Metric SCSPs: Partial Constraint Satisfaction via Semiring CSPs augmented with metrics

Aditya Ghose and Peter Harvey

Decision Systems Laboratory
School of IT and Computer Science
University of Wollongong
NSW 2522 Australia

Abstract. The Semiring CSP (SCSP) framework is a popular and robust approach to solving partial constraint satisfaction problems which generalizes several other schemes such as fuzzy CSP, weighted CSP etc. We argue in this paper that it is useful to augment the SCSP framework such that each constraint specifies, in addition, a metric on the semiring values. The additional knowledge of distances between 'preference values' (the elements of the semiring) permits us to define a notion of parameterized solving of SCSPs where we can seek solutions with a preference value no worse than a given value.

1 Introduction

There has been considerable recent interest in Semiring Constraint Satisfaction Problems (SCSPs) [BMR97]. SCSPs provide an elegant generalization of CSPs where instead of specific descriptions of preference levels or preference values (for example, the interval $[0, 1]$ for fuzzy CSPs) we have an abstract notion of preference values viewed as elements of a semiring[1]. The SCSP framework supports a fine-grained representation of preferences over tuples in finite-domain constraints, which may also be viewed as local specifications of optimization criteria. This offers a rich language for modelling a variety of different applications.

The notion of distance of a candidate solution to a partial constraint satisfaction problem from a given constraint (or combined over a set of constraints) has been considered in systems such as HCLP [WB93]. The integration of such notions with the SCSP framework can open up a variety of useful possibilities. We explore such an integration in this paper, specifically by extending the notion of a constraint in the SCSP scheme to include a metric on the semiring.

One way to view the formulation of classical CSP in the SCSP scheme [BMR97] is to think of a threshold of **1** (the 'best' value in the semiring) delineating the subset of tuples in the SCSP solution that are of interest. In other words, we wish to consider only solution tuples that have a semiring value no less than **1**. In general, an SCSP might not admit a solution containing tuples for which the corresponding semiring value is the

[1] The semirings of interest are *c-semirings* defined later in the paper, but we shall often informally refer to these simply as semirings.

‘best’ value. In response, Bistarelli et al [BMR97] define the notion of an *abstract solution*, which is the set of those solution tuples for which tuples with a ‘better’ semiring value do not exist. Thus a given problem might generate an abstract solution set consisting of multiple tuples, possibly with distinct semiring values which are all incomparable under the partial order generated by the semiring. We can consider this to be using **0** (the ‘worst’ value in the semiring) as the threshold, finding the ‘best’ solutions but able to accept any.

We propose instead to generalise the notion of a threshold which determines the solution tuples of interest by allowing any arbitrary semiring value α to be specified in place of **0**. The problem thus becomes one of seeking solution tuples whose semiring values are ‘no worse’ than α, with a system of relaxation in the case where the threshold is not achievable. Such an approach is useful if we are interested in solutions which may potentially violate[2] some constraints, with the caveat that the degree of violation (informally speaking) is minimized. We are thus able to answer the following question: if the given problem were to be minimally changed to obtain at least one solution tuple with a semiring value no worse than α, which tuple(s) would meet this requirement? Note that by adding a parameter α, we have *not* converted the soft constraint solving problem (fundamentally an optimization problem) to a satisfaction problem. This is because we are still interested in seeking out the ‘best’ amongst those solutions that satisfy the threshold (should some exist), or the ‘best’ relaxation of the problem that enables us to satisfy the threshold.

If a given problem does in fact have solution tuples with semiring values no worse than α, the problem is easily solved. Things become difficult when there are no solution tuples which satisfy this requirement. We are then interested in identifying solution tuples which, if ‘promoted’ to be assigned the higher semiring value α would result in minimum deviation (in a sense similar to the intuitions use in the HCLP framework [WB93]) from the specified constraints. The SCSP formulation only offers a partial order for comparing semiring values, which is inadequate for defining a meaningful notion of minimum deviation. This motivates the extension to Metric SCSPs that we propose in this paper.

In Metric SCSPs a constraint is a triple consisting of the constraint signature, a function that maps each value in the cartesian product of the domains of the variables in the signature to a semiring value and a metric on the semiring. The ability to represent distances between semiring values permits the definition of a variety of distinct notions of what constitutes a solution to a given problem. We speculate that for many interesting instances of Metric SCSPs, it might be possible to translate the problem in polynomial time into a traditional SCSP, but observe that this is not true in general (this has to do with the our framework permitting non-binary functions for combining measures of deviation of individual constraints from candidate solutions). We also note that the polytime reduction is possible only for a fixed threshold α, whereas our motivation is to develop a system that permits exploration of the search space with alternative threshold values (we discuss later how it might be possible to define incremental resolving strategies in such settings).

