Assistance in making decisions to promote planting and conservation of maize in the state of Puebla of Mexico through Answer Set Programming

Luis A. Montiel[1], Claudia Zepeda[2], Pedro Wesche[1], and Javier A. Cervantes[2]

[1] Universidad de las Américas, CENTIA
[2] Benemérita Universidad Autónoma de Puebla
luis.montielmo@udlap.mx

Abstract We present some examples which show how preference ordered disjunction adds useful features to answer set programming, as well as an interesting example of how it can be useful for modeling a real problem about agriculture. In this paper, we present a mathematical formalization for the problem of making decisions to promote planting and conservation of maize in some geographical zone inside the country of Mexico; we present modeling semantics that we consider necessary for the mathematical formalization of this problem using Answer Set Semantics.

Keywords: Agriculture, Ordered Disjunction, Preferences, Answer Set Programming, Data Mining.

1 Introduction

At the present time Maize is one of the most important crops around the world, it is a major source of food for humanity and the livestock sector [5]. Currently it is very important to support technological development in agriculture in order to increase yields (production per hectare) of crops that are important to satisfy the food demand of humanity in the world [20].

A problem present in Mexico is that the total production of maize in the country is not enough to meet food demand of all Mexican population; Mexico has faced the need to import foreign corn mainly from U.S. and Canada since it is cheaper to import corn than to produce national maize in Mexico [20][18].

Given experts recommendations and needs [20][18], and experts works [8][15][14] we identify that is possible to define *The maize fitting zone problem* and give a solution to it in order to support agriculture in the area of corn:

The maize fitting zone problem (MaizeFitZone problem). *This problem is defined by the following tasks and subproblems:*

1. *To conduct a study on planting and harvesting methodologies employed for corn in a given geographical area in recent years.*
2. *To identify biological properties for each variety of native maize in a given geographical area.*

3. *Given a geographical area for each land zones belonging to it, to identify the attributes (geographical, climatic, technological, economic, soil, hydrological land use and vegetation) that define it.*
4. *To relate this land attributes in function of classes, races and varieties of maize and their yields produced.*
5. *To determine which land zones report high or low yields for each variety of corn and to explain the necessary and sufficient conditions to obtain high yields in each land zone.*
6. *To produce corn in land zones selecting the best race of maize that fits each land zone to obtain high yields of production preserving an ecosystem that does not harm nature and that satifies the nutrimental needs of producers and population of a given geographical area.*
7. *To replicate these conditions in land zones with low yields of corn*
8. *Continuous monitoring of implementation of the model in reality for different classes of producers by experts, and feedback to the mathematical model with the results. Taking care not to affect the quality of life for producers who have their sustenance on their own crops.*

Recently work [8][15][14] has been done trying to solve the problem *The maize fitting zone problem*. There is great diversity of climatic and geographic attributes, attributes that represent the technologies necessary to grow, control pests and diseases of maize, and attributes of phenotypes according to breeds and varieties of existing corn and classes to which they belong. The data type for these attributes are strings, integers, real numbers, percentages, however in the works we have reviewed percentage and numerical attributes are handled through ranges between values in order to facilitate the categorization of these attributes, because of this it is possible to map, one to one, quantitative information to qualitative data. Due to the complexity in the relation between the attributes involved in these problems we require to handle incomplete knowledge. The problems show variables of uncertainty for example climate and pest or disease control attributes for the corn, and also we require the use of preferences for example preferences according to the classes of seeds to grow depending on the environment. Because we are dealing with these attributes and the type of rules that will arise in these problems, we identified that it is possible to use Answer Set Programming [10][9] as a means to formalize this problem in an expressive language. We identified the need of using an ASP extension that use preferred rules [7][4][21] for handling preference conditions.

There is now software that efficiently computes ASP [12][19], this is the reason we believe it is possible to propose a real solution that can be implemented based on this programming paradigm.

2 Background

In this section we introduce all the necessary terminology and relevant definitions in order to make this paper self-contained.[1]

[1] We assume that the reader has familiarity with basic concepts of *classical logic*, *logic programming*, *answer set semantics*, and *lattices*. For details the reader can refer to [1,6,13].

