emeric ROVERC’H contribution to the project

For this project, I first try to define and present a clear macro presentation of the project. This would be used to present our project as well as facilitate the communication and interaction of the different people working on different parts of the project. I created the global application scheme below:



Then, once we had the first working version of the application, I started to gather the information from the teams working on each parts of the application. All the gathered information had then to be reformulated to be understandable by people outside of the project. Moreover, this allowed the team to stand back on the ongoing project and to refocus the efforts on the most important points of the project, rather to persist on none crucial features/actions.

In parallel with those reporting/macro tasks, I also helped the teams (in charge of distinct parts of the application) that were sometimes lacking of time/manpower to stay on track with the schedule we had defined.

For example, I helped Pierre BOUTRY with the implementation of the *smart* heuristic that has been presented in the report above. I also finished to write some prolog function such has CanBeOn, which is used to check if an object can be put on top of a stack (possibly empty), or RetrieveGoalElements, which allow to get from the interpreter output what object we want to interact with, what action we want to perform and we the object should be placed.

I also performed the benchmark of the different implementations we have created during this entire project. This benchmark was based on the use of different worlds (small, medium and complex) as well as different heuristic we have implemented.