Bounds and Pre-processing for Local Computation of Semiring Valuations[1]

Nic Wilson[2]

1 Introduction

Local Computation based on hypertree/join tree decompositions [10, 8, 5, 7] is a very general computational framework which leads to effective computational procedures in many different formalisms. The inputs are a multiset of objects called *valuations*. Each valuation is associated with a set of variables (its *scope*). In this paper we focus on semiring valuations: they are functions from the product set domain of this set of variables, which assign each such tuple an element in a semiring. A semiring is a set with two operations, one labelled 'addition', the other 'multiplication', which are both commutative and associative and are such that the multiplication distributes over the addition.

This is still general enough to include a wide range of different formalisms, including CSPs, Bayesian networks, and a general class of soft CSPs. Computation is often exponential in the size of the largest node of the join tree. This makes many classes of problems infeasible for the standard algorithms.

We consider different approaches to improve the efficiency of the computation. In section 4 a pre-processing scheme is described, where constraints propagation is performed first, which may make the overall computation easier, as it often reduces the number of non-zero values in the inputs.

We consider approaches to generating upper and lower approximations in section 5. There is a natural pre-order on a semiring which generates a natural pre-order on valuations with the same scope; combination and marginalisation respect this pre-order. This enables us to replace the harder computations (usually those on the largest nodes of the join tree) with simpler computations generating upper and/or lower bounds, which may be sufficient to answer a query of interest. A special type of lower bound is where we set some (of the smaller) values of the input valuations to be zero. We show that under certain conditions we still get useful information from the subsequent propagation.

2 Semirings

Tuple $\mathcal{A} = \langle A, \oplus, \otimes \rangle$ is said to be a *semiring* if A is a set and $\oplus$ and $\otimes$ are operations on A satisfying the following properties: $\otimes$ and $\oplus$ are both associative and commutative, and $\otimes$ distributes over $\oplus$ i.e., for all $\alpha, \beta, \gamma \in A$, $\alpha \otimes (\beta \oplus \gamma) = (\alpha \otimes \beta) \oplus (\alpha \otimes \gamma)$.

Element $\mathbf{0}$ of A is said to be a *zero element* if $\mathbf{0}$ is an additive identity element (so that for all $a \in A$, $a \oplus \mathbf{0} = a$) and for all $a \in A$, $a \otimes \mathbf{0} = \mathbf{0}$. A can contain at most one zero element.

[1] This work has received support from Science Foundation Ireland under Grant 00/PI.1/C075

[2] Cork Constraint Computation Centre, Department of Computer Science, University College Cork, Cork, Ireland, n.wilson@4c.ucc.ie

We can add a zero element to any semiring. Let $\langle A, \oplus, \otimes \rangle$ be a semiring. Add an extra element $\mathbf{0}$ to A and extend $\oplus$ and $\otimes$ to $A \cup \{\mathbf{0}\}$ by, for all $a \in A \cup \{\mathbf{0}\}$, $a \oplus \mathbf{0} = \mathbf{0} \oplus a = a$, and $a \otimes \mathbf{0} = \mathbf{0} \otimes a = \mathbf{0}$. It can easily be confirmed that $\langle A \cup \{\mathbf{0}\}, \oplus, \otimes \rangle$ is a semiring.

Element $\mathbf{1}$ of A is said to be a *unit element* if it is a multiplicative identity element, i.e., if for all $a \in A$, $a \otimes \mathbf{1} = a$. There can be at most one unit element in a semiring.

We may write a semiring $\mathcal{A}$ with a zero and unit element as $\langle A_{\mathcal{A}}, \mathbf{0}_{\mathcal{A}}, \mathbf{1}_{\mathcal{A}}, \oplus_{\mathcal{A}}, \otimes_{\mathcal{A}} \rangle$, or just $\langle A, \mathbf{0}, \mathbf{1}, \oplus, \otimes \rangle$.

Let $\mathcal{A} = \langle A, \oplus, \otimes \rangle$ be a semiring. We define relation $\preceq_{\mathcal{A}}$ on A (abbreviated to $\preceq$) by: $a \preceq b$ if and only if either $a = b$ or there exists $c \in A$ with $a \oplus c = b$.