2.1 Extended Logic Programs

We consider extended logic programs which have two kinds of negation, strong negation $\neg$ and default negation *not*. A signature $\mathcal{L}$ is a finite set of elements that we call atoms, where atoms negated by $\neg$ are called extended atoms. Intuitively, *not* a is true whenever there is no reason to believe a, whereas $\neg a$ requires a proof of the negated atom. In the following we use the concept of atom without paying attention if it is an extended atom or not. A *literal* is either an atom a called *positive literal*, or the negation of an atom *not* a called *negative literal*. Given a set of atoms $\{a_1, ..., a_n\}$, we write *not* $\{a_1, ..., a_n\}$ to denote the set of atoms $\{not\ a_1, ..., \ not\ a_n\}$. An ***extended*** *normal rule* (rule, for short) r is a rule of the form $a \leftarrow b_1, \ldots, b_m, \ not\ b_{m+1}, \ldots, \ not\ b_{m+n}$ where a and each of the b_i are atoms for $1 \leq i \leq m + n$. If $m + n = 0$ the rule is an abbreviation of $a \leftarrow \top$ such that $\top$ is the proposition symbol that always evaluates to true; the rule is known as a *fact* and can be denoted just by a. If $n = 0$ the rule is an extended definite rule. We denote a rule r by $a \leftarrow \mathcal{B}^+, \ not\ \mathcal{B}^-$ where the set $\{b_1, \ldots, b_m\}$ and the set $\{b_{m+1}, \ldots, b_{m+n}\}$ are denoted by $\mathcal{B}^+$ and $\mathcal{B}^-$ respectively. A *constraint* is a rule of the form $\leftarrow \mathcal{B}^+, \ not\ \mathcal{B}^-$. We denote by $head(r)$ the head a of rule r and by $body(r)$ the $\mathcal{B}^+, \ not\ \mathcal{B}^-$ of the rule r. An *extended normal logic program* P is a finite set of extended normal rules and/or constraints. By $\mathcal{L}_P$ we denote the signature of P, i.e. the set of atoms that appear in the rules of P. If all the rules in P are extended definite rules we call the program P *extended positive logic program*. In our logic programs we will manage the strong negation $\neg$ as it is done in Answer Set Programming (ASP) [1]. Basically, each atom $\neg a$ is replaced by a new atom symbol a' which does not appear in the language of the program and we add the constraint $\leftarrow a, a'$ to the program. For managing the constraints in our logic programs, we will replace each rule of the form $\leftarrow \mathcal{B}^+ \ not\ \mathcal{B}^-$ by a new rule of the form $f \leftarrow \mathcal{B}^+, \ not\ \mathcal{B}^-, \ not\ f$ such that f is a new atom symbol which does not appear in $\mathcal{L}_P$.

2.2 Logic Programs with Ordered Disjunction

Logic programs with ordered disjunction (LPODs) are extended logic programs augmented by an ordered disjunction connector $\times$ which allows to express qualitative preferences in the head of rules [3]. A LPOD is a finite collection of rules of the form $r = c_1 \times \ldots \times c_k \leftarrow b_1, \ldots, b_m, \ not\ b_{m+1}, \ldots, \ not\ b_{m+n}$ where c_i (for $1 \leq i \leq k$) and each of the b_j (for $1 \leq j \leq m + n$) are atoms. The rule r states that if the body is satisfied then some c_i must be in the answer set, if possible c_1, if not then c_2, and so on, and at least one of them must be true. Each of the c_i represents alternative, ranked options for problem solutions the user specifies according to a desired order. If $k = 1$ then the rule is an extended normal rule. The semantics of LPODs is based on the following reduction.

Definition 1 ($\times$-reduction). [3] *Let* $r = c_1 \times \ldots \times c_k \leftarrow b_1, \ldots, b_m, \ not\ b_{m+1}, \ldots, \ not\ b_{m+n}$ *be an ordered disjunction rule and* M *be a set of atoms. Let* P *be an LPOD and* M *be a set of atoms. The* $\times$*-reduct* $r_\times^M$ *is defined as*

$$r_\times^M := \{c_i \leftarrow b_1, \ldots, b_m | c_i \in M \ and\ M \cap (\{c_1, \ldots, c_{i-1}\} \cup \{b_{m+1}, \ldots, b_{m+n}\}) = \emptyset\}.$$

The $\times$*-reduct* $P_\times^M$ *is defined as* $P_\times^M = \bigcup_{r \in P} r_\times^M$.