---

[2] We say that an SCSP constraint is violated when we assign to a tuple a semiring value higher than that assigned by the constraint itself

Recall that a metric space consists of a set A and a function $d : A \times A \rightarrow \mathbb{R}^+$. To be seen as representing 'distance', the function d must satisfy 3 properties (a definition of a metric using only 2 properties is possible, but for clarity we use 3). These are $d(\alpha,\beta) = 0 \Leftrightarrow \alpha = \beta$, $d(\alpha,\beta) = d(\beta,\alpha)$, and $d(\alpha,\gamma) \leq d(\alpha,\beta) + d(\beta,\gamma)$.

Also recall that a c-semiring is defined as a tuple $S = \langle A, \oplus, \otimes, \mathbf{0}, \mathbf{1} \rangle$ satisfying the following properties (where $\alpha,\beta,\gamma \in A$):

- A is a set with $\mathbf{0}, \mathbf{1} \in A$
- $\oplus$ is commutative, associative, $\alpha \oplus \alpha = \alpha$, $\alpha \oplus \mathbf{1} = \mathbf{1}$ and $\alpha \oplus \mathbf{0} = \alpha$
- $\otimes$ is commutative, associative, binary, distributive ($\alpha \otimes (\beta \oplus \gamma) = (\alpha \otimes \beta) \oplus (\alpha \otimes \gamma)$), $\alpha \otimes \mathbf{1} = \alpha$ and $\alpha \otimes \mathbf{0} = \mathbf{0}$

We can derive the partial ordering $\leq_S$ from a c-semiring by $(\alpha \leq_S \beta) \Leftrightarrow (\alpha \oplus \beta = \beta)$. As a result of this definition, $\oplus$ and $\otimes$ are both monotone on $\leq_S$, $\langle A, \leq_S \rangle$ is a complete lattice and $\alpha \oplus \beta = \mathrm{lub}(\alpha,\beta)$. For a more complete analysis of c-semirings and SCSP see [BMR97].

Augmenting the SCSP formulation with metrics is a non-trivial exercise. We begin by postulating some intuitive properties that metrics on semirings must satisfy - these properties delineate the class of metrics that are in some sense 'compatible' with a given semiring. We provide examples of useful metrics which satisfy these properties and identify some additional (and potentially useful) properties that follow from our initial axioms. We then define two alternative notions of a parameterized solution to a Metric SCSP. In the first, we seek solution tuples which, if assigned a semiring value no worse than α, would result in minimum deviation from the original set of constraints. Minimum deviation is determined via the metrics specified in each constraint and any combination function from a broad-ranging class of combination functions that we delineate. We further narrow down this set of solution tuples by seeking the equivalent of an abstract solution within them (i.e., seeking those tuples for which no tuple with a 'better' semiring value exists within the set). In the second approach, we use the metrics and the combination function to filter out from the set of possibly many abstract solution tuples those which represent minimum deviation from the original problem.

Example 1. Consider the following informal constraint satisfaction problem:
Three people wish to purchase a car together (hatchback, sedan, stationwagon, or 4-wheel drive). We wish to find the type of car which satisfies the most people, taking the approach that constraints supplied by each person will create an over-constrained problem. Each person writes down a single constraint they place on the type of car to be purchased. This constraint must reflect their opinions on the different types of cars, expressed as answers to the following questions:

1. Sufficient style
2. Sufficient passenger space
3. Sufficient luggage space

To model this problem, we use a single variable X with domain $D = \{$ hatchback, sedan, stationwagon, 4-wheel-drive $\}$, and will ask each persion to supply a single preference constraint representing their responses to the above questions. However, we must first specify a c-semiring to use such that each user can express themselves sufficiently.

One reasonable option for modelling this problem is to provide a simple linear c-semiring with the exact semantics of fuzzy constraint satisfaction. However, the result is an immediate loss of information as the answers to each of the above questions is mapped (rather arbitrarily) into a linear domain. Essentially, the relaxation strategy has been determined (by the mapping into a linear domain, specifying which trade-offs should occur first) before the constraints were written. The constraints are no longer expressions of preferences but are directives specifying a problem-specific input to the relaxation strategy.

Another reasonable option for modelling this problem is to use the c-semiring formed as the product of 3 of the above c-semirings (where the product is that described in [BMR97]). Each component c-semiring corresponds to exactly one of the above questions. In this way we do not lose any information (the answers to each question are clearly available) but two flaws present themselves. Again the relaxation strategy has become tied to the constraints themselves, evidenced by the fact that after the constraints have been written the relative 'weights' of these constraints cannot be easily adjusted. Additionally, this model provides no facilities for the types of prioritisation seen in the previous model (questions cannot be given greater priorities). This shortcoming would be present even in the case of a non-idempotant combination operator.