Proposition 1 *For any semiring $\mathcal{A} = \langle A, \oplus, \otimes \rangle$, the associated relation $\preceq$ satisfies the following properties:*

(i) *Relation $\preceq$ is a pre-order, i.e., a reflexive transitive relation;*
(ii) *$a \preceq b$ and $a' \preceq b'$ implies $a \oplus a' \preceq b \oplus b'$ and $a \otimes a' \preceq b \otimes b'$;*
(iii) *if for all $i = 1, \ldots, k$, $a_i \preceq b_i$, then $\bigoplus_{i=1}^k a_i \preceq \bigoplus_{i=1}^k b_i$ and $\bigotimes_i a_i \preceq \bigotimes_i b_i$;*
(iv) *If $a \oplus a = a$ and $b \oplus b = b$ then $a \preceq b$ if and only if $a \oplus b = b$.*

Part (iv) implies that $\preceq$ partially orders the elements which are $\oplus$-idempotent. In particular, if $\oplus$ is idempotent then $a \preceq b$ if and only if $a \oplus b = b$, and so $\preceq$ is a partial order.

There is a sense in which we can always make $\preceq$ into a partial order. Let $\equiv$ be defined by $a \equiv b$ if and only if $a \preceq b$ and $b \preceq a$. By (i), $\equiv$ is an equivalence relation. Furthermore, by (ii), $\oplus$ and $\otimes$ respect this equivalence: if $a \equiv b$ and $a' \equiv b'$ then $a \oplus a' \equiv b \oplus b'$ and $a \otimes a' \equiv b \otimes b'$. This means that we can define a *quotient semiring* $\mathcal{A}/\equiv$ to be $\langle A/\equiv, \oplus, \otimes \rangle$, where $A/\equiv$ is the set $\{[a] : a \in A\}$ of $\equiv$-equivalence classes, and $[a]$ is the equivalence class containing a. Operations $\oplus$ and $\otimes$ on $A/\equiv$ are defined by $[a] \oplus [b] = [a \oplus b]$ and $[a] \otimes [b] = [a \otimes b]$ (these are well-defined because the operations respect equivalence). The order $\preceq$ associated with $\mathcal{A}/\equiv$ is a partial order.

3 Pointwise Valuations and Semiring Valuations

Let V be a finite set of variables. For each $X \in V$ let $\underline{X}$ be the domain (i.e., the set of possible assignments) of X. For $U \subseteq V$ let the set of partial tuples $\underline{U} = \prod_{X \in U} \underline{X}$ be the set of possible assignments to set of variables U. A complete assignment is an element of $\underline{V}$. For $u \in \underline{U}$ and $W \subseteq U$, $u^{\downarrow W}$ is the projection of u to variables W, so that for all $X \in W$, $u^{\downarrow W}(X) = u(X)$.

3.1 $\mathcal{A}$-Pointwise Valuations

Consider a set A with operations $\oplus$ and $\otimes$ on it, where $\oplus$ is assumed to be commutative and associative, and write $\mathcal{A}$ as the triple $\langle A, \oplus, \otimes \rangle$

An $\mathcal{A}$-valuation ϕ associates a value in A with each of a particular set of partial tuples: ϕ is a function from $\underline{V_\phi}$ to A, where set of variables $V_\phi \subseteq V$ is called the *scope* of ϕ.

Remark. If $u \in \underline{U}$ is an assignment to set of variables U containing V_ϕ, we may write $\phi(u)$ as an abbreviation for $\phi(u^{\downarrow V_\phi})$.

The operations in $\mathcal{A}$ allow one to define a combination and a marginalisation operation on $\mathcal{A}$-pointwise valuations over V. The combination $\phi \otimes \phi'$ of $\mathcal{A}$-valuations ϕ and ϕ' is defined to have scope $V_\phi \cup V_{\phi'}$, and by, $(\phi \otimes \phi')(u) = \phi(u) \otimes \phi'(u)$, i.e., $\phi(u^{\downarrow V_\phi}) \otimes \phi'(u^{\downarrow V_{\phi'}})$, for each assignment u of $V_\phi \cup V_{\phi'}$.

