Probabilistic Reasoning with Continuous Constraints

E. Carvalho, J. Cruz and P. Barahona

Centro de Inteligência Artificial, Universidade Nova de Lisboa, Caparica, Portugal

Abstract. Continuous constraint reasoning assumes the uncertainty of numerical variables within given bounds and propagates such knowledge through a network of constraints, reducing the uncertainty. In some problems there is also information about the plausibility distribution of values within such bounds. However, the classical constraint framework cannot accommodate that information. This paper describes how the continuous constraint programming paradigm may be extended, in order to accommodate some probabilistic considerations, bridging the gap between the pure interval-based approach, that does not consider likelihoods, and the pure stochastic approach, that does not guarantee the safety of the results obtained.

INTRODUCTION

In many areas of application, namely engineering systems [1], models must include quantitative information and hence, a set of equations and inequations on variables, ranging over bounded intervals in the reals. These constraints are subject to several sources of uncertainty, both on parameter values and on *states of nature* values for which only statistical information is available. This uncertainty may be arbitrarily magnified in models based on non linear constraints where even small variations in some variables may lead to very large variations of other variables.

Uncertainty can be handled in distinct ways. When safety is a major concern, computational models must predict any possible behavior. For that purpose, rather then associate approximate values to variables, intervals can be used to include all their possible values. This is the approach that initially led to the development of interval arithmetic [2] which has, more recently, been adopted in the continuous constraint framework [3, 4, 5, 6].

Nevertheless, this approach has a major pitfall as it considers the same likelihood for all values in the intervals. This may lead to great inefficiency if some costly decisions are taken due to very unlikely values of some variables. Hence, the alternative use of stochastic models, which aim at predicting the probability of the different scenarios. But the computation of these probabilities can be quite elusive. In simpler systems, intervals may be partitioned in sub-intervals and joint probabilities assigned to the sets of sub-intervals. These joint probabilities may even be simplified (as in Bayesian networks [7]) if some conditional independence assumptions are acceptable. However, this approach is neither practical nor feasible in more complex, non linear models.

This paper describes how the continuous constraint programming paradigm may be extended in order to accommodate some probabilistic considerations, bridging the gap between the pure interval-based approach that does not consider likelihoods and the pure stochastic approach that does not guarantee the safety of the results obtained. First the constraint programming framework and the basic ideas of probabilistic reasoning are presented. Then the proposed method is explained and, finally, conclusions and future work are discussed.

CONTINUOUS CONSTRAINT REASONING

A Constraint Satisfaction Problem (CSP) [8] is defined by a triple (X;D;C) where X is a set of variables, each with an associated domain of possible values in D, and C is a set of constraints on subsets of the variables. A constraint specifies which values from the domains of its variables are compatible. A solution to the CSP is an assignment of values to all its variables, which satisfies all the constraints. In continuous CSPs (CCSPs) [3, 4, 5] variable domains are continuous real intervals and constraints are equalities and inequalities. The space of possibilities is represented by boxes, i.e., the Cartesian product of real intervals. The CCSP framework is powerful enough to model a wide range of problems. In particular, engineering systems with components described as sets of continuous valued variables and relations defined by numerical equalities or inequalities, eventually with uncertain parameters. Continuous constraint reasoning aims at eliminating value combinations from the initial search space (the Cartesian product of the initial domains), without loosing solutions. It combines pruning and branching steps until a stopping criterion is satisfied.

The pruning of the variable domains is based on constraint propagation. The main idea is to use the partial information expressed by a constraint to eliminate some incompatible values from the domain of its variables. Once the domain of a variable is reduced, this information is propagated to all constraints with that variable in their scopes. The process terminates when the domains can not be further reduced by any constraint. Safe narrowing functions (mappings between boxes) are associated with constraints, to eliminate incompatible value combinations. Efficient methods from interval analysis (e.g. the interval Newton [2]) are often used to implement efficient narrowing functions which are correct (do not eliminate solutions) and contracting (the box obtained is smaller or equal than the original one).

Constraint propagation is a local consistency algorithm for pruning the variable domains, which is often insufficient to support safe decisions. To obtain better pruning, it is necessary to split the boxes and reapply constraint propagation to each sub-box. Such branch and prune process enforces a stronger, non local, consistency criterion. Several consistency criteria have been proposed [3, 4, 9], offering distinct trade-offs between efficiency and pruning.