For both of these models there is no way an individual can express the interchangability of two properties. For example, a user cannot express that (for a specific car) they are satisfied with the *total space* available, viewing the passenger and luggage space as equivalent. A more detailed explanation of this problem is presented below in Table 3.

We will now present an alternative construction using metrics which seperates the relaxation strategy from the 'constraints' themselves. To clearly demonstrate the motivation for our system we have chosen a simpler c-semiring, $S = \langle A, \oplus, \otimes, \mathbf{0}, \mathbf{1} \rangle = \left\langle \{0,1\}^3, \max, \min, \langle 0,0,0 \rangle, \langle 1,1,1 \rangle \right\rangle$. [3] Each element of the c-semiring represents a yes or no answer to each of the above questions, with $\langle 1,1,1 \rangle$ being 'yes' to all questions. The operations max and min operate on each element of the tuples independently, so for example $\max(\langle 0,0,1 \rangle, \langle 1,0,0 \rangle) = \langle 1,0,1 \rangle$ and $\min(\langle 0,1,1 \rangle, \langle 1,1,0 \rangle) = \langle 0,1,0 \rangle$.

We then modify the normal definition of a constraint in SCSP by including a metric. As a result each of the three constraints $\{c_i : 1 \leq i \leq 3\}$ (one constraint per person) contains:

- A set con_i indicating the variables for this constraint (in this case $\{X\}$).
- A function $def_i : D \to A$ expressing the individuals opinion for each type of car. These are specified in Table 1.
- A metric $d_i : A \times A \to \mathbb{R}^+$ expressing their perceived difference between c-semiring values.

We can summarise each individuals preferences as follows:

def_1: Is satisfied with a small car for transporting people and luggage, and is concerned that larger cars are not stylish.

def_2: Needs a larger car for transporting luggage. A sedan would be sufficient for transporting people, but not luggage.

---

[3] To model this problem more accurately it is expected that multiple values (for example the range $[0,1]$ as used in fuzzy CSP models) would be used per question.

Table 1. Constraint definitions

X	def_1	def_2	def_3
hatchback	$\langle 1,1,1\rangle$	$\langle 1,0,0\rangle$	$\langle 0,0,1\rangle$
sedan	$\langle 1,1,1\rangle$	$\langle 1,1,0\rangle$	$\langle 1,1,1\rangle$
stationwagon	$\langle 0,1,1\rangle$	$\langle 1,1,1\rangle$	$\langle 0,1,1\rangle$
4-wheel-drive	$\langle 0,1,1\rangle$	$\langle 1,1,1\rangle$	$\langle 0,1,1\rangle$

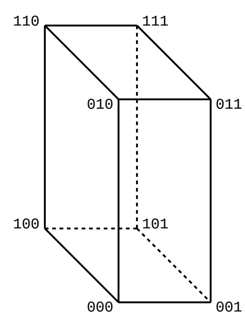
def_3: Considers a sedan to be the only stylish car. A hatchback has enough space for luggage, but is too small for passengers.

However, there is more information which can be revealed by having each person supply a metric over the c-semiring. We will supply two metrics which enhance the knowledge of a person's constraints, and explain the differences inferred by each.

Table 2. Metric used by first person

d_1	$\langle 0,0,0\rangle$	$\langle 0,0,1\rangle$	$\langle 0,1,0\rangle$	$\langle 0,1,1\rangle$	$\langle 1,0,0\rangle$	$\langle 1,0,1\rangle$	$\langle 1,1,0\rangle$	$\langle 1,1,1\rangle$
$\langle 0,0,0\rangle$	0	1	3	4	3	4	6	7
$\langle 0,0,1\rangle$	1	0	4	3	4	3	7	6
$\langle 0,1,0\rangle$	3	4	0	1	6	7	3	4
$\langle 0,1,1\rangle$	4	3	1	0	7	6	4	3
$\langle 1,0,0\rangle$	3	4	6	7	0	1	3	4
$\langle 1,0,1\rangle$	4	3	7	6	1	0	4	3
$\langle 1,1,0\rangle$	6	7	3	4	3	4	0	1
$\langle 1,1,1\rangle$	7	6	4	3	4	3	1	0

Fig. 1. Graphical representation



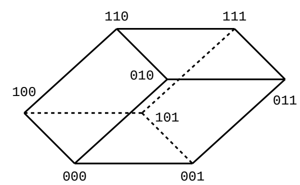
The metric d_1, detailed in Table 2, indicates that the person cares equally about style and passenger space, but cares little for luggage space. For example, the difference (or distance) between $\langle 0,0,0\rangle$ (not satisfied with anything) and $\langle 0,1,0\rangle$ (satisfied only with the passenger space) is 3. In contrast, the difference (or distance) between $\langle 0,0,0\rangle$ (not satisfied with anything) and $\langle 0,0,1\rangle$ (satisfied only with the luggage space) is 1, indicating that this person can see little difference between a car with insufficient luggage space and one with sufficient luggage space. This indifference may arise as the person does not carry luggage often, making the distinction between car luggage space possible but of little importance.