Let ϕ be an $\mathcal{A}$-valuation, and let U be a subset of V_ϕ. The marginalisation $\phi^{\downarrow U}$ of ϕ to U is defined by, for $u \in \underline{U}$, $\phi^{\downarrow U}(u) = \bigoplus \{c(w) : w \in \underline{V_\phi}, w^{\downarrow U} = u\}$.

Axioms for local computation The following axioms have been shown to be sufficient to perform local computations based on a join tree decomposition of valuations [10, 8, 7].

- **(LC1)** Combination of $\mathcal{A}$-valuations is commutative and associative.
- **(LC2)** Consonance/transitivity of Marginalisation: for any $\mathcal{A}$-valuation ϕ. if $T \subseteq U \subseteq V_\phi$ then $(\phi^{\downarrow U})^{\downarrow T} = \phi^{\downarrow T}$.
- **(LC3)** Distributivity of Marginalisation over Combination: $(\phi \otimes \psi)^{\downarrow T} = \phi \otimes \psi^{\downarrow V_\psi \cap T}$ for T, ϕ and ψ such that $V_\phi \subseteq T \subseteq V_\phi \cup V_\psi$.

We say that e.g., LC1 is satisfied for $\mathcal{A}$ if for any set of variables V then the combination of $\mathcal{A}$-valuations over V is commutative and associative.

Theorem 1 *For $\mathcal{A}$-valuations,*

(i) Combination of $\mathcal{A}$-valuations is commutative if and only if $\mathcal{A}$ operation $\otimes$ is commutative; combination of $\mathcal{A}$-valuations is associative if and only if $\otimes$ is associative; so axiom (LC1) is satisfied (for $\mathcal{A}$-valuations) if and only if the $\mathcal{A}$ operation $\otimes$ is commutative and associative;

(ii) Marginalisation is transitive, i.e., axiom (LC2) is always satisfied for $\mathcal{A}$-valuations;

(iii) Marginalisation is distributive over combination (axiom (LC3)) is satisfied if and only if $\mathcal{A}$ operation $\otimes$ distributes over $\oplus$.

This theorem shows that for pointwise valuations, that $\mathcal{A}$ is a semiring is a necessary and sufficient condition for the local computation axioms to be verified.

Many computational problems can be expressed as computing a marginalisation of a combination $\mathbf{C} = \bigotimes C$ of a (multi-)set C of semiring valuations. A straightforward approach to such a computation is exponential. However, the local computation approaches allow efficient computation under certain conditions.

3.2 Examples of semiring valuations

We list some examples of systems of semiring valuations.

CSPs. Finite CSPs can be expressed as a set of $\mathcal{A}$-valuations, where $\mathcal{A} = \langle \{0, 1\}, 0, 1, \max, \times \rangle$, so that $\otimes = \times$ (i.e., $\min$) and $\oplus$ is $\max$. A constraint on set of variables U is represented as a semiring valuation $\phi : \underline{U} \to \{0, 1\}$, where for assignment u of U, $\phi(u) = 1$ if and only if u satisfies the constraint. Let C be a set of $\mathcal{A}$-valuations representing a set of constraints, and let $\mathbf{C} = \bigotimes C$. Then, for assignment v of the set of all variables V, v is a solution of the CSP if and only if $\mathbf{C}(v) = 1$. Furthermore, there exists a solution if and only if $\mathbf{C}^{\downarrow \emptyset} = 1$, and there exists a solution with variable X assigned to x if and only if $\mathbf{C}^{\downarrow \{X\}}(x) = 1$.

Counting solutions of a CSP. Here we use semiring $\mathcal{A} = \langle I\!N \cup \{0\}, 0, 1, +, \times \rangle$, and use the same representation of constraints as given above, so we again have v is a solution of the CSP if and only if $\mathbf{C}(v) = 1$. But now $\mathbf{C}^{\downarrow \emptyset} = \sum_{v \in \underline{V}} \mathbf{C}(v)$ which is equal to the number of solutions of the CSP. Computing $\mathbf{C}^{\downarrow \{X\}}(x)$, for example, will give the number of solutions satisfying the assignment $X = x$.

