|  |
| --- |
| Chalmers university of technology |
| Artificial intelligence project report |
| Group 20 |
|  |
| **Pierre BOUTRY, Dan DOLONIUS, Julien MICHELET, Emeric ROVERC'H** |
| **21/05/2014** |

|  |
| --- |
| [Tapez le résumé du document ici. Il s’agit généralement d’une courte synthèse du document. Tapez le résumé du document ici. Il s’agit généralement d’une courte synthèse du document.] |



Table of content

[1. Project Presentation 2](#_Toc388118226)

[2. The World 3](#_Toc388118227)

[3. Improvements of the grammar 4](#_Toc388118228)

[3.1. Where 4](#_Toc388118229)

[3.2. What 4](#_Toc388118230)

[3.3. Count 4](#_Toc388118231)

[3.4. Other improvements 4](#_Toc388118232)

[4. Interpreter 5](#_Toc388118233)

[5. Planner 7](#_Toc388118234)

[5.1. Terminal cases 7](#_Toc388118235)

[5.2. Complex cases 7](#_Toc388118236)

[5.3. Heuristic 8](#_Toc388118237)

[6. Ambiguities handling 9](#_Toc388118238)

[7. Output 10](#_Toc388118239)

[8. Appendix 11](#_Toc388118240)

# 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 work the global application.

On client side:

The user type 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 sever 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 handle using prolog.

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

Once the goal has been defined, we need to find a resolution plan. All the functions we need to perform objects move have been implemented in the planner. This will give us a list of all the action 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 all the instructions required for the graphics rendering. The output is then sent back on client side for the graphics animations of all the steps.

The schema below give a macro view of the steps of the applications:



# 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 of this space is represented has a column, so that the world can be describe as a list of n columns of objects 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).

Bellow the list of forms, colours and sizes available in “our” world:

Forms: bricks, planks, balls, pyramids, boxes, tables.

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

Sizes: large, medium, small.



This world is described using a JSON file describing the world as a list of columns of objects.

# 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 improvements we implemented are the possibility for the user to ask 3 new types of questions: where, what and count.

## Where

The “where” question allow 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 aim to permit the user to get answer to requests such as “count the boxes in the world.” or “how many ball 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 stack 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 query such as: "take(relative\_entity(the,object(small,black,ball),relative(leftof,basic\_entity(the(big,-,-)))

We make 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 that 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) :-

10 interpret(relative(rightof,X), World, Holding, Objects, SelectedObject) :-

11 interpret(relative(above,X), World, Holding, Objects, SelectedObject) :-

12 interpret(relative(ontop,X), World, Holding, Objects, SelectedObject) :-

13 interpret(relative(ontop,floor), World, \_Holding, \_Objects, SelectedObject) :-

14 interpret(relative(under,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) :-

17 interpret(absolute(leftof,basic\_stack(N)), World, Holding, Objects, SelectedObject) :-

18 interpret(absolute(rightof,basic\_stack(N)), World, Holding, Objects, SelectedObject) :-

19 interpret(absolute(above,basic\_stack(N)), World, Holding, Objects, SelectedObject) :-

20 interpret(absolute(ontop,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)) :-

24 interpret(move(X,relative(beside, Y)), World, Holding, Objects, movebeside(SelectedObject,RelativeObject)) :-

25 interpret(move(X,relative(leftof, Y)), World, Holding, Objects, moveleft(SelectedObject,RelativeObject)) :-

26 interpret(move(X,relative(rightof,Y)), World, Holding, Objects, moveright(SelectedObject,RelativeObject)) :-

27 interpret(move(X,relative(above, Y)), World, Holding, Objects, moveabove(SelectedObject,RelativeObject)) :-

28 interpret(move(X,relative(ontop, Y)), World, Holding, Objects, moveontop(SelectedObject,RelativeObject)) :-

29 interpret(move(X,relative(under, Y)), World, Holding, Objects, moveunder(SelectedObject,RelativeObject)) :-

30 interpret(move(X,relative(inside, Y)), World, Holding, Objects, moveinside(SelectedObject,RelativeObject)) :-