This particular choice of metric does not exhibit any additional information that we could not include in a SCSP formulation using a more sophisticated c-semiring. However, observe that in this metric each dimension of the c-semiring is treated as independent - the rating of a car's passenger space cannot be used to infer any rating on the luggage space. This is highlighted in Fig. 1 where the c-semiring is represented as a cube, with proportions according to the metric. This is important in light of the next metric which does not treat each dimension of the c-semiring as independent.

Table 3. Metric used by second person

d_2	$\langle 0,0,0\rangle$	$\langle 0,0,1\rangle$	$\langle 0,1,0\rangle$	$\langle 0,1,1\rangle$	$\langle 1,0,0\rangle$	$\langle 1,0,1\rangle$	$\langle 1,1,0\rangle$	$\langle 1,1,1\rangle$
$\langle 0,0,0\rangle$	0	1	1	2	1	2	2	3
$\langle 0,0,1\rangle$	1	0	1	1	2	1	2	2
$\langle 0,1,0\rangle$	1	1	0	1	2	2	1	2
$\langle 0,1,1\rangle$	2	1	1	0	3	2	2	1
$\langle 1,0,0\rangle$	1	2	2	3	0	1	1	2
$\langle 1,0,1\rangle$	2	1	2	2	1	0	1	1
$\langle 1,1,0\rangle$	2	2	1	2	1	1	0	1
$\langle 1,1,1\rangle$	3	2	2	1	2	1	1	0

Fig. 2. Graphical representation



By the metric d_2, detailed in Table 3, the 'rating' of a vehicle with sufficient luggage space and insufficient passenger space is similar to the rating of a vehicle with sufficient passenger space but insufficient luggage space. A possible explanation is that this person can see only a small difference between luggage space and passenger space - in many cases, luggage can be placed in what is normally deemed as passenger space. The areas deemed as passenger space and luggage space overlap in the mind of this person, and so the dimensions of the c-semiring are not seen as independent.

It is extremely difficult to capture the differences between these two constraints using existing partial constraint satisfaction schemes. Although SCSP can represent c_1, it would be difficult to represent c_2 using the same c-semiring. Specifically, the preference value $\langle 0,1,1\rangle$ has a different meaning in c_1 than in c_2. The differences between the interpretation of the preference values made in each constraint become particularly important when the threshold parameter α is set to a value other than $\langle 1,1,1\rangle$.

2 Metrics for a c-semiring

The system we propose centres on a function representing the distance of each value α in the c-semiring from each other value β. Such functions are termed metrics [Gil87]. This function is used for determining the 'size' of the differences in elements of a c-semiring. It is trivial to define such a function for a c-semiring. However, this function should meet certain goals:

– It should not be possible to define a function which could be used to map an n-dimensional c-semiring into an (n-1)-dimensional c-semiring.
– The distance between two distinct c-semiring values should not be zero. To say that the distance between two distinct c-semiring values is zero would undermine the purpose of the c-semiring.
– The function is intended to describe the difference between c-semiring values. Relationships between c-semiring values (derived from $\oplus$ and $\leq_S$) should be reflected in an intuitive way in the function, ensuring the 'meaning' of each value is kept.

We will call a function which provides the concept of 'distance' or 'difference' in a c-semiring a *c-metric*. Below we state and describe the formal properties of a c-metric $d : A \times A \rightarrow \mathbb{R}^+$:

M1: $(d(\alpha,\beta) = 0) \Leftrightarrow (\alpha = \beta)$

The differences between two c-semiring values should be zero iff the two values are the same. It may seem useful to relax this property, so as to indicate that two c-semiring values may be 'interchanged' without penalty. However, if it were relaxed an entire region of the c-semiring would become 'interchangeable' (due to properties P1 and P2 below). This region would contain its upper bound, and so the question would arise: why has the upper bound not been used in place of all other values in the region?

M2 (Symmetry): $d(\alpha,\beta) = d(\beta,\alpha)$

The function d should be defined to be the same in either 'direction'. This is obvious as d should indicate the differences between c-semiring values, which do not change if the direction of measurement is changed. Care must be taken to avoid considering the function d as a measure of 'penalty' for moving a semiring value 'upwards' in the ordering, in which case the direction of measurement would change the result.

M3 (Triangle inequality): $d(\alpha,\gamma) \leq d(\alpha,\beta) + d(\beta,\gamma)$

The shortest 'path' from a c-semiring value α to another value γ is the most direct path. By providing a geometric interpretation for d, this property helps ensure more intuitive results later.