Computation in Bayesian networks. Let $\mathcal{A} = \langle I\!R^+, 0, 1, +, \times \rangle$, where $I\!R^+$ is the set of non-negative real numbers. As is well known, computation in Bayesian networks [9, 10] can be performed by computing marginalisations of combinations of such $\mathcal{A}$-valuations, where each $\mathcal{A}$-valuation represents a conditional probability table of a variable given its parents in the Bayesian network. In particular the marginal probability that variable X is assigned to x is given by $\mathbf{C}^{\downarrow \{X\}}(x)$.

If we use the same $\mathcal{A}$-valuations as in the previous case, but redefine $\oplus$ to be $\max$ instead of $+$, we get, for example, that $\mathbf{C}^{\downarrow \emptyset}$ is equal to maximum probability of any complete assignment.

Weighted CSPs. Let $\mathcal{A} = \langle I\!N \cup \{0, \infty\}, \infty, 0, \min, + \rangle$, so that $\oplus$ is $\min$, $\otimes$ is $+$, $\mathbf{0}_\mathcal{A}$ equals ∞ and $\mathbf{1}_\mathcal{A}$ equals 0. $\mathcal{A}$-valuations then represent weighted constraints. $\mathbf{C}^{\downarrow \emptyset}$ gives the weight of the best solution. Optimal solutions $v \in \underline{V}$ are those with $\mathbf{C}(v) = \mathbf{C}^{\downarrow \emptyset}$. If we allow weights to be real valued, using $\mathcal{A} = \langle I\!R^+ \cup \{\infty\}, \infty, 0, \min, + \rangle$, then this system can be considered as equivalent to the previous one (maximum probability in Bayesian networks), by using a $(-)\log$ transformation.

Another related systems is generated by setting $\mathcal{A} = \langle I\!R \cup \{\infty\}, \infty, 0, \min, + \rangle$. The GAI utility decomposition of (Bacchus and Grove, 95) [1] can then be expressed as a multiset of such $\mathcal{A}$-valuations.

Semiring-based CSPs. This elegant framework [3, 4] generalises soft constraints formalisms such as weighted CSPs, fuzzy CSPs, probabilistic CSPs, lexicographic CSPs and set-based CSPs. A c-semiring $\mathcal{A} = \langle A, \mathbf{0}, \mathbf{1}, \oplus, \otimes \rangle$ is defined to be a semiring (as defined above) such that $\oplus$ is idempotent and $\mathbf{1}$ is an absorbing element of $\oplus$, i.e., for all $\alpha \in A$, $\alpha \oplus \mathbf{1} = \mathbf{1}$. $\mathcal{A}$-valuations then correspond precisely with (semiring) *constraints* in the associated semiring-based CSP [4, 3]. If C is a multiset of constraints, and U is a subset of the variables, then the *solution* of a *constraint problem* $\langle C, U \rangle$ is $(\bigotimes C)^{\downarrow U}$, the marginalisation of a multiset of $\mathcal{A}$-valuations.

4 Pre-processing by first propagating constraints

We use the term 'constraint' for a $\mathcal{A}$-valuation which takes values only in $\{\mathbf{0}, \mathbf{1}\}$. For $\mathcal{A}$-valuation ϕ, define $\phi^{\mathbf{1}}$ by $\phi^{\mathbf{1}}(s) = \mathbf{0}$ if $\phi(s) = \mathbf{0}$; otherwise $\phi^{\mathbf{1}}(s) = \mathbf{1}$.

Suppose $\phi = \psi_1 \otimes \cdots \otimes \psi_k$, Let $\psi_i^{\mathbf{1}}$ be the constraint associated with ψ_i, and let constraint ψ_i^* be $(\psi^*)^{\downarrow T_i}$, where $\psi^* = \bigotimes_i \psi_i^{\mathbf{1}}$, T_i is the scope of ψ_i, and the marginalisation is marginalisation of constraints. Furthermore, let θ_i^* be a constraint which is an upper bound for ψ_i^*, i.e., whenever θ_i^* is $\mathbf{0}$, ψ_i^* is also $\mathbf{0}$. Then

