

Development of an expert system to support a real estate using CLIPS

Department of Computer Science: University of Turin

Authors: Alessandro Saracco, Leonardo Magliolo, Mattia Marra



Artificial intelligence and laboratory

A.Y. 2021/2022

Domain knowledge

The goal of the project was to develop an expert system that would support a real estate agent in his or her work: to suggest to a prospective client a small number of accommodations for sale to visit.

To this end, domain knowledge was modeled by considering that an apartment is generally characterized by a set of properties.

To perform its function, the system asks questions about properties that the desired property owns, and in addition, it also asks questions about the customer's own profiling.

The knowledge gained will be leveraged to suggest to the client a subset of real estate with an attached degree of certainty regarding suitability of the proposal itself.



Knowledge modeling: real estate properties



```
1 (deftemplate immobile
2   (slot name (type STRING))
3   (slot type (type SYMBOL) (allowed-symbols Alloggio Attico Villa))
4   (slot square-meters (type INTEGER))
5   (slot state (type SYMBOL) (allowed-symbols Nuovo Ristrutturato Ristrutturare))
6   (slot num-of-rooms (type INTEGER))
7   (slot floor (type SYMBOL) (allowed-symbols Terra Intermedio Ultimo))
8   (slot city (type SYMBOL))
9   (slot district (type SYMBOL)) ;quartiere
10  (slot elevator (type SYMBOL) (allowed-values Si No))
11  (slot box-auto (type SYMBOL) (allowed-values Si No))
12  (slot box-mq (type INTEGER))
13  (slot balcony (type SYMBOL) (allowed-values Si No))
14  (slot backyard (type SYMBOL) (allowed-values Si No))
15  (slot price (type INTEGER))
16  (slot energy-class (type INTEGER) (allowed-values 1 2 3))))
```



Knowledge modeling: real estate properties



```
1 (deftemplate district
2   (slot name (type SYMBOL))
3   (slot city (type SYMBOL))
4   (slot transports (type SYMBOL) (allowed-values Si No))
5   (slot station (type SYMBOL) (allowed-values Si No))
6   (slot gym (type SYMBOL) (allowed-values Si No))
7   (slot supermarket (type SYMBOL) (allowed-values Si No))
8   (slot school (type SYMBOL) (allowed-values Si No))
9   (slot park (type SYMBOL) (allowed-values Si No))
10  (slot hospital (type SYMBOL) (allowed-values Si No))
11 )
```



```
1 (deftemplate region
2   (slot name (type SYMBOL))
3   (multislot cities (type SYMBOL)))
4
5
6 (deftemplate city
7   (slot name (type SYMBOL))
8   (slot region (type SYMBOL)))
```



Knowledge modeling: customer profiling

In order to profile the client, the following questions were asked, with domain knowledge attached:

- Do you have children? **-yes->** School or Park
- Are you a sports person? **-yes->** Park or Gym
- Do you have a car? **-no->** Transportation or Supermarket or Station
-yes-> Box-car
- Are you over 60 years old? **-yes->** Hospital or Supermarket or Park



Knowledge modeling: customer profiling

Profiling operations are performed with the support of specific weights established a priori, with ranges between 1 (not very relevant) and 10 (very relevant):

```
1 (deffacts weights-children
2   (weights-children
3     (school 5)
4     (park 3)
5   )
6 )
7
8 (deffacts weights-sporty
9   (weights-sporty
10    (gym 4)
11    (park 4)
12  )
13 )
```

```
1 (deffacts weights-old
2   (weights-old
3     (hospital 7)
4     (supermarket 5)
5     (park 3)
6   )
7 )
8
9 (deffacts weights-car
10  (weights-car
11    (transports 7)
12    (supermarket 2)
13    (station 4)
14    (box-auto 3)
15  )
16 )
```



Knowledge modeling: immobile questions

In order to determine the properties most akin to the client's requirements, the following questions were asked, with attached valid answers:

- How much do you want to spend? [30000,...,1000000]
- What kind of property are you looking for? {Housing,Penthouse,Villa,Indifferent}
- How many square meters do you want the house to be? [40,...,400]
- What kind of property do you prefer? {New,Renovated,Renovating,Indifferent}
- In which floor do you prefer to live? {Earth,Middle,Last,Indifferent}
- How many rooms do you want the house to be? [2,...,10]
- Is the presence of an elevator required? {Yes,No}
- Is the presence of a balcony required? {Yes,No}
- Is the presence of a garden required? {Yes,No}
- Energy class of the house?(numerical) [(1->A),...,(7->G)]
- In which city do you want to buy a house? {Novara, Turin, Rome, Civitavecchia, Palermo, Catania, Indifferent}