P1 ($\leq_S$ Consistency): $(\alpha \leq_S \beta \leq_S \gamma) \Rightarrow (d(\alpha,\beta) \leq d(\alpha,\gamma))$

This property ensures that the interpretation of d as a measurement of distance is consistent with the $\leq_S$ ordering. Note that this property only guarantees that a sequence of increasing c-semiring values do not become closer to any lesser value.

P2 ($\oplus$ Consistency): $d(\alpha \oplus \gamma, \beta \oplus \gamma) \leq d(\alpha,\beta)$

This property guarantees that relationships between elements of the c-semiring (which indicate their similarity) are not discarded. To explain this property, we must discuss the purpose of $\oplus$, and how this relates it to d.

1. The $\oplus$ operation is intended to provide a projection of a constraint onto a smaller set of variables, constructing 'summary information'. This information is used in constraint propagation algorithms (and possibly other techniques such as heuristic search). To obtain 'maximum utility' from $\oplus$ the result of $\alpha \oplus \beta$ must provide the value which is most closely related to both α and β, yet which 'includes' (absorbs under addition) both.
2. Our use of metrics is based on the assumption that minimal deviation from constraints to reach a fixed level of consistency is our aim. Thus, our c-metric must be considered as a measurement of difference between two values, or decribing the total amount of 'trade-off' which occurs between two c-semiring values. This is a natural interpretation given the properties M1, M2 and M3.

Consider now the situation where $\alpha \oplus \gamma$ and $\beta \oplus \gamma$ differ more than α and β. For such a situation to arise, additional unrelated information must have been introduced by $\oplus$ when adding γ. Such additional information reduces the quality of the 'summary information' generated by $\oplus$, contrary to our assumption that $\oplus$ provides 'maximum utility'.

By this reasoning we can see that $\alpha \oplus \gamma$ and $\beta \oplus \gamma$ should differ less than α and β (the addition of γ should only increase the similarity). As our metric represents the 'difference' between elements we can conclude that $d(\alpha \oplus \gamma, \beta \oplus \gamma) \leq d(\alpha,\beta)$.

Although this property may seem unusually general, it does not severely restrict the choice of metrics that we would regard as reasonable.

The first 3 properties (M1, M2 and M3) are standard properties for a metric. They ensure the function d correctly encapsulates the standard notion of a distance for a set of values (in this case, the set A). The remaining properties (P1 and P2) provide consistency with the $\leq_S$ ordering, and the $\oplus$ operation. We have considered the purpose for $\oplus$ in the construction of these last 2 properties.

Definition 1. *A function $d : A \times A \rightarrow \mathbb{R}^+$ is a c-metric for the c-semiring $\langle A, \oplus, \otimes, \mathbf{0}, \mathbf{1} \rangle$ iff it satisfies the properties M1, M2, M3, P1, and P2.*

2.1 Examples of c-metrics

In this part of the paper we show that the restrictions (P1 and P2) placed on c-metrics do not block the usage of common metrics.

We cannot assume any restrictions on the type of c-semiring, and so any valid c-semiring (as defined in [BMR97]) should have at least one usable c-metric. Recall that the discrete metric can be described as $d(\alpha, \beta) = 0$ if $\alpha = \beta$, $d(\alpha, \beta) = 1$ if $\alpha \neq \beta$ [Gil87]. Using this definition it is possible to prove the following result.

Theorem 1. *The discrete metric is a c-metric for any c-semiring.*

In addition to the discrete metric, other known metrics also satisfy the conditions of P1 and P2. In particular, there are the large class of metrics for $\mathbb{R}^n$ derived from the p-norm. The p-norm [Gil87] is written $\| x \|_p$, where x represents a vector $\langle x_1, x_2, \ldots, x_n \rangle \in \mathbb{R}^n$ and $1 \leq p \leq \infty$. For any such vector x, its 'length' measured according to the p-norm is $\| x \|_p = \sqrt[p]{|x_1^p| + |x_2^p| + \ldots + |x_n^p|}$. So, we can form a standard metric using the concept of 'length' between two points: $d(\alpha, \beta) = \| \alpha - \beta \|_p$. We show that a c-semiring suitable for multi-criteria optimization can use this class of metrics.

Theorem 2. *For a c-semiring $\langle [0,1]^n, \sup, \inf, 0, 1 \rangle$, any metric $d(\alpha, \beta) = \| \alpha - \beta \|_p$ (where $1 \leq p \leq \infty$) is a c-metric.*

As d (defined using the p-norm) is a standard metric, properties M1, M2 and M3 are already proven as satisfied [Gil87]. For space considerations we have omitted the proof[4] for P1 and P2.