In the classical CCSP framework, the uncertainty associated with the problem is modelled by using intervals to represent the domains of the variables. Constraint reasoning reduces uncertainty, by *reshaping* the search space to become a safe approximation of the solution space. However, in many cases, safe reasoning is useless, intervals are often very wide, and subsequent constraint propagation is not able to narrow them. In fact, an uncertain value may range over a wide interval but a much narrower interval may include the most likely values.

In some problems, the plausibility distribution of values within the bounds of an uncertain parameter, is also known. For instance, uncertainty due to measuring errors may be naturally associated with an error distribution. However, the traditional CCSP framework cannot accommodate such information and thus, for each variable, all values in its domain are considered equally plausible. Consider the problem characterized by the constraints (from [10]):

$$x^2y^2 + x^2 + y^2 + k_1 = k_2xy \qquad x^2z^2 + x^2 + z^2 + k_1 = k_2xz \qquad y^2z^2 + y^2 + z^2 + k_1 = k_2yz$$

where $k_1 = 13$, $k_2 = 24$ and variables x, y, and z range within the interval $[0,100]$. Using the CCSP framework the following solution space enclosure is obtained, where $I_1 = [0.3,0.4]$, $I_2 = [0.7,0.8]$, $I_3 = [4.6,4.7]$ and $I_4 = [10.8,10.9]$:

$$\langle x,y,z\rangle \in \{\langle I_1,I_3,I_3\rangle, \langle I_2,I_2,I_2\rangle, \langle I_2,I_2,I_4\rangle, \langle I_2,I_4,I_2\rangle, \langle I_3,I_1,I_3\rangle, \langle I_3,I_3,I_1\rangle, \langle I_3,I_3,I_3\rangle, \langle I_4,I_2,I_2\rangle\}$$

For illustration purposes, the obtained intervals width is 0.1, but the framework is powerfull enough to narrow them up to a 10^{-12} precision. If some uncertainty on the constant values is considered, it could only be represented by an interval, even when a plausibility distribution is available. Such uncertainty implies a wider solution space. Assuming $12 \leq k_1 \leq 14$ and $23 \leq k_2 \leq 25$, the possible values, for each variable, are dispersed over a wider range (between 0 and 11), and there is no way to distinguish the values there within.

PROBABILISTIC REASONING

Probability provides a classical model for dealing with uncertainty. The basic element of probability theory is the random variable, which plays a similar role to that of the CSP variables. Each random variable has a domain where it can assume values. In particular, continuous random variables assume real values. A possible world, or atomic event, is an assignment of values to all the variables of the model. An event is a set of possible worlds. The complete set of all possible worlds in the model is the sample space. If all the random variables are continuous, the sample space is the hyperspace obtained by the Cartesian product of the variable domains, and the possible worlds and events are, respectively, points and regions from such hyperspace.

Probability measures may be associated with possible worlds or events. In the continuous case, an assignment of a probability to a point, is representative of the likelihood in its neighborhood. A probabilistic model is an encoding of probabilistic information, allowing to compute the probability of any event, in accordance with the axioms of probability. The usual method for specifying a probabilistic model employs, either explicitly or implicitly, a full joint probability distribution, which assigns a probability measure to each possible world.

Probabilistic reasoning aims at incorporating new information, known as evidence, by updating an *a priori* probability into an *a posteriori* probability given the evidence. The *a priori* probability is a description of what is known in the absence of the evidence. For incorporating this evidence, conditioning is used. Conditional probability $P(A|B)$ is the probability of some event A, given the occurrence of some other event B. The *a posteriori* probability is the conditional probability when the relevant evidence is taken into account.

Probabilistic graphical models [11] (Markov networks and Bayesian networks [7]) provide a powerfull framework for efficient probabilistic reasoning. The idea is to use a probabilistic network that captures the structural properties of the probabilistic model (such as conditional independence) and defines an implicit full joint probability distribution.

Given new evidence (information about some nodes), belief propagation [7] is one of the most efficient inference algorithms to compute *a posteriori* probabilities for all the non-evidence nodes in the network.