Knowledge modeling: immobile questions

Property suggestion operations are performed with the support of specific weights established a priori, with ranges between 1 (not very relevant) and 10 (very relevant):

```
1 (deffacts weights-immobile
2 (weights-immobile
3   (type 9)
4   (square-meters 10)
5   (state 8)
6   (num-of-rooms 7)
7   (floor 3)
8   (city 6)
9   (elevator 3)
10  (box-auto 3)
11  (box-mq 2)
12  (terrace 2)
13  (balcony 3)
14  (furnished 3)
15  (cellar 2)
16  (pool 1)
17  (heating-system 3)
18  (backyard 2)
19  (price 9)
20  (energy-class 5)
21  (region 8)
22  (district 5))
23 )
```



Calculation of Certainty Factors

In order to be able to propose to the client a list of properties as suitable as possible for his or her needs, it was chosen to calculate a Certainty Factor (between 0 and 100) for each desired attribute in relation to the weights established a priori by the domain expert.

Methodologies were then identified to be able to calculate the CF of a certain property from the CFs of the individual attributes that describe it.



Calculation of Certainty Factors: Types

In relation to the domain of attributes considered, it is possible to classify CFs into different categories:

- Elemental CFs: calculated from the weights given by the domain expert
 - Defined by means of a range: there is a sorting between the values that the reference attribute can take on
 - Non-range
- Aggregate CFs: calculated from elemental CFs



Calculation of CFs: elementary, not range

Assuming that an attribute A has eligible values belonging to $\{x_1, \dots, x_n\}$ and the client has indicated a preference for the value x_i , then:

- If the considered property has a value x_j of A, with $j \neq i$, then the calculated CF value will be equal to $100 - (w_A \cdot 10)$, with w_A weight associated with the attribute A
- If the property considered has a value x_j of A, with $j = i$, then the value of CF calculated will be equal to 100

Otherwise, if the client has not expressed a preference (indifferent), the calculated CF value will amount to 100 for any value of x_j



Calculation of CFs: elementary, range

Given an attribute A having range $[L_r, M_r]$ and value v with a score belonging [1,10], also considered the user's desired value V belonging to $[L_r, M_r]$

If $|M_r - V| > |L_r - V|$ then we associate value of CF with the attribute having value v equal $f(v)$ with $R = M_r$ otherwise we associate value $f(v)$ with $R = L_r$.

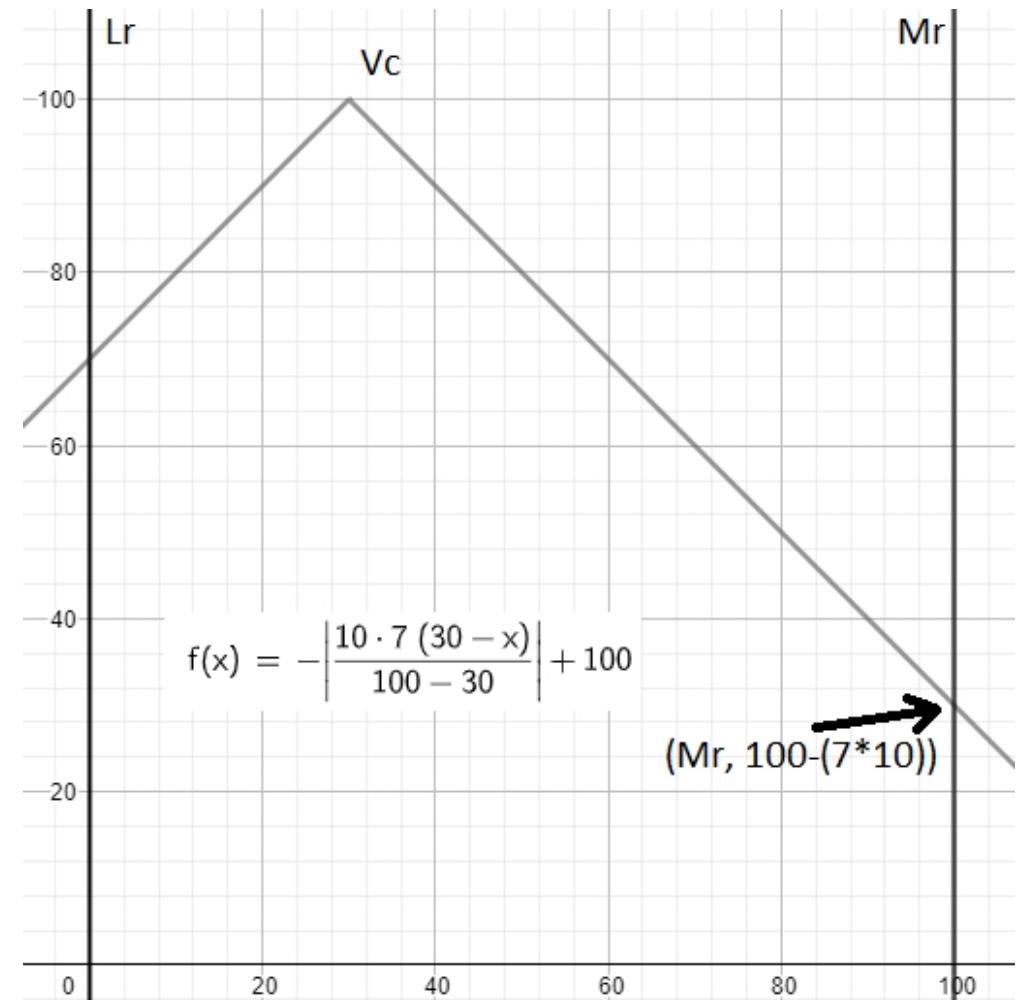
$$f(x) = - \left| \frac{10 \cdot w_A \cdot V - x}{R - V} \right| + 100$$