Definition 2. [3] *Let P be an LPOD and M a set of atoms. Then, M is an answer set of P if and only if M is a minimal model of $P_{\times}^{M}$. We denote by $SEM_{LPOD}(P)$ the mapping which assigns to P the set of all answer set of P.*

One interesting characteristic of LPODs is that they provide a mean to represent preferences among answer set by considering the rule satisfaction degree [3].

Definition 3. [3] *Let M be an answer set of an LPOD P. Then M satisfies the rule $r = c_1 \times \ldots \times c_k \leftarrow b_1, \ldots, b_m,\ not\ b_{m+1} \ldots,\ not\ b_{m+n}$:*
– to degree 1 if $b_j \notin M$ for some j $(1 \leq j \leq m)$, or $b_i \in M$ for some i $(m+1 \leq i \leq m+n)$,
– to degree j $(1 \leq j \leq k)$ if all $b_l \in M$ $(1 \leq l \leq m)$, $b_i \notin M$ $(m+1 \leq i \leq m+n)$, and $j = min\{r \mid c_r \in M, 1 \leq r \leq k\}$.

The degrees can be viewed as penalties: the higher the degree the less satisfied we are. If the body of a rule is not satisfied, then there is no reason to be dissatisfied and the best possible degree 1 is obtained [3]. The satisfaction degree of an answer set M w.r.t. a rule, denoted by $deg_M(r)$, provides a ranking of the answer set of an LPOD, and a preference order on the answer set can be obtained using some proposed combination strategies [3].

3 Important computational aspects for MaizeFitZone problem

In order to show that it is possible to model *MaizeFitZone problem* througth Answer Set Programming we need to focus our attention in subproblems (5) and (6). Subproblem (5) presents the problem of determining for a geographical zone its agricultural production potential. Subproblem (6) presents the problem of locating seeds candidates who best fit the needs of farmers based on the agricultural production potential of the geographical zone where they are located. Subproblems (5) and (6) describes the use of an integrated database *MaizeBioGeoClimAgriTechDB* obtained from subproblems (1) to (4). Subproblems (7) and (8) describes the solution in real life once the mathematical model of the main problem has been implemented.

***MaizeFitZone problem* (1) to (4).** We identify that it is possible to link Geographical, Weather and Agricultural techonology databases in order to cross information on the different corn seeds (Creole or Improved), their yields in each municipality and locality, and the relevant attributes (climatic, soil, etc.) involved to obtain certain level of production (low, medium, high, very high for example). Let us call this linked information GeoClimAgriTechDB.

From *GeoClimAgriTechDB* it is possible relate land attributes in function of maize classes and their yields produced, however to relate land attributes in function of races and varieties of maize and their yields produced it is necessary to integrate *GeoClimAgriTechDB* database with Biological properties of maize database, so it is possible to know in a general level the candidates of races and varieties of maize seeds that could be presented in the land zone and the probability of appearing in it. Let us call to this integrated database ***MaizeBioGeoClimAgriTechDB***.

To see more detalied information about the *MaizeBioGeoClimAgriTechDB* database see technical report[16].

***MaizeFitZone problem* (5).** Recent work addresses this problem [8][15][14]. The computer systems developed by INIFAP Puebla [8] and Chiapas [15][14] calculate agricultural production potential in given geographic areas, that is, the level of fitness for a plant to be cultivated successfully. The Puebla system provides the potential for various crops (monocultures), however there is no documentation about the mathematical model employed. The Chiapas system provides the potential for varieties of native and improved maize seeds (also monocultures) and uses climatic and soil information; documentation about the mathematical model employed exists [15][14].

In order to discern from classes(Creole or Improved), races and varieties of corn seeds yields produced for each municipality and locality the relevant attributes involved to obtain certain level of production (low, medium, high, very high for example) and to explain the necessary and sufficient conditions to obtain high yields in each land zone from *MaizeBioGeoClimAgriTechDB*, it is possible to run datamining algorithms to discover patterns in the information such as C4.5 [17] or ID3 [2], expose these patterns to agronomists experts to discern interesting patterns.

There exists efficient software that implements data mining algorithms such as ID3 and C4.5 that it is easy to use such as WEKA software [11].

Let us call to the combination of the subproblems *MaizeFitZone problem (4)* and *MaizeFitZone problem (5)* the ***maize zone production potential problem***.