$$\phi = \bigotimes_{i=1}^{k} (\psi_i \otimes \theta_i^*).$$

This can be shown as follows. The right-hand-side of the displayed formula can be written as $\left(\bigotimes_i \psi_i\right) \otimes \bigotimes_i \theta_i^*$. For v such that $\left(\bigotimes_i \theta_i^*\right)(v) = \mathbf{0}$ there exists i with $\theta_i^*(v) = \mathbf{0}$, and so $\psi_i^*(v) = \mathbf{0}$. This implies that $\psi^*(v) = \mathbf{0}$, since ψ_i^* is a marginalisation of ψ^*. Then there exists some j with $\psi_j^{\mathbf{1}}(v) = \mathbf{0}$, and so $\psi_j(v) = \mathbf{0}$. This implies that $\phi(v) = \mathbf{0}$, and we also have $\left(\bigotimes_{i=1}^k (\psi_i \otimes \theta_i^*)\right)(v) = \mathbf{0}$.

On the other hand, consider v such that $\left(\bigotimes_i \theta_i^*\right)(v) \neq \mathbf{0}$, so equals $\mathbf{1}$. Then $\left(\bigotimes_{i=1}^k (\psi_i \otimes \theta_i^*)\right)(v) = \left(\bigotimes_{i=1}^k \psi_i\right)(v) \otimes \left(\bigotimes_i \theta_i^*\right)(v) = \left(\bigotimes_{i=1}^k \psi_i\right)(v) = \phi(v)$, as required.

The result that $\phi = \bigotimes_{i=1}^k (\psi_i \otimes \theta_i^*)$ means that we can use (sound but incomplete) methods for propagating constraints to simplify the valuation propagation. In particular we can use arc consistency propagation; or we could use the type of upper approximation described in the next section for difficult computations. After doing this, the remaining propagation may be much easier because $\psi_i \otimes \theta_i^*$ may have many fewer non-zero tuples; the efficiency of the combination operation is related to the number of non-zero tuples in the input valuations. This idea is related to the notion of *shrinking* in [2].

5 Use of Upper and Lower Bounds

The semiring relation $\preceq_{\mathcal{A}}$ enables us to define a relation on semiring valuations. We define relation $\preceq$ on $\mathcal{A}$-valuations, by $\phi \preceq \psi$ if ϕ and ψ have the same scope U and for all $u \in \underline{U}$, $\phi(u) \preceq_{\mathcal{A}} \psi(u)$.

Proposition 2 *Let $\mathcal{A} = \langle A, \mathbf{0}, \mathbf{1}, \oplus, \otimes \rangle$ be a semiring.*

(i) $\preceq$ on $\mathcal{A}$-valuations is reflexive and transitive and hence a pre-order. If $\preceq_{\mathcal{A}}$ is a partial order then so is $\preceq$.
(ii) If for all $i = 1, \ldots, k$, $\phi_i \preceq \psi_i$ then $\bigotimes_i \phi_i \preceq \bigotimes_i \psi_i$.
(iii) if $\phi \preceq \psi$, where both have scope S, then if $R \subseteq S$, $\phi^{\downarrow R} \preceq \psi^{\downarrow R}$.

The propagation algorithms involve sequences of combinations and marginalisations. The above results imply that if at any point we replace any $\mathcal{A}$-valuation by an $\preceq$-upper bound of it, the result will be an upper bound of the correct result. Similarly with lower bounds. (Even though $\preceq$ is not necessarily a partial order, but only a pre-order, we still say "ϕ is a lower bound of ψ", and "ψ is a lower bound of ϕ" when $\phi \preceq \psi$; similarly for $\preceq$ on A.)

If it is helpful computationally, we can replace semiring $\mathcal{A}$ by the quotient semiring $\mathcal{A}/\!\equiv$, replacing each $\mathcal{A}$-valuation by the corresponding $\mathcal{A}/\!\equiv$-valuation. This will lead to equivalent upper and lower bounds, since if if $a \equiv b$ then $a \preceq_{\mathcal{A}} c$ if and only if $b \preceq_{\mathcal{A}} c$, and $c \preceq_{\mathcal{A}} b$ if and only if $c \preceq_{\mathcal{A}} b$.