Calculation of CFs: elementary, range

A graphical representation of the function described on the previous slide considering is shown:

- an attribute A whose range is in the range [0,100]
- A weight w_A attributed to A equal to 7
- A value V_c suggested by the customer of 30



Calculation of CFs: elementary, range

Given an attribute A, chosen a value of it V by the client, it is plausible to assume, in most cases, that the CF to be attributed with respect to a property with a value of $A = V_i$ may still either amount to 100 if $V_i \geq V$ or $V_i < V$, depending on the type of attribute being treated:

- If $V_i \leq V$ the CF does not decrease

- Price
 - Energy class

- If $V_i \geq V$ the CF does not decrease

- square meters
 - number of rooms
 - presence of the elevator
 - presence of the terrace
 - presence of the balcony
 - presence of the garden



Calculation of CFs: elementary, range

Given the domain assumptions described in the previous slide, it was useful to use the function described above to define two functions described by parts:

■ Greater Than function:

- Equal to $f(x)$ for each $x < V_c$
- Equal to 100 for each $\geq V_c$

■ Less Than function:

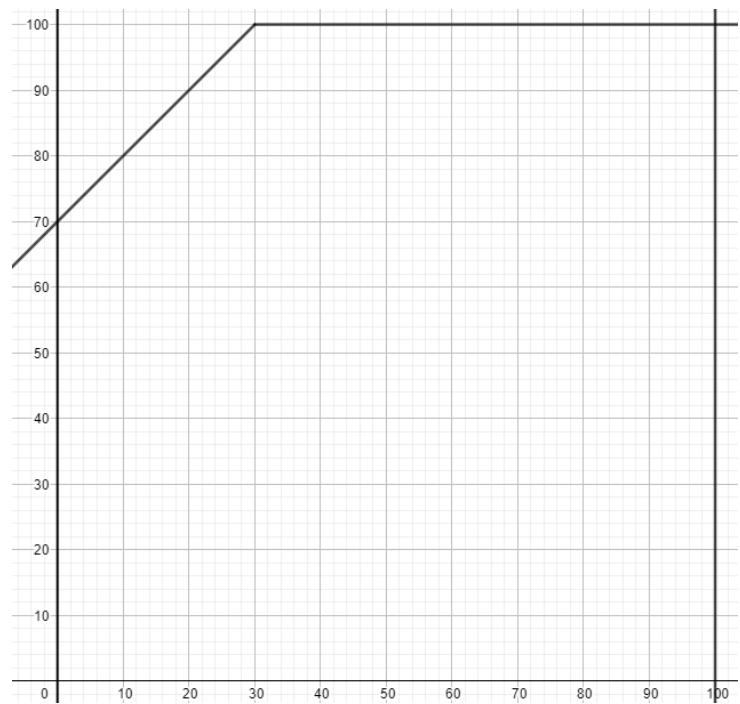
- Equal to $f(x)$ for each $x > V_c$
- Equal to 100 for each $x \leq V_c$