***MaizeFitZone problem* (6).** Once the potential of agricultural production (the level of fitness for a plant to be cultivated successfully) has been computed, in the Chiapas System [15][14] a suggestion of improved seeds candidates who best fit the needs of the user is given, this computed from a set of geographical, climatic, and technological attributes values given as input by the user.

Once a solution is given for *maize zone production potential problem* and based on this solution, given a geographical zone, let us call ***maizes best fitting zone problem*** to the problem of selecting the races and varieties of maize to be planted that fits each land zone to obtain high yields of production preserving an ecosystem that does not harm nature and that satifies the nutrimental needs of producers and population of a given geographical area.

4 Computational modeling Answer Set Programming approach for *maize fitting zone problem*

This section determines the knowledge that is required to provide a knowledge modeling, and using the model of recent work [15][14] we show that it is possible to model the knowledge that is required to provide the knowledge modeling of *MaizeFitZone problem.*

4.1 Knowledge Sources

We define the **maize agricultural knowledge** to cultivate maize plants as the union of the following knowledge:

1. *Agronomists experts' knowledge.*

2. *The knowledge generated by centuries-old traditions and experience of the rural poor*
3. *Patterns and knowledge discovered from MaizeBioGeoClimAgriTechDB database using knowledge discovering techinques such as data mining.*

in order to generate a mathematical model to solve *MaizeFitZone problem* efficiently and accurately.

Now we present the argumentation that shows that it is possible to model *MaizeFitZone problem* using Answer Sets programming using extended semantics handling preferences in disjunctive rules.

To model Agronomist experts' knowledge and experience of the rural poor represent the same problem since experience of the rural poor can be structured in the same way Agronomist experts' knowledge have been structured.

It is possible to provide a mathematical model for *maizes best fitting zone problem* using Answer Sets[10,9] based on work and documentation form INIFAP Chiapas [15,14] to show how it is possible to model Agronomist experts' knowledge.

It is possible to run datamining algorithms to discover patterns from the information contained in *MaizeBioGeoClimAgriTechDB* such as C4.5 [17] or ID3 [2], and from this kind of algorithms is possible to generate a logic program, for instance an ASP program for example from a classifictation tree to generate in a direct way, at least an ordered disjunctive logic program.

In order to model *maize fit problem* first part of this problem (*maize zone production potential problem*)is modeled from natural language to ASP rules, and the second part, *maizes best fitting zone problem* is modeled using a C4.5 tree obtained from implicit knowledge in [15][14] in order to generate a preferred ordered disjunctive program.

4.2 Modeling of *maize zone production potential problem*: Answer Sets approach

The works [15][14] present the scientific and technological agronomist experts' knowledge on how is determined for a given geographical zone if the potential of production in order to cultivate native maize seeds is very good, good, intermediate or low.

As input of the problem, part of extensional data base (this knowledge is included in *MaizeBioGeoClimAgriTechDB* knowledge base), we have the following facts:

- From processing and analysis of climate we have the data:
  - `precipitation`$(Zone, Day, Value)$.
  - `evaporation`$(Zone, Day, Value)$.
  - `temperature`$(Zone, Day, Value)$.
  - `precipitationByYear`$(Zone, Year, Value)$.
  - `evaporationByYear`$(Zone, Year, Value)$.
  - `temperatureByYear`$(Zone, Year, Value)$.
- From generation and classification of images (geographical data):
  - `zone`$(Zone)$.
  - `locality`$(Locality)$.
  - `localityInZone`$(Zone, Locality)$.
  - `terrainSlope`$(Zone, Value)$.
  - `soilDepth`$(Zone, Value)$.
  - `soilTexture`$(Zone, Vale)$.
- From user input attributes for a locality:
  - `organicMatter`$(L, Value)$.
  - `pending`$(L, Value)$.

- $\texttt{rainfedQuality}(L, Value)$.
- $\texttt{altitude}(L, Value)$.
- $\texttt{cultureCycle}(L, Value)$.
- $\texttt{moistureRegime}(L, Value)$.
- $\texttt{soilTexture}(Zone, Vale)$.