A valuable use of upper approximation is based on the following proposition:

Proposition 3 *For $i = 1, \ldots, k$, let ϕ_i have scope S_i, let $S = \bigcup_{i=1}^k S_i$, let R be a subset of S and define $R_i = R \cap S_i$. Then*

$$\Big(\bigotimes_i \phi_i\Big)^{\downarrow R} \preceq \bigotimes_i \phi_i^{\downarrow R_i}.$$

This result means we can use an upper approximation to reduce the number of variables involved in a combination, in particular, in computing a message in a join tree propagation. We might also choose the join tree to make use of this, to avoid performing any computation which is too big. If we want, for example, to avoid doing a combination involving more than r variables (for some given number r), we can generate a join tree which has no separator (i.e., the intersection of neighbouring nodes) which has more than r variables. Then, for any node which involves more than r variables, we use the above proposition to generate an upper approximation of each message.

The above algorithm generalises the upper bound obtained with the mini-buckets approach of (Dechter and Rish, 2003) [6] for the case of computing MPE (most probable explanation).

Computing upper approximations has an added benefit in situations where $\otimes$ is idempotent, as a pre-processing step. We can compute upper approximations for marginalisations, and then combine them with the input valuations. This new multiset of valuations will have the same combination as the original, but may be simpler.

When $\oplus$ is not idempotent we can compute a closer upper bound, which applies when one of the valuations, say ϕ_1, is such that its scope S_1 contains the set of variables $S - R$ being eliminated (this condition holds, in fact for all valuations, when we want to apply these bounds, e.g., in applying repeated elimination of variables in the fusion algorithm [11]). The following proposition gives a stronger result than previous one for this situation, since $\phi_i^{\downarrow R_i}$ satisfies the condition for θ_i, i.e., it is an upper bound for ϕ_i in the defined sense.

Proposition 4 *For $i = 1, \ldots, k$, let ϕ_i have scope S_i, let $S = \bigcup_{i=1}^k S_i$, let R be a subset of S and define $R_i = R \cap S_i$. Suppose also $S_1 \supseteq S - R$, and for all $i = 2, \ldots, k$, there exists semiring valuation θ_i with scope R_i which is an upper bound for ϕ_i, i.e., for all $w \in \underline{S_i}$, $\phi_i(w) \preceq \theta_i(w^{\downarrow R_i})$. Then*

$$\Big(\bigotimes_{i=1}^k \phi_i\Big)^{\downarrow R} \preceq \phi_1^{\downarrow R_1} \otimes \bigotimes_{i=2}^k \theta_i.$$

The same idea can be applied to give a lower bound, if instead the functions θ_is are all lower bounds rather than upper bounds.

In particular, when $\preceq_{\mathcal{A}}$ is a total order then we can define the upper bound functions θ_i using max, i.e., for $t \in \underline{S_i}$, $\theta_i(t) = \max\{\phi_i(w) : w^{\downarrow S_i} = t\}$, and the lower bound functions using min. When this is applied to belief updating for a Bayesian network, based on semiring $\mathcal{A} = \langle I\!R^+, 0, 1, +, \otimes \rangle$, this approximation then corresponds to the mini-buckets upper and lower bounds for belief updating [6].

Setting some semiring values to 0

The element $\mathbf{0}$ is a lower bound for every other element a in the semiring, since $\mathbf{0} \oplus a = a$. So a particular case of a lower bound is when we replace certain semiring values used in the input valuations to $\mathbf{0}$. This has computational advantages, as the efficiency of the computation is somewhat related to the number of non-zero values in the input valuations, and constraints propagation approaches can be used as discussed in section 4. We will consider the effect of choosing a subset P ($\not\ni \mathbf{0}$) of the semiring A, and replacing semiring values in the input valuations which are not in P by $\mathbf{0}$. With appropriate semiring and choice of P, it can be shown that this does not affect the answer to certain kinds of queries. This is related to the notion of *sinking* in [2].

Consider $\phi = \psi_1 \otimes \cdots \otimes \psi_k$. Define $M = \{\psi_i(u_i) : i = 1, \ldots, k, u_i \in \underline{V_{\psi_i}}\}$ to be the set of all semiring values taken by any of the input valuations. Define $M^{\otimes}$ to be closure of M under the $\otimes$ operation.

We consider input $\mathcal{A}$-valuations and set P satisfying the following condition:

$(*)$ If $a, b \in M^{\otimes}$ and $a \otimes b \in P$ then $a, b \in P$.