This theorem is useful for showing that the extra properties P1 and P2 do not exclude the usual set of metrics that could be used with such spaces. Although the theorem is stated for a Cartesian product of a fixed-size interval in $\mathbb{R}$, it applies equally for differing-sized and/or discrete intervals. This result further establishes that a range of standard c-metrics exist for multi-criteria optimization.

By Theorems 1 and 2 we have at least two possible c-metrics for the c-semiring $S = \langle [0,1], \sup, \inf, 0, 1 \rangle$. The discrete metric represents a simple, uniform penalty for any deviation from a specified point in $[0,1]$. In contrast, a metric based on a p-norm

[4] Note that proofs omitted from this paper are available in a longer technical version

represents a proportional scaled penalty for any such deviation. A linear combination of these two metrics would thus represent a uniform penalty (for the fact that deviation from a specified point has occured), plus an extra penalty based on the size of the deviation. By generating additional c-metrics from existing c-metrics on the same c-semiring, we significantly increase the possible choice of c-metrics for a given c-semiring. The most obvious (and useful) constructive methods for generating more c-metrics are the $+$, $\times$ and max of existing c-metrics.

Theorem 3. *Given any two c-metrics d_1 and d_2 for a c-semiring S, the combinations $d_1 + d_2$, $d_1 \times d_2$, and $\max(d_1, d_2)$ are also c-metrics for S.*

As d_1 and d_2 are both metrics, we know (by [Cop68]) that each of the above combinations are also metrics. Each proof that they satisfy P1 and P2 requires only a small sequence of algebraic manipulations and well-known inequalities.

In addition to constructing c-metrics from existing c-metrics defined over the same c-semiring, it is possible to construct c-metrics for the combination (Cartesian product) of two different c-semirings. We follow the method of combination of c-semirings set forth in [BMR97] and show that any weighted sum of c-metrics (valid for their respective c-semirings) forms a new c-metric for the combined c-semiring. This result allows the usual form of handling multi-criteria objectives (linear combination of numeric measures) to be simulated.

Theorem 4. *Given two c-semirings S_1 and S_2, and corresponding c-metrics d_1 and d_2, then a c-metric for the combined semiring S can be found by $d = xd_1 + yd_2$, where $x, y > 0$.*

As d_1 and d_2 are metrics it is already known that d is a metric, and so satisfies M1, M2 and M3. The proof that d satisfies P1 and P2 is based on simple linear inequality results. It is, however, too long to present here in detail and so has been omitted.

2.2 Properties of c-metrics

We will now detail certain properties resulting from the criteria for c-metrics, which will clarify the types of metrics suitable for use on c-semirings. We assume the use of a c-semiring $S = \langle A, \oplus, \otimes, \mathbf{0}, \mathbf{1} \rangle$ with a c-metric d. The majority of the resultant properties describe the distances between c-semiring values and higher c-semiring values.

We must define certain notation to be used; this notation is not unusual, but must be explicitly defined. In particular, Definition 3 below may appear obvious but alternative definitions can be found in other literature.

Definition 2. *If $S = \langle A, \oplus, \otimes, \mathbf{0}, \mathbf{1} \rangle$ is a c-semiring then we define $\hat{\beta} = \{\gamma \in A : \beta \leq_S \gamma\}$, where $\beta \in A$.*

Definition 3. *The distance from α to a region $\hat{\beta}$ is written $d(\alpha, \hat{\beta})$ and is defined as $d(\alpha, \hat{\beta}) = \inf\{d(\alpha, \gamma) : \gamma \in \hat{\beta}\}$ [Cop68].*

In this paper we have asserted that the c-semiring value $\alpha \oplus \beta$ should be that value greater than both α and β and which most accurately represents each one. Also, d has been characterised as the 'difference' between any two c-semiring values. It is therefore natural that, of all the values greater than β, $\alpha \oplus \beta$ should be that one closest to α.

Lemma 1. *The smallest distance from the c-semiring value* α *to any value in* $\hat{\beta}$ *is* $d(\alpha, \alpha \oplus \beta)$.

We can now define $d(\alpha, \hat{\beta}) = d(\alpha, \alpha \oplus \beta)$. For relaxation strategies in PCSP, we expect that solution tuples assigned lower c-semiring values are not given priority over solution tuples with higher c-semiring values. Translated into c-metric terms, we would expect that given two values α and γ with $\gamma \leq_S \alpha$, α would be at least as close to $\hat{\beta}$ as γ.

Theorem 5. *Given any c-semiring values* α, β *and* γ, *if* $\gamma \leq_S \alpha$ *then* $d(\alpha, \hat{\beta}) \leq d(\gamma, \hat{\beta})$.

3 Metric SCSP

We now investigate the application of a c-metric to the solutions of a SCSP. First we require an additional definition for a function used to combine the distances measured by c-metrics. To allow for many different relaxation strategies, this definition places few restrictions on the function.