Where *value* is a cualitative value or an integer in a range inside $[0, 1000]$

The rules that gives insigth to model this problem are based in this expert knowledge[15][14]:

$$\begin{aligned}
\texttt{potentialProductivity}(Z, veryGood) \leftarrow\ & \texttt{zone}(Z), \texttt{growingSeasonPeriod}(Z, middle),\\
& \texttt{soilDepthCentimeter}(Z, veryDeep),\ not\ \texttt{abnormal}(Z).\\
\texttt{potentialProductivity}(Z, good) \leftarrow\ & \texttt{zone}(Z), \texttt{growingSeasonPeriod}(Z, long),\\
& (\texttt{soilDepthCentimeter}(Z, shallow)\vee\\
& \texttt{soilDepthCentimeter}(Z, veryDeep)),\\
& not\ \texttt{abnormal}(Z).\\
\texttt{potentialProductivity}(Z, intermediate) \leftarrow\ & \texttt{zone}(Z), \texttt{growingSeasonPeriod}(Z, middle),\\
& \texttt{soilDepthCentimeter}(Z, shallow),\\
& not\ \texttt{abnormal}(Z).\\
\texttt{potentialProductivity}(Z, intermediate) \leftarrow\ & \texttt{zone}(Z), \texttt{growingSeasonPeriod}(Z, short),\\
& \texttt{soilDepthCentimeter}(Z, veryDeep),\\
& not\ \texttt{abnormal}(Z).\\
\texttt{potentialProductivity}(Z, low) \leftarrow\ & \texttt{zone}(Z), (\texttt{growingSeasonPeriod}(Z, short)\vee\\
& \texttt{soilDepthCentimeter}(Z, shallow)\\
& \vee\texttt{abnormal}(Z)).\\
\texttt{abnormal}(Z) \leftarrow\ & \texttt{zone}(Z), \texttt{erosion}(Z).\\
\texttt{abnormal}(Z) \leftarrow\ & \texttt{zone}(Z), \texttt{naturalResourcesDegradation}(Z).\\
\texttt{abnormal}(Z) \leftarrow\ & \texttt{zone}(Z), \texttt{droughtChance}(Z).
\end{aligned} \tag{1}$$

For example the rule:

$$\begin{aligned}
\texttt{potentialProductivity}(Z, veryGood) \leftarrow\ & \texttt{zone}(Z), \texttt{growingSeasonPeriod}(Z, middle),\\
& \texttt{soilDepthCentimeter}(Z, veryDeep),\\
& not\ \texttt{abnormal}(Z).
\end{aligned} \tag{2}$$

Can be read as: the agricultural potential of production for a given zone Z is *very good* if there is evidence that there exists in the zone Z a *middle* growing season perdiod and *very deep* soil depth, and there is no evidence that in zone Z there exists a risk of erosion, and there is no evidence that in zone Z there exists natural resources degradation risk, and there is no evidence that in zone Z there exists drought risk.

For the following rule:

$$\begin{aligned}
\texttt{potentialProductivity}(Z, low) \leftarrow\ & \texttt{zone}(Z),\\
& (\texttt{growingSeasonDays}(Z, short) \vee\\
& \texttt{soilDepthCentimeter}(Z, shallow)\\
& \vee\ \texttt{abnormal}(Z)).
\end{aligned} \tag{3}$$

It can be shown that $c \leftarrow l_1, \ldots, l_m, (l_{m+1} \vee \ldots \vee l_{m+n})$ where l_i are literals for $n, m \geq 0$ can be replaced by the rules contained in the set $\{c \leftarrow l_1, \ldots, l_m, l_{m+j} | 1 \leq j \leq n\}$.

If we have the predicate $\texttt{growingSeasonDays}(Z, LengthDays)$ it is possible to create a rule that discern the value for $\texttt{growingSeasonPeriod}(Z, value)$ from this predicate. For example:

$$\begin{aligned}
\texttt{growingSeasonPeriod}(Z, long) \leftarrow\ & \texttt{Zone}(Z),\\
& \texttt{growingSeasonDays}(Z, LengthDays),\\
& LengthDays \geq 146.
\end{aligned} \tag{4}$$

For the sentence [15][14]:

We determined the growing season (duration, start and end), which is the number of days during the year in which there is availability of water or moisture and a suitable temperature for crop development.

we propose the rule:

$$\begin{aligned}\texttt{growingSeasonDays}(Z, LDaysGSD) \leftarrow\ & LDaysGSD := max\{LengthDays|\texttt{Zone}(Z),\\ & \texttt{growingSeason}(Z, StartDay, FinishDay),\\ & LengthDays := FinishDay - StartDay,\\ & \ not\ \texttt{waterAvailabilityInadequate}(Z,\\ & StartDay, FinishDay),\\ & \ not\ \texttt{temperatureInadequate}(Z,\\ & StartDay, FinishDay)\}.\end{aligned} \tag{5}$$

where the last rule can be read as: the growing season days length for a given zone Z is the maximum growing season days period length for a zone Z from the starting day to the finishing day of the period, in which for this zone Z and this period, there is no evidence that there exists water availability inadequate, and there is no evidence that there exists temperature innadequate.