Condition $(*)$ implies that if elements a_i of $M^{\otimes}$, for $i = 1, \ldots, k$ are such that their combination $\bigotimes_{i=1}^{k} a_i$ is in P then a_i is in P for all $i = 1, \ldots, k$, . Hence we have for all $v \in \underline{V}$, if $\phi(v)$ is in P then $\psi_i(v)$ is in P for all $i = 1, \ldots, k$.

Condition $(*)$ is satisfied if M and P satisfy the pair of conditions (i) if $a \in M$ then $a \preceq \mathbf{1}$, and (ii) if $a \in P$ and $a \preceq b \preceq 1$ then $b \in P$.

An important particular case of this is when, for some $a \in A$, $P = P_a$ which is defined to be $\{b \in A : b \succeq a\}$. Condition $(*)$ is then satisfied as long as (i) is satisfied: the input semiring values are all bounded above by $\mathbf{1}$.

Consider P satisfying $(*)$. Define ψ_i^P by $\psi_i^P(u_i) = \psi_i(u_i)$ if $\psi_i(u_i) \in P$; otherwise $\psi_i^P(u_i) = \mathbf{0}$. Let $\phi' = \psi_1^P \otimes \cdots \otimes \psi_k^P$. By the above remarks, if $\phi(v) \in P$ then $\phi'(v) \in P$.

Suppose we are interested in finding complete assignments v whose combined semiring value is in P; for example, if we are only interested in v whose semiring value has a lower bound of a, we could use $P = P_a$. The above argument shows that we can use ϕ' instead of ϕ, without changing the result. This can sometimes greatly improve efficiency, as the components of ϕ' can be much smaller (i.e, have many fewer non-zero values) than those of ϕ.

A development of this idea is to define constraints ψ_i^* to be the constraint associated with ψ_i^P (see section 4) by $\psi_i^*(u_i) = \mathbf{1}$ if $\psi_i(u_i) \in P$; otherwise $\psi_i^P(u_i) = \mathbf{0}$. Then ψ_i^P can be decomposed as $\psi_i \otimes \psi_i^*$. We can then propagate these constraints, as in section 4, potentially simplifying the problem further.

Ensuring the equivalence $\phi(v) \in P \iff \phi'(v) \in P$. Condition $(*)$ ensures that the new combined valuation ϕ' will maintain any semiring values of complete assignment in P: if $\phi(v) \in P$ then $\phi'(v) \in P$. We can also consider sufficient conditions for the converse to hold: $\phi(v) \in P \iff \phi'(v) \in P$.

We consider semiring, valuations ϕ_i with associated $M^{\otimes}$ and set of semiring values P satisfying:

$(\dagger)$ If $a, b \in M^{\otimes}$ then $a \otimes b \in P$ if and only if $a, b \in P$.

This condition implies that if, for $i \in \{1, \ldots, k\}$, each a_i is an element of $M^{\otimes}$ then the following holds: the combination $\bigotimes_{i=1}^{k} a_i$ is in P if and only if each a_i is in P. Hence we have for all $v \in \underline{V}$, $\phi(v)$ is in P if and only if for all i, $\psi_i(v)$ is in P. Again let $\phi' = \psi_1^P \otimes \cdots \otimes \psi_k^P$. We have $\phi(v) \in P$ if and only if $\phi'(v) \in P$.

Condition $(\dagger)$ is satisfied, in particular, when (i) if $a \in M$ then $a \preceq \mathbf{1}$, (ii) there exists a_0 such that $P = P_{a_0}$, i.e., $\{a : a \succeq a_0\}$, and (iii) $\otimes$ is idempotent. This is because, $a, b \succeq a_0$ implies $a \otimes b \succeq a_0 \otimes a_0 = a_0$, so $a \otimes b \in P$, and hence P is closed under $\otimes$.