Definition 4. *A function* $f : (\mathbb{R}^+)^m \to \mathbb{R}^+$ *combines distances if* $f(x_1, \ldots, x_m) = 0 \Leftrightarrow \forall i, x_i = 0$ *and it is monotonic increasing in each argument.*

We can now define a Metric SCSP $P = \langle C, con, f \rangle$ where con is a set of variables, and $C = \{c_1, c_2, \ldots, c_m\}$ is a set of constraints over a predefined c-semiring S (note that we permit multiple constraints on the same signature). Each constraint is a tuple $c_i = \langle con_i, def_i, d_i \rangle$ defining the variables to be operated on, a function mapping tuples to S, and a c-metric mapping $S \times S$ to $\mathbb{R}^+$. The function $f : (\mathbb{R}^+)^m \to \mathbb{R}^+$ is used for combining the results of d_i. Our aim is to find the solution(s) such that minimal deviation is required from the SCSP $\langle C, con \rangle$ while ensuring they are assigned a c-semiring value in $\hat{\alpha}$.

We assume that we have n variables which all take values from some set D. Thus the value for a solution $t \in D^n$, as defined for SCSPs, is $def(t) = def_1(t \downarrow^{con}_{con_1}) \otimes \ldots \otimes def_m(t \downarrow^{con}_{con_m})$.[5] Normally we would select the solutions $ASol(P) \subseteq D^n$ such that there exists no $u \in D^n$ assigned a greater value. We, however, are searching for those solutions needing 'minimal deviation'. From the definition of def we can prove certain results which make finding the solutions needing 'minimal deviation' easier.

Theorem 6. $(\alpha \leq_S def(t)) \Rightarrow \left(\alpha \leq_S def_i(t \downarrow^{con}_{con_i})\right)$

This theorem comes immediately from the fact that $\otimes$ is intensive. From this theorem we know that if $\alpha \leq_S def(t)$ then $d_i(t \downarrow^{con}_{con_i}) = 0$, for all i. As we expect $\otimes$ in our modified system to be idempotent, we will prove further results with that assumption.

---

[5] The notation $(t \downarrow^{con}_{con_i})$ is read as 'the projection of tuple t from the set con to the set con_i'

Theorem 7. *If $\otimes$ is idempotent,* $\left(\forall i, \alpha \leq_S def_i(t \downarrow^{con}_{con_i})\right) \Rightarrow (\alpha \leq_S def(t))$

If $\otimes$ is idempotent, then to ensure the value $def(t)$ is in $\hat{\alpha}$ we need only ensure that all $def_i(t \downarrow^{con}_{con_i})$ are also within $\hat{\alpha}$. We can thus determine a measure of the deviation from P (so that $def(t) \in \hat{\alpha}$) by measuring the distance from $def_i(t \downarrow^{con}_{con_i})$ to $\hat{\alpha}$.

Definition 5. *Define* $f_\alpha(t) = f(d_1(def_1(t \downarrow^{con}_{con_1}), \hat{\alpha}), \ldots, d_m(def_m(t \downarrow^{con}_{con_m}), \hat{\alpha}))$.

The function f_α determines (using the function f) the size of the deviation from P required to move $def(t)$ into the region $\hat{\alpha}$. The value of $f_\alpha(t)$ can also be interpreted as the amount of *belief change* required to allow solution t to be accepted with a level of consistency α.

Definition 6. *Define* $m_\alpha(P) = \min\{f_\alpha(u) : u \in D^n\}$

Definition 7. *Define* $m^*_\alpha(P) = \min\{f_\alpha(u) : u \in ASol(P)\}$

$m_\alpha(P)$ represents the minimum deviation from the problem P required to find any solution with a c-semiring value in $\hat{\alpha}$. Alternatively, $m^*_\alpha(P)$ represents the minimum deviation from the problem P required to find any abstract solution with a c-semiring value in $\hat{\alpha}$.

Definition 8. *The set of solutions needing least deviation from P to have values in $\hat{\alpha}$ is written as $ASol_\alpha(P)$, and is defined as follows (with $t, u \in D^n$):*

$$ASol_\alpha(P) = \{t : (f_\alpha(t) = m_\alpha(P)) \wedge (\nexists u, (def(t) <_S def(u)) \wedge (f_\alpha(u) = m_\alpha(P)))\}$$

Definition 9. *The set of abstract solutions needing least deviation from P to have values in $\hat{\alpha}$ is written as $ASol^*_\alpha(P)$, and is defined as follows (with $t, u \in ASol(P)$):*

$$ASol^*_\alpha(P) = \{t : (f_\alpha(t) = m^*_\alpha(P)) \wedge (\nexists u, (def(t) <_S def(u)) \wedge (f_\alpha(u) = m^*_\alpha))\}$$

If we require a specific level of consistency α for the problem P, then $ASol_\alpha(P)$ contains the 'best' solutions we can find for that purpose. Similarly, $ASol^*_\alpha(P)$ contains the 'best' abstract solutions we can find. To determine $ASol^*_\alpha(P)$ requires prior calculation of $ASol(P)$, reducing the complexity required to be similar to the existing SCSP framework. If $\otimes$ is idempotent we are able to utilise all of the local consistency results from [BMR97] to assist in finding $ASol(P)$. Alternatively, to determine $ASol_\alpha(P)$ we can use partial constraint satisfaction search techniques such as branch-and-bound search.