Equation (5) can be easily translated to the following rules:

$$\begin{aligned}\texttt{setConditions}(Z, LengthDays) \leftarrow\ & \texttt{Zone}(Z),\\ & \texttt{growingSeason}(Z, StartDay, FinishDay),\\ & FinishDay := LengthDays + StartDay,\\ & \ not\ \texttt{waterAvailabilityInadequate}(Z, StartDay, FinishDay),\\ & \ not\ \texttt{temperatureInadequate}(Z, StartDay, FinishDay).\\ \texttt{betterSetConditions}(Z, LnDays) \leftarrow\ & \texttt{setConditions}(Z, LnDays),\\ & \texttt{setConditions}(Z, Y), LnDays < Y.\\ \texttt{growingSeasonDays}(Z, LnDays) \leftarrow\ & \texttt{setConditions}(Z, LnDays),\\ & \ not\ \texttt{betterSetConditions}(Z, LnDays).\end{aligned} \tag{6}$$

The condition about water availability inadequate is present for a period of time $[StartDay, FinishDay]$ and a zone Z if there exist evidence that for zone Z water resources are inadequate, or for zone Z in the period $[StartDay, FinishDay]$ the humidity is inadequate. So the following rule is generated:

$$\begin{aligned}\texttt{waterAvailabilityInadequate}(Z, StartDay, FinishDay) \leftarrow\ & \texttt{Zone}(Z),\\ & (\texttt{waterResoursces Inadequate}(Z) \lor\\ & \texttt{humidityInadecuate}(Z, StartDay, FinishDay)).\end{aligned} \tag{7}$$

The definition of temperature inadequate given for a zone Z for a period of time $[StartDay, FinishDay]$ is defined using the given temperature for every day in the given zone and period. The temperature for a given day and zone is part of extensional data base.

Humidity inadecuate given for a zone Z for a period of time $[StartDay, FinishDay]$ is present if for the zone Z and period of time $[StartDay, FinishDay]$ there exist excess of humidity or deficit of humidity. This gives the following rule:

$$\begin{aligned}\texttt{humidityInadecuate}(Z, StartDay, FinishDay) \leftarrow\ & \texttt{Zone}(Z),\\ & (\texttt{excessHumidity}(Z,\\ & StartDay, FinishDay) \lor\\ & \texttt{deficitHumidity}(Z,\\ & StartDay, FinishDay)).\end{aligned} \tag{8}$$

For the sentence [15][14]: *We estimated the water balance resulting from dividing the precipitation between the ETP, as support for determining the growing seasons and periods of excess and water deficit.*

$$\begin{aligned}\texttt{waterBalance}(Z, Day, Value) \leftarrow\ & \texttt{Zone}(Z),\\ & \texttt{potentialEvapotranspiration}(Z, Day, PETValue),\\ & \texttt{precipitation}(Z, Day, PrecipitationValue),\\ & PrecipitationValue := Value * PETValue.\end{aligned} \tag{9}$$

Rule defined in (9) there is a division that can be rewritten in terms of multiplication, in order to simulate the multiplication in answer set, multiplication can be replaced by a predicate `multiplication(A,B,AB)` in the extensional data base where the third argument is the result of multiplicate the first two arguments.

Definition of predicates `growingSeason`$(Zone, StartDay, FinishDay)$, `excessHumidity`$(Zone, StartDay, FinishDay)$ and `deficitHumidity`$(Zone, StartDay, FinishDay)$ depends on the definition of `waterBalance`$(Zone, Day, Value)$ (rule (9)). Works [8][15][14] does not provide the way how they are related, however we believe that is easy to construct the definition of these predicates if they are defined as an statistical function.

For the sentence [15][14]:

For each year level and estimated daily potential evapotranspiration (ETP) from the evaporation data multiplication by a constant factor of 0.75. The ETP is multiplied by the constant factor of 0.5. We identified periods of moisture surplus or deficit (duration, start and end).