Bayesian networks require the full specification of a conditional probability at each node of the network. These can be defined by a conditional probability table (CPT) when all the random variables are discrete. If the variables are continuous, such specification cannot be made using a table, since there would be infinite combinations of values. One technique used to handle continuous variables is discretization, by dividing each domain into a fixed set of intervals. However, discretization is often intractable, since it results in considerable loss of accuracy and very large CPTs. The problem is aggravated when there is a non linear relation between the variables since, small changes in the value of one variable, may induce large changes in the others. In this case the accuracy is dramatically lost when, instead of real values, intervals are considered (even very narrow intervals).

Other techniques to handle continuous variables are based on standard families of probability density functions (e.g. Gaussian distribution) that are specified by a finite number of parameters (e.g. mean and variance). However, when nodes are related by complex non linear relations, such approach cannot be adopted.

PROBABILISTIC REASONING WITH CONTINUOUS CONSTRAINTS

A Probabilistic Continuous Constraint Satisfaction Problem (PCCSP) is an extension of a CCSP, defined by (X;D;F;C), where X is a set of continuous random variables, each with an associated interval domain of possible values in D, distributed accordingly to the corresponding probability density function (p.d.f.) in F, and C is a set of constraints on subsets of the variables. Given a point in the domain of a random variable, its p.d.f. is representative of the *a priori* probability that the variable takes a value in the neighborhood of that point, without considering the relations between the variables. It is assumed that all relevant relations between variables are expressed by the constraints of the model. Thus, when the constraints are not accounted for, the variables are independent.

The initial search space represents a probability space, characterized by a joint p.d.f. which, due to the independence assumption, is implicitly defined by the product of the individual p.d.f.s of the random variables. In the process of reducing uncertainty, there is a combination of continuous constraint reasoning and probabilistic reasoning. While the first reduces uncertainty by *reshaping* the search space, the second *redefines* the search space *a priori* probability distribution by computing an *a posteriori* distribution, based on the constraint reasoning outcome.

The constraints are the new information that is incorporated in the probabilistic model. The solution space is the event containing all possible worlds that satisfy the constraints. Through constraint reasoning an approximation (enclosure) of the solution space is obtained. Therefore the *a posteriori* probability is computed as a conditional probability, given the evidence represented by the approximation of the solution space. This probability is calculated by the conditional probability rule $P(A|B) = P(A \cap B)/P(B)$. The probability of region A given the evidence, is the probability of the subregion of A contained in the approximation of the solution space, divided by a normalizing factor.

The quality of the solution space approximation depends on the consistency and stopping criteria used in the constraint reasoning process. If some regions of the search space were not pruned during constraint reasoning they may contain solutions, although there is no guaranty that they do. In fact, there is no knowledge why such regions are maintained. Was it due to lack of further exploration of this regions or did they contain solutions? Normally, the process of constraint reasoning, leads to non uniform sizes of the boxes that represent the solution space approximation. Nevertheless, for reasoning with probabilistic information, some kind of fairness in the exploration of the search space must be guaranteed, so that the obtained *a posteriori* distribution is not biased by heterogeneous search.

In the present PCCSP implementation the stopping criterion is based on the maximum width of the intervals that constitute a box. Assuming ε_1 is the maximum width allowed then, when all the intervals of a box are smaller or equal to ε_1, that box is no further explored. When all the boxes meet this criterion the search stops. The consistency adopted is based on the interval Newton method. The stopping criterion assures some uniformity of the solution space approximation. However, some heterogeneity is still present due to the consistency enforcement algorithm, because the narrowing ability of any such algorithm differs between distinct regions of the search space.

To maintain a generic non parametric representation of the *a posteriori* marginal p.d.f.s, some kind of discretization must be assumed. This is achieved by considering an hypergrid of width ε_2, i.e., a grid where the dimension is the number of variables in the PCCSP, and each grid unit has width ε_2 in all dimensions. The hypergrid allows to transform the non uniform solution space approximation, resulting from constraint reasoning, in a uniform one, providing a fair computation of the marginal p.d.f.s. The transformation is achieved by overlaying the hypergrid upon the solution space approximation, enforcing a *snap to grid* to this region. The new approximation is the set of grid hypercubes that intersect with the original one. For every random variable, the marginal p.d.f. is discretized accordingly to the hypergrid. Each function segment thus obtained, has a probability value computed by summing up the contribution of all hypercubes that are aligned with that segment (normalized by the sum of all hypercube contributions). The contribution of an hypercube is obtained by computing the joint p.d.f. value, for the event that it represents.