When $\oplus$ is $\max$

In this subsection we'll consider a special case of semiring: such that that for all $a, b \in A$, either $a \oplus b = a$ or $a \oplus b = b$. We call this the *addition-is-max* property. Any *valuation structure*, as defined in definition 11 of [4] gives rise to a semiring satisfying addition-is-max, by using the order relation to define $\oplus$. The semiring order $\preceq$ is then a total order. We will consider $P = P_a = \{b \in A : b \succeq a\}$. Given the input semiring values are all bounded above by $\mathbf{1}$ (as is always the case for valuations structures [12, 4]), condition $(*)$ then holds. So if $\phi(v) \in P$ then $\phi'(v) \in P$.

Furthermore, if $\phi^{\downarrow T}(t)$ is in P then there exists some $v \in \underline{V}$ extending t with $\phi(v) \in P$. Hence $\phi'(v) \in P$, and so $(\phi')^{\downarrow T}(t) \in P$. Therefore computing marginalisations of ϕ' will retain the semiring values of any tuples with values $\succeq a$. So if we want to compute projections of combinations of $\mathcal{A}$-valuations then we can use the reduced representation ϕ' (only keeping input semiring values $\succeq a$) if we are only interested in partial tuples with (output) semiring values in P.

6 Summary

Systems of semiring valuations include several of the most important systems of valuations, including CSPs, Bayesian networks and systems of soft CSPs. A natural pre-ordering on semiring valuations can be defined which is respected by combination and marginalisation, allowing methods for computing upper and lower bounds. This paper explores various ways of doing this, including a generalised form of the mini-buckets approach, where bounds are computed in such a way that the hardest computations are greatly simplified. Upper bounds can be especially useful when the multiplication is idempotent, as upper bounds can be combined with the input information as a pre-processing stage in an exact computation. We also show how constraints propagation can be used in a similar way as a pre-processing stage.

REFERENCES

[1] Bacchus, F., and Grove, A., Graphical models for preference and utility, *Proc. UAI'95*, 3–10 (1995).

[2] Bistarelli, S., Fung, S. K. L., Lee, J.H.M. Leung, H.F., A local search framework for semiring-based constraint satisfaction problems, in *Proc. CP2003 Workshop on Soft Constraints (Soft-2003)*, (2003).

[3] Bistarelli, S., Montanari, U., Rossi, F., Semiring-based Constraint Solving and Optimization, *Journal of the ACM*, vol. 44, no. 2, 201–236 (1997).

[4] Bistarelli, S., Montanari, U., Rossi, F., Schiex T., Verfaillie, G., and Fargier, H., Semiring-based CSPs and Valued CSPs: Frameworks, Properties and Comparison, *CONSTRAINTS: An international journal*, Vol.4, N.3, Kluwer (1999).

[5] Dechter, R., Bucket Elimination: A Unifying Framework for Reasoning, *Artificial Intelligence*, 113, 41–85 (1999).

[6] Dechter, R., and Rish, I., Mini-Buckets: A General Scheme for Approximating Inference, To be published in *Journal of the ACM* (2003).

[7] Kohlas, J., *Information Algebras*, Springer, London 2003.

[8] Kohlas, J., and Shenoy, P., Computation in Valuation Algebras. *Volume 5: Algorithms for Uncertainty and Defeasible Reasoning* J. Kohlas, S. Moral (eds.), *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, (Series eds.: D. Gabbay, P. Smets), Kluwer Academic Publishers, 5–39 (2000).

[9] Pearl, Judea, 88a, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers Inc. (1988).

[10] Shenoy, P. P., and Shafer, G., Axioms for Probability and Belief Function Propagation, in *Uncertainty in Artificial Intelligence 4*, R. D. Shachter, T. S. Levitt, L. N. Kanal, J. F. Lemmer (eds.), North Holland, also in *Readings in Uncertain Reasoning*, Shafer, G., and Pearl, J., (eds.) Morgan Kaufmann, San Mateo, California, 575–610 (1990).

[11] Shenoy, P. P., Valuation-Based Systems: A Framework for Managing Uncertainty in Expert Systems, pages 83–104, in *Fuzzy Logic for the Management of Uncertainty*, Zadeh, L.A., & Kacprzyk, J., (eds.), John Wiley & Sons (1992).

[12] Schiex T., Fargier, H., and Verfaillie, G., Valued constraint satisfaction problems: hard and easy problems, *Proc. IJCAI-95*, Montreal, Quebec, 631–637, (1995).