We model the rule:

$$\begin{aligned}\texttt{potentialEvapotranspiration}(Z, Day, Value) \leftarrow\ & \texttt{Zone}(Z)\\ & \texttt{evaporation}(Z, Day, EvaValue),\\ & Value := EvaValue * 0.75 * 0.5.\end{aligned} \tag{10}$$

In (10) there is a multiplication of rational numbers, since $0.75 \cdot 0.5 = 0.375$, it is possible to handle rational numbers knowing the range value of evaporation for every day in a similar way that in rule (9) and mapping each rational number to an integer proportional value. Potential evapotranspiration for a year is defined in terms of potential evapotranspiration of a day by a statistical function.

For the sentences [15][14]:

- *If the slope is less than 15 percent to prevent soil erosion and degradation of natural resources in general.*
- *If there are no droughts lasting over 45 days.*
- *If the probability of occurrence of a drought is not greater than 66 percent.*

the following rules are generated:

$$\begin{aligned}\texttt{soilErosion}(Z) \leftarrow\ & \texttt{zone}(Z)\\ & \texttt{terrainSlopePercent}(Z, S), S > 15.\\ \texttt{naturalResourcesDegradation}(Z) \leftarrow\ & \texttt{zone}(Z)\\ & \texttt{terrainSlopePercent}(Z, S), S > 15.\end{aligned} \tag{11}$$

$$\begin{aligned}\texttt{droughtChance}(Z) \leftarrow\ & \texttt{zone}(Z)\\ & \texttt{longerDroughtLengthDays}(Z, S), L > 45\\ & \texttt{droughtProbabilityPercent}(Z, E), E > 66.\end{aligned} \tag{12}$$

Rules in (11) are defined in terms of terrain slope, however this definition can be extended with more predicates related with this definition. In (12) definition of `longerDroughtLengthDays` can be declared in a similar way than in rule (6).

The predicate droughtProbabilityPercent can be defined as the top value of the range probability for the number of years a drought may occur for a given zone Z. For example:

$$\begin{aligned}\texttt{droughtProbabilityPercent}(Z,45) \leftarrow\ & \texttt{zone}(Z),\\ & \texttt{numberOfYearsDroughtMayOccur}(Z,4).\end{aligned} \tag{13}$$

4.3 Modeling of maizetech fit zone problem using Answer Set

In [15][14] once is determined for a given geographical zone the potential of production in order to cultivate native maize seeds, given some diagnosis data by the user (culture cycle, moisture regime, altitude, rainfed quality, pending, organic matter, soil texture), the system computes what seeds are more appropriate to plant in the geographical zone. Since in [15][14] is not clearly stated how this part is modeled, we evaluated of possible chances from the system by manual inspection and using data mining techniques (C4.5 algorithm) a decision tree, from this tree we can build rules as the following one:

$$\begin{aligned}\texttt{seed}(v537c) \times \texttt{seed}(v538c) \times \texttt{seed}(hv521c) \leftarrow\ & \texttt{locality}(L),\\ & \texttt{organicMatter}(L, OrgMatPercent),\\ & OrgMatPercent > 3,\\ & \texttt{pending}(L, PenPercent),\\ & PenPercent < 5,\\ & \texttt{rainfedQuality}(L, highRiskOfDrought),\\ & \texttt{altitude}(L, AltValue),\\ & AltValue < 1200,\\ & \texttt{cultureCycle}(L, springSummer).\end{aligned} \tag{14}$$

Note that there is a preference depending of the class of maize seed.

5 Conclusions

In this work we provide ASP modeling for AgriFitZone problem. We propose to use an ASP approach using preferences. Once necessary and sufficient conditions have been determined to give solution for AgriFitZone, it is possible to provide assistance to farmers in making decisions to promote planting and conservation of maize through the implementation of the mathematical model (in this case using ASP) according to the classes of producers, environmental conditions and geographical location. In future work we plan to research about planning on agricultural cultivation of maize using possibilistic logic programs.

References

1. Chitta Baral. *Knowledge Representation, reasoning and declarative problem solving with Answer Sets*. Cambridge University Press, Cambridge, 2003.
2. Elizabeth M. Boll and Daniel C. St. Clair. Analysis of rule sets generated by the cn2, id3, and multiple convergence symbolic learning methods. In *CSC '95: Proceedings of the 1995 ACM 23rd annual conference on Computer science*, pages 48–55, New York, NY, USA, 1995. ACM.
3. Gerhard Brewka, Ilkka Niemelä, and Tommi Syrjänen. Logic Programs with Ordered Disjunction. *Computational Intelligence*, 20(2):333–357, 2004.