31 interpret(move(X,absolute(beside, basic\_stack(N))), World, Holding, Objects, movebesidestack(SelectedObject,[N])) :-

32 interpret(move(X,absolute(leftof, basic\_stack(N))), World, Holding, Objects, moveleftstack(SelectedObject,[N])) :-

33 interpret(move(X,absolute(rightof,basic\_stack(N))), World, Holding, Objects, moverightstack(SelectedObject,[N])) :-

34 interpret(move(X,absolute(above, basic\_stack(N))), World, Holding, Objects, moveabovestack(SelectedObject,[N])) :-

35 interpret(move(X,absolute(ontop, basic\_stack(N))), World, Holding, Objects, moveontopstack(SelectedObject,[N])) :-

36 interpret(where(X), World, Holding, Objects, where(SelectedObject)) :-

37 interpret(count(X,relative(beside, Y)), World, Holding, Objects, countbeside(SelectedObject,RelativeObject)) :-

38 interpret(count(X,relative(leftof, Y)), World, Holding, Objects, countleft(SelectedObject,RelativeObject)) :-

39 interpret(count(X,relative(rightof,Y)), World, Holding, Objects, countright(SelectedObject,RelativeObject)) :-

40 interpret(count(X,relative(above, Y)), World, Holding, Objects, countabove(SelectedObject,RelativeObject)) :-

41 interpret(count(X,relative(ontop, Y)), World, Holding, Objects, countontop(SelectedObject,RelativeObject)) :-

42 interpret(count(X,relative(under, Y)), World, Holding, Objects, countunder(SelectedObject,RelativeObject)) :-

43 interpret(count(X,relative(inside, Y)), World, Holding, Objects, countinside(SelectedObject,RelativeObject)) :-

44 interpret(count(X,absolute(beside, basic\_stack(N))), World, Holding, Objects, countbesidestack(SelectedObject,[N])) :-

45 interpret(count(X,absolute(leftof, basic\_stack(N))), World, Holding, Objects, countleftstack(SelectedObject,[N])) :-

46 interpret(count(X,absolute(rightof,basic\_stack(N))), World, Holding, Objects, countrightstack(SelectedObject,[N])) :-

47 interpret(count(X,absolute(above, basic\_stack(N))), World, Holding, Objects, countabovestack(SelectedObject,[N])) :-

48 interpret(count(X,absolute(ontop, basic\_stack(N))), World, Holding, Objects, countontopstack(SelectedObject,[N])) :-

49 interpret(count(X,absolute(inside, world)), World, Holding, Objects, countinsidestacks(SelectedObject,N)) :-

50 interpret(what(relative(beside, Y)), World, Holding, Objects, whatbeside(RelativeObject)) :-

51 interpret(what(relative(leftof, Y)), World, Holding, Objects, whatleft(RelativeObject)) :-

52 interpret(what(relative(rightof,Y)), World, Holding, Objects, whatright(RelativeObject)) :-

53 interpret(what(relative(above,Y)), World, Holding, Objects, whatabove(RelativeObject)) :-

54 interpret(what(relative(ontop, Y)), World, Holding, Objects, whatontop(RelativeObject)) :-

55 interpret(what(relative(under, Y)), World, Holding, Objects, whatunder(RelativeObject)) :-

56 interpret(what(relative(inside, Y)), World, Holding, Objects, whatinside(RelativeObject)) :-

57 interpret(what(absolute(beside, basic\_stack(N))), World, Holding, Objects, whatbesidestack([N])).

58 interpret(what(absolute(leftof, basic\_stack(N))), World, Holding, Objects, whatleftstack([N])).

59 interpret(what(absolute(rightof,basic\_stack(N))), World, Holding, Objects, whatrightstack([N])).

60 interpret(what(absolute(above, basic\_stack(N))), World, Holding, Objects, whatabovestack([N])).

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

0-1 : Get object satisfying 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 object which satisfies relation to "object X"

16-21: Get object which satisfies 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. besides 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  "trueness" of X being on top of Y.