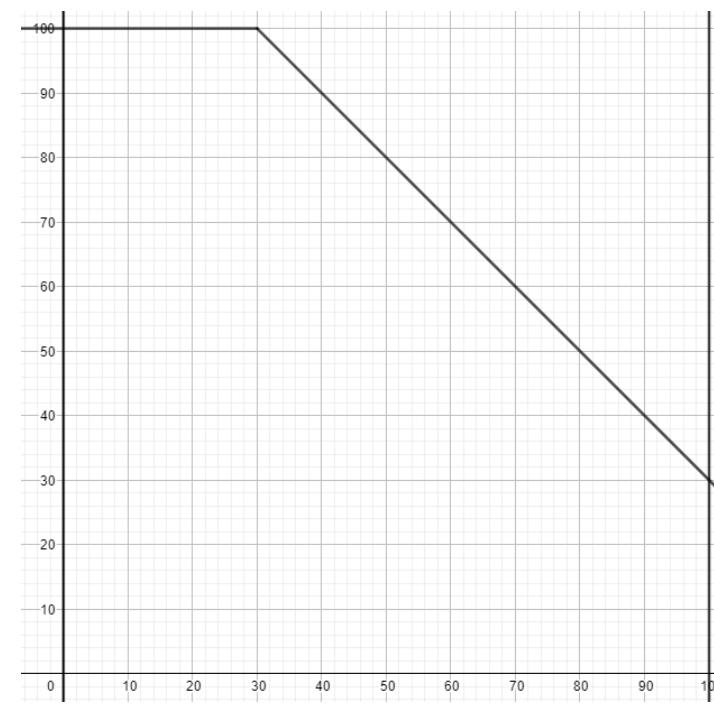
Calculation of CFs: elementary, range

Two examples of the functions defined by parts are given below, calculated starting with the example function introduced in slide 13:

Greater Than (GT)



Less Than (LT)



Calculation of CFs: aggregates

The aggregate CF of a given property I was calculated by considering the following information:

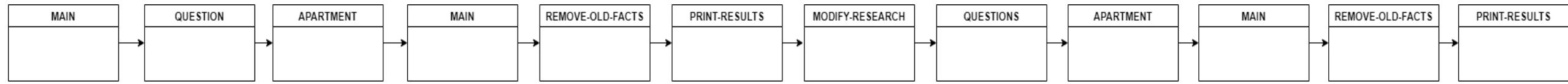
- A set of attributes characterizing the property: $\{A_1, A_2, \dots, A_n\}$
- A set of weights associated with each attribute: $\{W_1, W_2, \dots, W_n\}$
- A set of CFs calculated from attribute values: $\{CF_1, CF_2, \dots, CF_n\}$

Through the following formula:

$$CF_I = \frac{\sum_{i=1}^n CF_i \cdot W_i}{\sum_{i=1}^n W_i}$$



Organization of module stack



Description of Main module

The Main module is initially concerned with defining the organization of the stack
Of the forms through the Focus instruction.

It is invoked several times within the module stack to calculate CFs
aggregates from CF asserted by other modules.



Description of Form Question

The Question module deals with the management of questions, specifically.
Performs the following tasks:

1. Print the text of the questions contained in the knowledge base
2. Get the answer of the question mediante keyboard input
3. Performs checks regarding the eligibility of the value entered against the valid answers
4. If necessary, it asks the user if the value entered is the only acceptable one (soft-hard)
5. Asserts an attribute fact as a pair <attribute_name, value>



Description of the Apartment module

The Apartment module takes care of the calculation of elementary CFs (not calculated from other CFs, only from the weights defined a priori) based on the desired and user-indicated attributes asserted previously by the Question module.

The calculation of CFs is done by calling the functions **calculate-cf-no-range** and **calculate-cf-range** in relation to the type of attribute considered.

```
● ● ●  
1  (defrule APARTMENT::calculate-cf-city  
2    (question (attribute city))  
3    (immobile (name ?name)(city ?c))  
4    (attribute (name city) (value ?city))  
5    (weights-immobile (city ?w))  
6    =>  
7    (assert (attribute (name city_) (value ?c) (estate ?name) (weight ?w)  
8          (certainty (calculate-cf-no-range ?city ?c ?w))  
9          ))  
10 )
```



Description of the Remove-Old-Facts module.