The sequence of images in Fig. 1 illustrates the described probabilistic constraint reasoning process.

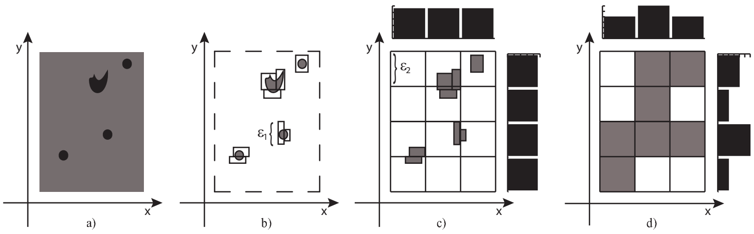

FIGURE 1. Process of probabilistic constraint reasoning. (a) Initial search space and solution space; (b) Solution space approximation; (c) Hypergrid and *a priori* marginal p.d.f.s; (d) Snap to grid and *a posteriori* marginal p.d.f.s.

In the problem previously introduced, the PCCSP framework can accommodate the plausibility distribution and, the obtained results, are far more informative. Figure 2a, illustrates the *a posteriori* distributions obtained with constant k_1 and k_2 and x, y and z *a priori* uniformly distributed. The bars height reflect the solution density on the corresponding range. Figure 2b, illustrates the *a posteriori* distributions obtained when *a priori* triangular distributions are added to model k_1 and k_2 uncertainty. The bars identify the more promising regions, within the narrowed domains.

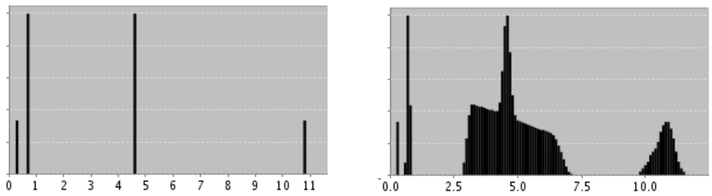

FIGURE 2. Preliminary results. (a) Uniform distributions and constant k_1, k_2; (b) Triangular distributions for k_1 and k_2

CONCLUSIONS AND FUTURE WORK

This paper presents a new framework that combines continuous constraint reasoning with probabilistic reasoning. Constraint reasoning reduces uncertainty by *reshaping* the search space. Probabilistic reasoning *redefines* an *a priori* probability into an *a posteriori* probability, based on the constraint reasoning outcome. Preliminary results are encouraging, illustrating the approach potential, in cases where classical alternatives could not be easily accommodated. We aim at extending the expressive power of the proposed framework, to cope with preference reasoning. Furthermore, we intend to develop an interactive prototype, to improve usability and fully explore the framework capabilities.

REFERENCES

1. J. Bellone, A. Chamard, and A. Fischler, "Constraint Logic Programming Decision Support Systems for Planning and Scheduling Aircraft Manufacturing at Dassault Aviation," in *Proc. of the 3rd ICPAP*, Paris, 1995, pp. 63–67.
2. R. Moore, *Interval Analysis*, Prentice-Hall, Englewood Cliffs, 1966.
3. O. Lhomme, "Consistency Techniques for Numeric CSPs," in *Proc. of the 13th IJCAI*, 1993, pp. 232–238.
4. F. Benhamou, D. McAllester, and P. van Hentenryck, "CLP(intervals) revisited," in *ISLP*, MIT Press, 1994, pp. 124–138.
5. D. Sam-Haroud, and B. Faltings, *Constraints* **1**, 85–118 (1996).
6. J. Cruz, and P. Barahona, *Applied Numerical Analysis Computational Mathematics* **1**, 140–154 (2004).
7. J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann, 1988.
8. U. Montanari, *Information Sciences* **7**, 95–132 (1974).
9. J. Cruz, and P. Barahona, "Maintaining Global Hull Consistency w/ Local Search for CCSPs," in *COCOS*, 2002, pp. 178–193.
10. Frisco project (1999), URL `www-sop.inria.fr/saga/POL/BASE/2.multipol/cyclohexan.html`.
11. S. L. Lauritzen, *Graphical Models*, Oxford Statistical Science Series 17, Clarendon Press, Oxford, 1996.