## Quantifiers

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

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]) |

# 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

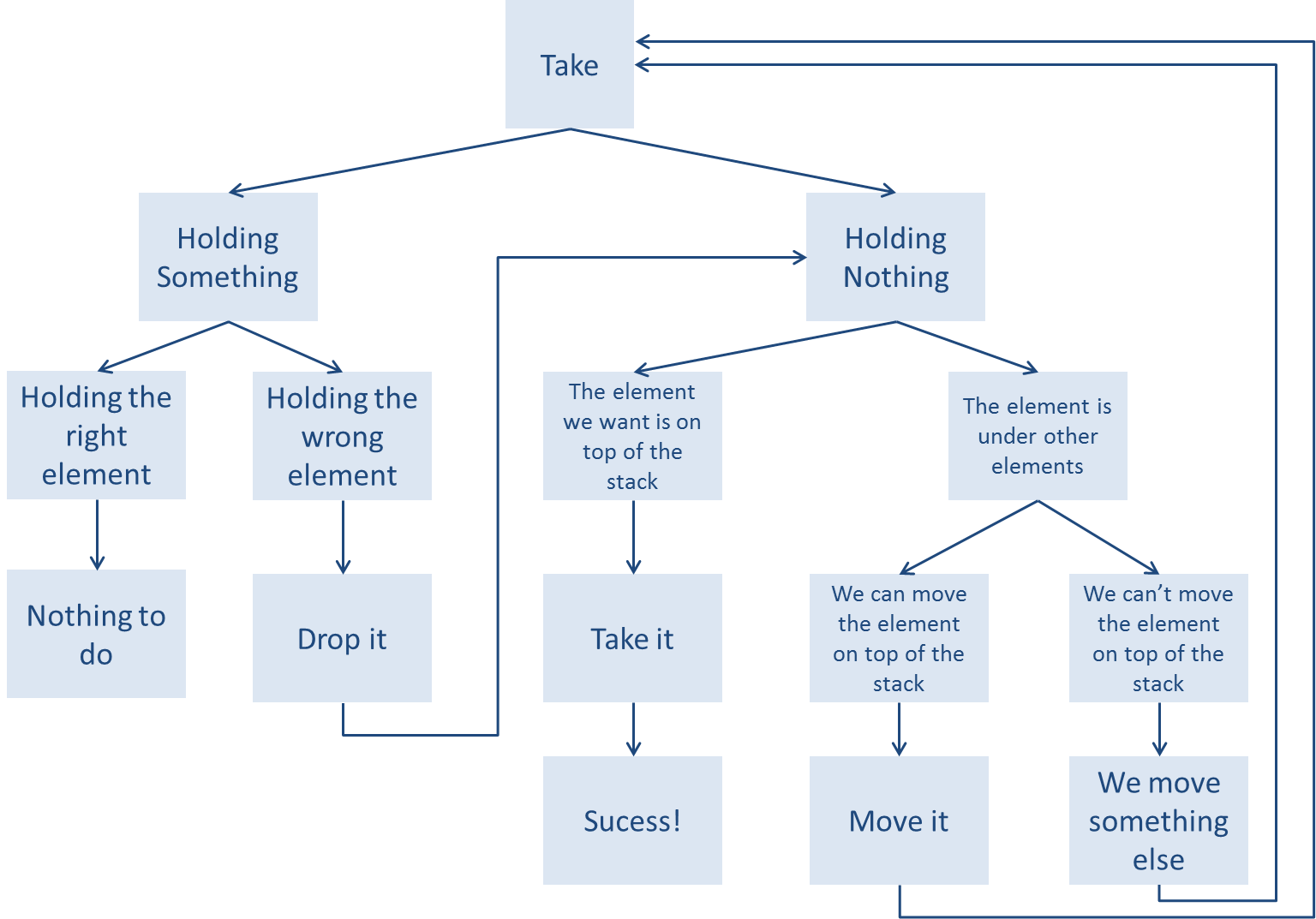
When the robot as to deal with not straight forward case, such as moving object that are not on top of a stack or moving a ball on top of a table, we need the heuristic in order to optimize handling of those case and to move objects in a better way.

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.