The Remove-Old-Facts module takes care of removing the attribute indicating the CF of a certain property that violates hard constraints (if any) imposed by the user .

In this way, the form that later has to indicate a list of properties to be proposed to the client in relation to their CF will not propose properties that violate Hard constraints.



Description of the Print-Results module

The Print-Results module takes care of printing the first five properties in order of CF calculated.

In addition to their name, the following attributes are also specified:

- Price in Euro
- Type
- City
- Status of the property
- Square meters
- Factor of certainty

In case at least one Hard constraint has been set and there are no properties that do not violate those constraints, an inability to propose a property with the required characteristics will be reported.



(N.B. : The printed table data is correctly indented by ClipsDOS but not by ClipsIDE, as it does not reserve the same amount of pixels for each character).

Description of the Print-Results module

Example of output printed by the Print-Results module:

Immobile	Prezzo(Euro)	Tipologia	Citta'	Stato	mq	CF
Villa da 8 stanze	334000	Villa	Novara	Nuovo	160	91
Villa nuova con giardino	640000	Villa	Catania	Nuovo	230	85
Villa con giardino	370000	Villa	Novara	Nuovo	135	84
Villa da 10 stanzze	575000	Villa	Novara	Ristrutturare	200	82
Villa unifamiliare da ristrutturare	530000	Villa	Palermo	Ristrutturare	175	82
CLIPS>						



Description of the Modify-Search module

The Modify-Search module takes care of proposing to the client the possibility of partially retracting the values entered by the Question module (especially one of the attributes between *Price*, *Status* and *Type*), it also performs the following operations:

- In case the user chooses to retract an attribute, it places *the already-asked* attribute of the associated question equals False
- Takes care of removing the fact that indicates whether the value of the attribute entered by the user is soft or hard
- Deletes the value of the attribute previously entered by the user
- Deletes CF values calculated from elemental CFs



Scenarios investigated

The test scenarios investigated were formulated by means of a combined:

- Three fictitious requests were described regarding the properties sought
- Three fictitious users were profiled
- For each of the nine possible <furniture, user> pairs, the 5 properties suggested by the expert system were reported



Scenarios investigated: real estate#1

Property #1	Budget (€)	Type	m ²	Status	Plan	No. rooms	Elevator	Balcony	Garden	Class Energetics	City
	30.000	Penthouse	40	New	Earth	2	Yes	No	No	1	Turin (hard)

User #1	> 60 years old	Car	Sportsman	Children
	Yes	No	Yes	No

User #2	> 60 years old	Car	Sportsman	Children
	No	Yes	Yes	Yes

User #3	> 60 years old	Car	Sportsman	Children
	No	No	No	No

I#1, U#1	Price (€)	Type	City	Status	m ²	CF
Accommodation from 4 rooms	95.000	Accommodation	Turin	New	95	88
Penthouse in center 1	65.000	Penthouse	Turin	Renovated	48	84
Accommodation for 2 rooms 2	55.000	Accommodation	Turin	Renovate	55	81
Accommodation restructured	99.000	Accommodation	Turin	Renovated	75	81
Villa from 7 rooms	450.000	Villa	Turin	Renovated	130	78

I#1, U#2	Price (€)	Type	City	Status	m ²	CF
Penthouse in center 1	65.000	Penthouse	Turin	Renovated	48	85
Accommodation from 4 rooms	95.000	Accommodation	Turin	New	95	84
Accommodation with garden private	68.000	Accommodation	Turin	Renovate	59	77
Accommodation in the Cathedral	74.000	Accommodation	Turin	Renovate	45	77
Accommodation for 2 rooms 2	55.000	Accommodation	Turin	Renovate	55	76

I#1, U#3	Price (€)	Type	City	Status	m ²	CF
Penthouse in center 1	65.000	Penthouse	Turin	Renovated	48	89
Accommodation from 4 rooms	95.000	Accommodation	Turin	New	95	87
Accommodation with garden private	68.000	Accommodation	Turin	Renovate	59	79
Accommodation in the Cathedral	74.000	Accommodation	Turin	Renovate	45	78
Accommodation for 2 rooms 2	55.000	Accommodation	Turin	Renovate	55	77