The definitions of $ASol_\alpha$ and $ASol^*_\alpha$ have been formulated to give two different methods of using c-metrics with SCSPs. However, we can still make the observation that $ASol^*_{\mathbf{0}}(P) = ASol_{\mathbf{0}}(P) = ASol(P)$ That is, if we can accept any level of consistency, then we should accept the standard notion of abstract solutions $ASol(P)$ as being the best solutions to our problem.

Unfortunately, due to the dependence of $ASol_\alpha$ on the parameter α we cannot possibly guarantee, for an arbitrary problem P, that $ASol_\alpha(P) \subseteq ASol(P)$. The elements of each set are determined by different methods (by the use of c-metrics, and by the direct use of the c-semiring respectively). These results simultaneously reduce the chance of applying methods for SCSPs to Metric SCSPs, and indicate that Metric SCSPs provide a facility not available otherwise.

Presume that f can be represented as repeated applications of a commutative binary operation. For example, if $f = \sum$ then it can instead be considered as repeated applications of $+$ to all it's arguments. In such a situation, and with a fixed α, a Metric SCSP reduces easily to a Valued CSP, with an additional selection scheme for the solutions with the minimal value. This result allows the use of constraint propagation techniques [Sch00] for Valued CSPs [BFM$^+$96] in the solving of a wide class of Metric SCSPs. Note that the ability to reduce to a Valued SCSP does not infer that Metric SCSPs provide no new expressive power. In the above result we made the assumption that α is fixed. If over time we wish to vary the value of α it is not easy to reduce a Metric SCSP to a Valued CSP. Herein lies the additional expressiveness and power of Metric SCSPs.

Presume that the value α represents a minimum value for multiple objective functions, where the objective functions are represented by the c-semiring. Presume also that the value α' represents a *new* minimum value which we would wish to use. Let V be the Valued CSP constructed using α and V' be the Valued CSP constructed from α'. As both are constructed using the same Metric SCSP it appears likely that incremental resolving techniques can be used to rapidly find a new solution to V' from the existing solutions to V. This can be considered similar to the modification of the objective function in linear programming where the previous solution provides a good starting position for the simplex method. Such a technique would be particularly useful in the light of the obvious practical applications of Metric SCSPs.

Metric SCSPs can be viewed as defining a class of Valued CSPs with intuitive and structured relationships between them. They allow the writing of 'constraints' in the highly expressive SCSP framework independent of the final goal or objective function. Existing CSP frameworks focus on the solution to a problem with a specific goal or objective. The Metric SCSP framework allows the variance of the goal in an intuitive fashion with sufficient restrictions to make incremental resolving techniques likely.

The application of these intuitions to parameterized notions of constraint theory maintenance in dynamic settings (e.g., parameterized constraint retraction) is being addressed in a separate paper. An implementation of a prototype Metric SCSP solver is also being planned.

References

[BFM$^+$96] S. Bistarelli, H. Fargier, U. Montanari, F. Rossi, T. Schiex, and G. Verfaillie. Semiring-based CSPs and Valued CSPs: Basic Properties and Comparison. In M. Jampel, E. Freuder, and M. Maher, editors, *Over-Constrained Systems (Selected papers from the Workshop on Over-Constrained Systems at CP'95, reprints and background papers)*, volume 1106, pages 111–150. 1996.

[BMR97] Stefano Bistarelli, Ugo Montanari, and Francesca Rossi. Semiring-based constraint satisfaction and optimization. *Journal of the ACM*, 44(2):201–236, 1997.

[Cop68] E. T. Copson. *Metric Spaces*. Cambridge University Press, 1968.

[Gil87] John R. Giles. *Introduction to the Analysis of Metric Spaces*. Number 3 in Australian Mathematical Society Lecture Series. Cambridge University Press, 1987.

[Sch00] Thomas Schiex. Arc consistency for soft constraints. In *Principles and Practice of Constraint Programming*, pages 411–424, 2000.

[WB93] Molly Wilson and Alan Borning. Hierarchical constraint logic programming. Technical Report TR-93-01-02, 1993.