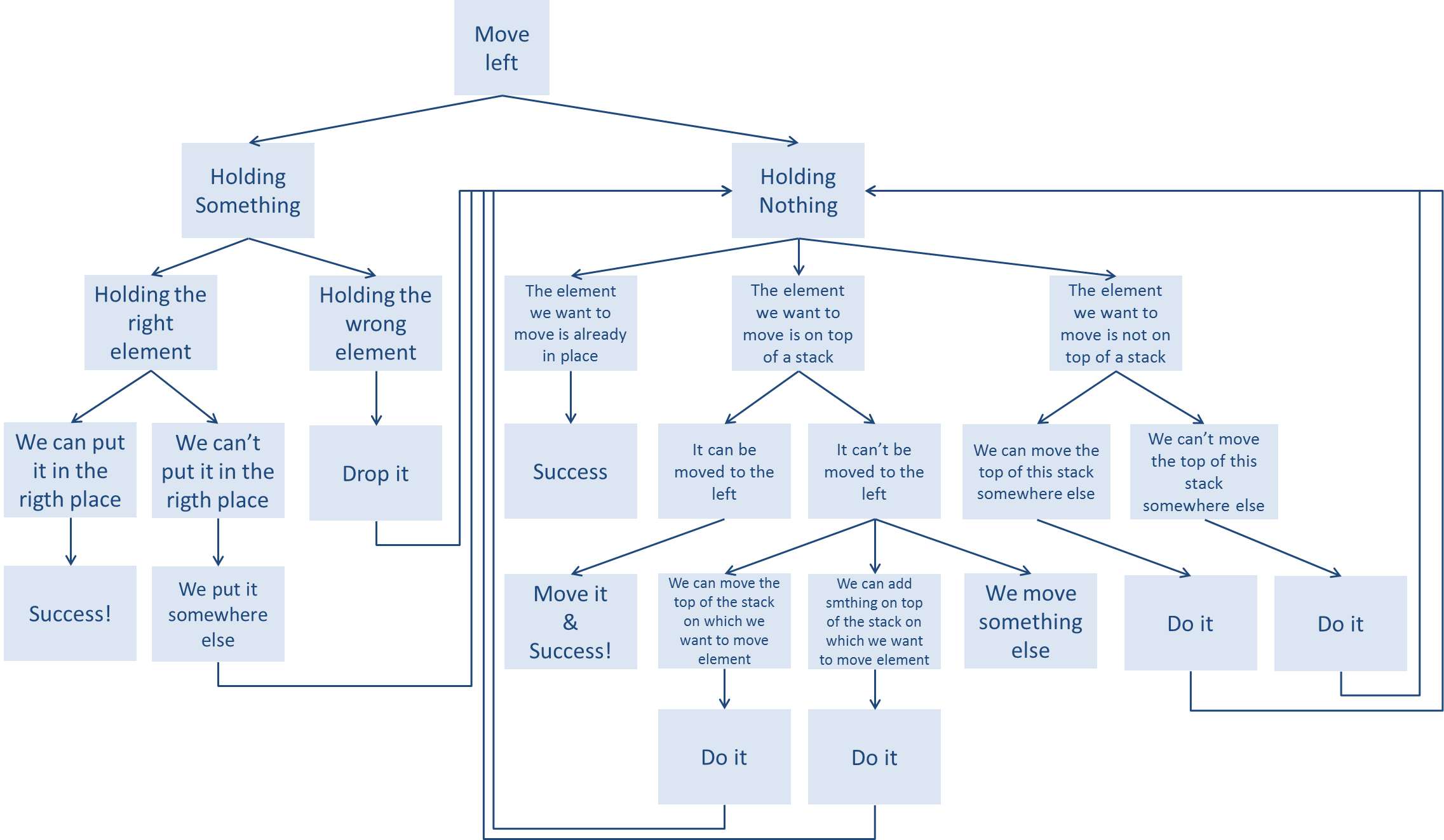
To check be sure that the robot is not looping (if the resulting state of the world has not been encountered before), we incremented a list of all the encountered states of world we reached and checking if the state we are moving to is not in the list.

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 all the possible goal he has identified. It then select the unique matching solution (if it exists).

If an ambiguity still occurs 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.

# Appendix