Scenarios investigated: real estate#2

Property #2	Budget (€)	Type	m ²	Status	Plan	No. rooms	Elevator	Balcony	Garden	Energy Class	City
	500.000	Villa	180	Indiff.	Indiff.	5	No	Yes	Yes	3	Rome

User #1	> 60 years old	Car	Sportsman	Children
	Yes	No	Yes	No

User #2	> 60 years old	Car	Sportsman	Children
	No	Yes	Yes	Yes

User #3	> 60 years old	Car	Sportsman	Children
	No	No	No	No

I#2, U#1	Price (€)	Type	City	Status	m ²	CF
Villa from 7 rooms	450.000	Villa	Turin	Renovated	130	90
Villa with garden	370.000	Villa	Novara	New	135	86
10-room villa	575.000	Villa	Novara	Renovate	200	86
Single-family villa to be renovated e	530.000	Villa	Palermo	Renovate	175	86
Villa from 8 rooms	334.000	Villa	Novara	New	160	90

I#2, U#2	Price (€)	Type	City	Status	m ²	CF
Villa from 10 rooms	575.000	Villa	Novara	Renovate	200	91
Villa from 8 rooms	334.000	Villa	Novara	New	160	90
Single-family villa to be renovated e	530.000	Villa	Palermo	Renovate	175	89
Villa single family re	314.000	Villa	Civitavecchia	New	105	88
New villa with garden	640.000	Villa	Catania	New	230	87

I#2, U#3	Price (€)	Type	City	Status	m ²	CF
Villa single-family house re to be renovated re	530.000	Villa	Palermo	Renovate	175	93
Villa from 10 rooms	575.000	Villa	Novara	Renovate	200	92
Villa from 8 rooms	334.000	Villa	Novara	New	160	90
Villa from 7 rooms	450.000	Villa	Turin	Renovated	130	89
Villa bifamily re	1.000.000	Villa	Civitavecchia	Renovate	400	85



Scenarios investigated: real estate#3

Property #3	Budget (€)	Type	m ²	Status	Plan	No. rooms	Elevator	Balcony	Garden	Energy Class	City
	800.000	Indiff.	250	Renovated	Indiff.	8	No	Yes	Yes	8	Catania

User #1	> 60 years old	Car	Sportsman	Children
	Yes	No	Yes	No

User #2	> 60 years old	Car	Sportsman	Children
	No	Yes	Yes	Yes

User #3	> 60 years old	Car	Sportsman	Children
	No	No	No	No

I#3, U#1	Price (€)	Type	City	Status	m ²	CF
Villa from 7 rooms	450.000	Villa	Turin	Renovated	130	84
Villa renovated with garden	850.000	Villa	Palermo	Renovated	350	84
Accommodation with sea view	178.000	Accommodation	Catania	Renovated	60	80
Villa semi-detached	1.000.000	Villa	Civitavecchia	Renovate	400	79
Accommodation renovated	99.000	Accommodation	Turin	Renovated	75	78

I#3, U#2	Price (€)	Type	City	Status	m ²	CF
Villa restructured with garden	850.000	Villa	Palermo	Renovated	350	89
New villa with garden	640.000	Villa	Catania	New	230	83
Villa from 10 rooms	575.000	Villa	Novara	Renovate	200	80
Villa from 7 rooms	450.000	Villa	Turin	Renovated	130	79
Villa semi-detached	1.000.000	Villa	Civitavecchia	Renovate	400	77

I#3, U#3	Price (€)	Type	City	Status	m ²	CF
Villa restructured with garden	850.000	Villa	Palermo	Renovated	350	89
Villa duplex	1.000.000	Villa	Civitavecchia	Renovate	400	82
Villa from 7 rooms	450.000	Villa	Turin	Renovated	130	82
New villa with garden	640.000	Villa	Catania	New	230	81
Accommodation with sea view	178.000	Accommodation	Catania	Renovated	60	80



Conclusions on the goodness and limitations of the solution

Regarding the proposed solution, some design choices should be considered:

- Weights were defined according to individual attributes, not n-uples of attributes:
 - The amount of weights has been extremely reduced, simplifying the task of the expert of the domain
 - This implies less accuracy on the part of the expert system.
Moreover, as the amount of attributes describing a certain property increases, the final contribution of individual attributes to CF calculation decreases
- Reduced domain knowledge, not experimentally verified
- It is possible to retract only the value of a restricted amount of attributes
- For simplicity, all attributes characterizing the proposed properties are not reported
- Database provided with representative examples, but small size