4. Gerhard Brewka, Ilkka Niemelä, and Mirosław Truszczyński. Preferences and Nonmonotonic Reasoning. *AI Magazine*, 29(4):69–78, 2008.
5. Pedro Reyes C. *El Maiz y su cultivo*. AGT Editor, S. A., Mexico, D.F., first edition, 1990.
6. B. A. Davey and H. A. Priestly. *Introduction to Lattices and Order*. Cambridge University Press, second edition, 2002.
7. James Delgrande, Torsten Schaub, Hans Tompits, and Kewen Wang. A classification and Survey of Preference Handling Approaches in Nonmonotonic Reasoning. *Computational Intelligence*, 20(2):308–334, 2004.
8. Hilario García-Nieto, Oscar J. García-Silva, Ceferino Ortiz-Trejo, Román de Jesús Barajas, and Germán Sánchez-Altamirano. *Sistema de Consulta del Potencial Agroproductivo del Estado de Puebla*, 2002. December, 2002. INIFAP and FUPPUE and SDR.
9. Michael Gelfond and Nicola Leone. Logic programming and knowledge representation - the a-prolog perspective. *Artificial Intelligence 138(1-2)*, pages 3–38, 2002.
10. Michael Gelfond and Vladimir Lifschitz. The stable model semantics for logic programming. In R. Kowalski and K. Bowen, editors, *5th Conference on Logic Programming*, pages 1070–1080. MIT Press, 1988.
11. Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. The weka data mining software: an update. *SIGKDD Explor. Newsl.*, 11(1):10–18, 2009.
12. Nicola Leone, Thomas Eiter, Cristinel Mateis, Gerald Pfeifer, Georg Gottlob, Simona Perri, Wolfgang Faber, Christoph Koch, and Francesco Scarcello. The dlv system for knowledge representation and reasoning, 2002. `http://www.dbai.tuwien.ac.at/proj/dlv/`.
13. John W. Lloyd. *Foundations of logic programming*. Springer-Verlag New York, 1987.
14. Walter Lopez-Baez and et al. Zonas potenciales y recomendaciones tecnicas para producir maiz en chiapas, mexico. Technical report, INIFAP Chiapas, Mexico, March 2007.
15. Walter Lopez-Baez, Aurelio Lopez-Luna, Bulmaro Coutiño-Estrada, Robertony Camas-Gomez, Bernardo Villar-Sanchez, Jaime Lopez-Martinez, Victor Serrano-Altamirano, Ausencio Zamarripa-Moran, Pedro Cadena-Iñiguez, and Carlos Sandoval-Morado. *Sistema de Zonas Potenciales y Recomendaciones tecnicas para la Produccion Sustentable de Maiz en el estado de Chiapas, Mexico*, 2007. INIFAP and Fundacion Produce Chiapas, A. C.
16. Luis A. Montiel, Claudia Zepeda, Pedro Wesche, and Javier A. Cervantes. Assistance in making decisions to promote planting and conservation of maize in the state of puebla of mexico through asp. Technical report, Universidad de las Americas Puebla, August 2010.
17. J. Ross Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, CA, 1993.
18. Gerardo Reyes, Jimena Guerra, and Gilberto Calderón. Condiciones de cultivo del maíz criollo en comunidades de puebla, tlaxcala e hidalgo: Un análisis de las economias de autoconsumo. In BUAP, editor, *Aportes*, volume 10, pages 63–82. Benemerita Universidad Autónoma de Puebla, Puebla, Mexico, first edition, May 2005.
19. Tommi Syrjnen and Ilkka Niemel. The smodels system, 2001. `http://www.tcs.hut.fi/Software/smodels/`.
20. D. Vega-Valdivida and P. Ramírez-Moreno. Situación y perspectivas del maíz en méxico. Research report, Universidad de Chapingo, March 2004.
21. Claudia Zepeda, Mauricio Osorio, Juan Carlos Nieves, Christine Solnon, and David Sol. Applications of Preferences using Answer Set Programming. In *Proc. of the 3rd Intl. Workshop in Advances in Theory and Implementation*, volume 142 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2005.